Comparative seasonal abundance and diversity of populations of the *Pseudomonas syringae* and Soft Rot *Pectobacteriaceae* species complexes throughout the Durance River catchment from its French Alps sources to its delta

Morris, C.E.[1*], Lacroix, C.[1], Chandeysson, C.[1], Guilbaud, C.[1], Monteil, C.[1], Piry, S.[1], Rochelle Newall E.[2], Fiorini S.[3], Van Gijsegem, F.[2], Barny, M.A.[2], Berge, O.[1]

[1]INRAE, Pathologie Végétale, F-84140, Montfavet, France

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

[3]Ecole Normale Supérieure, PSL Research University, CNRS, Centre de recherche en écologie expérimentale et prédictive (CEREEP-Ecotron Ile De France), UAR 3194, 11 chemin de Busseau, F-77140 Saint-Pierre-lès-Nemours, France

*Corresponding author Cindy E. Morris, INRAE, UR0407 Pathologie Végétale, 67 allée des Chênes, 84140 Montfavet, France, tel: +33 04 32 72 28 86, cindy.morris@inrae.fr

ABSTRACT

Flowing surface waters (rivers, creeks, streams) are integrators of biological, chemical and physical processes occurring in a catchment and they link catchment land cover from the headwaters to the outlet. The dynamics of human and animal pathogens in catchments have been widely studied in a large variety of contexts allowing the optimization of disease risk reduction. In parallel, there is an emerging awareness that crop pathogens might also be disseminated via surface waters especially when they are used for irrigation. However, there are no studies on the extent to which potential plant pathogens are present – nor about their dynamics - along the full course of a catchment. We do not know if they are confined to regions close to crops or if they are present throughout a catchment. Here we have compared the seasonal dynamics of populations of the *Pseudomonas syringae* (Psy) and the Soft Rot *Pectobacteriaceae* (SRP) species complexes along a 270 km stretch of the Durance River from the upstream alpine reaches to the downstream agricultural production areas at the confluence with the Rhone River at Avignon. Among 168 samples collected at 21 sites in fall, winter, spring and summer of 2016 and 2017, Psy strains were detected at all sampling sites and in 156 of the samples at population densities up to 10^5 bacteria L^{-1}. In contrast, SRP strains were detected in 98 of the samples, mostly from the southern part of the river, at population densities that did not exceed 3×10^4 bacteria L^{-1}. Among the aquatic parameters that were characterized at each sampling site (total culturable bacteria, temperature, conductivity, concentrations of dissolved organic carbon (DOC), PO_4^{3-}, NH_4^+, NO_2^- and NO_3^-), temperature was the only factor that explained a significant amount of the variability in population size for both species complexes. Psy densities decreased with increasing temperature whereas SRP densities increased with increasing temperature. River-borne populations of SRP were composed mainly of *Pectobacterium versatile* and *P. aquaticum* that have little known epidemiological importance. Only a few strains of *Pectobacterium* and *Dickeya* species reputed for their epidemiological impact were observed. In contrast, Psy populations at all sites were dominated by a genetic lineage of phylogroup 2 known from other studies for its broad host range and its geographic and habitat ubiquity. These results suggest that strains of SRP with pathogenic potential to plants have lower competence for saprophytic survival (in freshwaters) than do potentially pathogenic strains of Psy and that their presence in river water is probably much more dependent on specific vegetative sources than are pathogenic strains of Psy. Nevertheless, their vegetative sources have not yet been identified. We discuss how to incorporate knowledge of the abundance and diversity of these two groups of plant pathogens in river water into a strategy for anticipating risk for disease outbreaks on crops in a catchment.

INTRODUCTION

Surface waters are vital components of agro-systems. They provide water for irrigation and industrial processing of foods as well as being important for other uses such as for drinking, generation of electricity, recreation and navigation. Lakes and rivers are defining features of landscape topography and they influence the fertility and humidity of soils in their proximity. Flowing surface waters (rivers, creeks, streams) are physical links between agricultural production fields and other land covers both up- and down-stream as they transport particles and various chemicals that enter rivers along their paths from source to sink. In this light, the paucity of information about plant pathogens in surface waters is remarkable. The available data on plant pathogens in surface waters concern some bacterial plant pathogens (24), the so-called water molds and a few other fungi published prior to a 20-year hiatus in studies of the aquatic phases of plant pathogen ecology (53).

The paucity of information about plant pathogens in surface waters is in stark contrast with the abundance of reports on the presence and dynamics of human and animal pathogens in these habitats (51). The abundant data on human and animal pathogens in river catchments has led to models of the dynamics of the populations of these microorganisms along the course of rivers. These models are used to assess where water poses risks for human health and where and when to optimize interventions to reduce these risks (46; 51). With sufficient knowledge bases, similar applications would be possible for plant health including assessing where use of river water for irrigation poses the greatest risk of plant diseases, conceiving interventions to reduce the risk, and orienting surveillance of river water quality to validate the efficiency of those interventions. Nevertheless, there are additional challenges to assessing the sources and dynamics of plant pathogens in river water. Firstly, there is the daunting question of where to start given that the list of plant pathogens possibly in river water is very long. Secondly, due to the saprophytic capacities of many plant pathogens, there are likely to be multiple, diffuse sources of these microorganisms in landscapes rather than discrete sources that can be surmised. Therefore, data are needed all along the course of a river to give models the power to infer sources. Finally, as for human and animal pathogens, data are needed on the regular occurrence of plant pathogens in river water to assure that observations are not anecdotal and that the organism is sufficiently frequent to foster the modeling of its dynamics. In this light, the *Pseudomonas syringae* (Psy) and the Soft Rot *Pectobacteriaceae* (SRPs) species complexes stand out in terms of multiple previous reports of their presence in surface waters (10; 11; 17; 29; 31; 36; 39; 41).

52 Here, we have mapped the abundance of two groups of plant pathogenic bacteria in a 270 km stretch of the Durance River,
53 several tributaries and a canal in Southern France. Situated in a Mediterranean fruit and vegetable production region, the
54 Durance River drains over 14000 km^2 of which 20% is agricultural production (2). This river has been exploited since the 1100's
55 for irrigation, milling, navigation, drinking water, mining of sediment, generation of electricity and recreation. This has involved
56 the creation of canals and dams, restructuration of banks and dredging of sediments leading to changes in flow rates (2). Land
57 use and ground cover in the Durance River catchment are influenced by the topography of the basin with recreation, pastures
58 and nature reserves mostly in the mountainous zone from its source to the Lake Serre Ponçon reservoir (that retains 1.2 billion
59 m^3 and is the second largest reservoir in Europe). Downstream of the lake crop cultivation and large urban zones dominate (2).
60 This river basin is in a typically Mediterranean region. Therefore, it is subjected to the vicissitudes of climate leading to
61 landslides, flooding and droughts that alter the flow and particle content of the river and that complement the seasonal water
62 discharge dynamics that are mostly influenced by snowmelt.
63 The objective of this work was to compare the abundance and reoccurrence across seasons of two groups of plant pathogenic
64 bacteria – the *Pseudomonas syringae* and the Soft Rot *Pectobacteriaceae* (SRPs) species complexes - along the stretch of the
65 Durance River from alpine regions to the agricultural production region where the Durance joins the Rhone River at Avignon.
66 These two plant species complexes are classified within different orders of the gamma-Proteobacteria, the Pseudomonadale
67 order for *P. syringae* and the Enterobacterale order for the SRP. Furthermore, they differ in the mechanisms by which they cause
68 disease. While the SRP secrete a large cocktail of plant cell wall degrading enzyme to destroy the plant cell and recover nutrients,
69 *P. syringae*'s main virulence weapon is a type III secretion system known to inject a battery of effector proteins into plant cells
70 that collectively allow suppression of plant defenses and gain of access to nutrients (27; 41). *Pseudomonas syringae* is a species
71 complex composed of numerous phylogroups (PG) and clades (5) with a few having recognized taxonomic status as species. In
72 studies that quantify its abundance in the environment (30; 37; 42; 48) members of this complex are identified based on
73 phylogenetic affiliation according to a partial nucleotide sequence of the citrate synthase housekeeping gene (*cts*). Sequence
74 analysis based on this portion of the *cts* gene allows strain identification and placement in the context of the phylogeny that
75 accounts for the broadest scope of genetic diversity of this group (5). Strains in the *P. syringae* group are present in fresh waters
76 and have been isolated previously from sources and tributaries of the Durance River (37). However, their abundance along the
77 full course of the Durance River and across seasons has not been assessed. Species of the SRP complex can be quantitatively

78 isolated from environmental sources on a medium that reveals their capacity to degrade pectin (3) and they can be identified
79 based on phylogenetic affiliation according to partial sequences of the housekeeping gene glyceraldehyde-3-phosphate
80 dehydrogenase A (*gapA*) (9). The occurrence of species representing the SRPs throughout the Durance River has been reported
81 recently but not quantitatively (3). Nevertheless, this first report suggests that, despite the capacity of SRPs to proliferate as a
82 saprophyte on decaying plant material making it likely for them to be ubiquitous in rivers (17; 21; 41; 44), the SRPs seem to
83 be markedly different in their population dynamics in river water compared to the ubiquitous *P. syringae* group. Here, we have
84 compared the spatial and temporal dynamics of the populations of these two groups of bacteria to identify the environmental
85 factors and adaptive features that could distinguish them in terms of their capacity to establish reservoirs in river water and
86 especially in rivers used for irrigation of crops.
87
88 **RESULTS**
89 **Populations of *P. syringae* and SRP species complexes are present throughout the Durance River catchment but differ**
90 **in size and frequency of occurrence**
91 Bacterial population sizes were evaluated for 21 sites throughout the Durance River catchment (Tab. 1, Fig. 1). Strains in the *P.*
92 *syringae* complex (referred to collectively from here on as Psy) were detected at all sampling sites and almost all dates
93 throughout the catchment at population densities up to 10^5 bacteria L^{-1} (Fig. 1). Population densities of this bacterial group
94 were below the detection threshold (10 - 40 bacteria L^{-1}) in only 12 (7%) of the 168 water samples analyzed in this study. In
95 contrast, members of the SRP species complex (referred to collectively from here on as SRP) were less frequently detected than
96 Psy and were most often detected at the sites in the southern-most end of the catchment but rarely in the northernmost reaches
97 of the catchment. SRP population densities were under the detection threshold in 70 (42%) of the samples. SRP and Psy co-
98 occurred in 87 (52%) of the samples. When there was co-occurrence of SRP and Psy, SRP population densities were equal to or
99 exceeded those of Psy in 25% (22) of those samples by up to about one order of magnitude; Psy population densities exceeded
100 those of SRP in 75% (65) of those samples by up to nearly four orders of magnitude (Supp. Fig. 1).
101 Among the different sampling dates and sites, total culturable populations ranged from 10^5 to 5×10^7 bacteria L^{-1} (Supp. Tab.
102 1). Although there was an overall positive trend in the correlation between the densities of total culturable bacterial populations
103 and those of Psy or SRP, the statistical significance (at the 5% level) of the correlation depended on the geographic situation

104 according to the three basins of the catchment (delimited in Fig. 1). In the upper, northernmost basin, densities of Psy and SRP
105 were each significantly correlated with total population densities (R = 0.275 and 0.222 for Psy and SRP respectively; p = 0.009
106 and 0.038, respectively). In the southernmost, lower basin the population densities of neither bacterial group were significantly
107 correlated with total population density (R = 0.187 and 0.052 for Psy and SRP respectively; p = 0.304 and 0.779, respectively).
108 The middle basin differentiated Psy from SRP where total population densities were significantly correlated with Psy densities
109 (R = 0.340, p = 0.018) but not with SRP densities (R = 0.245, p = 0.094). These results suggest that the factors that influence
110 the densities of Psy and SRP populations are likely to be somewhat different from each other and not completely correlated
111 with the factors influencing the abundance of total culturable bacteria.
112
113 **Among variables describing the physical-chemical conditions of river water, temperature has the greatest predictive**
114 **power for population sizes of Psy and SRP with inverse effects on these two species complexes.**
115 Seven variables describing the physical-chemical characteristics of the water at each sampling time and according to the
116 geographical context of the site (altitude, longitude and latitude) were measured. Temperature, conductivity and dissolved
117 organic carbon (DOC) concentration were measured in 2016 and 2017 and, in addition, concentrations of PO_4^{3-}, NH_4^+, NO_2^- and
118 NO_3^- were determined in 2017. The ranges of values for these variables are presented in Fig. 2 and are indicative of an alpine
119 catchment with increasing influence of human activities and agriculture as altitude decreases. Variables for water conditions
120 assessed in 2017 had varying degrees of inter-correlations (Tab. 2). Therefore, to assess the influence of the ensemble of the
121 physical-chemical properties on bacterial population size, we used Principle Component Analysis (PCA) to construct composite
122 factors that accounted for the importance of each of the seven individual physical-chemical variables determined in 2017 for
123 the overall variability of water conditions (Tab 3).
124 The PCA led to the construction of seven composite factors (*F*1 – *F*7) for the 79 observations in 2017, each based on a complete
125 set of observations for all water variables. Water temperature contributed ca. 10% to 30% of the variability of six of the factors
126 and the other water variables contributed to the same extent of variability for four or fewer of the factors (Tab. 3A). A multiple
127 regression of the population sizes of either Psy, SRP or total mesophilic bacteria against all seven factors revealed a significant
128 contribution of *F*1, *F*2 and *F*4 to the variability of Psy population sizes; a significant contribution of *F*1, *F*3 and *F*6 to the variability
129 of SRP population sizes; and a significant contribution of *F*1 and *F*2 to the variability of total bacterial population sizes

130 throughout the catchment and across seasons in 2017 (Tab. 4). To identify the water variables that contributed the most to the
131 variability of Psy, SRP and total bacterial populations, we ranked the contribution of each water variable for each *F* (Tab 3A) and
132 calculated the cumulative contribution to the variability of each *F* with decreasing rank. For *F*1, *F*3, *F*4 and *F*6, three water
133 variables explained at least 80% of the variability of the factors; and for F2, four water variables explained at least 80% of the
134 variability of each factor (here, we refer to these as the "top explanatory variables"). Among the top explanatory variables, only
135 temperature was common to all the *F* that had significant effects in the regressions (i.e., *F*1, *F*2, *F*3, *F*4 and *F*6). For the *F* that
136 were significant for Psy populations, conductivity and the concentration of NH_4^+ were common to two factors (*F*2, *F*4); likewise
137 for SRP the concentration of NO_3^- was common to two factors (*F*1, *F*6). Otherwise, there were no other top explanatory variables
138 that were consistently common to the significant *F* factors.
139 In light of the dominant correlation of water temperature with the size of Psy, SRP and total bacterial populations, we
140 determined to what extent this variable alone could explain the variability in population size. Simple linear regressions of
141 bacterial population size vs temperature revealed that temperature alone significantly explained about 20-40% of the variability
142 of the population sizes of Psy and SRP in 2017 ($R^2 = 0.187$, $p = 0.000$ for Psy; $R^2 = 0.393$, $p = 0.000$ for SRP). When both
143 2016 and 2017 were considered together, temperature explained about 25-35% of the variability of the sizes of these bacterial
144 populations ($R^2 = 0.249$, $p = 0.000$ for Psy; $R^2 = 0.360$, $p = 0.000$ for SRP). In contrast, temperature alone had no significant
145 explanatory power for the variability of total bacterial population sizes in 2017 ($R^2 = 0.018$, $p = 0.223$) and explained only 5%
146 of the variability of total bacterial populations when 2016 and 2017 sampling campaigns were considered together ($R^2 =$
147 0.050, $p = 0.004$).
148 Whereas temperature explained about the same amount of variability of Psy and SRP population sizes, it had inverse effects on
149 population size (Fig. 2). For both Psy and SRP, a change of 10°C was associated with roughly a change in population size by a
150 factor of 10. In the case of Psy, populations increased with decreasing temperature; in the case of SRP, populations decreased
151 with decreasing temperature within the range of temperatures observed in this study. The regression for SRP population sizes
152 predicts that populations would be below the detection level when water temperatures are less than 7°C (Fig. 2). In this study,
153 there were 55 observations where water temperature was colder than 7°C. For these 55 cases, SRP populations were below the
154 detection level for 37 cases whereas Psy populations were detected for all of these cases.

155 The overriding correlation of temperature with densities of Psy and SRP populations might be due in part to the effect of the
156 wide range of temperatures that are accounted for when data were pooled from sites across the three basins (Fig. 1) spanning
157 altitudes from 39 m to 2090 m (Tab. 1). Pooling data from sampling sites across the three basins might also mask local effects
158 of other water variables that are affected by increasing anthropogenic activities along the land use from the source to the delta
159 of the Durance River catchment. Therefore, we assessed the correlations of bacterial populations with water variables for each
160 of the three basins separately (basin attribution is indicated in Tab. 1). The mean temperature of the water in the upper basin
161 during the sampling campaigns was about 7 °C cooler than that of the middle and lower basins (Fig. 3). Nevertheless,
162 population densities of SRP were positively and significantly ($p < 0.05$) correlated with temperature in each of the three basins
163 (Fig. 3). Likewise, population densities of Psy were negatively correlated with temperature in each of the three basins;
164 correlations were significant at the 5% level for the upper and middle basins and at the 10% level for the lower basin. Among
165 the top explanatory variables identified above via PCA, NH_4^+ concentrations were significantly and positively correlated with
166 Psy population densities in the middle and lower basins in spite of the similar concentrations of this compound across the three
167 basins (Fig.3). Although NO_3^- concentrations were identified in PCA as one of the top explanatory water variables for SRP
168 densities, there were no significant correlations at the 5% level in any of the three basins. In the PCA, neither conductivity nor
169 DOC were identified as important explanatory factors for the variability of bacterial populations. Nevertheless, conductivity was
170 positively correlated ($p < 0.05$) with densities of Psy in the upper basin and in the lower basin. These basins were markedly
171 different in the range of conductivity values observed. In the middle basin, where the range of conductivity was similar to that
172 of the lower basin, this variable was negatively correlated ($p < 0.05$) with SRP densities but had no significant correlation with
173 Psy densities. DOC was positively correlated ($p < 0.05$) with Psy densities in the upper basin and with SRP densities in the
174 middle basin but not elsewhere.
175 The correlations of total population sizes with water conditions were in marked contrast to those for Psy and SRP populations.
176 When assessed according to the individual basins, there were no consistent correlations of total populations with any water
177 variables in a basin with the exception of NO_3^- concentrations (Fig. 3). Total bacterial population densities were positively
178 correlated with NO_3^- concentrations in both the upper and lower basins.
179

Populations of *P. syringae* and SRP species complexes in the Durance River catchment are composed of both pandemic and endemic genotypes representing bacterial groups with and without known pathogenic potential

As previously reported (3), the 582 SRP strains isolated from the catchment and identified at species level based on MLST analysis were composed of *Pectobacterium* (94% of SRP strains from the Durance catchment) and *Dickeya* (6%) species. *Pectobacterium* populations were dominated by species that have no reported epidemiological importance including *P. versatile* (known to be associated with a wide range of plants) and *P. aquaticum* (not known to be pathogenic) constituting 47% and 40%, respectively, of the *Pectobacterium* strains isolated (3). In contrast, important *Pectobacterium* pathogens described on crop such as *P. atrosepticum* or *P. brasiliense* were rarely detected or absent. For *Dickeya* populations (6% of the SRP population), *D. oryzae* (pathogenic mainly on monocots but also on potato) constituted 72% of the *Dickeya* isolates. Among the few *Dickeya* strains, all belong to species of known epidemiological importance including *D. oryzae*, *D. fangzhongdai*, *D. solani*, *D. dianthicola*, and *D. dadantii*.(3).

For Psy, identification was based on phylogroups (PG) and haplotypes within PGs. For these PGs and haplotypes we could then associate them with likely epidemiological behaviors based on previous descriptions. Phylogenetic characterization was conducted for 5436 colonies isolated here that were putative Psy. For these colonies, based on criteria described in material and methods, 2628 could be attributed to known phylogroups of Psy based on comparison with a 388 bp segment of the *cts* gene for 910 strains in the reference data set, and were used to study Psy diversity. The strains that were not attributed to known phylogroups of *P. syringae* might indeed be within the *P. syringae* complex but they were not included in the analyses here because of current taxonomic uncertainties. Strains were identified as belonging to PG01 (9.34 % of all strains), PG02 (45.04 %), PG03 (0.16 %), PG04 (1.30 %), PG07 (13.68 %), PG09 (8.00 %), PG10 (11.60 %), PG12 (0.05 %), PG13 (10.12 %) and PG15 (0.57 %). Unlike *SRP* species where only *P. versatile* was distributed throughout the catchment and other species were mostly in the southern part of the catchment (3), six PGs of Psy (PG01, 02, 7, 9, 10 and 13) were detected at 19-21 of the 21 sampling sites. The other PGs that each constituted only about 1% or less of the Psy population were found at fewer sites (at 13 sites for PG04 and PG015; two sites for PG03 and PG12).

The 2628 strains of Psy represented 291 different *cts* sequences (referred to here as haplotypes). Nearly half (128) of these haplotypes appeared endemic and were found at only one of the 21 sampling sites (Fig. 4). Nevertheless, these rare haplotypes only accounted for 5% (154 strains) of the 2628 strains assigned to known PG. Overall, 18 haplotypes accounted for 50% of

206 these strains and each were detected at 15 or more sampling sites. Among these haplotypes, one (referred to here as DD.1) was
207 detected at all 21 sampling sites and represented 10% of all of the strains attributed to known PG in this study. Among
208 individual samples, the fraction of the total population of Psy that was constituted by DD.1 was very consistent and showed a
209 strong positive correlation between the size of the Psy population and that of DD.1 (Spearman Rank correlation coefficient =
210 0.917, p = 0.000).
211 We found that haplotype DD.1 corresponds to a *cts* haplotype of PG02 (in the PG02b clade) that is very widespread when
212 compared to previous reports (35; 37). This *cts* haplotype was the same (100% identity of the 388 bp segment. The *cts* sequence
213 is listed among the amplicon sequence variants in Supp. Tab. 2) as the dominant Psy haplotype found among the 236 strains
214 isolated from river headwaters sampled in the US (Montana and Wyoming), Europe (mostly France and Italy) and New Zealand,
215 representing 39 of the water strains and being the only haplotype found on all three continents and at 11 of the 13 sites
216 sampled (37). When compared to reference strains in the study of headwaters by Morris and colleagues (37), the haplotype
217 that dominated in headwaters is also the same as the haplotype of 15 reference strains from crops used in that study. These
218 included the type strain of *P. syringae* pv. *syringae* (CFBP1392) isolated from lilac in the UK in 1950 (14), strains 601 and
219 CFBP1906 of *P. syringae* pv. *aptata* isolated from sugar beet in Japan in 1966 (47) and in France in 1979 (16), respectively, and
220 strains involved in an epidemic of bacterial blight of cantaloupe that was described to have emerged in France as of 1993 (34)
221 (strains CC0001, CC0037, CC0125, CC0354, CC0440, CC0441, CC0457). Additional strains in this haplotype were identified
222 in a subsequent study on host range of Psy (35) showing that strains with the same *cts* haplotype as DD.1 were involved in
223 diseases of apricot in France (strain 41A, according to strain names indicated by Morris and colleagues (35)), of cantaloupe and
224 squash in New Zealand (CFBP 1788, ICMP 3390, ICMP 7501), and of sugar beets in the Netherlands (CFBP 2471 and CFBP
225 2507) and Serbia (P004 – P102). By comparing the 388 bp *cts* sequence of haplotype DD.1 to the full GenBank database (BLAST
226 search, https://www.ncbi.nlm.nih.gov/) we also found 100% identity with strains from a freshwater lake in Virginia (strains
227 CLC07, CLC10 (42)), from freshly fallen snow collected at a high altitude meteorological observatory in Switzerland (JFJ-0007,
228 JFJ-0043 (48)), from blighted leaves of pea in Japan (H5E3 (47)) and from home garden philodendrons (several IZB1 and IZB2
229 strains (20)) as documented examples that expand the sites and substrates of isolation compared to the information described
230 earlier in this paragraph.
231

232 **Discussion**

233 Although there have been previous reports of *P. syringae* and Soft Rot *Pectobacteriaceae* complexes in surface waters (37; 49)
234 here we have made comparisons of their prevalence and abundance across a range of sites representing the diverse
235 environmental conditions across a catchment. This comparison is intended to lead us to identify what typifies each system and
236 what trends are shared. In our effort to make quantitative comparisons of the two groups of bacteria, we faced an initial
237 challenge due to the difference in magnitude of the number of isolates we could collect for each group of bacteria. In a previous
238 study, we verified that differences in abundance were not due to differences in isolation efficiency of the two media used in this
239 work (40). Whereas the hundreds of SRP collected could be characterized as individual strains (sequencing of single or multiple
240 housekeeping genes or full genomes), the thousands of Psy isolates encountered led us to adopt a high-throughput MiSeq
241 sequencing strategy of amplicons of a single housekeeping gene on the basis of Psy-like colonies isolated on KBC growing
242 media. By overcoming this technical challenge, we have shown that for both bacterial species complexes there are genetically
243 diverse populations present throughout the full expanse of the Durance River catchment from near its source - across varying
244 topography, altitude and land use - to the delta where it converges with the Rhone River. Among the SRP, *Pectobacterium* spp.
245 were the most frequently encountered representatives with *P. versatile* being present throughout the catchment (3). Psy
246 populations were dominated by a *cts* haplotype that accounts for 10% of populations at all sites and sampling dates. The
247 structure of Psy populations in the Durance River catchment is similar to that of *Listeria monocytogenes* in surface waters (lakes,
248 rivers, ponds) along the Central California Coast that is dominated overall by a clonal line that constituted 27% of the 1200
249 strains isolated from these waters (15). However, the specific clonal dominating *L. monocytogenes* populations in waters or
250 other environmental reservoirs differed among the different geographic locations studied (15). In contrast, we observed that
251 Psy populations in distant geographical locations are dominated by the same haplotype that is dominating the Durance River
252 populations overall (37), thus illustrating the efficiency of Psy dissemination and the overriding capacity for the DD.1 haplotype
253 to dominate Psy populations.

254 Populations of Psy were detectable at nearly all sites and all seasons during the two years of sampling whereas SRP populations
255 were frequently below the detection threshold and especially in the upper basin of the catchment where they could not be
256 detected in more than half of the samples. In light of the marked saprophytic capacity of many of the soft rotting bacteria (41)
257 and numerous reports of their presence in surface waters (49) it could be considered surprising that we did not detect SRP more

258 frequently than Psy. However, our observations suggest that temperature adaptation has a critical role in the ecology of the Psy
259 and SRP species complexes. Our observations also reinforce the idea that Psy is particularly well adapted to freshwater habitats
260 as well as the various other habitats and substrates (plants, precipitation, litter) that are linked via the dissemination of *P.*
261 *syringae* through the hydrological cycle (36) . In contrast, SRP are likely to be more dependent on proximity of and seasonality
262 of external plant sources (49). In comparison, for *Listeria monocytogenes* that has saprophytic as well as human pathogenic
263 potential (15), the relative importance of anthropogenic *vs.* natural sources for populations in rivers is unknown. Its saprophytic
264 capacity could allow for establishment of "natural" reservoirs but it could also be leaked into rivers from anthropogenic sources.
265 Both cases beckon the need for further research to find environmental sources to improve the understanding of disease
266 epidemiology.

267 Although it is likely that there is run-off of these two groups of bacteria into the Durance River from vegetation, we observed
268 that water temperature is strongly correlated with the densities of the populations of Psy and SRP in river water: population
269 densities were positively correlated with temperature for SRP and negatively correlated with temperature for Psy. Temperature
270 was correlated with altitude of the site – and this could reflect differences in land use and vegetation type along the banks of
271 the rivers. Nevertheless, the trend with temperature observed throughout the entire Durance catchment was also observed
272 within each of the three basins (upper, middle and lower) when considered separately. This further strengthens the hypothesis
273 that temperature is a critical factor and not simply a reflection of its correlation with other gradients across the entire catchment.
274 Temperature seems to influence the abundance of these two groups of bacteria whatever the context of the basin and the
275 associated sources of bacteria that the river encounters as it crosses different land uses from pastures, to fruit tree production
276 and to vegetable crops with their varying anthropogenic characteristics. The influence of temperature on Psy and SRP appears
277 to be much stronger than on the total culturable bacterial population. This is probably due to differences among the component
278 species in their sensitivity to environmental factors. When they were detected, Psy populations constituted only 10^{-5} % to less
279 than 4% of the total bacterial population and SRP constituted only 10^{-5} % to less than 0.7% of the total population. Therefore, it
280 is reasonable to assume that the perceptible effect of environmental factors on total populations is strongly influenced by the
281 major species components rather than by Psy and SRP. Indeed, the bacterial assemblages in Durance river water are highly
282 diverse when assessed by metagenomics that target 16S rDNA (40). The diversity of total bacteria suggests that there is not only
283 a range of environmental tolerances among the bacteria in the river system that makes it difficult to identify overriding

correlations with environmental factors, but that there are also opportunities for competition and antagonistic interactions. However, when there were significant correlations between densities of total bacterial populations and those of Psy or SRP, they were positive suggesting that increasing densities of bacteria that co-occur with Psy or SRP in river water were not detrimental to the populations of these latter two groups of bacteria. Such positive correlations between Psy and total bacteria in water were also observed in a previous study in the Durance catchment (31). Furthermore, a metagenomic analysis of samples from three sites along the Durance catchment representing the upper-, mid- and lower basins (40) showed the same trend for the *Pseudomonas* genus as we observed for the *P. syringae* complex suggesting that *P. syringae* might be representative of the genus as a whole in terms of its population dynamics in river water.

A remarkable observation for both Psy and SRP is that river water contains a diversity of populations of these groups of bacteria beyond what is known to be associated with disease on crops. This raises the intriguing question of the origin of these bacteria in river water. For example, river water harbors genetic groups of Psy and SRP with no known epidemiological importance - PG10 and PG13 for Psy (5) and *P. aquaticum* and *P. quasiaquaticum* for SRP (4; 39; 43). Some strains of Psy PG10 and PG13 have been found in association with plants (6), but the vast part of their diversity has been found in water elsewhere (5). The recently discovered *P. aquaticum* and *P. quasiaquaticum* found in the Durance catchment have been reported only from aquatic environments. Perhaps it is autochthonous in water but the low prevalence of SRP in water suggests it has other yet-to-be-discovered habitats that serve as sources for populations in the river. Notably, *P. aquat*icum was found mostly in the lower half of the Durance catchment with an important occurrence on a limited number of sites suggesting its requirement for either very specific conditions or its association with a limited number of sources. This is in contrast to *P. versatile*, known to be associated with a wide range of plant species (crops and ornamentals, for example), that was detected throughout the catchment.

River water also contained strains that are likely to be of epidemiological importance – but not necessarily in the Durance River catchment or on crops. We detected *P. peruviense* (3), a species that has only been reported at high altitudes in South America as a pathogen of potato (50). Its presence in the Durance River suggests that it has a natural but previously unknown ubiquity in the environment or that there was a rare and unrecognized dissemination event from South America. The presence of a few strains of *P. atrosepticum*, a species mostly recorded on potato (3), also raises the question of its origin – either from disease on the very small surface of potato crops in the Durance catchment or the association of this bacterium with wild solanaceous plants or a few brassicas (49). Other plant sources are unlikely for this species in light of its very narrow host range (28; 49).

310 Furthermore, we wonder if the presence in river water of *D. oryzae* (3), known to be a pathogen of rice but also pathogenic on
311 potato, maize and several other crops (19), might have its origin in disease on the small amount of regional potato crop or an
312 association with wild grasses that has yet to be described. For Psy, its population is dominated by a ubiquitous subgroup of
313 PG02 (here named *cts*-haplotype DD.1) that has been associated with numerous crop disease epidemics. Nevertheless, the
314 epidemics linked to the DD.1 haplotype of Psy have occurred mostly outside of the Durance River basin with the closest known
315 epidemic in Southwestern France (34). For another haplotype in PG02 closely related to DD.1 , quasi-clonal lines from
316 epidemics of cantaloupe blight in southwestern France, from snowfall in the French Vercors Massif and from a pristine creek on
317 the south island of New Zealand have been identified (32) supporting the hypothesis that long distance movement of bacteria
318 – even between the northern and southern hemispheres – does occur. Hence, there are indeed mechanisms for long distance
319 movement - most likely via the atmosphere - that can link rivers with cropped fields elsewhere.

320 To understand the potential epidemiological significance of the presence of diverse Psy and SRP throughout the Durance
321 catchment, we need to identify the processes that have contributed to this state of the microbiology of river water. We wonder
322 if the assemblages of Psy and SRP populations in river water are the result of rivers being simply collectors of bacteria from the
323 local landscape (from run-off, for example) and from more distant sources (via rain and snowfall, for example). If run-off is the
324 main process leading to the abundance and diversity of Psy and SRP in river water, it would be very important to identify all of
325 the potential sources including prairies, pastures and wild plant stands in addition to known crop hosts for disease. It is also
326 important to consider if Psy and SRP simply survive or if there is multiplication and diversification. Interestingly, pathogenicity
327 tests with *D. dianthicola*-like strains isolated from river water in Finland revealed that water-borne strains were more aggressive
328 than strains of *D. dianthicola* isolated from potato (25). Furthermore, *D. aquatica* isolated in Finnish rivers were later found to
329 be aggressive on acidic fruits such as tomato or cucumber (10). The lack of xylanases and xylose degradation pathways in *D.*
330 *aquatica* could reflect adaptation to aquatic charophyte hosts which, in contrast to land plants, do not contain xyloglucans. This
331 suggests that water-borne species have experienced some selective pressures that lead to adaptations that could, in turn, be
332 useful in causing disease to crops.

333 Our results point to the need to clarify the role of temperature in influencing population densities of Psy and SRP. The
334 differential effect of temperature on population sizes of Psy and SRP could be due to effects on growth and/or die-off – both
335 processes being important in structuring the gene pool of these populations. In laboratory tests of growth of SRP strains

336 inoculated into filter- and autoclave-sterilized river water, 100-fold increases in populations of *Dickeya* and *Pectobacterium*
337 strains were observed over 10 days at 20°c. However, at 8°C growth was lower for *Pectobacterium* and die-off was observed for
338 *Dickeya* (3). This implies that, under natural conditions, there are stresses caused by a fluctuating environment that maintain
339 SRP populations at low levels. However, the nature of these stresses is not clear from our work. For Psy, a previous study
340 suggested that populations in rivers did not necessarily multiply (31). The authors of that study noted that the similarities in
341 population structure between rain, snow melt and headwaters in France could be attributed to effective transportation of Psy
342 strains with snow melt and rain water infiltrating through the soil of subalpine grasslands. However, in a study of headwaters
343 in France, the USA and New Zealand it was observed that about half of the populations at the headwater sites were composed
344 of *cts* haplotypes that were unique to the region from which they were sampled (43% for New Zealand headwaters, 67% for USA
345 headwaters and 70% for French headwaters) (37), implying the existence of a local diversification process. Preliminary
346 laboratory experiments show that growth in river water is possible (Berge, unpublished) thereby suggesting that this could
347 contribute to diversification.

348 A critical epidemiological aspect of the regular occurrence of Psy and SRP in the Durance River catchment is the potential of
349 river-borne bacteria to cause disease to crops irrigated with river water. This concern brings to the forefront the questions of
350 how to assess the epidemiological potential of river-borne bacteria and how to anticipate disease outbreaks. The
351 epidemiological potential of Psy and SRP strains in river water could be addressed via pathogenicity tests such as those
352 conducted for the *D. dianthicola*-like and *D. aquatica* strains isolated from river water in Finland (25). However, in the case of
353 Psy, the choice of pertinent hosts to test against strains in the dominant DD.1 haplotype is complicated by its variable and
354 potentially broad host range (35). To anticipate disease outbreaks, data on epidemiological potential needs to be set in the
355 context of rate of exposure of crop plants and the local environmental conditions. Exposure of plants due to irrigation with river
356 water could be estimated. For the main departments of France that irrigate with water from the Durance River catchment (Alpes
357 de Haute Provence, Hautes Alpes and Vaucluse) there are > 55000 ha of agriculture that could be irrigated including fruits and
358 vegetables, pastures and cereals (8). In the case of lettuce – a vegetable crop produced in abundance in the Durance River
359 catchment - plants require a total of about 30 mm (30 L m^{-2}) in the few days following planting (about a week) and about 5 or
360 6 subsequent irrigations of about 15 mm each during the 2 to 3 months of culture afterwards for a total of about 110 mm (110
361 L m^{-2}) during the life cycle of the plant (26). For a lettuce field planted at a density of 150000 plants ha^{-1} and a mean bacterial

362 population density of 10^3 Psy or SRP L^{-1} of river water, each plant on average could potentially be in contact with more than 10^3
363 bacteria belonging to the Psy or SRP species complex from water during the first few days after planting and about 10^4 Psy or
364 SRP bacteria throughout the period of culture due to irrigation. At a first glance, this might seem to be cause for alarm. However,
365 the fate of these bacteria is unknown. We do not know if they survive, if they are physiologically competent, or if they are
366 compatible with the crop hosts they encounter. We should keep in mind that only a fraction of the Psy or SRP strains – or perhaps
367 none at all - that contact the plants via irrigation will have pathogenic potential for the crop that they encounter. Overall, this
368 speculation points out why it is important to have quantitative data on bacterial population size that allows for estimations of
369 exposure - and that go beyond the uncertainties of risk assessment based simply on presence and prevalence.

370 Historical epidemiological information for the Durance River catchment does not point to Psy and SRP as re-occurring pathogens
371 of crops in this region - with the exception of bacterial canker of apricot (38) and leaf spot of lettuce (1), both caused by Psy. This
372 could suggest that the range of environmental conditions, the historical land use and the intensity of agriculture up to present
373 are within the spectrum of conditions that do not generally favor epidemics by Psy or SRP. It could also suggest that the bacteria
374 present in the two species complexes studied here are not well adapted to the cultivated crops cultivated in the Durance River
375 catchment. In light of these observations, we can make recommendations for a first approach to developing indicators to survey
376 for estimating the risk posed by river-borne populations of Psy and SRP. These indicators would account for total population
377 sizes of Psy and SRP (or the dominant genetic lines), for water temperature and for the various chemical conditions in each of
378 the three basins that we determined to be correlated with Psy and SRP population densities. Risk alerts could be developed to
379 express the deviation from the trends we observed here. Disease risk could also be evaluated for any major changes in
380 agricultural land use that can be anticipated - such as changes in the geographic ranges of certain crops or the introduction of
381 new crops into the region. This would require assessment of their sensitivity to diseases caused by Psy and SRP in the
382 environmental conditions associated with the anticipated changes.

383 This work raises the general question of how river water reflects the diversity of plant-associated microorganisms beyond what
384 is reflected by populations associated with crops or other vegetation. We raise the critical question of how this diversity be used
385 for anticipating disease emergence and the need to elucidate the underlying processes that connect these populations to
386 epidemics. These processes would indeed be targets for management. Integrating nonagricultural reservoirs of plant
387 pathogens – such as river water – into a more comprehensive vision of pathogen ecology and life history could improve

388 forecasting disease risk and anticipating epidemics in the face of changes in land use and climate. Although various bacteria
389 and fungi have been detected in irrigation water (24; 53), some of these might be present only when inoculum reservoirs in
390 diseased plants are nearby and they might not be able to persist in river water. To develop such an integrative approach, it will
391 be important to distinguish pathogens with the capacity to thrive in environmental reservoirs vs. those whose presence in the
392 environment represents transient residues from agriculture. This approach would open new directions in disease surveillance
393 that would allow for anticipation on larger scales of space and time and could foster better adaptation of land use in the face of
394 changing climate.

395

396 **Experimental Procedures**

397 ***Sample collection and handling***. Water was collected from 21 sites representing the three hydrological sections of the Durance
398 catchment (23), of which 8 were along the main course of the river, 11 were from 9 different tributaries that flow into the main
399 river, and 2 were from a major man-made canal (Tab. 1). This canal is a managed distributary of the Durance River and its
400 floodway inlet is located in Mallemort, FR (43.73267° N, 5.18599° E). Water was collected at each site at 8 dates to represent
401 four seasons across two years. Sampling campaigns were conducted in 2016 on 1-17 Feb., 13-19 May, 24-28 Aug. and 18-21
402 Nov; and in 2017 on 3-8 Feb., 4-6 June, 21-25 Aug. and 8-13 Nov. These dates will be referred to, respectively, as Winter-16,
403 Spring-16, Summer-16, Fall-16, Winter-17, Spring-17, Summer-17 and Fall-17.
404 Surface water was collected at several meters distance from the banks at each site with a 12-L bucket attached to a rope. For all
405 sampling dates, each site was represented by a single bulk sample, resulting in 168 water samples (hereafter referred to as
406 "main experiment"). In Fall-17, triplicate samples at three different times in the day were collected at sites R02 and R08 to
407 assess representativeness of the bulk samples (hereafter referred to as "variability experiment"). The bucket was rinsed twice
408 with water from the sample site before each sample was collected. About 1.5 L of water was collected into sterile plastic bottles
409 from the bucket. With the water remaining in the bucket, temperature and electrical conductivity were measured using a Multi
410 Probe System (YSI 556 MPS, YSI, Yellow Springs, USA) and water turbidity was measured using a EUTECH Instruments turbidity
411 meter (TN100, Paisley, Scotland). Samples were maintained in a cooler (ca. 15°C) for no more than 24 h until further processing
412 for chemical and microbiological analyses. To prepare samples for microbiological analyses, 500 mL were filtered across 0.2
413 µm porosity cellulose acetate filters (Sartorius, 11107-47-ACN, Goettingen, Germany). The bacteria retained on the filter were

414 suspended in 1 mL of sterile distilled water. This suspension, concentrated by a factor of 500 compared to the original sample,
415 was immediately used for subsequent bacterial isolation and quantification. The filtrate was collected for nutrient and dissolved
416 organic carbon analysis as detailed elsewhere (3). Methods for determination of the concentration of DOC, nitrates, nitrites,
417 ammonium, ortho-phosphates, and total dissolved nitrogen and phosphorus were as previously described (3).
418
419 ***Quantification of total culturable bacteria.*** The concentrated suspension was dilution-plated on 10% tryptic soy agar (37).
420 Plates were incubated at ambient temperature (18 to 25°C) for 2 to 4 days when total plate counts were recorded.
421
422 ***Isolation and quantification of Soft Rot Pectobacteriaceae***. The bacterial suspensions were serially diluted in water and
423 plated on crystal violet pectate (CVP) medium plates, a semi-selective medium containing pectin that is widely used for the
424 isolation of pectinolytic *Pectobacterium* and *Dickeya* (13; 18). Plates were incubated at 28°C for 2 days and the number of
425 colonies forming deep pits in the CVP medium typical of *Pectobacteriaceae* were recorded. For each treated sample, up to 30
426 pit-forming colonies were purified on CVP medium and further streaked on LB medium for conservation. Qualitative description
427 of the purified strains has been published recently (3). In the present paper, we evaluated the quantity of recovered SRP by
428 counting the deep pits formed on plates and analyzed these data with regard to other variables measured in the course of this
429 study.
430
431 ***Isolation and quantification of bacteria in the P. syringae complex***. The concentrated suspension was dilution-plated as
432 previously described (37) on King's medium B supplemented with cephalexin, boric acid and cycloheximide (referred to as KBC
433 medium). Two to three replicates of each dilution were plated to assure that when possible at least 30 colonies suspected to be
434 *P. syringae* ("putative" *P. syringae* based on colony traits) could be isolated for each site at each date. After 3 to 7 days incubation
435 of KBC plates at room temperature (~20-25°C) the numbers of putative *P. syringae* colonies were recorded. Based on our
436 previous work with the diversity of the *P. syringae* complex (5), few phenotypic traits are reliable for screening colonies to hone
437 in specifically on *P. syringae*. Non-putative colonies were eliminated according to pigmentation, pin-point colony size and
438 ornate or crusty colony appearance. From the remaining, 30 or more putative colonies (or all putative colonies if there were
439 fewer than 30) were randomly selected for each sample and streaked onto a plate of King's medium B (22) to increase the

440 number of bacterial cells per colony for further characterization but not to purify strains. Each isolate was then introduced into
441 the well of 96-well plate (i.e. initial plates) previously loaded with 150 µl of demineralized water kept at 4°C until being
442 identified on the basis of high-throughput MiSeq sequencing of a fragment of the *cts* (citrate synthase) gene and bioinformatic
443 analysis (described below). The sizes of the populations of *P. syringae* in each water sample were calculated by adjusting the
444 number of putative *P. syringae* colonies per each sample according to the percentage that were identified as *bona fide* members
445 of the *P. syringae* complex (5) through the MiSeq sequencing approach (cf. Supp. Tab. 1 and 3.).
446
447 ***Preparation of the material for MiSeq sequencing including, putative* P. syringae *isolates, replicates and controls.***
448 Among the 5436 isolates introduced into the well of 96-well plates (i.e. initial plates) a total of 537 isolates were introduced
449 twice, i.e. into the well of two different 96-well PCR plates, which were further analyzed as replicates to assess the reproducibility
450 of bacterial identification through MiSeq sequencing approach. In addition, each of the initial 96-well PCR plate contained at
451 least one replicate of five different types of controls. Pure colonies of a known *P. syringae* (CC94, phylogroup 02 [PG02]) and
452 *Pseudomonas tolaasi* (CFBP2068) strains were separately introduced into 80 wells of the initial plates (i.e. 160 wells total).
453 Respectively, these two types of positive controls were used to assess the efficiency and repeatability of *P. syringae* identification,
454 and the level of biological or sequencing/bioinformatics contamination across wells. Three types of negative controls were also
455 included in the analysis. A total of 80 and 414 wells were filled only with ultrapure water or PCR mix, respectively (see below),
456 while 142 wells were left empty during the *cts* PCR amplification (described below).
457
458 ***Amplicon production, preparation of MiSeq libraries and sequencing.*** The resulting 6769 wells were subjected to PCR
459 amplification targeting a 388 pb fragment (primers excluded) of the *cts* (citrate synthase) gene. PCR was performed on groups
460 of four plates, using one of twenty specifically designed forward primers, and a single common reverse primer (cf. Supp. Tab.
461 4). Each forward and reverse primers was composed of the binding site for *cts* gene amplification, and of an adapter used in the
462 further steps of MiSeq library preparation for adding the Illumina indexes. In addition, each forward primer included a different
463 6 nucleotide tag in order to be able to assign output sequences to each initial bacterial isolate during the bioinformatics
464 analyses. The PCR mix was composed of 3.4 µl of 5X Q5 reaction buffer (New England Biolabs), 0.14 µl of 25 mM dNTPs
465 (Promega), 0.68 µl of 10µM of each of the forward and reverse primer, 0.17 µl of Q5 hot start HF DNA polymerase (New England

466 Biolabs), and was adjusted to a final volume of 15 µl with ultrapure water, to which 2µl of the initial material (i.e. isolates,
467 replicates or controls) was added. After a 30 min 98°C activation period of the Taq polymerase, DNA fragments were amplified
468 following 30 cycles of denaturation (10 s at 98°C), annealing (30 s at 55°C) and extension (30 s at 72°C). After a final extension
469 time of 2 min at 72°C, DNA amplicons were stored at –20°C until use. After PCR amplification, four groups of 20 PCR plates
470 were pooled each into a single plate. To this end, 5 µl of the identical well of each of 20 plates were mixed together. The four
471 resulting plates (i.e. pooled plates) were sent to GeT-PlaGe core facility (INRAE, Toulouse, France) where the final MiSeq libraries
472 were prepared. The libraries were run using an Illumina MiSeq pair-end 2*250 pb sequencing technology.
473
474 ***Bio-informatic processing of raw sequences.*** Raw sequences were processed in R (version 4.0.3) (45) with the FASTA program
475 version 36.3.8h (https://fasta.bioch.virginia.edu/fasta_www2/fasta_list2.shtml), and the packages ShortRead, DADA2 and
476 ggplot2 (7; 33; 52). The amplicon and MiSeq library preparation strategy resulted in both forward and reverse reads being
477 present in the R1 and R2 files associated with each well of the pooled plates. Reads pairs were labelled as forward and reverse
478 complement based on the comparison of their sequence with the one of a reference *P.syringae* strain (CVB0016, phylogroup
479 02 [PG02]; Supp. Tab. 5). Then, read pairs with e-value >= 10^{-40} were removed. Paired reads were sorted such as to include
480 forward and reverse complement reads in final R1 and R2 fastq files, respectively. All reads were demultiplexed using the tags
481 included in the forward primers (cf. Supp. Tab. 4) in order to separate and assign raw sequences to each well of the initial plates.
482 Only sequences that contained exactly matching tags were kept. Reads that were too short, relatively to the required
483 length for merging R1 and R2 reads with an adequate overlap (i.e. 25), were removed. Then, tags and *cts* primer
484 sequences were removed. The quality of sorted and demultiplexed reads was checked and plotted. Reads were not trimmed as
485 the observed error rates were similar to the estimated ones, and as the expected overlap length between the paired forward
486 and reverse reads was relatively short (i.e. 25). Reads that included at least one unidentified nucleotide or which sequence
487 matched the phiX genome were discarded. Then, the Amplicon Sequence Variants (ASVs) were inferred, paired reads were
488 merged, the sequence table was constructed and chimeras were removed.
489
490 ***Identification of*** Amplicon Sequence Variants (***ASVs)***
491 ASVs were identified through a blastn+ with the sequence of 910 reference, mostly *P. syringae*, strains (cf. Supp. Tab. 5) using
492 the FROGS Affiliation OTU (12) available on the Genotoul-Sigenae Galaxy server (https://vm-galaxy-prod.toulouse.inra.fr/). The

493 ASVs which percentage identity with the closest reference strains was lower than 98.2% were removed. This value corresponds
494 to the similarity threshold determined previously for accurate clade affiliations within the *P. syringae* species complex (5).
495
496 ***Verification of controls, filters of ASV, and analysis of replicates***
497 A sequence of the expected strain (either *P. syringae* CC94 or *P. tolaasi* CFBP 2068) was identified in 156 out of the 160 wells
498 corresponding to positive controls. No sequence was detected in 3 positive control wells. Some ASVs were identified non-
499 expectedly in one positive control well, and 54 negative control wells, with copies number ranging from 1 to 389. Therefore, a
500 conservative approach was taken whereby each ASV was considered as positively detected in the wells of the initial plates, and
501 thus assigned to the corresponding 5436 isolates, if the copy number was higher than 400. This resulted in the final
502 identification of 291 *P. syringae* ASV (Supp. Tab 2), which copy number ranged from 400 to 11055 in each initial well where it
503 was detected (hereafter referred as haplotypes). Out of the 537 isolates that were included as duplicates in the initial plates,
504 392 (73%) yielded in similar results. Specifically, no *P. syringae* haplotype was detected for each of the duplicates of 154
505 isolates, and a sequence identified as being a *bona fide* member of the *P. syringae* species complex was detected for the
506 duplicates of 238 isolates. The remaining 145 isolates corresponded to cases where a *P. syringae* haplotype was detected for
507 one duplicate but not in the other. Hence, the MiSeq isolate identification approach described here might have led to an under-
508 estimation of the number of *P. syringae* colonies in water samples, but could not have led to the wrong identification (i.e. false
509 positive) of a colony as *bona fide* member of the *P. syringae* species complex.
510
511 ***Statistical analyses.***

512 The representativeness of single bulk water samples was determined through the analysis of the variability in total Psy and SRP
513 population sizes based on triplicates of water samples taken at three different times of the same day in Fall 2017 at two sites
514 (R02 and R08, Supp. Tab. 3). The analysis was conducted for each site in R (version 4.1.1) with a linear regression using the
515 $\log_{10}$ transformed total number of Psy and SRP colonies per L of water as response variable. The assumption of normality was
516 verified using a Shapiro test ($p > 0.62$ for site R02 and $p = 0.67$ for site R08 for Psy; and $p > 0.36$ for site R02 and $p = 0.49$ for
517 site R08 for SRP). The variability in population size for both organisms across sampling times within a day (Supp. Fig. 2) was

non-significant for both sites (p=0.32 for site R02, p = 0.2 for site R08 for Psy; and p=0.19 for site R02, p = 0.32 for site R08 for SRP).

Statistical analyses of total *P. syringae* population size across the Durance catchment (cf. Supp. Tab. 1) and of associated water chemical characteristics (cf. Supp. Tab. 6) were conducted with modules in the Statistica 10 package (StatSoft www.statsoft.fr, accessed 27 Aug 2019). This included the characterization of water variables via Principle Component Analysis leading to the construction of composite Principle Component Factors and calculation of the correlations among water variables according to Spearman's Rank correlation. This statistical package was also used to calculate parameters of regressions of the observed values of bacterial population sizes at each site and date (expressed as $\log_{10}$ bacteria L^{-1}) against single water variables or composite Principle Component factors. Significant effects were reported if p-values were < 0.05.

Acknowledgments

This work was possible due to funding contract ANR-17-CE32-0004 for the project on *Strategic preemptive pathogen surveillance of air and water to anticipate plant disease emergence in scenarios of changing land use.* We thank Karine Berthier (INRAE-PACA-Pathologie Végétale) for valuable insights on bioinformatics analyses. We thank Ariane Toussaint for her interest in the project and her logistic support in Val-des-Prés during the sampling campaigns in the upper Durance. We thank the GeT-PlaGe core facility (INRAE, Toulouse, France) for their collaboration, and especially Olivier Bouchez, Arien Castinel, Marie Gislard and Lisa Gil.

References

1. Allex D, Rat B. 1990. Les bactérioses des salades : un problème omniprésent. *PHM-Rev. Hort.* 310:45-50.
2. Andrew JT, Sauquet E. 2017. Climate change impacts and water management adaptation in two Mediterranean-climate watersheds: Learning from the Durance and Sacramento Rivers. *Water* 9:doi: 10.3390/w9020126
3. Ben Moussa H, Bertrand C, Rochelle-Newall E, Fiorini S, Pédron J, Barny MA. 2022. The diversity of soft rot Pectobacteriaceae along the Durance River stream in the south-east of France revealed by multiple seasonal surveys. *Phytopathology* DOI: 10.1094/PHYTO-12-21-0515-R
4. Ben Moussa H, Pédron J, Bertrand C, Hecquet A, Barny M-A. 2021. *Pectobacterium quasiaquaticum* sp. nov., isolated from waterways. *International Journal of Systematic and Evolutionary Microbiology* 71
5. Berge O, Monteil CL, Bartoli C, Chandeysson C, Guilbaud C, et al. 2014. A user's guide to a data base of the diversity of *Pseudomonas syringae* and its application to classifying strains in this phylogenetic complex. *PLoS ONE* 9:(9): e105547. doi:10.1371/journal.pone.010554

549 6. Borschinger B, Bartoli C, Chandeysson C, Guilbaud C, Parisi L, et al. 2015. A set of PCRs for rapid identification and
550 characterization of *Pseudomonas syringae* phylogroups. *Journal of Applied Microbiology* 120:714-23
551 7. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: High-resolution sample
552 inference from Illumina amplicon data. *Nature Methods* 13:581-3
553 8. Chambre_d'Agriculture_PACA. 2014. SRHA Provence Alpes Côte d'Azur: Stratégie Régionale Hydraulique Agricole.
554 https://paca.chambres-agriculture.fr/fileadmin/user_upload/National/FAL_commun/publications/Provence-Alpes-
555 Cote_d_Azur/rdiagnostic-agriculture_irriguee_paca_2014.pdf (accessed 1 April 2022)
556 9. Cigna J, Dewaegeneire P, Beury A, Gobert V, Faure D. 2017. A *gapA* PCR-sequencing assay for identifying the *Dickeya*
557 and *Pectobacterium* potato pathogens. *Plant Disease* 101:1278-82
558 10. Duprey A, Taib N, Leonard S, Garin T, Flandrois J-P, et al. 2019. The phytopathogenic nature of *Dickeya aquatica* 174/2
559 and the dynamic early evolution of *Dickeya* pathogenicity. *Environmental Microbiology* 21:2809-35
560 11. Eayre CG, Bartz JA, Concelmo DE. 1995. Bacteriophages of *Erwinia carotovora* and *Erwinia ananas* isolated from
561 freshwater lakes. *Plant Disease,* 79:801-4.
562 12. Escudié F, Auer L, Bernard M, Mariadassou M, Cauquil L, et al. 2018. FROGS: Find, Rapidly, OTUs with Galaxy Solution.
563 *Bioinformatics* 34:1287-94
564 13. Faye P, Bertrand C, Pédron J, Barny M-A. 2018. Draft genomes of "*Pectobacterium peruviense*" strains isolated from
565 fresh water in France. *Standards in Genomic Sciences* 13:27
566 14. Gardan L, Cottin S, Bollet C, Hunault G. 1991. Phenotypic heterogeneity of *Pseudomonas syringae* van Hall. *Res.*
567 *Microbiol.* 142:995-1003.
568 15. Gorski L, Cooley MB, Oryang D, Carychao D, Nguyen K, et al. 2022. Prevalence and clonal diversity of over 1,200
569 *Listeria monocytogenes* isolates collected from public access waters near produce production areas on the Central
570 California Coast during 2011 to 2016. *Appl Environ Microbiol* 88:e0035722
571 16. Guillorit-Rondeau C, Malandrin L, Samson R. 1996. Identification of two serological flagellar types (H1 and H2) in
572 *Pseudomonas syringae* pathovars. *EuropeanJournal of Plant Pathology* 102:99-110
573 17. Harrison M, Franc G, Maddox D, Michaud J, Mccarter-Zorner N. 1987. Presence of *Erwinia carotovora* in surface water
574 in North America. *Journal of Applied Bacteriology* 62:565-70
575 18. Hélias V, Hamon P, Huchet E, van der Wolf J, Andrivon D. 2012. Two new effective semiselective crystal violet pectate
576 media for isolation of *Pectobacterium* and *Dickeya*. *Plant Pathology* 61:339-45
577 19. Hugouvieux-Cotte-Pattat N, Van Gijsegem F. 2021. Diversity within the *Dickeya zeae* complex, identification of *Dickeya*
578 *zeae* and *Dickeya oryzae* members, proposal of the novel species *Dickeya parazeae* sp. nov. *International Journal of*
579 *Systematic and Evolutionary Microbiology* 71
580 20. Ivanović Ž, Blagojević J, Nikolić I. 2018. Leaf spot disease on *Philodendron scandens, Ficus carica* and *Actinidia*
581 *deliciosa* caused by *Pseudomonas syringae* pv. *syringae* in Serbia. *European Journal of Plant Pathology* doi:
582 10.1007/s10658-018-1437-4
583 21. Jorge PE, Harrison MD. 1986. The association of *Erwinia carotovora* with surface water in northeastern Colorado. I. The
584 presence and population of the bacterium in relation to location, season and water temperature. *American Potato*
585 *Journal* 63:517–31
586 22. King EO, Ward MK, Raney DE. 1954. Two simple media for the demonstration of pyocyanin and fluorescein. *J. Lab. &*
587 *Clin. Med.* 44:301-7
588 23. Kuentz A. 2013. *Un siècle de variabilité hydro-climatique sur le bassin de la Durance : Recherches historiques et*
589 *reconstitutions*. AgroParisTech, Paris. 374 pp.
590 24. Lamichhane JR, Bartoli C. 2015. Plant pathogenic bacteria in open irrigation systems: what risk for crop health? *Plant*
591 *Pathology* 64:757-66
592 25. Laurila J, Ahola V, Lehtinen A, Joutsjoki T, Hannukkala A, et al. 2008. Characterization of *Dickeya* strains isolated from
593 potato and river water samples in Finland. *European Journal of Plant Pathology* 122:213-25
594 26. Lecompte F. 2012. Management of soil nitrate heterogeneity resulting from crop rows in a lettuce–tomato rotation
595 under a greenhouse. *Agronomy for Sustainable Development* 32:811-9
596 27. Lindeberg M, Cunnac S, Collmer A. 2012. *Pseudomonas syringae* type III effector repertoires: last words in endless
597 arguments. *Trends in Microbiology* 20:199-208
598 28. Ma B, Hibbing ME, Kim HS, Reedy RM, Yedidia I, et al. 2007. Host range and molecular phylogenies of the soft rot
599 enterobacterial genera *Pectobacterium* and *Dickeya*. *Phytopathology* 97:1150-63

29. McCarter-Zorner NJ, Franc GD, Harrison MD, Michaud JE, Quinn CE. 1984. Soft rot Erwinia bacteria in surface and underground waters in southen Scotland and in Colorado, United-States. *J. Appl. Bacteriol.* 57:95-105
30. Monteil CL, Bardin M, Morris CE. 2014. Features of air masses associated with the deposition of Pseudomonas syringae and Botrytis cinerea by rain and snowfall. *ISME J* 8:2290-304
31. Monteil CL, Lafolie F, Laurent J, Clement J-C, Simler R, et al. 2013. Soil water flow is a source of the plant pathogen *Pseudomonas syringae* in subalpine headwaters. . *Environ. Microbiol.* 16:2038–52
32. Monteil CL, Yahara K, Studholme DJ, Mageiros L, Méric G, et al. 2016. Population-genomic insights into emergence, crop-adaptation, and dissemination of *Pseudomonas syringae* pathogens. *Microbial Genomics* doi: 10.1099/mgen.0.000089
33. Morgan M, Anders S, Lawrence M, Aboyoun P, Pagès H, Gentleman R. 2009. ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics* 25:2607-8
34. Morris CE, Glaux C, Latour X, Gardan L, Samson R, Pitrat M. 2000. The relationship of host range, physiology, and genotype to virulence on cantaloupe in *Pseudomonas syringae* from cantaloupe blight epidemics in France. *Phytopathology* 90:636-46
35. Morris CE, Lamichhane JR, Nikolić I, Stanković S, Moury B. 2019. The overlapping continuum of host range among strains in the *Pseudomonas syringae* complex. *Phytopathology Research* 1:4
36. Morris CE, Monteil CL, Berge O. 2013. The life history of *Pseudomonas syringae*: linking agriculture to Earth system processes. *Annu. Rev. Phytopathol.* 51:85-104
37. Morris CE, Sands DC, Vanneste JL, Montarry J, Oakley B, et al. 2010. Inferring the evolutionary history of the plant pathogen *Pseudomonas syringae* from its biogeography in headwaters of rivers in North America, Europe and New Zealand. *mBio* 1(3): e00107-10-e00107-20
38. Parisi L, Morgaint B, Blanco-Garcia J, Guilbaud C, Chandeysson C, et al. 2019. Bacteria from four phylogroups of the *Pseudomonas syringae* complex can cause bacterial canker of apricot. *Plant Pathology* 68:1249-58
39. Pédron J, Bertrand C, Taghouti G, Portier P, Barny M-A. 2019. *Pectobacterium aquaticum* sp. nov., isolated from waterways. *International Journal of Systematic and Evolutionary Microbiology* 69:745-51
40. Pédron J, Guyon L, Lecomte A, Blottière L, Chandeysson C, 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
41. Pérombelon MCM, Kelman A. 1980. Ecology of the soft rot Erwinias. *Ann. Rev. Phytopathol.* 18:361-87
42. Pietsch RB, Vinatzer BA, Schmale DG. 2017. Diversity and abundance of ice nucleating strains of *Pseudomonas syringae* in a freshwater lake in Virginia, USA. *Frontiers in Microbiology* 8
43. Portier P, Pédron J, Taghouti G, Dutrieux C, 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
44. Potrykus M, Golanowska M, Sledz W, Zoledowska S, Motyka A, et al. 2015. Biodiversity of *Dickeya* spp. isolated from potato plants and water sources in temperate climate. *Plant Disease* 100:408-17
45. R_Core_Team. 2020. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria.* https://www.R-project.org/
46. Rankinen K, Butterfield D, Faneca Sànchez M, Grizzetti B, Whitehead P, et al. 2016. The INCA-Pathogens model: An application to the Loimijoki River basin in Finland. *Science of The Total Environment* 572:1611-21
47. Sarkar SF, Guttman DS. 2004. Evolution of the core genome of *Pseudomonas syringae*, a highly clonal, endemic plant pathogen. *Appl. Environ. Microbiol.* 70:1999-2012
48. Stopelli E, Conen F, Guilbaud C, Zopfi J, Alewell C, Morris CE. 2017. Ice nucleators, bacterial cells and *Pseudomonas syringae* in precipitation at Jungfraujoch. *Biogeosciences* 14:1189-96
49. Toth IK, Barny M-a, Brurberg MB, Condemine G, Czajkowski R, et al. 2021. *Pectobacterium* and *Dickeya*: Environment to Disease Development. In *Plant Diseases Caused by Dickeya and Pectobacterium Species*, ed. F Van Gijsegem, JM van der Wolf, IK Toth:39-84. Cham: Springer International Publishing. Number of 39-84 pp.
50. Waleron M, Misztak A, Waleron M, Franczuk M, Wielgomas B, Waleron K. 2018. Transfer of *Pectobacterium carotovorum* subsp. *carotovorum* strains isolated from potatoes grown at high altitudes to *Pectobacterium peruviense* sp. nov. *Systematic and Applied Microbiology* 41:85-93

650 51. Whitehead PG, Leckie H, Rankinen K, Butterfield D, Futter MN, Bussi G. 2016. An INCA model for pathogens in rivers
651 and catchments: Model structure, sensitivity analysis and application to the River Thames catchment, UK. *Science of*
652 *The Total Environment* 572:1601-10
653 52. Wickham H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Nes York: Springer-Verlag.
654 https://ggplot2.tidyverse.org. pp.
655 53. Zappia RE, Hüberli D, Hardy GESJ, Bayliss KL. 2014. Fungi and oomycetes in open irrigation systems: knowledge gaps
656 and biosecurity implications. *Plant Pathology* 63:961-72

657

658 **Table 1**. Sampling sites in the Durance River catchment.
659

Site code[a]	Basin[b]	Latitude °N	Longitude °E	Altitude (m)	Distance from main river (km): upstream from confluence	downstream from divergence	Description
C01a	Lower	43.756378	5.150282	106	-	5	Carpentras Canal, a managed distributary of Durance River, near Logis Neuf village
C01b	Lower	43.820640	5.082083	98	-	15	Carpentras Canal, a managed distributary of Durance River, near Les Taillades village
R01	Upper	45.024096	6.564294	1813	0	-	Main course of the river, historically called Clarée River, at this location named "Pont de la Souchère"
R02	Upper	44.924983	6.67987	1363	0	-	Main course of the river, historically called Clarée River, at this location named "Pont des amoureux"
R03	Upper	44.704668	6.60111	907	0	-	Durance River near St. Crépin village
R04	Upper	44.550722	6.484659	790	0	-	Durance River at Embrun city, the entrance to Serre Ponçon Lake
R05	Middle	44.475576	6.112166	620	0	-	Durance River at "Archidiacre" site, not far downstream from the Serre-Ponçon Dam.
R06	Middle	44.212041	5.939179	459	0	-	Durance River, at « Plan de la Baume » village near Sisteron
R07	Middle	43.804113	5.825458	291	0	-	Durance River at Manosque
R08	Lower	43.667174	5.490215	188	0	-	Durance River near Pertuis
T01	Upper	44.924603	6.679891	1363	at confluence	-	A tributary of the main river, historically called Durance River, at this location named "Pont des amoureux" and upstream of this confluence
T02	Upper	45.015400	6.459957	1659	20	-	Guisane River, a tributary of the Durance river, not far from the "Col du Lautaret"

T03	Upper	44.877228	6.47717	1294	12	-	Ailefroide Stream near Pelvoux village, tributary of the Gyr river
T04	Upper	44.681169	6.696794	1066	9	-	Guil Stream, a tributary of Durance River, site near the place named "La Maison du Roi",
T05a	Upper	44.536205	6.703007	2090	55	-	Riou Mounal Creek, tributary of Ubaye River, site near the "Col de Vars"
T05b	Upper	44.514918	6.75597	1443	50	-	Ubaye River, a tributary of Durance River, site near Saint-Paul-Sur-Ubaye village
T05c	Upper	44.397539	6.480858	968	11	-	Ubaye River, a tributary of Durance River, site near Le Martinet village, just upstream of its entrance into Serre Ponçon Lake
T06	Middle	44.201151	5.928711	459	1	-	Buëch River, a tributary of Durance river, site near Sisteron
T07	Middle	44.042293	6.040222	438	4	-	Bléone River, tributary of Durance River, upstream of a small dam
T08	Middel	43.728389	5.816048	274	6	-	Verdon River, tributary of Durance River, site at Vinon-sur Verdon
T09	Lower	43.884328	4.892864	39	5	-	Grand Anguillon River, tributary of Durance River, site at Noves where the river becomes a canal

[a]The codes for the sites indicate where water was collected from the main course of the river (R), a tributary flowing into the main course of the river (T) or a canal used to distribute water from the main course of the river for agricultural and other uses (C).

[b]The three basins are depicted in Figure 1.

Table 2. Spearman rank correlations (and associated p-values) among variables describing physical-chemical conditions of water collected at 8 dates at each of 21 sites throughout the Durance River catchment. Significant correlations ($p < 0.05$) are in bold face.

variable	**Temperature °C**	**Conductivity (µS)**	**Dissolved organic carbon (mgL^{-1})**	**PO_4^+ (μgL^{-1})**	**NH_4^+ (μgL^{-1})**	**NO_2^- (μgL^{-1})**
Conductivity (µS)	0.011 (0.921)					
Dissolved organic cabon (mgL^{-1})	**0.428 (0.000)**	0.160 (0.157)				
PO_4^+ (μgL^{-1})	0.183 (0.098)	-0.002 (0.987)	-0.039 (0.732)			
NH_4^+ (μgL^{-1})	-0.142 (0.200)	0.155 (0.162)	0.119 (0.294)	-0.002 (0.984)		
NO_2^- (μgL^{-1})	**0.426 (0.000)**	0.128 (0.250)	**0.257 (0.022)**	0.195 (0.079)	-0.138 (0.215)	
NO_3^- (μgL^{-1})	0.161 (0.147)	**0.404 (0.000)**	**0.575 (0.000)**	-0.151 (0.174)	0.116(0.296)	**0.402 (0.000)**

Table 3. Description of composite factors (*F*) from Principal Component Analysis of seven variables of the physical-chemical conditions of water collected at 8 dates at each of 21 sites throughout the Durance River catchment in terms of (A) the contribution of each individual variable to the variability within each *F* and (B) the correlation of each water variable with each *F*.

A. Contribution of each water variable to the variability within each *F*							
variable	***F1***	***F2***	***F3***	***F4***	***F5***	***F6***	***F7***
Temperature °C	0.165	0.117	0.136	0.121	0.005	0.357	0.098
Conductivity (µS)	0.045	0.234	0.025	0.618	0.000	0.008	0.070
Dissolved organic cabon (mgL^{-1})	0.312	0.000	0.098	0.083	0.146	0.002	0.359
PO_4^+ (μgL^{-1})	0.007	0.104	0.570	0.002	0.233	0.080	0.003
NH_4^+ (μgL^{-1})	0.008	0.357	0.038	0.144	0.066	0.384	0.003
NO_2^- (μgL^{-1})	0.140	0.125	0.108	0.004	0.533	0.032	0.058
NO_3^- (μgL^{-1})	0.323	0.064	0.024	0.027	0.017	0.137	0.409
B. Correlation between the water variable and each *F*							
Temperature °C	-0.566	0.421	0.382	-0.315	-0.060	-0.471	-0.178
Conductivity (µS)	-0.296	-0.594	-0.163	-0.711	-0.011	0.072	0.150
Dissolved organic carbon (mgL^{-1})	-0.778	0.003	0.324	0.260	0.323	0.038	0.341
PO_4^+ (μgL^{-1})	-0.116	0.396	-0.781	-0.042	0.409	-0.223	0.030
NH_4^+ (μgL^{-1})	-0.127	-0.734	-0.203	0.344	-0.217	-0.488	0.029
NO_2^- (μgL^{-1})	-0.522	0.435	-0.340	0.057	-0.617	0.140	0.137
NO_3^- (μgL^{-1})	-0.792	-0.310	-0.160	0.149	0.111	0.291	-0.363

13 **Table 4**. Parameters from the multiple regression of population sizes of Psy or SRP (expressed as $\log_{10}$ bacteria L^{-1}) vs seven
14 composite factors (*F*) from Principal Component Analysis (*c.f.* Tab. 3). Significant values are in bold face.
15

	Dependent variable: $\log_{10}$ Psy L^{-1} $R = 0.541$, $R^2 = 0.293$ p-value$_{regression}$ = **0.000**		Dependent variable: $\log_{10}$ SRP L^{-1} $R = 0.628$, $R^2 = 0.394$ p-value$_{regression}$ = **0.000**		Dependent variable: $\log_{10}$ Total L^{-1} $R = 0.491$, $R^2 = 0.241$ p-value$_{regression}$ = **0.004**	
F	Slope (b)	p-value	Slope (b)	p-value	Slope (b)	p-value
1	**0.183**	**0.034**	**-0.583**	**0.000**	**-0.125**	**0.000**
2	**-0.352**	**0.000**	0.046	0.594	**-0.103**	**0.010**
3	-0.092	0.422	**0.265**	**0.011**	0.012	0.792
4	**0.380**	**0.005**	-0.152	0.195	0.016	0.768
5	-0.021	0.882	-0.156	0.213	0.088	0.124
6	0.201	0.183	**-0.293**	**0.031**	0.008	0.892
7	0.214	0.304	-0.338	0.072	0.036	0.673

16

17

Figure 1. Population densities (log_{10} bacteria L^{-1}) of *Pseudomonas syringae* and Soft Rot *Pectobacteriaceae* (SRP) in water in the Durance River basin at eight sites in the main river (R), eleven sites in tributaries (T) and two sites in irrigation canals (C) in four seasons in each of 2016 and 2017. All values for population density are based on culturable bacteria isolated from single samples at each date and site except for values for sites R02 and R08 in 2017 that were means of triplicate samples. Error bars for those values represent the standard error. The map portrays the sampling sites along the full expanse of the Durance River basin from its most northern reaches in the Hautes Alpes department southward through the departments of Alpes d'Haute Provence and Vaucluse. The black contours of the map represent the three hydrological sections of the river basin (23).

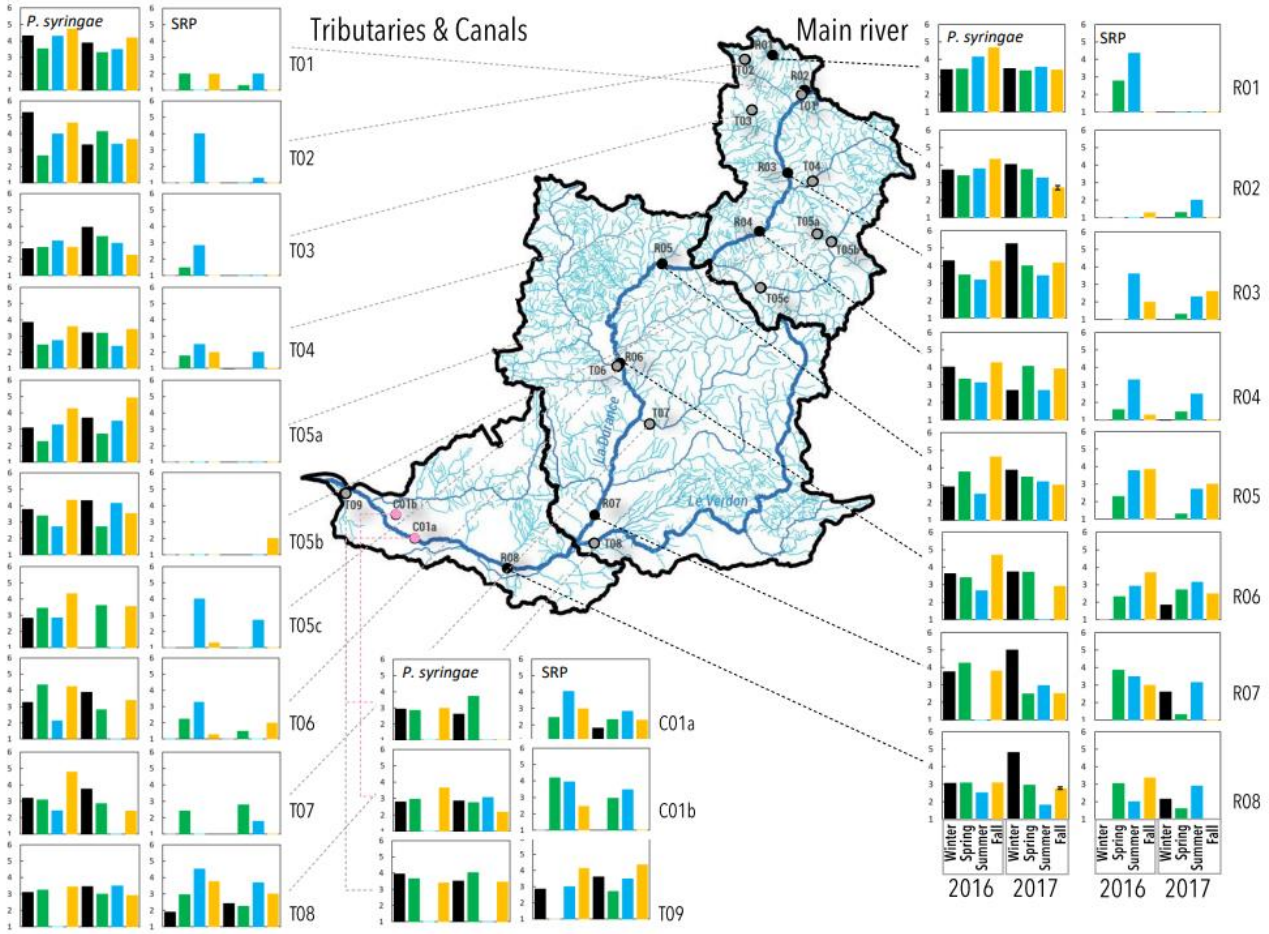
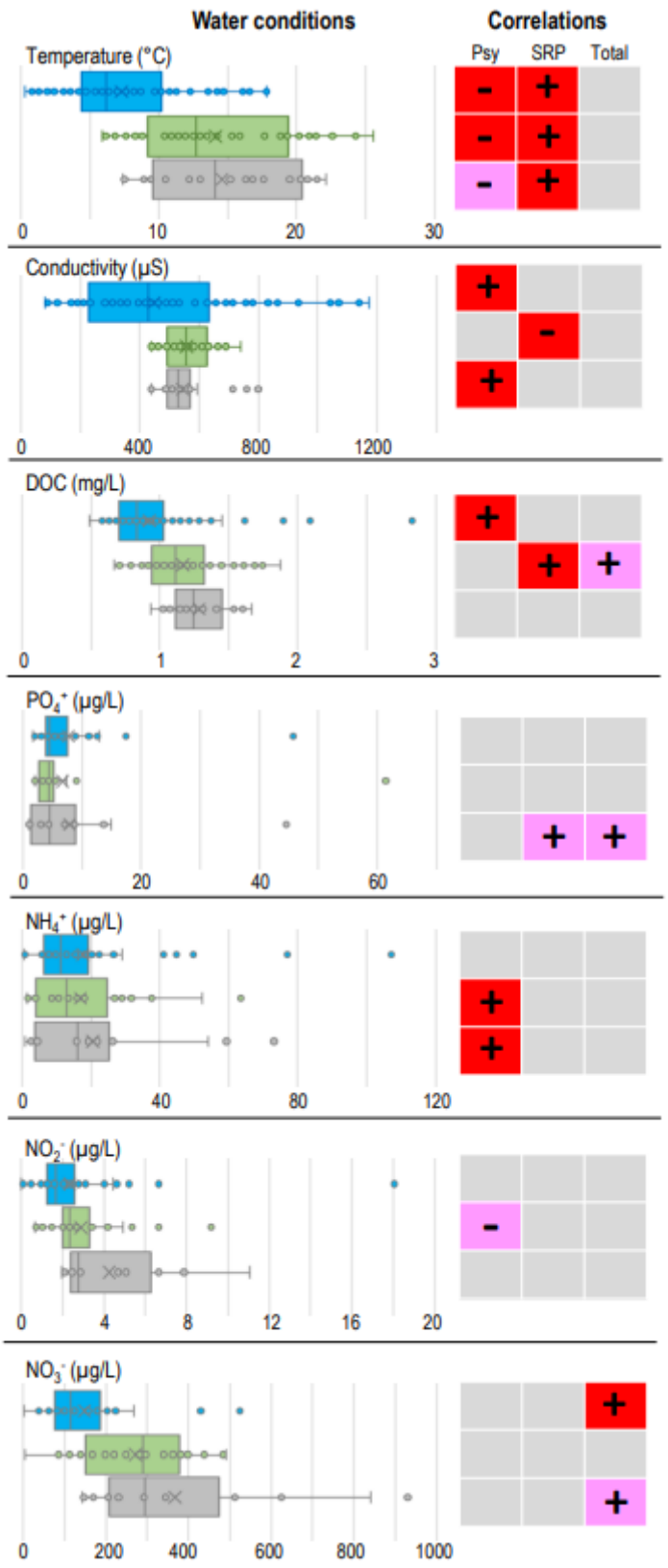
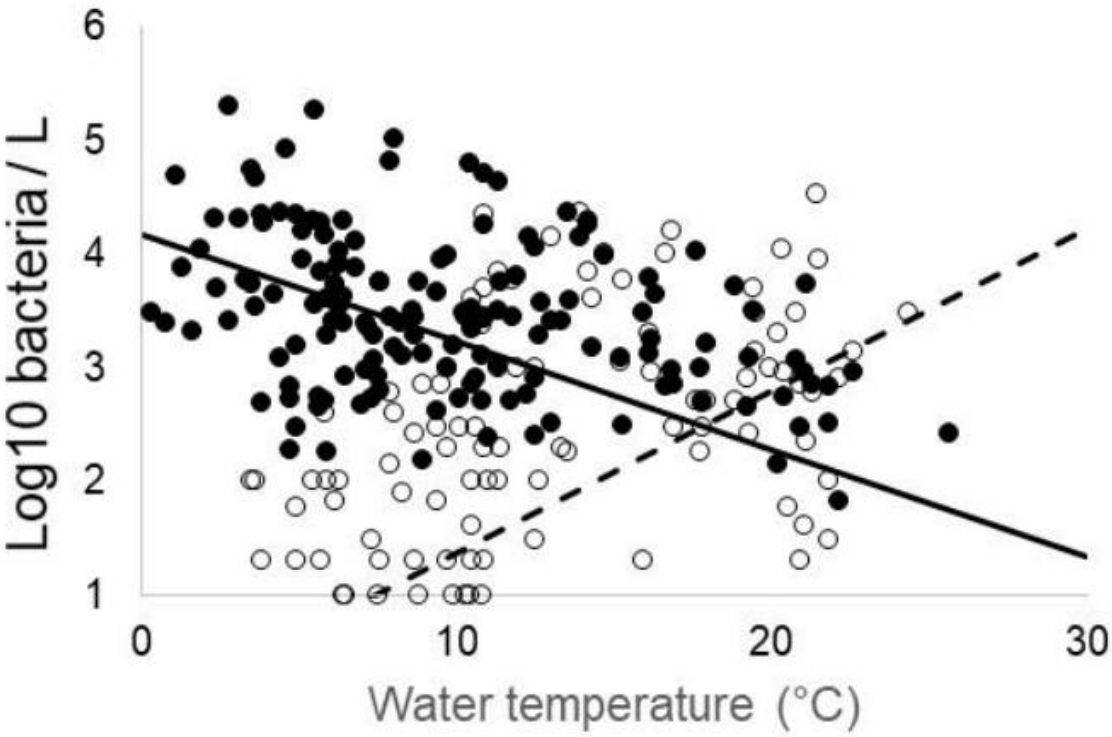


Figure 2. Correlation of *Pseudomonas syringae* (Psy), Soft Rot *Pectobacteriaceae* (SRP) and total mesophilic (Total) bacterial population densities with water conditions in the three basins of the Durance River catchment. The left-hand panel indicates the water conditions (box plots including a presentation of all data values) in the three basins (as depicted in Fig. 1) (upper in blue, middle in green and lower in grey). The right hand panel indicates whether the values of the Spearman Rank correlation between the water conditions and each of the bacterial population densities were positive (+) or negative (-), and if they were significant according to $p < 0.05$ (red background) or $0.05 > p < 0.10$ (pink background). Grey backgrounds indicate that $p > 0.10$ for this statistical test.



53 **Figure 3.** The relationship of bacterial population sizes with water temperature in the Durance River catchment. Water
54 temperature accounted for about 30% of the variability in population size of Psy (solid symbols) ($R^2 = 0.277$, $p_{regression} = 0.000$)
55 and SRP (open symbols) ($R^2 = 0.317$, $p_{regression} = 0.000$) according to linear regressions for data from both 2016 and 2017
56 combined. The linear regressions are represented by a solid line for Psy (Log_{10} Psy $L^{-1} = 4.248 - 0.100 \times °C$) and a dotted line
57 for SRP (Log_{10} SRP $L^{-1} = 0.346 - 0.136 \times °C$).
58



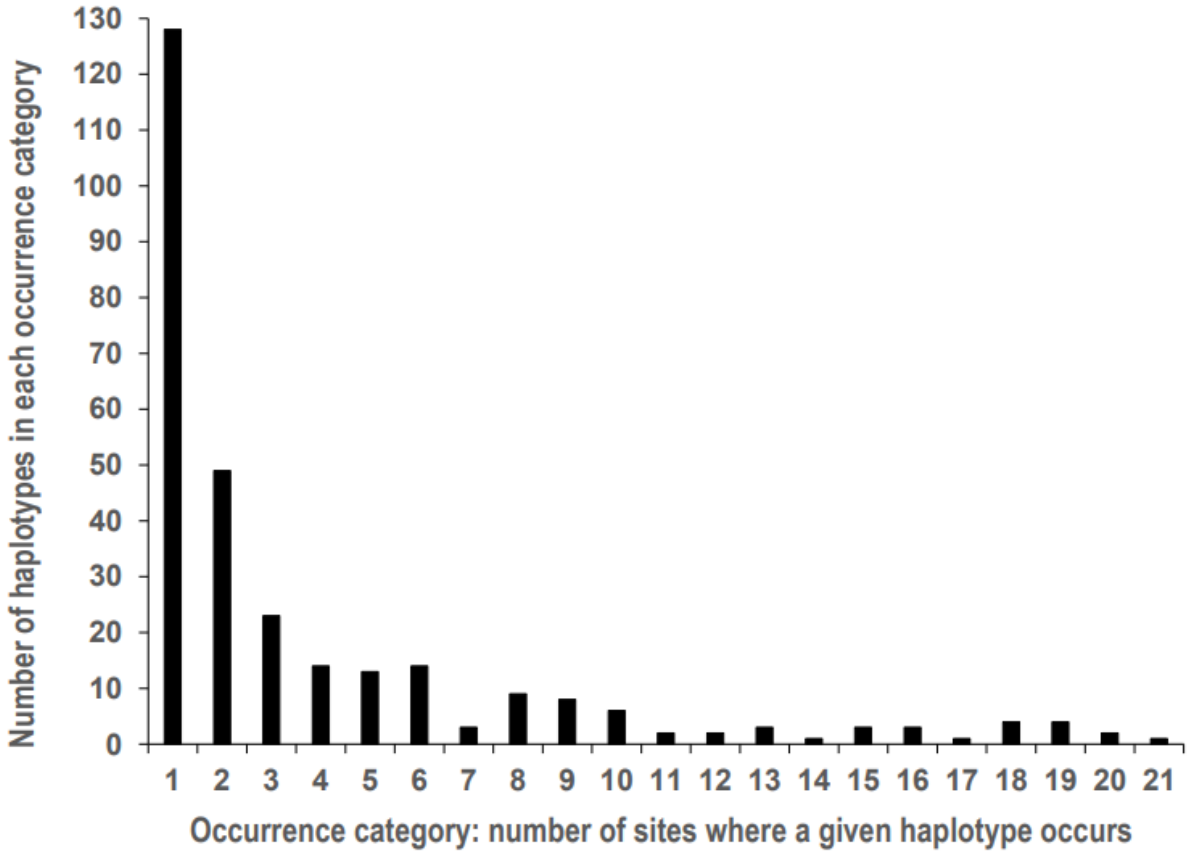
59

60

61 **Figure 4.** Frequency of occurrence of the 291 haplotypes of *P. syringae* throughout the Durance River basin. Only one haplotype

62 of *P. syringae* (DD.1) was detected at all 21 sampling sites whereas 128 haplotypes were detected at only 1 site during the two

63 years of sampling.



64

65

66 ***Supplementary Information***

67

68 **Supplementary Table 1:** Population sizes of bacteria in the Durance river, tributaries and canals

69 **Supplementary Table 2:** Description of amplicon sequence variants.

70 **Supplementary Table 3:** Variability of population sizes of bacteria at two sampling site along the Durance river.

71 **Supplementary Table 4:** Primers used for NGS for *P. syringae*

72 **Supplementary Table 5:** *Cts* sequences of reference strains of *Pseudomonas syringae* used in this study

73 **Supplementary Table 6:** Values for water physical-chemistry variables.

74 **Supplementary Figure 1**: Relationship between population densities of *Pseudomonas syringae* and Soft Rot

75 Pectobacteriaceae (SRP) species complexes in Durance River

76 **Supplementary Figure 2**: Variability in densities of *Pseudomonas syringae* and Soft Rot Pectobacteriaceae (SRP) at three

77 sampling times within the same day at two sites along the Durance River catchment.

78

79