



**HAL**  
open science

## **Cyanobacteria and Algae in Clouds and Rain in the Area of puy de Dôme, Central France**

Kevin P Dillon, Florence Correa, Celine Judon, Martine Sancelme, Donna E  
Fennell, Anne-Marie Delort, Pierre Amato

► **To cite this version:**

Kevin P Dillon, Florence Correa, Celine Judon, Martine Sancelme, Donna E Fennell, et al.. Cyanobacteria and Algae in Clouds and Rain in the Area of puy de Dôme, Central France. *Applied and Environmental Microbiology*, 2020, 87 (1), 10.1128/AEM.01850-20 . hal-03033301

**HAL Id: hal-03033301**

**<https://hal.science/hal-03033301v1>**

Submitted on 1 Dec 2020

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

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

1 **Cyanobacteria and algae in clouds and rain in the area of puy de Dôme, Central France**

2 Running Title: Cyanobacteria and algae in clouds and rain

3 Kevin P. Dillon<sup>a#</sup>, Florence Correa<sup>b</sup>, Celine Judon<sup>b</sup>, Martine Sancelme<sup>b</sup>, Donna E. Fennell<sup>a</sup>,

4 Anne-Marie Delort<sup>b</sup>, and Pierre Amato<sup>b</sup>

5 <sup>a</sup>Department of Environmental Sciences, Rutgers University, New Brunswick, NJ, USA

6 <sup>b</sup>Université Clermont Auvergne, CNRS, SIGMA Clermont, ICCF, F-63000 Clermont-Ferrand,

7 France

8

9 #Corresponding Author: Kevin P. Dillon, [dillon.kevin.p@gmail.com](mailto:dillon.kevin.p@gmail.com)

10 Present Address: 14 College Farm Rd., New Brunswick, NJ 08901, United States of America

11

12 **Keywords:** bioaerosols, dispersal, microalgae, Chlorellaceae, Trebouxiales,

13 *Chroococciopsidales*

14

15

16

17

18

19

20

21

22

23 **ABSTRACT:** The atmosphere contains diverse living microbes, of which the heterotrophic  
24 community has been the most studied. Microbes with other trophic modes, such as  
25 photoautotrophy, have received much less attention. Here, culture-independent and dependent  
26 methods were used to examine the presence and diversity of oxygenic photoautotrophic microbes  
27 in clouds and rain collected at or around puy de Dôme Mountain, central France. Cloud water  
28 was collected from the summit of puy de Dôme (1465 m above sea level (a.s.l.)) for cultivation  
29 and metagenomic analysis. Cyanobacteria, diatoms, green algae, and other oxygenic  
30 photoautotrophs were found to be recurrent members of clouds, while green algae affiliated with  
31 the Chlorellaceae were successfully cultured from three different clouds. Additionally, rain  
32 samples were collected below the mountain from Opme meteorological station (680 m, a.s.l.).  
33 The abundance of chlorophyll *a*-containing cells and the diversity of cyanobacteria and green  
34 algae in rain were assessed by flow cytometry and amplicon sequencing. The corresponding  
35 downward flux of chlorophyll *a*-containing organisms to the ground, entering surface ecosystems  
36 with rain, varied with time and was estimated between ~1 and >300 cells cm<sup>-2</sup> day<sup>-1</sup> during the  
37 sampling period. Besides abundant pollen from Pinales and Rosales, cyanobacteria of the  
38 *Chroococciopsidales* and green algae of the Trebouxiales were dominant in rain samples.  
39 Certain members of these taxa are known to be ubiquitous and stress-tolerant and could use the  
40 atmosphere for dispersal. Overall, our results indicate that the atmosphere carries diverse, viable  
41 oxygenic photoautotrophic microbes and acts as a dispersal vector for this microbial guild.

42 **IMPORTANCE:** Information regarding the diversity and abundance of oxygenic  
43 photoautotrophs in the atmosphere is limited. More information from diverse locations is needed.  
44 These airborne organisms could have important impacts upon atmospheric processes and on the  
45 ecosystems they enter after deposition. Oxygenic photoautotrophic microbes are integral to

46 ecosystem functioning and some have the potential to affect human health. A better  
47 understanding of the diversity and the movements of these aeolian dispersed organisms is needed  
48 to understand their ecology, as well as how they could affect ecosystems and human health.

## 49 **INTRODUCTION**

50         The atmosphere harbors a diverse microbial assemblage that can be free-floating, adhered  
51 to particles, or suspended in water droplets (e.g. clouds, fog, and rain). Bacteria are typically  
52 present at concentrations between  $10^4$  to  $10^5$  cells  $m^{-3}$  in air (1) and about  $10^3$  to  $10^4$  cells  $mL^{-1}$  of  
53 water in clouds (2). Microbes aerosolized from the surface environment including bodies of  
54 water, soils, plant surfaces, and other sources can rise in the atmosphere, be incorporated into  
55 clouds, and return to the surface via dry or wet deposition (for a general review, see Frölich-  
56 Nowoisky et al. (3)). Some can help to form cloud droplets and ice crystals by acting as nuclei  
57 for the condensation and freezing of water (4, 5). Over mid-latitude continental areas, the  
58 microbial assemblage is often dominated by heterotrophic bacteria such as *Pseudomonas* and  
59 *Sphingomonas* both in terms of abundance and activity (2, 6). These heterotrophic microbes have  
60 been studied for their interaction with atmospheric processes potentially affecting cloud  
61 chemistry (7). Other microbial guilds, such as oxygenic photoautotrophic microbes, have  
62 received less attention regarding their abundance, diversity, and activity in the atmosphere.

63         Based upon previous studies, the main airborne oxygenic photoautotrophic microbes are  
64 green algae, cyanobacteria, and diatoms (8). These are found in/on and can be aerosolized from  
65 surface environments including lakes, trees, and building facades (9-12). Their abundance in the  
66 air was reported to be affected by meteorological parameters such as wind direction and speed  
67 (13). The aerosols containing cyanobacteria and algae are important to study due to their impacts  
68 upon human health (11, 14), ice nucleation (15), and surface ecosystems by deposition and

69 colonization (16). In general, the atmospheric dispersal and cycling of oxygenic photoautotrophs  
70 is not well documented (17). Cyanobacteria have been detected in clouds at variable abundances  
71 as part of a few studies (2, 18, 19). However, the diversity of oxygenic photoautotrophs in clouds  
72 remains understudied. Most studies have been conducted on atmospheric cyanobacteria and  
73 algae in near-surface air environments, snow, and rain (20-22). For in-depth reviews about algae  
74 and cyanobacteria in the atmosphere, see Tesson et al. and Wiśniewska et al. (8, 23).

75         This study aimed to investigate the diversity and abundance of cyanobacteria and algae in  
76 clouds and rain at and around the puy de Dôme mountain (1465 m, 680 m above sea level (a.s.l.),  
77 respectively), a mid-altitude rural area in central France. We investigated metagenomes of cloud  
78 water and cultured cloud water samples for oxygenic photoautotrophic microbes by enrichment.  
79 Additionally, chlorophyll *a*-containing cells were enumerated in rain samples, the corresponding  
80 downward fluxes were estimated, and the diversity of cyanobacteria and algae was assessed by  
81 targeted amplicon sequencing.

## 82 **RESULTS**

### 83 **Cyanobacteria and green algae in clouds**

84         Cloud water from the summit of the puy de Dôme was collected and used directly to  
85 cultivate microbial oxygenic photoautotrophs. Cloud water was collected on 01/06/2010,  
86 08/06/2010, and 29/06/2017 (dd/mm/yyyy) for enrichment cultures of oxygenic photoautotrophs  
87 (referred to as 01062010, 08062010, and 29062017 cultures, respectively). The physical and  
88 meteorological characteristics for each cloud during the cloud water collection process are  
89 presented in Table S1. Diversity metrics for the enrichment culture communities are presented in  
90 Table S2. The main eukaryotic oxygenic photoautotrophic microbes cultured were aeroterrestrial  
91 epiphytic green algae of the Chlorellaceae (Figure 1A). The most dominant Chlorellaceae

92 operational taxonomic unit (OTU) in the 08062010 and 29062017 cultures, and second most  
93 abundant in the 01062010 culture, was most closely associated with *Chloroidium saccharophila*  
94 SAG 211-9a. The 01062010 culture was dominated by an unclassified eukaryote (Figure 1A),  
95 whose taxonomy could not be resolved further. Surprisingly, no cyanobacteria were detected in  
96 the oxygenic photoautotrophic cloud water enrichments (Figure 1B). Sphingomonadaceae,  
97 Chitinophagaceae, and Rhizobiaceae bacterial families were dominant in the 01062010,  
98 08062010, and 29062017 cultures, respectively (Figure 1B).

99         Along with the culture-dependent approach, we used metagenomics to explore cloud-  
100 borne oxygenic photoautotrophs. Clouds were collected on 11/10/2013, 14/10/2013, and  
101 05/11/2013 (denoted as 11102013, 14102013, and 05112013, respectively; Table S1). Cloud-  
102 borne oxygenic photoautotrophs were identified based upon taxonomy in six cloud  
103 metagenomes—three obtained for this study (11102013, 14102013, 05112013), and three that  
104 were previously described (24) (17112014-1, 17112014-2, 17112014-3). Oxygenic  
105 photoautotrophs were detected in all six cloud metagenomes, indicating that these were  
106 persistent members of the microbial assemblage in clouds (Table 1). Bacillariophyceae (diatoms)  
107 and Chrysophyceae (golden algae) were >1% of the SSU *rRNA* gene sequences across all six  
108 clouds, while the cyanobacteria were <1% of the sequences (Table 1).

109         In addition to taxonomy, we investigated functional genes related to photosynthesis.  
110 Overall, based on the UniProtKB identifiers and their associated gene ontology identifiers (GO  
111 IDs), the functional genes related to photosynthesis by cyanobacteria were rare in the  
112 metagenomes, ranging from 2 to 103 sequences in all clouds (Figure 2). The main genes  
113 identified were related to *psb* genes, which encode for components of photosystem II (Data Set  
114 S1). The dominant bacterial genera associated with the detected UniProtKB identifiers of the

115 photosynthesis-related genes were *Lyngbya*, *Synechococcus*, and *Pseudanabaena* (Figure 2). The  
116 11102013, 17112014-2, and 17112014-3 clouds only had 2 functional gene sequences associated  
117 with photosynthesis, and each were 100% associated with *Oscillatoria*, *Microcystis*, and  
118 *Crocospaera*, respectively (Figure 2). Functional genes associated with photosynthesis by  
119 eukaryotic photoautotrophs were not detected in any of the clouds.

120         Previously, Amato et al. (2) investigated the microbes that were potentially active in  
121 clouds. DNA and RNA were extracted from cloud water and amplicons of fragments of the *16S*  
122 *rRNA* and *18S rRNA* genes (for prokaryotes and eukaryotes, respectively) were generated. Based  
123 on the ratio of the relative abundance of a specific OTU in the *16S* or *18S* rRNA and *16S* or *18S*  
124 *rRNA* DNA of a sample, an OTU's activity can be gauged (25). This is not a perfect method of  
125 assessing activity, but it can be used as a qualitative metric (26). Here we re-examined the data  
126 specifically for different types of oxygenic photoautotrophic microbes. Eukaryotic algae had  
127 RNA/DNA average ratios ranging from 0.59 to 2.11 and likely were mainly inactive (Table 2).  
128 In contrast, the cyanobacteria appeared to be more active with RNA/DNA ratios averaging 41.69  
129 and reaching values as high as 1100 (Table 2) Unfortunately, the cyanobacteria that were  
130 detected were associated with unclassified phylotypes, and could not be classified further with  
131 the sequence data available.

### 132 **Chlorophyll *a*-containing cells in rain**

133         We collected 16 rain samples over a 3-week period at our rural sampling site (Opme) and  
134 assessed the chlorophyll *a*-containing cell abundance by flow cytometry. The main  
135 meteorological parameters and chemical composition of the rain samples are shown in Table S3.  
136 There were chlorophyll *a*-containing cells present in every rain event over the sampling period.  
137 Their concentration ranged from 9 to 2750 cells mL<sup>-1</sup> of rainwater, corresponding to downward

138 fluxes of ~1 to 328 chlorophyll *a*-containing cells cm<sup>-2</sup> day<sup>-1</sup> being deposited with rain (Figure  
139 3A). The chlorophyll *a* content in cells vary throughout their lives and environmental conditions  
140 (27) and cell membranes could rupture during manipulation, allowing chloroplasts to be  
141 detected. These factors could affect the enumeration of chlorophyll *a*-containing cells. In  
142 addition to these measurements, the diversity of cyanobacteria and green algae in the rain  
143 samples was assessed by amplicon sequencing of the *16S* and *18S rRNA* genes, specific for each  
144 respective group. The diversity data from the sequencing is presented in Table S4. There was a  
145 significant positive correlation between the Chao1 index and biomass (the number of chlorophyll  
146 *a*-containing cells;  $p = 0.04$ ,  $\rho = 0.52$ , Data Set S2). Additionally, there was a significant positive  
147 correlation between air masses coming from the west and the number of chlorophyll *a*-  
148 containing cells ( $p = 0.02$ ;  $\rho = 0.57$ , Data Set S2). Despite the variable concentrations/fluxes, we  
149 enriched oxygenic photoautotrophic microbes from each rain event by incubation in MWC (28)  
150 or BG11 medium (ATCC medium 616), as indicated by observation of green-tinted cultures.

### 151 **Oxygenic photoautotrophs in rain**

152 The dominant cyanobacteria present varied across rain events. Unclassified members of  
153 the *Nostocales* order were the most prominent cyanobacteria that range from not detected to  
154 100% of the OTUs in the rain samples (Figure 3B). The *Chroococcidiopsaceae* varied from not  
155 detected to 75%, being most represented by an OTU associated with an uncultured  
156 cyanobacterium. *Aliterellaceae* ranged from not detected to 50% of OTUs with the most  
157 dominant OTU being closely related to *Aliterella* CENA595 (29). The *Microcoleaceae* ranged  
158 from not detected to 8.9% of OTUs with the dominant OTU matched to *Tychonema* CCAP 1459-  
159 11B (30) (Figure 3B). The *Microcoleaceae* family's relative abundance was significantly  
160 positively correlated with pH ( $p = 0.036$ ,  $\rho = 0.53$ ; Data Set S2) and with the concentration of



161 chlorophyll *a*-containing cells ( $p = 0.028$ ,  $\rho = 0.55$ ; Data Set S2). These cyanobacteria might  
162 have been an important component of the group of chlorophyll *a*-containing cells that were  
163 detected. There were no cyanobacteria present on 25 May, 01 June, and 06 June (Figure 3B).

164         Among green algae, the Trebouxiales order dominated the rain samples accounting for  
165 0.1 to 6% of OTUs across the rain events with a peak in abundance on 25 May (Figure 3C). An  
166 uncultured *Trebouxia* alga was the most dominant sequence associated with OTUs matched to  
167 this order. Other Trebouxiophyceae varied in abundance from not detected to 0.76%. The other  
168 green algae detected were lower in abundance. Members of the order Chlorellales were only  
169 detected in some rain events in relative abundances up to about 0.3% (13 May; Figure 3C). The  
170 Chlamydomonadales detected in some of the rain samples peaked in relative abundance on 25  
171 May at about 0.11% (Figure 3C). Members of the Pedinomonadales were only detected on 3  
172 May at about 0.01% relative abundance (Figure 3C).

173         Other than cyanobacteria and algae, pollen was also detected in rain (Figure 3D).  
174 Notably, pollen from the order Pinales dominated the eukaryotic sequences at early time points,  
175 and then decreased in relative abundance across the sampling time frame (Figure 3D). The order  
176 Pinales was dominated by phylotypes associated with the genus *Pinus*, pine trees, which are  
177 abundant in the sampling region. During the final days of the sampling period the relative  
178 abundance of Pinales decreased and the relative abundances of the Rosales and Poales increased.  
179 Therefore, there were significant negative correlations between the relative abundances of  
180 Pinales and Rosales ( $p = 1.7 \times 10^{-9}$ ,  $\rho = -0.96$ ; Data Set S2) as well as Pinales and Poales ( $p =$   
181  $2.5 \times 10^{-6}$ ,  $\rho = -0.90$ ). There were no significant correlations between pollen sequences' relative  
182 abundances and chemical/meteorological parameters or geographic origins of the air masses  
183 (Data Set S2).

184 Finally, it should be mentioned that taxa other than oxygenic photoautotrophic microbes  
185 were abundant in the amplicons from rain samples, although the primers used to generate *16S*  
186 *rRNA* and *18S rRNA* gene amplicons were designed for excluding them (Cya 359f/Cya 781r and  
187 Euk528f/CHLO02r, respectively (31-33)). The use of these primers might need to be reexamined  
188 for future investigations.

## 189 **DISCUSSION**

### 190 **Biodiversity and sources of cyanobacteria and algae in clouds and rain**

191 Various genera of microalgae and cyanobacteria associated with the terrestrial and  
192 phyllosphere environments have previously been found in atmospheric samples (8). Microbes of  
193 the phyllosphere experience many of the same stressors as airborne microbes (34), so they might  
194 be better suited to living and surviving in the atmosphere than those from other sources. Overall,  
195 our data indicated that many of the microalgae and cyanobacteria detected in clouds and rain  
196 were of terrestrial origin, but some were likely of marine origin.

197 The BG11 media used for enrichment of cloud water selected for freshwater  
198 cyanobacteria and algae. The predominant photoautotrophic microbes enriched from clouds  
199 under these conditions were green algae of the Chlorellaceae (Figure 1B). The main OTU was  
200 associated with *Chloroidium saccharophila* SAG 211-9a. This organism was originally isolated  
201 from tree sap (35, 36), which suggests again that the phyllosphere could be an important source  
202 of atmospheric microalgae. Furthermore, *Chloroidium* is known to have a wide variety of  
203 habitats that range from freshwater to being lichen-associated (36). *Chloroidium* isolated from  
204 atmospheric samples were able to grow in a simulated freshwater environment (15).

205 We focused on enrichment of freshwater oxygenic photoautotrophs because these  
206 organisms might be able to colonize the local environment after atmospheric transport. The puy

207 de Dôme region is inland, far from oceanic sources, but it has been noted that the microbial  
208 communities in clouds at puy de Dôme often retain a marine signature (2, 24, 37, 38). Most of  
209 the clouds that were characterized for this study were of marine origin. However, the single  
210 cloud of continental origin (11102013) had similar relative abundances of SSU sequences of  
211 oxygenic photoautotrophic microbes compared to the marine clouds (Table 1).

212 Cyanobacteria were present in low relative abundances in the metagenomes across three  
213 clouds sampled for this study (11102013, 14102013, and 05112013) and the three clouds that  
214 were characterized previously (17112014-1, 17112014-2, 17112014-3) (24) (Table 1). Based  
215 upon the UniProtKB annotation of functional genes detected in the metagenomes, however, there  
216 were diverse cyanobacteria present. Different cyanobacterial genera associated with the marine  
217 environment were detected including *Lyngbya*, *Crocospaera*, and *Synechococcus*. *Lyngbya* is a  
218 common cyanobacterium found in microbial mats in intertidal areas and *Crocospaera* and  
219 *Synechococcus* are common in the ocean (39, 40). Additionally, *Synechococcus* and *Lyngbya*  
220 have both been found in air samples (10, 41). Besides marine cyanobacteria, there were also  
221 terrestrial cyanobacteria such as *Gloeocapsa* and *Gloeobacter*, which were all present in the  
222 14102013 metagenome (Table 1). *Gloeocapsa* has been found in the atmosphere previously (42,  
223 43) and *Gloeobacter* has been found in terrestrial environments including caves (44).  
224 Cyanobacteria such as *Lyngbya* and *Gloeocapsa* were found to have a wide distribution amongst  
225 various atmospheric locations in Asia, Antarctica, Polynesia, Nearctic, and the Palearctic (11), in  
226 addition to some cloud samples from this study. Their heterogeneous presence in this study could  
227 be due to various reasons relating to meteorological conditions of the different locations over  
228 which the air mass traveled.

229           Despite detecting a clear marine signature in the cloud metagenomes, many of the algae  
230 and cyanobacteria in rain at the sampling site/time could be from other environments. The major  
231 cyanobacterial phylotypes detected in rain were of the order *Chroococciopsidales*. This  
232 contains two families: *Chroococciopsaceae* and *Aliterellaceae*. A major OTU detected was  
233 related to the genus *Aliterella* whose type strain was isolated from the Atlantic Ocean off the  
234 coast of Brazil (29). This genus has also been found on rocks in the Atacama Desert (45) and in  
235 rain collected in Hawaii (43). It has a cosmopolitan distribution and might be stress tolerant (as  
236 per its presence in a desert and in the atmosphere). Uncultured members of the  
237 *Chroococciopsaceae* and unclassified members of the *Nostocales* were also detected in the  
238 samples. Members of these families have been previously detected in atmospheric samples in  
239 Hawaii (43). A member of the *Microcoleaceae*, matching with *Tychonema* CCAP 1459-11B, a  
240 freshwater cyanobacterium (30), was dominant in some of the rain samples.

241           There were diverse algae present in clouds. Chlorellales members were cultivated from  
242 clouds (Figure 1A) and detected in rain (Figure 3C). Previously, green algae such as *Chlorella*  
243 and *Chlorococcum* were cultured from rain (20). Overall, the most abundant group of green  
244 algae in our rain samples were of the Trebouxiales order (Figure 3C), with phylotypes associated  
245 with uncultured *Trebouxia* being the most dominant. Certain members of the *Trebouxia* are  
246 prominent phycobionts of lichen (46) and have been detected in air and rain samples in South  
247 Korea (47) and air samples in Spain (48). The Trebouxiophyceae, which includes Chlorellales  
248 and Trebouxiales, have also been found in the marine environment (49), suggesting a  
249 cosmopolitan distribution of this taxonomic class.

250           Other than microbial oxygenic photoautotrophs, pollen was detected in the rain samples.  
251 The pollen did not correlate with any of the physicochemical parameters we measured. This fact,

252 along with their large sizes, suggests the pollen grains were most likely derived from local  
253 sources, emphasizing the importance of local landscapes on the atmospheric biological content  
254 (50). The types of pollen that were detected varied over the three-week sampling period. In  
255 anemophilous species, pollen is obviously largely dispersed aurally with well-known seasonality  
256 (47). The beginning of our sampling (May) likely corresponded with pine pollen season and the  
257 end could have corresponded with the end of the pine pollen season (June).

258         Because there were conflicting correlations with a western origin of rain air masses with  
259 air temperature and the number of chlorophyll *a*-containing cells, respectively, it is not possible  
260 to discriminate between a geographical source and an actual environmental temperature effect  
261 with this data. In a recent study, which included our sampling area, local landscapes and  
262 meteorological conditions were significant factors affecting the composition of the airborne  
263 microbial community (50).

264         Based upon the taxonomic affiliations, the cyanobacteria and algae detected in rain were  
265 likely derived from a mixture of locations including marine, freshwater, and terrestrial habitats.  
266 Many of the marine species were probably transported by the clouds and some of the freshwater  
267 and terrestrial phylotypes were transported by a combination of the clouds and the locally  
268 generated bioaerosols washed-out of the air column by the rain.

### 269 **Survival and potential activity**

270         Since many of the oxygenic photoautotrophs in our rain samples have a cosmopolitan  
271 distribution, atmospheric transport could be a substantial dispersal mechanism for these  
272 organisms (51). However, for successful dispersal the organisms must survive atmospheric  
273 stressors, including osmotic shocks, light, oxidants, and freezing. Many Trebouxiophyceae  
274 tolerate desiccation and other stressors that they would experience in the atmosphere (52).

275 Osmotic shocks and dehydration stress occur in both terrestrial and phyllosphere habitats. It was  
276 speculated that airborne cyanobacteria were more tolerant to different relative humidities than  
277 algae (53). The cellular biochemical properties of cyanobacteria and algal cells may be key to  
278 their survival in the atmosphere. Algae of the genus *Chloroidium*, such as *Chloroidium sp.*  
279 UTEX 3007 (54), can potentially survive a variety of atmospheric stressors because of  
280 intracellular desiccation resistance-promoting sugars (36, 54). Some green algae isolated from  
281 the atmosphere have carotenoids, thick cells walls, and intracellular antioxidants which help  
282 them to survive atmospheric stressors (55). Many members of the Trebouxiophyceae have  
283 mycosporine-like amino acids in the cytoplasm which help to protect the algae from UV  
284 radiation and oxidants (56). Mycosporine-like amino acids are also found in cyanobacteria such  
285 as *Lyngbya purpurem* (57). Other cyanobacteria such as *Synechococcus*, in particular, are known  
286 to survive various osmotic, oxidative, and temperature stressors (40). The *Chroococcidiopsaceae*  
287 family was prominent in our rain samples and a member of this family, *Chroococcidiopsis*, is  
288 well-known to be highly tolerant to desiccation and radiation (58). The relative abundances of  
289 this family were negatively correlated with increasing air temperature (ranging from 14°C to  
290 22°C; Table S3), a trend that was previously reported for the abundance of various cultivated  
291 algae (59).

292         Amongst atmospheric stressors, freeze-thaw stress has been found to affect bacteria the  
293 most (60). The ability of microbial cells to tolerate freezing can be an important factor for  
294 survival during atmospheric transport. Algae isolated from atmospheric samples were able to  
295 survive freezing to a greater extent than aquatic algae (15). The 01062010 and 08062010 cloud  
296 water enrichments were revived after being stored in glycerol at -20°C for several years. This  
297 method for storing microorganisms could have an impact upon the mortality of cells and thus

298 potential composition of the microbial community (61). Green algae and cyanobacteria, the most  
299 abundant oxygenic photoautotrophs in the atmosphere (41), are more amenable to  
300 cryopreservation than other oxygenic photoautotrophs (62). The dominance of the Chlorellaceae  
301 in the cloud enrichments could be due to a larger number of organisms of this taxon present in  
302 the culture when it was frozen. This would increase the likelihood of survival of this taxon as  
303 opposed to other lower abundant taxa. However, when comparing 08062010 (stored at -20°C for  
304 several years) and 29062017 (which was never stored at -20°C) cultures, they shared a dominant  
305 OTU in the family Chlorellaceae that matched with *Chloroidium saccharophila* SAG 211-9a.  
306 The fact that this alga was the major phylotype in two cloud enrichments and survived  
307 cryopreservation in the 01062010 and 08062010 cultures supports its tolerance to atmospheric  
308 stressors.

309         In addition to their vegetative states, many cyanobacteria and algae go through different  
310 stages of development. Many of the resting states that the cells manifest, which are also part of  
311 their reproductive cycle, impart stress-tolerance and allow the microbe to survive harsh  
312 conditions (for a review see Ellegaard and Ribeiro (63)). For example, unequal autospores are  
313 used for reproduction in *Chloroidium* and small autospores could be amenable to aeolian  
314 transport (36). Small resting states would be more aerodynamically favorable for long-distance  
315 transport as compared to larger vegetative cells.

316         While transiting through the atmosphere, if oxygenic photoautotrophs can survive  
317 atmospheric stressors, they have the potential to be metabolically active. After re-assessing a  
318 previous study on RNA/DNA ratios using rRNA amplicons, the eukaryotic algae were found to  
319 be relatively inactive with low ratios of RNA/DNA (Table 2). This could be consistent with their  
320 presence as resting states (e.g. spores) (63). Alternatively, this could indicate that cellular

321 functions of the eukaryotic algae were greatly affected by atmospheric stressors. In contrast the  
322 cyanobacteria appeared to be more active with RNA/DNA ratios of not detected to 1100 (Table  
323 2). Thus, they could affect cloud water droplet chemistry via carbon fixation, carbon and  
324 nitrogen metabolism, or oxidant detoxification, similar to how heterotrophic bacteria were  
325 hypothesized to be affecting cloud droplets (24). Consistent with other data, different genes  
326 associated with oxygenic photosynthesis were persistently present in cloud metagenomes.  
327 Nevertheless, not all genes associated with oxygenic photosynthesis and carbon fixation were  
328 detected: for instance, no genes encoding for ribulose-1,5-bisphosphate carboxylase/oxygenase  
329 (RuBisCO) were annotated (Data Set S1).

### 330 **Potential impacts on ecosystems**

331 The oxygenic photoautotrophs that are suspended in the air will eventually return to  
332 different surface environments through dry or wet deposition (e.g. rain). Some eukaryotic algae  
333 can promote ice formation, and thus could promote their deposition to new surface environments  
334 by affecting cloud physics and precipitation patterns (15). All the rain samples analyzed  
335 contained viable, culturable, aerially dispersed oxygenic photoautotrophs. As a result, rain  
336 undoubtedly acts as a dissemination vector for oxygenic photoautotrophs to colonize new  
337 environments. Near the rain sampling site of Opme is Aydat Lake, a freshwater lake of  
338 approximately 65 hectares ( $6.5 \times 10^9 \text{ cm}^2$ ). The cyanobacteria and green algae that were  
339 transported through the atmosphere in this area were thus likely to be deposited in this lake  
340 ecosystem, as well as other locations. Extrapolating the observed flux of chlorophyll *a*-  
341 containing cells with precipitation to the surface area of the lake indicates that  $10^9$ - $10^{12}$   
342 chlorophyll *a*-containing cells enter the lake per each rainy day. These values are specific for this  
343 region and sampling time and would likely differ compared to other areas in the world and/or



344 other periods of the year. Whether the rain-borne microbes will be successful in their new  
345 environment will vary depending upon the microbe's viability and physiology, as well as  
346 atmospheric conditions (64). These new organisms could impact the local water quality and  
347 ecology, but the exact effects will depend upon their functional capabilities.

348         The ability to reproduce under different conditions is very important to dispersal success  
349 (15). *Trebouxia*, the most abundant green algae detected in our rain sequences, can be considered  
350 a “generalist” due to its ability to grow in freshwater and brackish water conditions (15). If these  
351 green algae can reproduce after atmospheric transport and deposition, their colonization could  
352 impact their new environments, just as aerially dispersed bacteria and fungi can affect the  
353 microbial composition and functioning of the environments they enter (65). Previously, the  
354 primary production rate of autotrophic communities in water was reported to be negatively  
355 impacted by aeolian-transported microbes (66). Notably, green algae were suggested to be better  
356 colonizers of water than aerially dispersed diatoms or cyanobacteria (67).

357         The transported microbes could also affect human health. It has been speculated that the  
358 allergenic alga *Gonyostomium semen* could be aerially dispersed, thus leading to its expansion  
359 into new ecosystems (8). Here, phylotypes associated with toxin-producing oxygenic  
360 photoautotrophs including *Microcystis*, *Scytonema*, *Tychonema*, and *Anabaena*, whose toxins  
361 can have varying effects upon humans and other microbes (68-70), were detected in clouds  
362 (Figure 2) and rain (Figure 3). This illustrates the aerial spread of microbes of health-related and  
363 ecological interests and the possibility that they enter ecosystems with precipitation. Once they  
364 enter a new environment such as Aydat Lake, they could produce toxins which could affect  
365 human health or even the health of other algae, which are important to primary production. Even  
366 if a microbe does not establish a lasting presence in a system, it still could have long-term

367 impacts upon the resident microbial community (71). Through these methods, aeolian  
368 transported oxygenic photoautotrophs could impact the health and biogeochemical functioning of  
369 a lake ecosystem.

## 370 **Conclusions and Perspectives**

371 Our study indicates that there are diverse oxygenic photoautotrophs among the microbial  
372 assemblages carried by the atmosphere. We detected cyanobacteria, diatoms, as well as red,  
373 green, and golden algae as persistent members of cloud microbial assemblages. Many of the  
374 organisms detected were associated with stress tolerant phylotypes that have widespread  
375 distributions, and some were toxin producers with potential ecological and health-related  
376 impacts. We were able to culture some of them, mostly the aeroterrestrial epiphytic green algae  
377 of the Chlorellaceae. Cyanobacteria were not detected in cloud water enrichments under the  
378 conditions utilized, but molecular approaches indicated that these are present and probably active  
379 in clouds. Genes encoding for part of photosystem II of various cyanobacteria were present in  
380 metagenomes. Rain acted to mediate the dispersal of many of these diverse oxygenic  
381 photoautotrophs, with members of *Chroococciopsidales* and Trebouxiales being notable, into  
382 new ecosystems. Despite small sample sizes and short time periods, our study provides key  
383 information on the aeolian transport of cyanobacteria and green algae that is currently missing.  
384 Microbiologists, ecologists, and meteorologists should work together to conduct long-term  
385 studies to identify the sources of the oxygenic photoautotrophic microbes in rain and determine  
386 their emission/depositional fluxes. Aeolian dispersal is more conducive to certain microbes and  
387 understanding the physicochemical effects on their dispersal patterns and the exact biological  
388 and ecological implications of their dispersal is an understudied aspect of aerobiology.

## 389 MATERIALS AND METHODS

### 390 Cloud and rain sampling

391 Clouds and rain were sampled from meteorological stations operated by the Observatoire  
392 de Physique du Globe de Clermont-Ferrand (OPGC) and parts of the national COPDD  
393 atmospheric survey network.

394 Clouds were sampled at the puy de Dôme (PDD) station (1465 m a.s.l., 45.7720° N,  
395 2.9655° E) on 01/06/2010, 08/06/2010, and 29/06/2017 (dd/mm/yyyy) as described previously  
396 (38) (see Table S1 for sampling details). Clouds were sampled again at the same site on  
397 11/10/2013, 14/10/2013, and 05/11/2013 (sample names 11102013, 14102013, 05112013,  
398 respectively), as described previously (2, 24). The cloud water samples were filtered through a  
399 47 mm, 0.22 µm porosity membrane (MoBio 14880-50-WF) in a laminar flow hood, cut in half,  
400 and half-filters were stored at -80°C prior to use for metagenomic analyses following previous  
401 protocols (24). Backward trajectories of the air masses were generated with HYSPLIT (Figure  
402 S1) (72).

403 Rain was collected at the Opme meteorological station (680 m a.s.l., 45.712500° N,  
404 3.090278° E) with an automated wet-deposition sampler (Eigenbrodt NSA 181/KHS), as  
405 previously described (73). Briefly, eight autoclaved 1-L glass bottles were kept at 4°C in the  
406 chamber with a new bottle designated for collection every 24 hours. Before putting new bottles  
407 in the carousel, the funnel was disinfected with 70% ethanol and rinsed with sterile water. The  
408 funnel was covered by a polytetrafluoroethylene (PTFE) lid that automatically opened when  
409 precipitation occurred, and which prevented dry deposition between rain sampling events. All  
410 rain water samples were collected and transported to the laboratory within 24 hours of the  
411 precipitation event. Sixteen rain samples were collected between 23 May, 2018 and 13 June,

412 2018. The rain water was filtered through 0.22  $\mu\text{m}$  porosity membranes (MoBio 14880-50-WF)  
413 and membranes were stored at  $-20^{\circ}\text{C}$  for later DNA extraction. The pH of filtered rain was  
414 determined and ion concentrations were measured by ion chromatography (Dionex, Sunnyvale,  
415 CA, USA) as previously described (2). The backward trajectories of the air masses (Figures S2)  
416 were generated with HYSPLIT (72).

### 417 **Cultivation of photosynthetic microbes**

418 Cloud water samples collected on 01/06/2010, 08/06/2010, and 29/06/2017  
419 (dd/mm/yyyy) were incubated in BG11 medium (ATCC 616) under constant light ( $75 \mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ )  
420 <sup>1</sup>) at  $20^{\circ}\text{C}$  in the INFORS HT Multitron II equipped with Sylvania Gro-Lux F15W/Gro lights  
421 while shaking at 80 rpm to produce cultures named 01062010, 08062010, and 29062017,  
422 respectively. We used BG11 medium, which is designed to favor the growth of cyanobacteria  
423 and chlorophytes from freshwater environments, not diatoms. The study site is far from the  
424 ocean, thus we focused on cultivating organisms that could successfully survive if deposited in  
425 the local environment. This excluded marine organisms. The 01062010 and 08062010 cultures  
426 were preserved in 15% glycerol and stored at  $-20^{\circ}\text{C}$  for about seven years. Once the 29062017  
427 culture was established, the others were revived, and all were incubated in BG11 media at  $15^{\circ}\text{C}$   
428 under constant illumination ( $75 \mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ ) at 80 rpm. After incubating for 7 days, 5 mL of each  
429 culture was centrifuged at  $8,000 \times g$  for 3 minutes and the cell pellet was stored at  $-20^{\circ}\text{C}$  until  
430 DNA extraction.

431 All rain samples were cultivated in MWC (28) or BG11 media. Visual observation of  
432 green tinted cultures indicated successful enrichment of oxygenic photoautotrophs.

### 433 **DNA Extractions**

434 DNA from the cloud enrichment cultures was extracted with a modified phenol-  
435 chloroform based protocol (74). A 50 mM glucose/10 mM EDTA/ 25 mM Tris-HCl solution was  
436 added to the cell pellet, then five freeze-thaws were performed by freezing at -80°C in ethanol  
437 (200 proof) and heating to 55°C. Lysozyme (0.4 mg L<sup>-1</sup>) and EDTA (0.5 M) were added, and  
438 samples were shaken at 100 rpm at room temperature for 5 min. Ten percent SDS was added and  
439 followed quickly by extraction with phenol: chloroform: isoamyl alcohol (25:24:1). The  
440 emulsion was separated by centrifuging at 16,000 × g for 3 min. The resulting aqueous phase  
441 was extracted with phenol: chloroform: isoamyl alcohol and centrifuged again. The aqueous  
442 phase from the second extraction was precipitated with 200 proof ethanol (1 mL) and 2 µg  
443 glycogen.

444 DNA was extracted from half of the cloud filters and from the whole rain filter using a  
445 MoBio PowerWater isolation kit (now, Qiagen, Hilden, Germany) following the manufacturer's  
446 protocol.

#### 447 **Cloud Enrichment PCR and Sequencing**

448 DNA extracts from cloud enrichments were sent for PCR amplification and sequencing  
449 by MR DNA (www.mrdnalab.com, Shallowater, TX, USA). The *16S rRNA* amplicons were  
450 generated using the primer set 515f/806r (75) with a barcode on the forward primer (Table S2).  
451 HotStarTaq Plus Master Mix Kit (Qiagen, USA) was used for amplification with the following  
452 conditions: 94°C for 3 min, followed by 30 cycles of 94°C for 30 sec, 53°C for 40 sec and 72°C  
453 for 1 min, and ending with a final elongation at 72°C for 5 min. To generate the *18S rRNA* gene  
454 amplicon, PCR using 1391f (76) and Eukbr (77), with a barcode on the forward primer (Table  
455 S2) was performed with HotStarTaq Plus Master Mix Kit (Qiagen, USA) under the following  
456 conditions: 94°C for 3 min, followed by 35 cycles of 94°C for 30 sec, 53°C for 40 sec and 72°C

457 for 1 min, and ending with a final elongation at 72°C for 5 min. The products were confirmed on  
458 a 2% agarose gel. The *16S rRNA* gene amplicons were sequenced with an Ion S5 XL instrument  
459 (Ion 530 chip) following the manufacturer's guidelines. The *18S rRNA* gene amplicons were  
460 combined in equal proportions and purified using calibrated Ampure XP beads (Beckman  
461 Coulter). These were used to prepare the DNA library for Illumina MiSeq 2×300 bp paired end  
462 sequencing following the manufacturer's protocols. The corresponding sequence files were  
463 deposited at the EMBL-EBI European Nucleotide Archive under the study accession number  
464 PRJEB35708.

### 465 **Rain Samples PCR and Sequencing**

466 For the DNA extracts from rain samples, portions of the *16S rRNA* genes targeting  
467 cyanobacteria (31) and of the *18S rRNA* genes targeting green algae (78) were PCR-amplified.  
468 The *16S rRNA* gene PCR reaction was composed of 34.3 µL sterile water, 5 µL 10X Green PCR  
469 buffer without Mg (ThermoFisher Scientific), 0.2 µL Platinum<sup>TM</sup> Taq DNA Polymerase (10 U  
470 µL<sup>-1</sup>; ThermoFisher Scientific), 3 µL 25 mM MgCl<sub>2</sub>, 1 µL 10 mM dNTPs, 2 µL 10 µM forward  
471 primer Cya359 F (31), 2 µL 10 µM reverse primer Cya 781 R modified from Nübel et al. (31),  
472 and 1 µL of template DNA (5-10 ng). The primer sequence of the modified reverse primer from  
473 Nubel et al. (31) was 5' GACTACWGGGGTATCTAATCCCWTT 3'. The conditions were as  
474 followed: 94°C for 2 min, then 35 cycles of 94°C for 30 sec, 55°C for 30 sec, 72°C for 30 sec,  
475 and a final extension of 72°C for 5 min. For the *18S rRNA* gene PCR 33.7 µL sterile water, 5 µL  
476 10X Green PCR buffer, 0.3 µL *Taq* DNA polymerase (5 U µL<sup>-1</sup>; Fischer BioReagents<sup>TM</sup>;  
477 ThermoFisher Scientific), 3 µL 50 mM MgCl<sub>2</sub>, 1 µL 10 mM dNTPs, 2 µL 10 µM EUK528f (32),  
478 2 µL 10 µM CHLO02r (33), and 3 µL of template DNA (5-10 ng) were mixed. The conditions  
479 (78) were as follows: 94°C for 5 min., then 35 cycles of 94°C for 30 s., 55°C for 30 s., 72°C for 1

480 min., and a final extension of 72°C for 5 min. Both sets of amplicons were purified with  
481 MinElute PCR purification kit (Qiagen). Negative controls and positive controls were processed  
482 alongside the samples. The positive control for the *16S rRNA* and *18S rRNA* gene amplifications  
483 were DNA extracts from cultures of *Microcystis aeruginosa* PCC 9443 and an enrichment  
484 culture of rain water collected on 27/05/2018 in MWC media (28) containing algae (27052018),  
485 respectively. Purified amplicons from the rain sample DNA extracts were sent to GenoScreen  
486 (Lille, France) and library preparation followed the Metabiote®v2.0 protocol (GenoScreen, Lille,  
487 France) using 5 ng amplicon DNA. Library preparation quality was confirmed by capillary gel  
488 electrophoresis. The samples were sequenced on one line with a 2×250-bp paired-end MiSeq  
489 Illumina instrument. The sequence files were deposited at the EMBL-EBI European Nucleotide  
490 Archive under the study accession numbers PRJEB35600, samples ERS4058723 to  
491 ERS4058754.

#### 492 **Bioinformatic analysis of PCR amplicons**

493 The *18S rRNA* amplicons from photosynthetic cloud enrichment cultures were  
494 demultiplexed with sabre (<https://github.com/najoshi/sabre>) and paired ends were assembled  
495 with the default parameters of the DADA2 (79) plugin in QIIME2 (v2020.2) (80). The *16S rRNA*  
496 amplicons from oxygenic photoautotrophic cloud enrichment cultures were demultiplexed with  
497 software from MR DNA LLC (<http://www.mrdnalab.com/mrdnafreesoftware/>). Both the *16S* and  
498 *18S rRNA* gene rain amplicons were demultiplexed with the CASAVA v1.0 software (Illumina)  
499 and the paired-ends were assembled with the FLASH tool (81) with trimming of poor-quality  
500 reads (phred score <30) and a 97% similarity requirement.

501 All fastq files were imported into QIIME2 (v2020.2) (80), trimmed and denoised with  
502 DADA2 (79) and grouped into amplicon sequence variants (ASVs). The ASVs were then

503 clustered into operational taxonomic units (OTUs) (97%) with the vsearch (82) plugin. A naive  
504 Bayes scikit-learn classifier (83, 84) for each primer set was trained using extracted reference  
505 full-length operon reads of the Silva v132 database (85) and used for classification. Taxa  
506 associated with negative controls and prokaryotic reads in eukaryotic-targeted amplicons (and  
507 vice versa) were filtered out from the samples. OTUs, that were of interest and lacked taxonomic  
508 resolution from the identification, were queried manually against NCBI BLAST  
509 (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) with the BLASTn algorithm in May 2020. Samples  
510 were rarefied and analyzed for Shannon's diversity, Chao1, and observed OTUs in QIIME2  
511 (v2020.2).

## 512 **Cloud metagenomes**

513 DNA extracts from cloud water samples 11102013, 14102013, and 05112013 were  
514 amplified by MDA using GenomiPHI (GE Healthcare) to aid in generating sequencing libraries.  
515 The libraries were prepared, sequenced, and analyzed as previously described (24) using an  
516 Illumina MiSeq 2 × 300 bp (Genoscreen, Lille, France). The sequence files were deposited at  
517 the EMBL-EBI European Nucleotide Archive under the study accession numbers PRJEB35574,  
518 samples ERS4058712 to ERS4058714. Methods for the taxonomic and functional annotation of  
519 the sequences have been described previously (24). In brief, taxonomic annotations were  
520 assigned based upon the SILVA database (85). The oxygenic photoautotrophs were identified  
521 based upon the taxonomy. The functional annotations were assigned based upon the UniprotKB  
522 protein database (86). With the UniprotKB identifiers, associated genera and gene ontologies  
523 identifiers (GO IDs) (87) were assigned. GO IDs associated with photosynthesis were selected.  
524 Additional cloud metagenomes obtained from Amato et al. (24) were reanalyzed to investigate



525 the presence of oxygenic photoautotrophic microbes (sequences files ENA accession numbers  
526 ERS2351639 to ERS2351641).

### 527 **Re-analysis of cloud RNA/DNA ratios**

528 Previously generated prokaryotic and eukaryotic ribosomal DNA and cDNA amplicons  
529 were re-assessed (2). Sequence files accession number in NCBI BioProject PRJNA380262 were  
530 reanalyzed to examine the potential activity of photosynthetic microbes in clouds. Known  
531 oxygenic photoautotrophs were manually selected from the OTUs based upon their taxonomic  
532 identification. The number of reads for each oxygenic photoautotrophic OTU were expressed as  
533 a percentage of the total community for the cDNA (representing rRNA) and DNA sequences.  
534 The RNA/DNA ratio for each OTU was calculated using the relative percentages. The OTUs  
535 were grouped at the phylum or class level and the RNA/DNA ratios were averaged for each  
536 taxonomic group. A taxon was considered to be active if the ratio was greater than 1, meaning  
537 that the relative abundance of the taxon was greater in the rRNA amplicons rather than in the  
538 *rRNA* gene amplicons. Phantom taxa, phylotypes present in the rRNA but not the *rRNA* genes  
539 (25), were excluded in this analysis.

### 540 **Enumeration of chlorophyll *a*-containing cells**

541 Triplicates of fresh rain water samples of 0.45 mL were fixed with a final concentration  
542 of 0.5% glutaraldehyde and stored at 4°C in the dark until flow cytometry analysis. Just before  
543 analysis, fixed samples were amended with 1 vol of Tris-EDTA buffer (40 mM Tris; 1 mM  
544 EDTA; pH 8.0) and analyzed by a BD FACSCalibur flow cytometer (Becton Dickinson,  
545 Franklin Lakes, NJ) to enumerate chlorophyll *a*-containing cells. Detection was on FL3  
546 (excitation 488 nm; emission 670 nm) with detection thresholds set at 500 mV for FL3 and 380  
547 mV for SSC. Acquisition time was 315-400 s. at a flow rate between 66 – 72  $\mu\text{L min}^{-1}$  (high

548 setting). The flow rate was determined precisely for each sample through the acquisition time  
549 and the gravimetric loss from the water sample. All data was analyzed in the software BD  
550 CellQuest Pro™ (Becton Dickinson, Franklin Lakes, NJ). Chlorophyll *a*-containing cells were  
551 expressed both as rainwater concentration (cells mL<sup>-1</sup>) and as a surficial flux (cells cm<sup>-2</sup> day<sup>-1</sup>).  
552 The flux was calculated from the total number of cells in a rain sample for a given day divided  
553 by the area of the collection surface.

#### 554 **Statistical Analyses**

555 Statistical analyses were conducted in OriginPro Version 2019 9.6 (OriginLab  
556 Corporation, Northampton, MA, USA). To correlate the presence of phylotypes with geographic  
557 origin, the two types of air masses were defined and statistically analyzed as described in Pouzet  
558 et al. (73).

#### 559 **Data Availability**

560 The sequence files were deposited to the European Nucleotide Archive with project  
561 accession numbers PRJEB35574 (cloud metagenomes; ERS2351639 to ERS2351641),  
562 PRJEB35708 (cloud enrichments), and PRJEB35600 (rain samples; ERS4058723 to  
563 ERS4058754).

#### 564 **ACKNOWLEDGEMENTS**

565 This research was funded by the National Science Foundation CBET award #1512511  
566 and a grant from Rutgers Global awarded to DEF, the CNRS EC2CO MicroBiEN program  
567 “FONCOMIC” awarded to PA, and a Graduate Student Grant-in-Aid from the New Jersey Water  
568 Resources Research Institute awarded to KPD. This material is based upon research supported by  
569 the Chateaubriand Fellowship of the Office for Science & Technology of the Embassy of France  
570 in the United States and the Centre National de la Recherche Scientifique (CNRS). KPD was

571 also supported by an Air Pollution Educational and Research Grant (APERG) from the Mid-  
572 Atlantic States Section of the Air and Waste Management Association (A&WMA). The authors  
573 are thankful to J.L. Baray, L. Deguillaume, and the Observatoire de Physique du Globe de  
574 Clermont Ferrand staff for operating the sampling site, allowing access, and sharing  
575 meteorological data. Thank you to A.H. Lejeune of Laboratoire Microorganismes: Génome et  
576 Environnement, UMR 6023 CNRS and the Université Clermont Auvergne for providing positive  
577 controls for PCR amplification of *rRNA* genes and Prof. P.J. Morin of Rutgers University for  
578 sharing equipment.

579 **Competing Interests:** The authors declare no competing interests.

## 580 REFERENCES

- 581 1. Lighthart B. 2000. Mini-review of the concentration of variations found in the al fresco  
582 atmospheric bacterial populations. *Aerobiologia* 16:7-16.
- 583 2. Amato P, Joly M, Besaury L, Oudart A, Taib N, Mone A, Deguillaume L, Delort A-M,  
584 Debroas D. 2017. Active microorganisms thrive among extremely diverse communities  
585 in cloud water. *PLoS ONE* 12:e0182869.
- 586 3. Frölich-Nowoisky J, Kampf C, Weber B, Huffmann A, Pöhlker C, Andreae M, Lang-  
587 Yona N, Burrows S, Gunthe S, Elbert W, Su H, Hoor P, Thines E, Hoffmann T, Després  
588 V, Pöschl U. 2016. Bioaerosols in the Earth system: Climate, health, and ecosystem  
589 interations. *Atmospheric Research* 182:346-376.
- 590 4. Möhler O, DeMott P, Vali G, Levin Z. 2007. Microbiology and atmospheric processes:  
591 the role of biological particles in cloud physics. *Biogeosciences* 4:1059-1071.
- 592 5. Morris C, Sands D, Bardin M, Jaenicke R, Vogel B, Leyronas C, Ariya P, Psenner R.  
593 2011. Microbiology and atmospheric processes: research challenges concerning the  
594 impact of airborne micro-organisms on the atmosphere and climate. *Biogeosciences* 8:17-  
595 25.
- 596 6. Barberán A, Henley J, Fierer N, Casamayor E. 2014. Structure, inter-annual recurrence,  
597 and global-scale connectivity of airborne microbial communities. *Sci Tot Environ*  
598 487:187-195.
- 599 7. Delort A-M, Vaitilingom M, Amato P, Sancelme M, Parazols M, Mailhot G, Laj P,  
600 Deguillaume L. 2010. A short overview of the microbial population in clouds: Potential  
601 roles in atmospheric chemistry and nucleation processes. *Atmos Res* 98:249-260.
- 602 8. Tesson SVM, Skjoth C, Santl-Temkiv T, Londahl J. 2016. Airborne Microalgae: Insights,  
603 Opportunities, and Challenges. *Appl Environ Microbiol* 82.
- 604 9. Murby A, Haney J. 2016. Field and laboratory methods to monitor lake aerosols for  
605 cyanobacteria and microcystins. *Aerobiologia* 32:395-403.

- 606 10. Schlichting H. 1969. The Importance Of Airborne Algae and Protozoa. *J Air Pollut*  
607 *Control Assoc* 19:946-951.
- 608 11. Sharma N, Rai A, Singh S, Brown R. 2007. Airborne Algae: Their Present Status and  
609 Relevance. *J Phycol* 43:615-627.
- 610 12. Gaylarde P, Gaylarde C. 2000. Algae and cyanobacteria on painted buildings in Latin  
611 America. *Int Biodeter Biodegr* 46:93-97.
- 612 13. Rosas I, Roy-Ocotla G, Mosino P. 1989. Meteorological effects on variation of airborne  
613 algae in Mexico. *Int J Biometeorol* 33:173-179.
- 614 14. Genitasaris S, Kormas K, Moustaka-Gouni M. 2011. Airborne Algae and Cyanobacteria:  
615 Occurrence and Related Health Effects. *Front Biosci* E3:772-787.
- 616 15. Tesson SVM, Santl-Temkiv T. 2018. Ice nucleation activity and aeolian dispersal success  
617 in airborne and aquatic microalgae. *Front Microbiol* 9:2681.
- 618 16. Genitsaris S, Moustaka-Gouni M, Kormas K. 2011. Airborne microeukaryote colonists in  
619 experimental water containers: diversity, succession, life histories and established food  
620 webs. *Aquat Microb Ecol* 62:139-152.
- 621 17. Curren E, Leong S. 2020. Natural and anthropogenic dispersal of cyanobacteria: a  
622 review. *Hydrobiol* doi:10.1007/s10750-020-04286-y.
- 623 18. Kourtev P, Hill K, Shepson P, Konopka A. 2011. Atmospheric cloud water contains a  
624 diverse bacterial community. *Atmos Environ* 45:5399-5405.
- 625 19. Xu C, Wei M, Chen J, Sui X, Zhu C, Li J, Zheng L, Sui G, Li W, Wang W, Zhang Q,  
626 Mellouki A. 2017. Investigation of diverse bacteria in cloud water at Mt. Tai, China.  
627 *Science of the Total Environment* 580:258-265.
- 628 20. van Overeem M. 1937. On green organisms occurring in the lower troposphere. *Rec Trav*  
629 *Botan Neerl* 3:389-439.
- 630 21. Carson JL, Brown RM. 1976. The correlation of soil algae, airborne algae, and fern  
631 spores with meteorological conditions on the island of Hawaii. *Pacific Science* 30:197-  
632 205.
- 633 22. Harding T, Jungblut A, Lovejoy C, Vincent W. 2011. Microbes in high arctic snow and  
634 implications for the cold biosphere. *Appl Environ Microbiol* 77:3234-3243.
- 635 23. Wiśniewska K, Lewandowska A, Śliwińska-Wilczewska S. 2019. The importance of  
636 cyanobacteria and microalgae present in aerosols to human health and the environment –  
637 Review study. *Environ Int* doi:10.1016/j.envint.2019.104964:104964.
- 638 24. Amato P, Besaury L, Joly M, Penaud B, Deguillaume L, Delort A-M. 2019.  
639 Metatranscriptomic exploration of microbial functioning in clouds. *Sci Rep*  
640 doi:10.1038/s41598-019-41032-4:4383.
- 641 25. Bowsher AW, Kearns PJ, Shade A. 2019. 16S rRNA/rRNA Gene Ratios and Cell  
642 Activity Staining Reveal Consistent Patterns of Microbial Activity in Plant-Associated  
643 Soil. *mSystems* 4.
- 644 26. Steven B, Hesse C, Soghlglan J, Gallegos-Graces L, Dunbar J. 2017. Simulated  
645 rRNA/DNA ratios show potential to misclassify active populations as dormant. *Appl*  
646 *Environ Microbiol* 83:e00696-17.
- 647 27. Geider R. 1987. Light and Temperature Dependence of the Carbon to Chlorophyll a  
648 Ratio in Microalgae and Cyanobacteria: Implications for Physiology and Growth of  
649 Phytoplankton. *New Phytol* 106:1-34.
- 650 28. Guillard R, Lorenzen C. 1972. Yellow-green algae with chlorophyllide c. *J Phycol* 8:10-  
651 14.

- 652 29. Rigonato J, Gama W, Alvaranga D, Branco L, Brandini F, Genuário D, Fiore M. 2016.  
653 *Aliterella atlantica* gen. nov., sp. nov., and *Aliterella antarctica* sp. nov., novel members  
654 of coccoid Cyanobacteria. *Int J Syst Evol Micr* 66:2853-2861.
- 655 30. Suda S, Watanabe M, Otsuka S, Mahakahant A, Yongmanitchai W, Nopartnaraporn N,  
656 Liu Y, Day J. 2002. Taxonomic revision of water-bloom-forming species of oscillatorioid  
657 cyanobacteria. *Int J Syst Evol Micr* 52:1577-1595.
- 658 31. Nubel U, Garcia-Richel F, Muyzer G. 1997. PCR Primers To Amplify 16S rRNA Genes  
659 from Cyanobacteria. *Appl Environ Microbiol* 63:3327-3332.
- 660 32. Elwood HJ, Olsen GJ, Sogin ML. 1985. The Small-Subunit Ribosomal RNA Gene  
661 Sequences from the Hypotrichous Ciliates *Oxytricha nova* and *Stylonychia pustulata*.  
662 *Mol Biol Evol* 2:399-410.
- 663 33. Simon N, Campbell L, Örnolfsdóttir E, Groben R, Guillou L, Lange M, Medlin LK.  
664 2000. Oligonucleotide Probes for the Identification of Three Algal Groups by Dot Blot  
665 and Fluorescent Whole-Cell Hybridization. *J Eukaryot Microbiol* 47:76-84.
- 666 34. Rastogi G, Coaker G, Leveau J. 2013. New insights into the structure and function of  
667 phyllosphere microbiota through high-throughput molecular approaches. *FEMS*  
668 *Microbiol Lett* 348:1-10.
- 669 35. Krüger W. 1894. Beitrage zur Kenntnis der Organismen des Saftflusses (sog.  
670 Scheimflusses) der Laubbaum. II. Ueber zwei aus Saftflüssen rein gezüchtete Algen.  
671 Zopf's Beitr Physiol Morph Organ 4:69-116.
- 672 36. Darienko T, Gustavs L, Mudimu O, Menendez C, Schumann R, Karsten U, Friedl T,  
673 Pröschold T. 2010. *Chloroidium*, a common terrestrial coccoid green alga previously  
674 assigned to *Chlorella* (Trebouxiophyceae, Chlorophyta). *Eur J Phycol* 45:79-95.
- 675 37. Amato P, Parazols M, Sancelme M, Mailhot G, Paolo L, Delort A-M. 2007. An important  
676 oceanic source of micro-organisms for cloud water at the puy de Dome (France). *Atmos*  
677 *Environ* 41:8253-8263.
- 678 38. Vaitilingom M, Attard E, Gaiani N, Sancelme M, Deguillaume L, Flossmann A, Amato  
679 P, Delort A-M. 2012. Long-term features of cloud microbiology at the puy de Dôme  
680 *Atmos Environ* 56:88-100.
- 681 39. Mares J, Johansen J, Hauer T, Zima J, Ventura S, Cuzman O, Tiribilli B, Kastovsky J.  
682 2019. Taxonomic resolution of the genus *Cyanothece* (Chroococcales, Cyanobacteria),  
683 with a treatment on *Gloeothece* and three new genera, *Crocospaera*, *Rippkaea*, and  
684 *Zehria*. *J Phycol* 55:578-610.
- 685 40. Ludwig M, Bryant D. 2012. *Synechococcus* sp. strain PCC 7002 transcriptome:  
686 acclimation to temperature, salinity, oxidative stress, and mixotrophic growth conditions.  
687 *Front Microbiol* 3:354.
- 688 41. Sharma N, Singh H. 2010. Differential Aerosolization of Algal and Cyanobacterial  
689 Particles in the Atmosphere. *Indian J Microbiol* 50:468-473.
- 690 42. Gregory P, Hamilton E, Sreeramulu T. 1955. Occurrence of the alga *Gloeocapsa* in the  
691 air. *Nature* 176:1270.
- 692 43. Sherwood A, Wade R, Conklin K. 2020. Seasonality of tropical airborne algae: a 16-  
693 month study based on high-throughput sequencing in the Hawaiian Islands. *Grana*  
694 doi:10.1080/00173134.2020.1738541.
- 695 44. Saw J, Schatz M, Brown M, Kunkel D, Foster J, Shick H, Christensen S, Hou S, Wan X,  
696 Donachie S. 2013. Cultivation and Complete Genome Sequencing of *Gloeobacter*

- 697 kilauensis sp. nov., from a Lava Cave in Kilauea Caldera, Hawai'i. PLoS ONE  
698 8:e76376.
- 699 45. Jung P, Schermer M, Briegel-Williams L, Baumann K, Leinweber P, Karsten U, Lehnert  
700 L, Achilles S, Bendix J, Büdel B. 2019. Water availability shapes edaphic and lithic  
701 cyanobacterial communities in the Atacama Desert. *J Phycol* 55:1306-1318.
- 702 46. Hájek J, Váczi P, Barták M, Jahnová L. 2012. Interspecific differences in cryoresistance  
703 of lichen symbiotic algae of genus *Trebouxia* assessed by cell viability and chlorophyll  
704 fluorescence. *Cryobiology* 64:215-222.
- 705 47. Dong K, Woo C, Yamamoto N. 2019. Plant assemblages in atmospheric deposition.  
706 *Atmos Chem Phys* 19:11969-11983.
- 707 48. Cáliz J, Triadó-Margarit X, Camarero L, Casamayor E. 2018. A long-term survey unveils  
708 strong seasonal patterns in the airborne microbiome coupled to general and regional  
709 atmospheric circulations. *Proc Natl Aca Sci* 115:12229-12234.
- 710 49. Tragin M, Vaultot D. 2018. Green microalgae in marine coastal waters: The Ocean  
711 Sampling Day (OSD) dataset. *Sci Rep* 8.
- 712 50. Tignat-Perrier R, Dommergue A, Thollot A, Keuschnig C, Magand O, Vogel TM, Larose  
713 C. 2019. Global airborne microbial communities controlled by surrounding landscapes  
714 and wind conditions. *Sci Rep* 9.
- 715 51. Hodač L, Hallmann C, Spitzer K, Elster J, Faßhauer F, Brinkmann N, Lepka D, Diwan V,  
716 Friedl T. 2016. Widespread green algae *Chlorella* and *Stichococcus* exhibit polar-  
717 temperate and tropical-temperate biogeography. *FEMS Microbiol Ecol* 92:fiw122.
- 718 52. Holzinger A, Karsten U. 2013. Desiccation stress and tolerance in green algae:  
719 consequences for ultrastructure, physiological, and molecular mechanisms. *Front Plant*  
720 *Sci* 4:327.
- 721 53. Ehresmann DW, Hatch MT. 1975. Effect of Relative Humidity on the Survival of  
722 Airborne Unicellular Algae. *Appl Microbiol* 29:352-357.
- 723 54. Nelson DR, Khraiwesh B, Fu W, Alseekh S, Jaiswal A, Chaiboonchoe A, Hazzouri KM,  
724 O'Connor MJ, Butterfoos GL, Drou N, Rowe JD, Harb J, Fernie AR, Gunsalus KC,  
725 Salehi-Ashtiani K. 2017. The genome and phenome of the green alga *Chloroidium* sp.  
726 UTEX 3007 reveal adaptive traits for desert acclimatization. *eLife* 6:e25783.
- 727 55. Chiu C-S, Chiu P-H, Yong T, Tsai H-P, Soong K, Huang H-E, Chen C-N. 2020.  
728 Mechanisms protect airborne green microalgae during long distance dispersal. *Sci Rep*  
729 10:13984.
- 730 56. Karsten U, Friedl T, Schumann R, Hoyer K, Lembcke S. 2005. Mycosporine-like amino  
731 acids and phylogenies in green algae: *Prasiola* and its relatives from the  
732 *Trebouxiophyceae* (Chlorophyta). *J Phycol* 41:557-566.
- 733 57. Chandra R, Pons-Faudoa F, Saldivar R, Rittmann B. 2020. Effect of ultra-violet exposure  
734 on production of mycosporine-like amino acids and lipids by *Lyngbya purpurem*.  
735 *Biomass Bioenerg* 134:105475.
- 736 58. Singh H. 2018. Desiccation and radiation stress tolerance in cyanobacteria. *J Basic*  
737 *Microbiol* 58:813-826.
- 738 59. Smith P. 1973. The effects of some air pollutants and meteorological conditions on  
739 airborne algae and protozoa. *J Air Pollut Control Assoc* 23:876-880.
- 740 60. Joly M, Amato P, Sancelme M, Vinatier V, Abrantes M, Deguillaume L, Delort A-M.  
741 2015. Survival of microbial isolates from clouds toward simulated atmospheric stress  
742 factors. *Atmos Environ* 117:92-98.

- 743 61. Hubálek Z. 2003. Protectants used in the cryopreservation of microorganisms.  
744 Cryobiology 46:205-229.
- 745 62. Day J. 2007. Cryopreservation of Microalgae and Cyanobacteria. In Day J, Stacey G (ed),  
746 Cryopreservation and Freeze-Drying Protocols. Humana Press Inc., Totowa, NJ.
- 747 63. Ellegaard M, Ribeiro S. 2018. The long-term persistence of phytoplankton resting stages  
748 in aquatic 'seed banks'. Biol Rev 93:166-183.
- 749 64. Alsvéd M, Holm S, Christiansen S, Smidt M, Ling M, Boesen T, Finster K, Bilde M,  
750 Löndahl J, Santl-Temkiv T. 2018. Effect of aerosolization and drying on the viability of  
751 *Pseudomonas syringae* cells. Front Microbiol 9:3086.
- 752 65. Evans S, Bell-Dereske L, Dougherty K, Kittredge H. 2019. Dispersal alters soil microbial  
753 community response to drought. Environ Microbiol doi:10.1111/1462-2920.14707.
- 754 66. Mescioglú E, Rahav E, Frada M, Rosenfeld S, Raveh O, Galletti Y, Santinelli C, Herut B,  
755 Paytan A. 2019. Dust-Associated Airborne Microbes Affect Primary and Bacterial  
756 Production Rates, and Eukaryotes Diversity, in the Northern Red Sea: A Mesocosm  
757 Approach. Atmosphere 10:358.
- 758 67. Chrisostomou A, Moustaka-Gouni M, Sgardelis S, Lanaras T. 2009. Air-dispersed  
759 phytoplankton in a Mediterranean River-Reservoir system (Aliakmon-Polyphytos,  
760 Greece). J Plankton Res 31:877-884.
- 761 68. Rastogi R, Madamwar D, Incharoensakdi A. 2015. Bloom Dynamics of Cyanobacteria  
762 and Their Toxins: Environmental Health Impacts and Mitigation Strategies. Front  
763 Microbiol 6:1254.
- 764 69. Tan K, Huang Z, Ji R, Qiu Y, Wang Z, Liu J. 2019. A review of allelopathy on  
765 microalgae. Microbiology 165:587-592.
- 766 70. Bubak I, Sliwínska-Wilczewska S, Glowacka P, Szczerba A, Mozdzen K. 2020. The  
767 Importance of Allelopathic Picocyanobacterium *Synechococcus* sp. on the Abundance,  
768 Biomass Formation, and Structure of Phytoplankton Assemblages in Three Freshwater  
769 Lakes. Toxins 12:259.
- 770 71. Amor D, Ratzke C, Gore J. 2020. Transient invaders can induce shifts between  
771 alternative stable states of microbial communities. Sci Adv 6:eaay8676.
- 772 72. Draxler RR, Rolph GD. 2013. HYSPLIT (HYbrid Single-Particle Lagrangian Integrated  
773 Trajectory) Model access via NOAA ARL READY Website, on NOAA Air Resources  
774 Laboratory. <https://ready.arl.noaa.gov/HYSPLIT.php>. Accessed
- 775 73. Pouzet G, Peghaire E, Agues M, Baray J-L, Conen F, Amato P. 2017. Atmospheric  
776 Processing and Variability of Biological Ice Nucleating Particles in Precipitation at  
777 Opme, France. Atmosphere 8.
- 778 74. Sakano Y, Kerkhof L. 1998. Assessment of Changes in Microbial Community Structure  
779 during Operation of an Ammonia Biofilter with Molecular Tools. Appl Environ  
780 Microbiol 64:4877-4882.
- 781 75. Caporaso JG, Lauber C, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ,  
782 Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of  
783 sequences per sample. Proc Natl Aca Sci 108:4516-4522.
- 784 76. Lane D. 1991. 16S/23S sequencing, p 115-175. In Stackebrandt E, Goodfellow M (ed),  
785 Nucleic Acid Technologies in Bacterial Systematics. Wiley, NY.
- 786 77. Medlin L, Elwood H, Stickel S, Sogin M. 1988. The characterization of enzymatically  
787 amplified eukaryotic 16S-like rRNA-coding regions. Gene 71:491-499.

- 788 78. Zhu F, Massana R, Not F, Marie D, Vaulot D. 2005. Mapping of picoeucaryotes in  
789 marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol Ecol*  
790 52:79-92.
- 791 79. Callahan B, McMurdie P, Rosen M, Han A, Johnson A, Holmes S. 2016. DADA2: High-  
792 resolution sample inference from Illumina amplicon data. *Nat Methods* 13:581-587.
- 793 80. Bolyen E, Rideout J, Dillon M, Bokulich N, Abnet C, Al-Ghalith G, Alexander H, Alm  
794 E, Arumugam M, Asnicar F, Bai Y, Bisanz J, Bittinger K, Brejnrod A, Brislawn C,  
795 Brown C, Callahan B, Caraballo-Rodriguez A, Chase J, Cope E, DaSilva R, Diener C,  
796 Dorrestein P, Douglas G, Durall D, Duvallet C, Edwardson C, Ernst M, Estaki M,  
797 Fouquier J, Gauglitz J, Gibbons S, Gibson D, Gonzalez A, Gorlick D, Gonzalez A,  
798 Gorlick K, Guo J, Hillman B, Holmes S, Holste H, Huttenhower C, Huttley G, Janssen S,  
799 Jarmusch A, Jiang L, Kaehler B, Kang K, Keefe C, Keim P, et al. 2019. Reproducible,  
800 interactive, scalable and extensible microbiome data science using QIIME 2. *Nat*  
801 *Biotechnol* 37:848-857.
- 802 81. Magoč T, Salzberg S. 2011. FLASH: fast length adjustment of short reads to improve  
803 genome assemblies. *Bioinformatics* 27:2957-2963.
- 804 82. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. 2016. VSEARCH: a versatile open  
805 source tool for metagenomics. *PeerJ* 4:e2584.
- 806 83. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M,  
807 Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D. 2011.  
808 Scikit-learn: Machine Learning in Python. *J Mach Learn Res* 12:2825-2830.
- 809 84. Bokulich N, Kaehler B, Rideout J, Dillon M, Bolyen E, Knight R, Huttley G, Caporaso J.  
810 2018. Optimizing taxonomic classification of marker-gene amplicon sequences with  
811 QIIME 2's q2-feature-classifier plugin. *Microbiome* 6.
- 812 85. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner F.  
813 2013. The SILVA ribosomal RNA gene database project: improved data processing and  
814 web-based tools. *Nucleic Acids Res* 41:D590-596.
- 815 86. Pundir S, Martin M, O'Donovan C. 2017. UniProt Protein Knowledgebase, p 41-55,  
816 Protein Bioinformatics doi:10.1007/978-1-4939-6783-4\_2. Humana Press, New York,  
817 NY.
- 818 87. Consortium TGO. 2017. Expansion of the Gene Ontology knowledgebase and resources.  
819 *Nucleic Acids Res* 45:D331-D338.

820

821

822

823

824

825



826 Table 1: Relative abundance of oxygenic photoautotrophic microbial small subunit (SSU)  
827 ribosomal sequences in cloud metagenomes. All clouds were of marine origin, based on back  
828 trajectories from NOAA HYSPLIT (Figure S1), except for sample 11102013, which was  
829 continental in origin.

Sample		11102013	14102013	05112013	17112014- 1	17112014- 2	17112014- 3
Reference	Common	This study	This study	This study	Amato et al. 2019	Amato et al. 2019	Amato et al. 2019
Taxa	Name	(%)	(%)	(%)	(%)	(%)	(%)
<b>Prokaryotes</b>							
Cyanobacteria	Blue-green algae	0.92	0.51	0.99	0.25	0.27	0.23
<b>Eukaryotes</b>							
Chlorophyta	Green algae	1.2	0.54	0.66	1.1	0.59	0.49
Trebouxiophyceae	Green algae	0.77	0.32	0.49	0.85	0.44	0.29
Rhodophyta	Red algae	0.15	0.054	0.066	0.040	0.022	n/a
Bacillariophyceae	Diatoms	1.8	2.4	1.5	1.5	1.4	1.7
Dinophyceae	Dinoflagellates	0.26	0.075	0.099	0.048	0.32	0.017
Chrysophyceae	Golden algae	1.6	3.1	1.8	2.3	2.6	2.8

830  
831  
832  
833  
834  
835  
836  
837  
838  
839

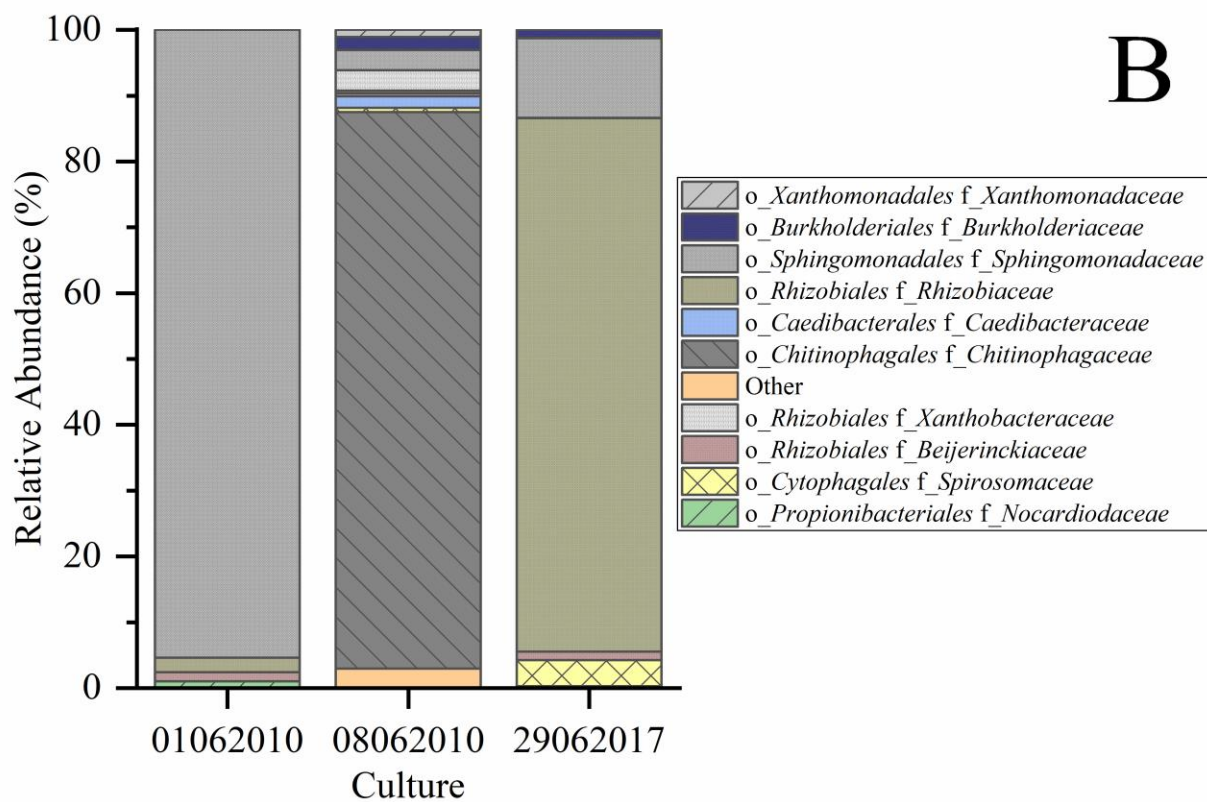
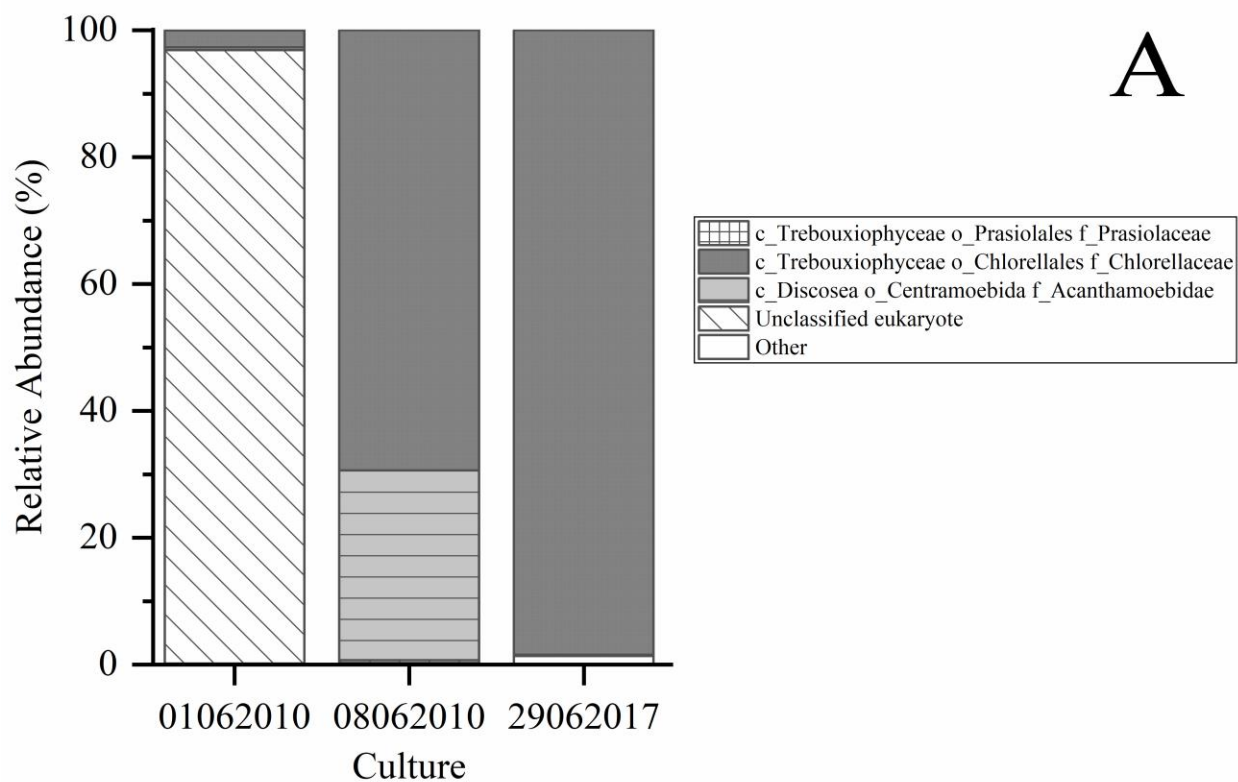
840 Table 2: Calculated RNA:DNA ratios of taxa including photosynthetic microbes in clouds  
 841 (original data from Amato et al. (2)). The relative abundances of the taxa in the RNA and DNA  
 842 reads were used to calculate ratios. The range is the minimum and maximum ratio observed for  
 843 the corresponding taxon in the sample.

Group	Polluted Cloud		Continental Cloud	
	Average $\pm$ Std Dev	Range	Average $\pm$ Std Dev	Range
<b>Prokaryotes</b>				
Cyanobacteria	41 $\pm$ 119	nd* - 1100	11 $\pm$ 23	nd - 120
<b>Eukaryotes</b>				
Chlorophyta	0.85 $\pm$ 0.94	nd - 2.9	1.2 $\pm$ 1.7	nd - 7.1
Cryptophyta	0.59 $\pm$ 0.82	nd - 2.9	0.67 $\pm$ 0.85	nd - 2.8
Bacillariophyta	2.1 $\pm$ 11.7	nd - 101	0.95 $\pm$ 1.6	nd - 10
Dinophyceae	1.9 $\pm$ 5.3	nd - 25	2.2 $\pm$ 3.0	nd - 11
Chrysophyceae	1.7 $\pm$ 5.4	nd - 40	1.3 $\pm$ 1.5	nd - 8.7
Phaeophyceae	1.6 $\pm$ 2.1	nd - 8.7	0.74 $\pm$ 0.57	nd - 2.3

844 *\*nd: not detected*

845

846



847

848 Figure 1: Microbial composition of oxygenic photoautotrophic enrichments. Cloud water was  
849 incubated with BG11 medium at 15°C under constant illumination. (A) The eukaryotic  
850 population of the three cloud enrichments at the family level and their relative abundances in *18S*  
851 *rRNA* gene reads. (B) The taxonomic classification and relative abundance of bacteria in *16S*  
852 *rRNA* gene reads in all three cloud enrichments at the family level.

853

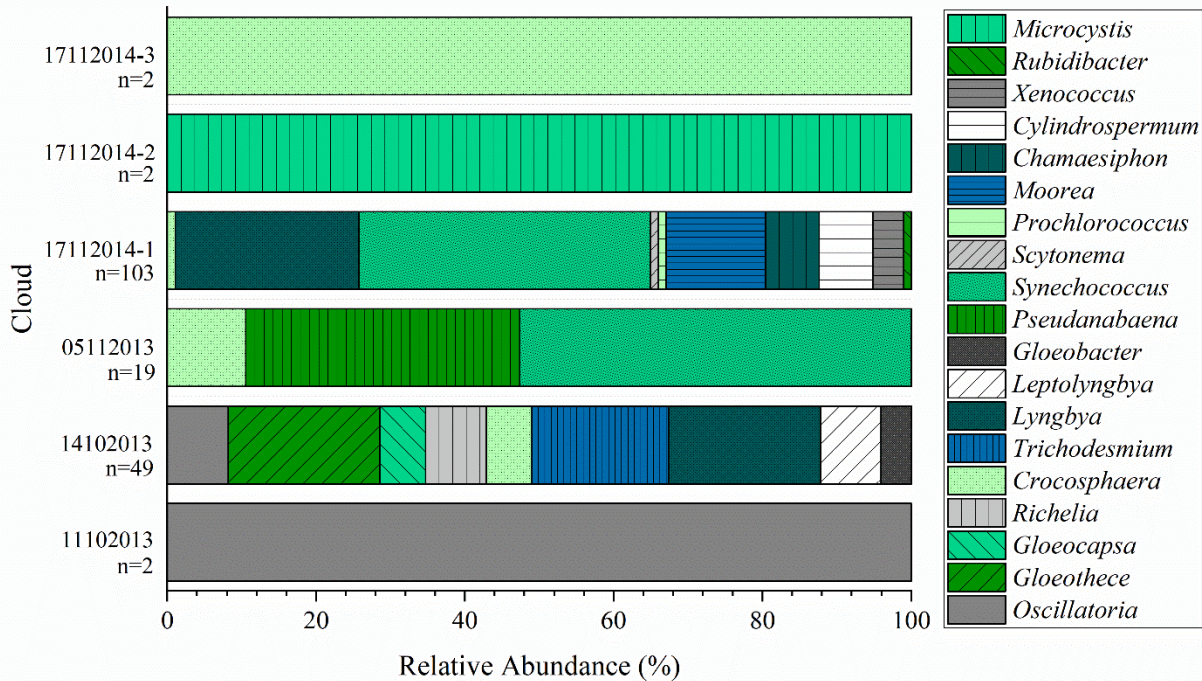
854

855

856

857

858



859

860 Figure 2: Relative abundance of bacterial genera associated with detected photosynthetic genes  
 861 in cloud water. The total number of reads of functional genes related to photosynthesis (n) in  
 862 each cloud metagenome and their associated taxonomy at the genus level were based upon the  
 863 UniProtKB database. All clouds were marine in origin except 11102013 which was continental  
 864 in origin.

865

866

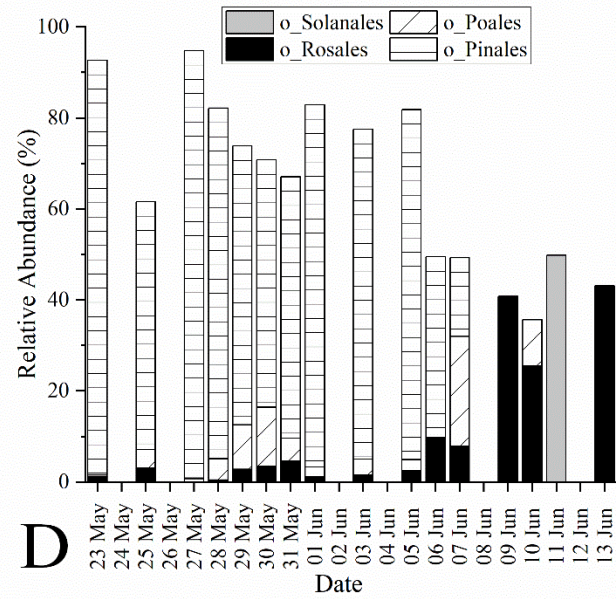
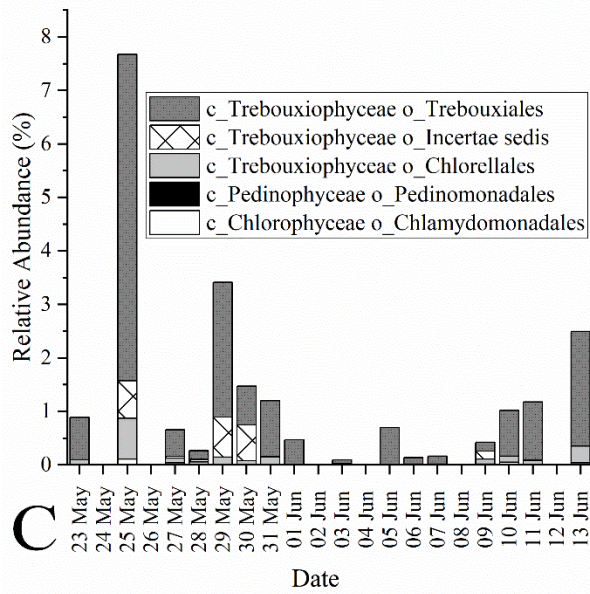
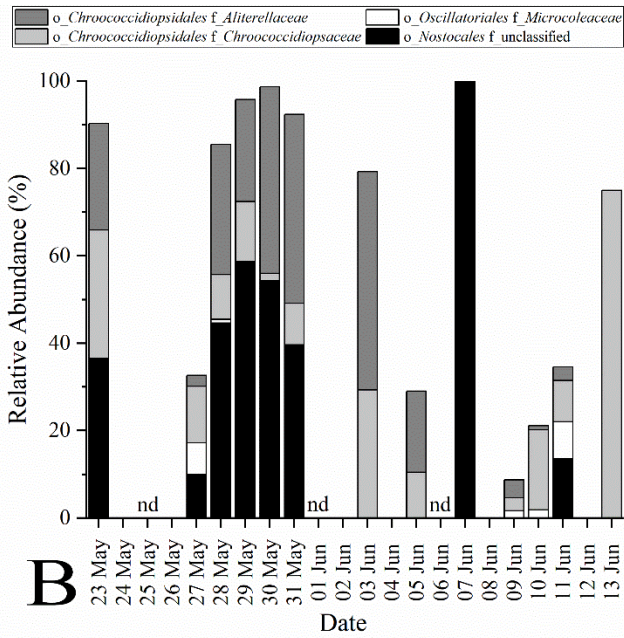
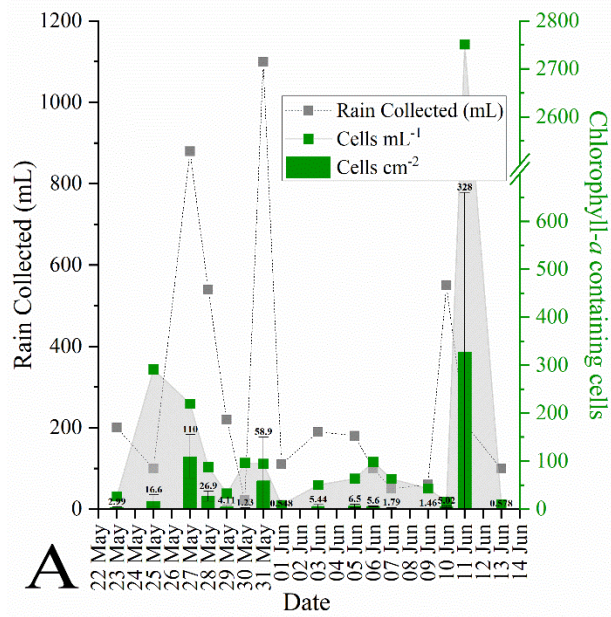
867

868

869

870





872 Figure 3: Abundance of chlorophyll *a*-containing cells in rain and the composition of  
873 cyanobacteria and green algae in rain from 23 May, 2018 to 13 June, 2018. (A) The  
874 concentration of chlorophyll *a*-containing cells in rain samples and the amount of rain collected.  
875 The downward flux is the amount of chlorophyll *a*-containing cells entering the surface  
876 environment per square centimeter of area, as inferred from concentration measurements. There  
877 were no rain events on 24 and 26 May or 02, 04, 08, and 12 June.

878 (B) The composition of *16S rRNA* partial amplicons with primers targeting cyanobacteria from  
879 fresh rain samples collected. The nd designation is for “not detected,” which were days when  
880 none of the phylotypes were detected in the microbial populations. On 01 June, only sequences  
881 associated with chloroplasts were detected. (C) The algal composition of *18S rRNA* partial  
882 amplicons with primers targeting green algae from fresh rain samples collected. (D) The pollen  
883 composition of *18S rRNA* partial amplicons with primers targeting green algae from fresh rain  
884 samples collected. The 48 hour back-trajectories of the air masses from NOAA HYSPLIT for all  
885 sampled rain is presented in Figure S2.