HAL
open science

Identification of a novel type of glucose dehydrogenase involved in the mineral weathering ability of Collimonas pratensis strain PMB3(1)

Laura Picard, Marie-Pierre M.-P. Turpault, P. Oger, Stephane Uroz

To cite this version:

Laura Picard, Marie-Pierre M.-P. Turpault, P. Oger, Stephane Uroz. Identification of a novel type of glucose dehydrogenase involved in the mineral weathering ability of Collimonas pratensis strain PMB3(1). FEMS Microbiology Ecology, 2021, 97 (1), pp.fiaa232. 10.1093/femsec/fiaa232 . hal-03000681

HAL Id: hal-03000681

https://hal.science/hal-03000681v1

Submitted on 5 Jan 2021

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.

TITLE: Identification of a novel type of glucose dehydrogenase involved in the mineral weathering ability of *Collimonas pratensis* strain PMB3(1)

Running title: Mineral weathering by a glucose dehydrogenase

AUTHORS: Picard, L.[1,2], Turpault, M-P. [2], Oger, P.M.[3], Uroz, S. [1,2#]

[1] Université de Lorraine, INRAE, UMR1136 « Interactions Arbres-Microorganismes », F-54280 Champenoux, France

[2] INRAE, UR1138 « Biogéochimie des écosystèmes forestiers », F-54280 Champenoux, France

[3] Université de Lyon, INSA de Lyon, CNRS UMR 5240, Lyon, France

 # Corresponding author: Mailing address:

Université de Lorraine, INRAE, "Interactions Arbres Microorganismes", UMR 1136, 54280 Champenoux, France. Phone: +33 (0)3 83 39 40 81, Fax: +33 (0)3 83 39 40 69. E-mail: stephane.uroz@inra.fr.

Key words: *Collimonas pratensis* PMB3(1), mineral weathering, Glucose/Methanol/Choline oxidoreductase, PQQ-independent glucose dehydrogenase, biotite

ABSTRACT

The exact molecular mechanisms as well as the genes involved in the mineral weathering (MW) process by bacteria remain poorly characterized. To date, a single type of glucose dehydrogenase (GDH) depending on a particular co-factor named pyrroloquinoline quinone (PQQ) is known. These enzymes allow the production of gluconic acid through the oxidation of glucose. However, it remains to be determined how bacteria missing PQQ-dependent GDH and/or the related *pqq* biogenesis genes weather minerals. In this study, we considered the very effective mineral weathering bacterial strain PMB3(1) of *Collimonas pratensis*. Genome analysis revealed that it does not possess the PQQ based system. The use of random mutagenesis, gene complementation and functional assays allowed us to identify mutants impacted in their ability to weather mineral. Among them, three mutants were strongly altered on their acidification and biotite weathering abilities (58 to 75% of reduction compared to WT) and did not produce gluconic acid. The characterization of the genomic regions allowed noticeably to the identification of a Glucose/Methanol/Choline oxidoreductase. This region appeared very conserved among collimonads and related genera. This study represents the first demonstration of the implication of a PQQ-independent GDH in the mineral weathering process and explains how *Collimonas* weather minerals.

INTRODUCTION

Minerals represent physical supports and reservoirs of nutritive elements for the development of the biosphere. Their dissolution, *i.e.*, mineral weathering (MW), participates in the release of their nutritive constituents. This essential process permits to maintain the relative fertility of the soils and to support the growth of the plants. In nutrient-poor and non-amended environments such as forests, minerals are considered as one of the most important stock of nutrients for the long-lasting development of the ecosystem (Ranger and Turpault, 1999). However, MW is a slow process, which depends on the intrinsic properties of those minerals (*i.e.*, dissolution constant) and extrinsic conditions that result from a combination of abiotic and biotic processes. For instance, in acidic draining medium, biotite dissolves faster than quartz because of their intrinsic properties (*i.e.*, chemical composition) (Brantley, 2008). In drained soils, water and acid circulation are known as important abiotic factors involved in the dissolution of minerals. Plants and soil microorganisms are also known to participate in this process (Gadd, 2007; Landeweer*t et a*l. 2001). Tree roots and fungal hyphae can physically weather minerals (*i.e.*, division) through the pressure applied (Bonneville*e et a*l. 2009). Plant development through nutrient pumping and rhizosphere acidification induces stronger MW than in the surrounding bulk soil (Calvaruso et al. 2009; Turpault et al. 2009; Hinsinger et al. 2003). The production of chelating and acidifying agents was also reported for a wide range of plants, symbiotic fungi and bacteria (Jongman*s et al*. 1997; Uro*z et a*l. 2009a).

Effective MW bacteria have been described in various environments, and especially in the rhizosphere of different plants (*i.e.*, trees, cactus) living in nutrient-poor environments (Puente et al. 2004). Notably, several studies evidenced a significant enrichment of these functional communities in the rhizosphere compared to the surrounding bulk soil (Collignon*et a*l. 2011; Calvarus*o et al*. 2010; Zhang *et* al. 2016). In addition, microcosm experiments revealed that

effective MW bacterial strains can improve plant growth. Indeed, the inoculation or co-inoculation of different MW bacterial strains was shown to increase the seedling biomass, the shoot length, the root biomass, the total root length and the total root surface area of *Pinus sylvestris* compared to non-inoculated plants (Calvarus*o et al*. 2006; Koel*e et a*l. 2009). To date, the ability to solubilize mineral has been reported for a wide range of bacterial genera (*i.e*., *Achromobacter, Agrobacterium, Bacillus, Collimonas, Pseudomonas, Rhizobium, Burkholderia, Microccocus, Aereobacter, Flavobacterium* or *Erwinia*) (Puent*e et a*l. 2004; , Collignon*et al*. 2011; Rodriguez and Fraga, 1999; Uro*z et a*l. 2011).

If the ability of the bacteria to weather minerals is well documented, our understanding of the molecular basis and especially of the genes involved in MW remains limited. This gap of knowledge has seriously hampered studies of the ecology, evolution and relative role of the mineral weathering bacteria. To date, no molecular tools exist to assess and quantify the occurrence, diversity, and activity of mineral weathering genes in natural bacterial populations. Heterotrophic bacteria are known to produce chelating agents (*i.e*., siderophores and organic acids) and acidifying metabolites (*i.e*., protons, organic acids), but the molecular demonstrations of their relative role in the MW process are limited as well as is our knowledge of the conservation among bacteria of the molecular markers involved. Considering the organic acid production, several studies showed that the direct oxidation pathway of glucose is involved and more specifically the steps producing gluconic acid (GA) and 2-ketogluconic acid (2KGA) through the activity of a glucose dehydrogenase (GDH) and a gluconate dehydrogenase (GADH), respectively (Goldstein, 1995). To date, the demonstration of the link between phosphate dissolution and the activity of the GDH activity was obtained on a wide range of bacterial genera. All these studies highlighted the implication of a specific type of GDH that depends on a particular cofactor, the pyrroloquinoline quinone (PQQ) (Goldstein *et al*. 2003; Babu-Kha*n et a*l. 1995; Li*u et al*. 1992; Wag*h et a*l. 2014).

91 The implication of a PQQ-dependent GDH enzyme in the dissolution of a complex mineral,
92 biotite, was recently evidenced for a strain of *Pseudomonas* (Wa*ng et a*l. 2020).
93 In this study, we investigated the molecular basis of MW in the model bacterial strain
94 *Collimonas pratensis* PMB3(1), which was shown to be very effective at MW (Uro*z et* al.
95 2007; Uro*z et a*l. 2009a). This bacterial genus is usually considered as member of the rare
96 biosphere due to its low representativeness in metagenomic data from soil (Levea*u et a*l.
97 2010). However, collimonads can represent dominant taxa in specific habitats such as on
98 mineral particles or in the mycorrhizosphere (Lepleux *et al.* 2012; Uroz *et al*. 2012, Colin *et*
99 *al*. 2017). Collimonads are recognized as particularly effective at mobilizing nutrients from
100 recalcitrant sources such as chitin and minerals, suggesting that they are particularly well
101 adapted to live in nutrient-poor conditions (Levea*u et a*l. 2010). Their strong MW
102 effectiveness was suggested to be linked to their ability to acidify the medium in presence of
103 glucose and to produce high amounts of gluconic acid (Uroz *et al.* 2009b). Using a random
104 mutagenesis approach, associated to a functional screening and a gene complementation and
105 chemical analyses, we identified the genes involved in the MW ability of strain PMB3(1) and
106 evidenced a novel MW pathway involving a PQQ-independent, synthesis of gluconic acid,
107 encoded by a glucose dehydrogenase (GDH).
108
109 **MATERIALS AND METHODS**
110
111 **Bacterial strains and growth media.** Bacterial strains and plasmids are listed in Table 1. The
112 model bacterial strain *Collimonas pratensis* PMB3(1) considered in this study was isolated
113 from oak (*Quercus petraea*)–*Scleroderma citrinum* ectomycorrhizae sampled in the organo-
114 mineral soil horizon from the long term experimental forest site of Breuil-Chenue located in
115 the Morvan region (France; for sampling details see Calvaruso et al. 2007). Strain PMB3(1)

116 was isolated from the symbiotic fungal mantle of the mycorrhizal fungus. Previous
117 experiments showed that the strain PMB3(1) was particularly effective at weathering mineral
118 (Uroz *et al.* 2007, 2009b). All strains were grown at 25°C, except the strains of *Escherichia*
119 *coli*, which were grown at 37°C. The media used were Luria-Bertani (LB), AB medium
120 (Chilto*n et a*l. 1974) and a modified version of the Bushnell-Haas (BHm; devoid of iron)
121 medium (Uro*z et a*l. 2007). The AB minimal (ABm) medium was supplemented with
122 mannitol (2 g/l final concentration) as carbon source. The BHm medium was supplemented
123 with glucose (2 g/l final concentration) as carbon source. Glucose was selected in our
124 experiments because it represents the dominant carbohydrate source found in forest soil and
125 in plant root exudates (Jolivet *et al*. 2006; Medeiros *et al*. 2006; Grayston and Campbell,
126 1996). Its presence is also explained by hydrolysis of various compounds such as starch,
127 saccharose (Derrien *et al.* 2004) or by hydrolysis of trehalose, a carbohydrate accumulated in
128 mycorrhizal roots (Martin *et al.* 1985). Antibiotics, when required, were added to the media at
129 the following final concentrations: Tetracycline 10 µg/ml, Kanamycin 100 µg/ml and
130 Ampicillin 100 µg/ml.
131
132 **Tn5-OT182 mutagenesis**. The wild-type strain PMB3(1) of *C. pratensis* was mutagenized
133 with Tn5-OT182 as described previously (Deshaze*r et a*l. 1997; Denni*s et a*l. 1998), with
134 minor modifications. The donor strain, *E. coli* S17-1λpir (pOT182), was grown at 37°C in
135 antibiotic-containing LB broth overnight, and the recipient strain, *C. pratensis* PMB3(1), was
136 grown at 25°C in ABm liquid medium for 2 days. The day of the bi-parental conjugation,
137 each culture (30 ml) was centrifuged and washed twice with sterile MQ water to remove any
138 antibiotic from the medium and mixed in a final volume of 5 ml of sterile MQ water. The
139 mixed suspension was spotted on agar LB plates (*i.e*., 10 µl drops). After 5h of incubation at
140 25°C, the bacterial lawn of each plate was harvested using sterile MQ water (5 ml), and 150

141 µl of this suspension were plated on tetracycline-containing ABm plates for isolation of the
142 Tn5 mutants of the strain PMB3(1). A total of 2,000 mutants able to grow on ABm
143 supplemented with mannitol and tetracycline were recovered after 5 days incubation at 25°C
144 and organized in 96-well microplates. To confirm the presence of the plasposon in the mutant
145 genome, a portion of the plasposon sequence was amplified by PCR using pOT-For and pOT-
146 Rev primers (Table 1). The quality of the mutant library was tested on a subset of Tn5
147 mutants and on the mutants of interest to determine whether the library was formed by
148 independent events.
149
150 **DNA manipulations and plasposon rescue.** Total DNA was extracted from the WT strain
151 PMB3(1) and the related Tn5 mutants using the protocol of Pospiech and Neumann (1995).
152 The flanking Tn5-OT182 integration regions were identified using the rescue method as
153 described previously (Deshazer et al. 1997). Briefly, 5 µg of DNA of each mutant was
154 digested 4 h with *Eco*RI-HF in 50 µl. After heat inactivation, the digested DNA was ligated
155 overnight in 25 µl. A volume of 10 µl was transformed into chemically competent *E. coli*
156 DH5α and the transformants recovered on the appropriate selective medium. The resulting
157 plasmids were extracted using the Miniprep kit from QIAGEN and sequenced with the primer
158 pOT-RT at Eurofins MWG Operon (https://www.eurofinsgenomics.eu/). All enzymes for
159 restriction digestion and ligation were purchased from New England Biolabs and were used
160 according to the manufacturers' instructions. When required, PCR amplicons were purified
161 using QIAquick PCR purification and gel extraction kits from QIAGEN. PCR was performed
162 using Taq polymerase from 5PRIME.
163
164 **Construction of pB2-*gdhC*, pB2-*resB* and pB2-*tatABC***. The ineffective MW phenotype of
165 three Tn5-OT182 mutants of the strain PMB3(1) were complemented by expressing each

166 gene or region of interest (*gdhC*, *resB* and *tatABC*) in trans on plasmids pB2-*gdhC*, pB2-*resB*
167 and pB2-*tatABC*. These genes/regions were amplified using specific primers (Table 1). The
168 conditions for PCR involved a 5min denaturation step at 94°C, followed by 30 cycles of 94°C
169 for 45s, melting temperature (Tm) for 45s, and 72°C for 2min30s. The Tm temperature was
170 adapted according to the gene or region of interest. Each PCR product was purified and
171 ligated into the pGEM-T Easy plasmid (Promega), and validated by Sanger sequencing (52).
172 The gene *gdhC* was cloned as a *Kpn*I - *Spe*I fragment, the gene *resB* as a *Sac*I - *Spe*I
173 fragment, and the *tatABC* genes as a *Sma*I - *Xba*I fragment into pBBR1MCS-2 (Kovach *et al*.
174 1995) to yield plasmids pB2-*gdhC*, pB2-*resB* and pB2-*tatABC* respectively. After verification
175 by sequencing, the resulting constructs were transferred in chemically competent *E. coli* S17-
176 1λpir and conjugated to the corresponding Tn5-OT182 mutant. The complemented mutants
177 were selected on agar ABm medium supplemented with tetracycline and kanamycin. The
178 presence of the intact gene in the complemented mutants was verified by a PCR using the
179 appropriate primers (Table 1), the WT strain was the positive control and the mutant was the
180 negative control. For each construction, a single transconjugant was conserved after
181 validation and used in the different bioassays described below.
182
183 **Analysis and quantification of organic acids.** After a 4-day culture (25°C; 200 rpm) in
184 liquid BHm amended with glucose (2 g/l) and devoid of iron, the supernatants of cultures of
185 the WT and Tn5-OT182 mutants were recovered by centrifugation (10,000g during 15min).
186 The supernatants were then filtered at 0.22 µm (GHP Acrodisc 25 mm syringe filter ; PALL)
187 and stored at -20°C. Organic acid analyses were performed on an ion chromatography with
188 conductivity detection (ICS 3000, Dionex Corp.) associated to an analytical column (IonPac®
189 AS 11 HC, Dionex corp.) according to Balland *et al.* (2010). The supernatants were eluted
190 with KOH solutions of varying concentrations (0.9-60 mM) over time (step gradient) with a

191 flow rate of 1.3 ml/min. Synthetic organic acids have used as references (sodium formate, D-
192 gluconic acid, sodium butyrate, pyruvic acid sodium salt, sodium citrate tribasic, sodium
193 oxalate, sodium propionate, sodium acetate, succinic acid disodium salt, DL-malic acid
194 disodium salt, sodium-L-lactate and malonic acid disodium salt).
195
196 **Analysis of the production of gluconate, 2KGA and 5KGA**. Qualitative analysis was
197 performed using thin layer chromatography (TLC) to identify which form of gluconic acid
198 was produced by the WT strain PMB3(1). To do it, cultures were done in ABm medium
199 supplemented with 10 g/l of glucose at 25°C and under shaking. After 4 days, the supernatant
200 of culture was recovered and filtered at 0.22 µm (GHP Acrodisc 25 mm syringe filter; PALL).
201 A volume of 4 µl for the standards and twice 4 µl for the culture supernatants were spotted on
202 a Silica Gel 60 plate (Merck). After drying at room temperature, migration was performed in
203 a solvent reagent containing ethyl acetate-acetic acid-methanol and deionized water
204 (6:1.5:1.5:1). When the TLC plate was dried, spread with a freshly prepared revelation
205 solution (diphenylamine (1 g) ; aniline (1 ml) ; acetone (50 ml) ; phosphoric acid (7.5 ml)).
206 The treated TLC plate was then incubated at 120°C for 20 min to allow apparition of coloured
207 spots. The results were analysed by measuring the retention factor (Rf) (corresponding to the
208 distance moved by the sample divided to the solvent front) and the observation of the spot
209 colour. The glucose, gluconate, 2KGA and 5KGA appeared as dark blue, pink, brown and
210 dark purple spots, respectively. Solutions of glucose, gluconate, 2-keto gluconate and 5-keto
211 gluconate adjusted at 100 mM were used as standard. In addition, gluconic acid was also
212 quantified using an enzymatic bioassay (kit 10428191035 from r-biopharm/Roche), according
213 to the manufacturer's instructions.
214

215 **Screening of the Tn5-pOT182 mutant library.** A first screening of the mutant library was
216 done in microplates using liquid ABm medium supplemented with tetracycline. After 3-days
217 culture, part of the cell suspension was spotted using a 96-well replicator on : i) solid
218 TriCalcium Phosphate (TCP) to determine the MW ability of each Tn5-mutant and ii)
219 modified liquid TCP medium (TCPm) to determine their acidification ability. These two
220 bioassays are commonly used to determine the mineral weathering potential of bacterial
221 strains (Calvaruso *et al.* 2007; Lepleux *et al.* 2013; Uroz *et al.* 2009b; Colin *et al.* 2017). In
222 addition, the homogeneity of the $Ca_3(PO_4)_2$ in the TCP assays allow for a high throughput
223 screening of mutant libraries. On the contrary, biotite is a heterogeneous material requiring
224 important replication of the assay, making its use incompatible with a first screening of the
225 mutant libraries. The solid TCP medium is composed per litter of : NH_4Cl 5 g ; NaCl 1 g ;
226 $MgSO_4$ 1 g ; Glucose 10 g ; $Ca_3(PO_4)_2$ 4 g and 16 g of agar (pH 6.5). After incubation at 25°C
227 for 7 days, the diameters of colonies and the diameter of discoloration zone (halo) were
228 measured to determine the ability of each Tn5-mutant to solubilize inorganic phosphorous.
229 The TCPm medium is modified version of the TCP devoid of $Ca_3(PO_4)_2$. The $Ca_3(PO_4)_2$ was
230 removed from the medium to allow a better determination of the protons produced. The pH
231 determination was performed after a 3-day incubation time. The supernatant was recovered by
232 centrifugation (8,000 g for 15 min) and a volume of 180 µl was transferred to new
233 microplates containing 20 µl of pH indicator (bromocresol green solution ; 1g/l), according to
234 Uroz *et al*. (2007). The pH of the supernatant was determined at 595 nm using an automatic
235 microplate reader (Bio-Rad, model iMark) and converted in pH values using a calibration
236 curve (Uroz *et al.* 2007).
237
238 **Functional assays done on Tn5-OT182 mutants impaired in their MW ability.** After the
239 first step of screening, all the Tn5-OT182 mutants presenting a significantly reduced halo of

240 solubilisation on solid TCP and/or a pH value significantly higher than the WT strain were
241 tested in a calibrated way on different bioassays in order to generate quantitative and
242 comparable data. The recovered Tn5-mutants were cultivated overnight in liquid LBm
243 medium supplemented with tetracycline. The WT strain was used as a positive control. Ten
244 milliliters of each culture was washed three times with sterile MQ-water to eliminate all
245 medium traces. The absorbance at 595 nm of each resulting suspension was then adjusted at
246 0.95±0.03.

247

248 *Phosphate solubilisation and solution acidification*

249 Ten microliters of each calibrated suspension were then dropped in triplicates on the surface
250 of the solid TriCalcium Phosphate (TCP) and in microplates containing the modified liquid
251 TCP medium (TCPm). Measures of P solubilisation and of acidification were performed as
252 described above (Colin *et al.* 2017).

253

254 *Biotite dissolution*

255 Five hundred microliters of the calibrated suspension of each strain was transferred into
256 sterile glass tubes containing 4.5 ml of BHm medium (20 mg/l KCl; 150 mg/l $MgSO_4$,$7H_2O$;
257 80 mg/l NaH_2PO_4,$2H_2O$; 90 mg/l Na_2HPO_4, $2H_2O$; 65 mg/l $(NH_4)_2SO4$, 100 mg/l KNO_3; 20
258 mg/l $CaCl_2$) buffered at pH 6.5 and supplemented with glucose (2 g/l) and 100 mg of sterile
259 biotite particles.

260 The biotite was obtained from Bancroft (Canada) and is a 2:1 phyllosilicate frequently
261 found in acidic forest soils. This mineral holds K, Mg, and Fe nutrients. It is a pure
262 homogeneous mineral, and its composition is as follows: 410.1 g/kg SiO_2, 109 g/kg Al_2O_3,
263 22.1 g/kg Fe_2O_3, 100.5 g/kg FeO, 2.7 g/kg MnO, 189 g/kg MgO, 4.1 g/kg Na_2O, 94.6 g/kg
264 K_2O, 22.8 g/kg TiO_2, 44.2 g/kg F, and 0.8 g/kg Zn. Its structural formula is (Si_3Al_1) $(Fe^{3+}_{0.12}$

265 $Fe^{2+}_{0.61}$ $Mg_{2.06}$ $Mn_{0.02}$ $Ti_{0.13}$) and $K_{0.88}$ $Na_{0.06}$ O_{10} ($OH_{0.98}$ $F_{1.02}$). The particle size of biotite was
266 calibrated from 200 to 500 µm.

267 Before use, all glass was rinsed with chloridric acid and deionised water to remove any
268 chemical traces. The tubes containing the culture medium with biotite were then sterilised by
269 autoclaving at 121°C. After seven days of incubation at 25°C (200 rpm), 1 ml of supernatant
270 was sampled and centrifuged at 11,000g for 15 min and filtered at 0.22 µm. Quantification of
271 the iron released from biotite and of the acidification of the medium were determined by
272 adding 180 µl of supernatant of each culture in a new microplate containing 20 µl of
273 ferrospectral or bromocresol green, respectively (Uro*z et a*l. 2007). The absorbance of these
274 suspensions was measured at 595 nm and the data transformed in pH and iron concentrations
275 using the colorimetric assays described above to determine the MW potential of each Tn5-
276 pOT182 mutant.

277 In addition, model chelating (citric acid) or acidifying (hydrochloric acid) molecules were
278 used as control in the same mineral weathering assay. Serial dilutions of citric acid (10^{-3} M;
279 pH adjusted to 6 to 2) and hydrochloric acid (concentration adjusted to be at pH 6 to 2) were
280 performed in BHm medium and incubated in the same condition than the bacterial strains
281 tested as described in Uroz *et al.* 2007.

282

283 *Cytochrome oxydase test*

284 To test if mutants are affected in their cytochrome oxydase activity, 100 µl of the calibrated
285 cell suspension were spotted on a filter paper and flooded in a 1% Tetramethyl-p-
286 phenylenediamine (TMPD) solution during 10 min. After this incubation period, the filter
287 paper was dried 30s and inspected. In the case of a positive reaction, the filter paper becomes
288 purple, while in the case of a negative reaction it remains white (Kovacs, 1956).

289

290 *Glucose dehydrogenase (GDH) test*

291 To test if mutants were affected in their GDH activity, the test of Matsushita and Ameyama
292 (1982) was used on the calibrated cell suspension (CS) or on cell crude extract (CE). The cell
293 crude extract was obtained by ultra sonication treatment (Vibra-Cell; Sonics, Newtown USA)
294 of the CS, followed by a centrifugation at 7000 g (15 min) to remove intact cells and a
295 filtration at 0.22µm. A volume of 100µl of CS or CE was incubated with 100 µl of Tris-HCL
296 (50mM; pH:7.5), 10 µl of 2,6-dichlorophenolindophenol (DCIP; 6.7mM), 10 µl of phenazine
297 methosulfate (PMS; 20mM), 10 µl of (FAD; 600µM), 100µl of glucose (1M), and MQ water
298 in a final volume of 300 µL. The reaction is initiated by the addition of D-glucose. GDH
299 activity is measured at 600 nm in microplate reader (Tecan infinite M200 pro). Controls
300 without glucose and/or FAD were performed. In addition, a control without the CS or CE was
301 performed for each of the condition tested (+/- FAD; +/- Glucose).

302

303 *Growth assay and carbon substrate use*

304 To determine whether the mutants were impaired in their growth ability, tests were performed
305 under nutrient-rich and nutrient-poor conditions. For the nutrient-rich condition, 10 µl of each
306 calibrated suspension (WT and mutants) were added to 190 µl of liquid LB medium in a
307 microplate. The microplate was incubated at 25°C and orbital shaking during 4 days in a
308 microplate reader (Tecan infinite M200 pro). The absorbance was measured at 600 nm each 3
309 hour. For the nutrient-poor condition, 10 µl of each calibrated suspension was added to 190 µl
310 of liquid BHm medium amended with a source of carbon (2 g/l). The carbon sources used
311 were glucose, gluconic acid, 2-Keto-gluconate, 5-Keto-gluconate. After 5 days of incubation
312 at 25°C, the pH was measured with bromocresol green as described above and the growth was
313 measured at 595 nm.

314

315 **Bioinformatic analyses.** Genome analysis was performed using NCBI and MAGE (Vallene*t*
316 *et a*l. 2009). Homology search analysis and alignment were performed with BLAST (BlastN
317 and BlastP ; Altschul *et al.* 1997). Localization of the Tn5-OT182 insertion was done on
318 MAGE. A research of Twin Arginine Translocation (TAT) signal carrying proteins was done
319 using the predictor servers Tatfind (Dilk*s et al*. 2003) and TatP 1.0 (Bendtse*n et a*l. 2005).
320 The two proteins lists were assembled and analysed. Analyses of the pfam domains and of the
321 conservation of the GdhL subunit sequence were done using pfam (El-Gebal*i et al*. 2019) and
322 Seaview (version 4.5.4) (Gou*y et al*. 2010). To study the conservation of the GdhL sequence,
323 a set of 22 different genera or species related to *Collimonas* and available on international
324 databases were analysed among them, 11 were assigned to *Collimonas* genus (*Collimonas* sp.
325 PA-H2, OK242, OK307 and OK607, *C. fungivorans* Ter331 and Ter6*, C. arenae* Cal35,
326 Ter10 and Ter282*, C. pratensis* Ter91, Ter291). Other genera taxonomically close to
327 *Collimonas* were chosen : 2 belonging to the *Glaciimonas* genus (*Glaciimonas* sp. PCH181
328 and GS1), one to *Janthinobacterium,* 2 to *Herbaspirillum* (*H. rubrisubalbicans* and *H.*
329 *seropidecae*) and one to *Burkholderia cepacia*. In addition, we selected the sequences of the
330 large subunit of a set of GMC oxidoreductase which function was demonstrated: a gluconate
331 dehydrogenase (GADH ; BAH80545.1) and a sorbitol dehydrogenase (BAD60913.1) from
332 *Gluconobacter frateurii* (Toyam*a et al*. 2005; Saichan*a et al*. 2009), a fructose dehydrogenase
333 large subunit of *Gluconobacter japonicas* (BAM93252.1) (Kawa*i et a*l. 2013). We also
334 considered the sequences of the large subunit of a putative sorbitol dehydrogenase from
335 *Gluconobacter oxydans* H24 (AFW02570.1) and a dehydrogenase subunit from *Pantoea*
336 *agglomerans* (AAF21261.1). The same genome resources were also used to recover the pqq-
337 genes as well as the pqq-dependent enzymes.
338

339 **Statistical analysis.** Statistical analyses were performed in R software. Data shown were
340 means of, at least, triplicates. Differences between sample's means were analysed by
341 ANOVA and Tukey HSD tests.
342
343 **Nucleotide sequence accession number.** The nucleotide sequence reported in this paper are
344 related to the genome sequence of the strain PMB3(1), which has been deposited in the
345 GenBank database under accession no. WXXL00000000.
346
347 **RESULTS**
348 ***1) Genomic and physiological characterization***
349 At the genomic level, the detailed analysis of the genome of strain PMB3(1) did not permit
350 to evidence any PQQ-dependent GDH nor the presence of the PQQ cofactor encoding genes.
351 Similarly, none of the *Collimonas* genome available in the international databases (*i.e.*,
352 *Collimonas* sp. PA-H2, *C. fungivorans* Ter331 and Ter6*, C. arenae* Cal35, Ter10 and Ter282*,*
353 *C. pratensis* Ter91, Ter291) possess PQQ encoding genes and/or PQQ-dependent enzymes,
354 suggesting that alternative MW pathways exist in collimonads.
355 At the chemical level, high concentrations of gluconic acid were detected in the supernatant
356 of cultures of strain PMB3(1) done with glucose (Table 2), evidencing that this strain was
357 capable of converting glucose to gluconic acid (GA) through an alternative, PQQ-independent
358 pathway. Noticeably, the thin layer chromatography analyses confirmed the large production
359 of gluconic acid, but also a weak production of 2 keto-gluconic acid (2KGA) by strain
360 PMB3(1) as evidenced by the comparison with the migration retention factor (Rf) and the
361 colour of the synthetic GA and 2KGA controls used (Fig. 1; Table 2). A kinetic analysis
362 confirmed the continuous accumulation of GA, while 2KGA was hardly detectable.
363 Interestingly, the growth assays done in minimal medium supplemented with a single source

364 of carbon (*i.e*., glucose, GA, 2KGA or 5KGA) confirmed that glucose, GA and 2KGA were
365 used as carbon source by strain PMB3(1), but not 5KGA. Among them, glucose was the sole
366 carbon which metabolism allowed an acidification of the culture supernatant (pH=3.5), while
367 the pH did not change in presence of GA and 2KGA (*i.e.,* pH=6.2).
368

369 ***2) Isolation of mutants affected in the ability to solubilize inorganic phosphorous***

370 The first high throughput screening of the mutant library allowed us to recover a total of 4
371 mutants (*i.e*., 25(10F), 31(11B), 34(11H), 59(7B)) impaired in their ability to solubilize
372 inorganic phosphorous (Pi).

373 The use of the calibrated procedure (see materials and methods) confirmed that these 4
374 mutants were significantly altered in their ability to solubilize Pi ($P<0.05$). Based on this
375 second screening, three of these mutants, 25(10F), 31(11B), 34(11H), did not present a
376 solubilisation halo (Table 2). For the mutant 59(7B), a reduction of 34% of the halo diameter
377 was observed compared to the WT strain (halo_WT: 0.79 cm ± 0.04 vs halo_59(7B): 0.52 cm
378 ± 0.1). All these mutants and WT presented similar colony diameters. The acidification assay
379 revealed that the ability of mutants 25(10F), 31(11B) and 34(11H) to acidify was significantly
380 reduced ($P<0.05$). Indeed, while the WT strain and mutant 59(7B) acidified the medium from
381 the initial pH 6.5 to pH 3.7 in three days, mutants 25(10F), 31(11B), 34(11H) acidified the
382 medium only to *c.a*. pH 4.7 (average value).
383

384 ***3) Ability to weather biotite***

385 The ability to weather a complex mineral was tested only for the three mutants totally
386 impaired in their ability to solubilize Pi (*i.e.*, 25(10F), 31(11B), 34(11H)) as well as for the
387 WT strain (Fig. 2). After a 7-day incubation in presence of glucose as carbon source, the WT
388 strain was able to release 1.29 ± 0.13 mg/l of iron in the solution. In comparison, the three

389 mutants 25(10F), 31(11B) and 34(11H) released significantly less iron (0.32 ± 0.08, 0.54 ±
390 0.04 and 0.39 ± 0.02 mg/l, respectively) ($P<0.05$) (Fig. 2A). Those values corresponded to a
391 maximal reduction of 75% for mutant 25(10F) and a minimal reduction of 58% for mutant
392 31(11B) compared to the WT strain. After the 7-day incubation, the WT strain decreased the
393 pH of the culture supernatant to pH=3.3 ± 0.1, while the mutants reached significantly higher
394 value (4.3 to 4.5; $P<0.05$) (Fig. 2B). The co-visualization of the quantity of iron released from
395 biotite and of the pH of the supernatant confirmed the strong perturbation of the mineral
396 weathering ability of the mutants compared to the WT strain (Fig. 2C).
397

398 ***4) Sequence analyses of the genomic region for the different mutants***

399 The genomic region flanking the plasposon insertion of each mutant was isolated by self-
400 cloning using the plasposon rescue method. For each mutant, 3 independent clones were
401 recovered, sequenced and blasted against the genome sequence of strain PMB3(1). This
402 revealed a single genomic region per mutant, demonstrating a unique plasposon insertion
403 event in each of the mutants considered [25(10F), 31(11B), 34(11H), 59(7B)]. The sequences
404 were also blasted against nucleic acid and protein databases of the NCBI for homologous
405 sequences.

406 In mutant 25(10F), the plasposon is inserted in a gene we named *gdhC* (NKI70796.1). This
407 gene presented high homology with a cytochrome c subunit of a Glucose/Gluconate/Sorbitol
408 2-dehydrogenase (Fig. 3, Table 3). Pfam analysis highlighted two conserved domains
409 (cytochrome c from 36 to 137 aa/ cytochrome CBB3 from 322 to 406 aa ; e-value = 6.6e-07
410 /5.4e-10, respectively) corresponding to a cytochrome c oxidase protein. This gene is
411 preceded by two genes presenting sequence homology with subunits of a
412 Glucose/Methanol/Choline (GMC) oxidoreductase. The first one presented high sequence
413 homology (98 %) with the large subunit (NKI70795.1; *gdhL*) belonging to the FAD

414 flavoprotein oxidoreductases, while the next gene presented high homology (80 %) with the
415 gamma subunit (NKI70794.1; *gdhS*) of a Sorbitol/Glucose/2-Keto-D-Gluconate
416 dehydrogenase. Pfam analysis revealed that this gamma subunit presented high sequence
417 similarities with the FAD-Sorbitol dehydrogenase family protein (e-value = $2.5.10^{-48}$) and
418 carried a potential TAT-signal.

419 In mutant 31(11B), the plasposon is inserted in a genomic region presenting high sequence
420 homology with a *resB* homologous gene (NKI71673.1 ; 94% identity with cytochrome c
421 biogenesis protein ResB in *Collimonas sp.* PA-H2) (Fig. 3, Table S1). This gene belongs to a
422 genetic structure composed of 3 genes (NKI71672.1; NKI71673.1 ; NKI71674.1) involved in
423 cytochrome c biogenesis /maturation. Among them, gene NKI71672.1 codes for a cytochrome
424 c4 immature protein and gene NKI71673.1 is gene *resC* coding for a cytochrome c biogenesis
425 protein. In the vicinity of the *resB* gene, two other genes, *yedY* and *yedZ* (NKI71676.1,
426 NKI71677.1), presenting high sequence homology with a sulfite oxidase were detected.

427 In mutant 34(11H), the plasposon is inserted in a genomic region presenting high sequence
428 homology with the Twin Arginine Translocation pathway (TAT) operon (*tatABC*). The
429 insertion was localized between the gene *tatB* (NKI71758.1; 93% identity with Sec-
430 independent protein translocase subunit TatB in *Collimonas sp.* PA-H2) and *tatC*
431 (NKI71759.1; 95% of identity with the twin-arginine translocase subunit TatC in *Collimonas*
432 *sp*. PA-H2) (Fig. 3, Table S2).

433 In mutant 59(7B), the plasposon is inserted in the gene *int* (NKI72639.1) presenting high
434 sequence homology with an apolipoprotein N-acyltransferase (92% of identity with the
435 apolipoprotein N-acyltransferase of *Collimonas sp.* PA-H2) (Fig. 3, Table S3). This enzyme is
436 responsible for the last step of lipoprotein maturation. In the vicinity of this gene, 5 genes
437 involved in metabolism and two genes involved in information storage and processing were
438 identified according to their COG class.

439

440 ***5) Identification of TAT-dependent proteins***

441 Due to the localization of the plasposon insertion of mutant 34(11H) in the TAT system
442 region, a bioinformatic analysis was done to identify all proteins presenting a putative TAT
443 signal in the genome of strain PMB3(1). Indeed, the perturbation of the MW ability may be
444 related to a protein carried by the TAT system. This screening allowed the identification of 74
445 TAT-proteins related to different Clusters of Orthologous Groups (COG) (metabolism,
446 cellular processes and signalling, information storage and processing, or poorly characterized
447 (Table S4; Fig. S1). Among them, we noticeably identified the protein *gdhS* that corresponds
448 to the gamma subunit of the Sorbitol/Glucose/2-Keto-D-Gluconate dehydrogenase gene
449 (NKI70794.1). This gene is in an operon along with gene *gdhC* (*i.e.*, cytochrome subunit of
450 the Sorbitol/Glucose/2-Keto-D-Gluconate dehydrogenase) interrupted in mutant 25(10F) (see
451 fig. 3) Other enzymes related to putative alkaline phosphatases (*i.e.*, PhoD like, NKI71510.1
452 and NKI72605.1), involved in cytochrome c biogenesis (NKI72023.1) or presenting an
453 oxidoreductase activity (NKI71676.1) were also identified in the list of TAT-proteins.

454

455 ***6) Functional and physiological characterization of mutants 25(10F), 31(11B) and 34(11H)***

456 *Cytochrome activity*

457 As mutants 25(10F) and 31(11B) were affected on genes encoding a cytochrome subunit or
458 involved in the biogenesis of cytochrome c, a cytochrome oxidase test was performed on all
459 mutants. This qualitative test based on the oxidation of Tetramethyl-p-phenylene diamine
460 (TMPD) revealed that the three mutants formed a purple pigment as the WT strain (Table 2),
461 revealing that the cytochrome c oxidase of the strain PMB3(1) was still functional in all the
462 mutants. This result highlights that strain PMB3(1) possesses several cytochrome c biogenesis

463 systems independent on the resB/resC system interrupted in mutant 31(11B). Indeed, a
464 genome analysis revealed that the strain PMB3(1) possesses two cytochrome biogenesis loci.
465
466 *Detection and quantification of organic acids*
467 Ionic chromatography analyses permitted to identify the main acids produced (i.e., Gluconate,
468 oxalate, pyruvate, formate, and succinate) by the WT strain and mutants in presence of
469 glucose (Table 2). Gluconate was the major organic acid detected in the WT culture
470 supernatant, reaching a concentration of 878.7±3.3 mg/l, while all other organic acids are
471 present at very low concentrations (pyruvate: 2.87 mg/l; succinate: 0.07 mg/l). Noticeably,
472 gluconate was not detected in any of the supernatants from the 3 mutants (*i.e.*, mutants
473 25(10F), 31(11B) and 34(11H); Table 2). Amongst other organics, only pyruvate presented
474 significant differences between the WT strain and mutants, but with very low concentrations
475 (WT: 2.87±0.63 mg/l; 25(10F): 1.08± 0.03 mg/l; 31(11B): 1.55±0.23 mg/l and 34(11H):
476 0.11±0.06 mg/l) ($P<0.05$).
477
478 *Growth in rich (LB) medium*
479 To get a better understanding of the importance of the gene(s) interrupted in the different
480 mutants, their growth was monitored for 120 hours in LB medium (Fig. 4; Table 2). No
481 difference between the WT and the 25(10F) mutant was observed. Indeed, the growth rate
482 (0.09 ± 0.02 h^{-1}; $P<0.01$), the lag time (ca. 10 h), the maximal growth yield (OD_{595nm} of $1.20 \pm$
483 0.05) as well as the time necessary to reach the stationary phase for the WT and mutant
484 25(10F) (*i.e.*, 40 h) were similar. In contrast, mutants 31(11B) and 34(11H) differed
485 significantly from the WT strain. Mutant 31(11B) presented a growth rate of 0.06 ± 0.01 h^{-1}, a
486 lag phase of 25 h, a maximal growth yield (OD_{595nm}) of 0.95 ± 0.08 and it required 50 h to
487 reach the stationary phase. Mutant 34(11H) presented the lowest growth rate of 0.02 ± 0.01 h^{-}

488 [1], a maximal growth yield (OD_{595nm}) of 1.08 ± 0.08 and it required 120 h to reach this
489 maximum.
490
491 *Impact of the carbon source on the acidification process*
492 As acidification plays an important role in the mineral weathering process by strain
493 PMB3(1), several carbon sources coming from the metabolism of glucose (*i.e.*, gluconate,
494 2KGA, 5KGA) were tested in minimal medium (BHm) (Fig. 5). Contrary to the strong
495 acidification observed with glucose, these growth assays revealed no difference of pH
496 between the WT and mutant strains when they were cultivated in minimal medium amended
497 with glucose derivatives such as GA or 2KGA (pH_{WT} = pHmutant = 6.2). No growth was
498 observed with 5KGA.
499
500 ***7) Complementation analysis of the mutants***
501 To confirm the involvement of the interrupted genes identified in the different mutants (*i.e.*,
502 25(10F), 31(11B) and 34(11H)) in the MW ability of strain PMB3(1), a functional
503 complementation was performed. For each mutant, the native gene(s) under a constitutive
504 promoter was cloned in the pBBR1MCS-2 vector. These constructions were then transformed
505 in the corresponding mutant and tested with the same set of functional assays as described
506 above. Noticeably, the quantitative assays performed on solid TCP medium revealed that the
507 complemented mutants 25(10F)(pB2-*gdhC*) and 34(11H)(pB2-*tatABC*) recovered fully their
508 mineral weathering ability to WT effectiveness (Table 2). The mutant 31(11B)(pB2-*resB*) was
509 restored in its ability to weather mineral, but presented a slightly lower efficiency than the
510 WT strain, when harbouring gene *resB* on the pBBR1MCS-2 plasmid. Regarding
511 acidification, upon complementation pH was further reduced from pH 5.1 ± 0.1 observed in
512 the mutants to 4.4 ± 0.1 in the complemented mutants after 3 days, which was not

513 significantly different from that observed with the WT strain harbouring the empty vector
514 pBBR1MCS2 (4.3 ± 0.1). In addition, the restoration of the glucose dehydrogenase activity
515 was observed in the complemented mutants. Last, complementation restored growth of all
516 mutants (Fig. 4).
517

518 ***8) Bioinformatic properties and conservation of GMC oxidoreductases in the collimonads***
519 ***and related genera***

520 For bacteria carrying a glucose dehydrogenase enzyme of the GMC oxidoreductase family,
521 the conversion of glucose to gluconic acid is mainly performed by the large subunit (*i.e.*,
522 *gdhL*) of the enzyme (Yamaoka *et al.* 2008). In this context, bioinformatic analyses were
523 performed on this subunit to analyse its properties and its conservation among collimonads
524 and related taxa.

525 In strain PMB3(1), this approach allowed to identify a single GMC oxidoreductase
526 organized in three subunits and corresponding to the genomic region described above for
527 mutant 25(10F). A second homologous of the large subunit (GdhL) was also identified in
528 strain PMB3(1), but with a low identity (20%). Contrary to the GMC oxidoreductase, this
529 homologue is organized as a single unit (*i.e.*, one gene) and encodes a putative choline
530 dehydrogenase. The pfam analysis done on the sequence of the large subunit of the strain
531 PMB3(1) evidenced the presence of a Rossmann-like domain (amino acid 11 to 41; e-value =
532 2.0e-5) and two GMC oxidoreductase domains (domain C: 405 to 525, e-value 1.5e-18;
533 domain N: 176 to 318, 1.5e-5), supporting by this way the function attributed to the gene
534 identified in our study (*i.e.*, a Glucose dehydrogenase from the GMC oxidoreductase family)
535 (Fig. S2). Notably, all the GMC oxidoreductases described in the literature present the 3
536 domains detected (Rossmann, GMC_N, GMC_C) and all are specifically associated to
537 enzymes binding FAD as cofactor (*i.e.*, gluconate dehydrogenase, GADH; glucose

538 dehydrogenase, GDH; fructose dehydrogenase, FDH; sorbitol dehydrogenase, SDH) (Kataoka
539 *et al.* 2015; Toyama *et al.* 2005).

540 The comparison of homologous sequences from the available *Collimonas* genome
541 sequences revealed a very good conservation of the different domains (Fig. S3). In the
542 putative FAD domain, only two positions on a total of 31 varied among the collimonads (Fig.
543 S2). In the GMC_C domain, 10 positions on a total of 86 varied, while 16 positions on a total
544 of 121 varied in the GMC_N domain (Fig. S2). Beside these pfam domains, we also detected
545 in the large subunit sequence of strain PMB3(1), the cysteine-rich region (cys208, cyst214,
546 cys218 in our alignment; Fig. S2) identified by Shiota *et al.* (2016) as specific of the FAD-
547 dependent dehydrogenases. This region is constituted by three conserved cysteine (Cys)
548 residues (**C**CGNNN**C**MPI**C**P), all conserved in the *Collimonas* genomes considered. This
549 particular site constitutes an iron-sulfur cluster essential for electron transfer from FAD to the
550 cytochrome c subunit (Shiot*a et a*l. 2016).

551 When the sequence comparison integrated neighbouring genera, it revealed a good
552 conservation of the large subunit sequence among the taxa tested and a stronger relatedness of
553 *Janthinobacterium* and *Glaciimonas* with *Collimonas*. In contrast, *Herbaspirillum* presented
554 important variations (Fig. S3). The sequence of the GdhL protein of strain PMB3(1) presented
555 ca. 90 % of identity with that of strains PCH181 and GS1 of *Glaciimonas* sp., 80 % with that
556 of *Janthinobacterium* and only 50 % with that of *Herbaspirillum seropedicae* and *H.*
557 *rubrisubalbicans*. The low relatedness of the GdhL of *Herbaspirillum* was also visible in the
558 sequence alignment, especially in the domain GMC_C.

559 Last, a sequence comparison between the large subunit of strain PMB3(1) and the large
560 subunit of FAD-dependent dehydrogenases for which the function and/or the FAD-
561 dependence were experimentally demonstrated (*i.e.*, a GADH from *Gluconobacter frateurii*,
562 two SDHs from *Gluconobacter oxydans* and *G. frateurii,* a FDH from *Gluconobacter*

563 *japonicas* and a GDH from *Burkholderia cepacia*) or suspected (choline dehydrogenase
564 (CDH) from *Pantoea agglomerans*) was performed (Fig. S3). This analysis revealed that the
565 most distantly related sequence was the GADH. The positions on the phylogenetic tree are
566 supported by the strong sequence variations observed in the sequence alignment. The large
567 subunit sequence of strain PMB3(1) appeared also poorly related to the FDH and SDH. The
568 stronger relationship was observed with the glucose dehydrogenase (GDH) from
569 *Burkholderia cepacia* (Tsuya *et al.* 2006).
570

571 **DISCUSSION**

572 The ability to weather minerals was evidenced for a broad range of bacterial genera, but
573 the genes involved have been poorly investigated and characterized. To date, the main system
574 described is related to the PQQ-dependent glucose dehydrogenases (PQQ-GDH), which
575 allows the production of protons and gluconic acid. The role of such enzymes and the related
576 *pqq* genes (*i.e.*, *pqqABCDEF*) in the solubilisation of highly weatherable minerals (*i.e.*, Pi)
577 was reported in several studies (Goldstein, 1995; Goldstei*n et a*l. 2003; Babu-Kha*n et* al.
578 1995), but only recently for a complex mineral (*i.e.*, biotite; Wang *et al.* 2020). In this study,
579 we investigated the molecular mechanisms and the genes explaining the high MW
580 effectiveness of strain PMB3(1) of *Collimonas* as well as their relative conservation among
581 collimonads.

582 Collimonads have been isolated from various nutrient-poor environments (*i.e.*, sand dune,
583 forests soil) and in fungal rich habitats (*i.e.*, mycorrhizosphere) (Colin *et al.* 2017; Uroz *et al.*
584 2007; Lepleux *et al.* 2012). They are recognized as particularly effective at mobilizing
585 nutrients using fungal derived metabolites such as glucose, trehalose, mannitol (Uroz *et al*.
586 2009b). The genome analysis performed in this study reveals that none of the sequenced
587 *Collimonas* to date possess the canonical *pqq* biosynthesis genes nor PQQ-dependent

588 enzymes, meaning that alternative mechanisms to the PQQ-dependent pathway are involved
589 in their MW ability. Those results confirmed the observations made on the Ter *Collimonas*
590 strains (Ter331, Ter14, Ter6, Ter91 or Ter10; Mela *et al.* 2012) or on strain PMB3(1) (Picard
591 *et al.* 2020). In this study, we confirmed that strain PMB3(1) is very effective at weathering
592 and at producing gluconic acid, and we showed for the first time that at least a part of this
593 gluconic acid is converted to 2-keto-gluconic acid (2KGA), evidencing the presence of a
594 glucose dehydrogenase (GDH) activity and of a gluconate dehydrogenase (GADH) activity.
595 Such conversion of glucose to gluconic acid and then to 2KGA corresponds to the direct
596 oxidative pathway described by Eber*t et a*l. (2011). The 2KGA is then converted into 6-P-
597 gluconate by a 2-keto gluconate kinase and a 2-keto gluconate 6-phosphate reductase prior to
598 entering the Pentose pathway or the Entner-Doudoroff pathway (Eber*t et a*l. 2011). Beside its
599 capability to produce gluconic and 2-keto-gluconic acids, we demonstrated that strain
600 PMB3(1) is capable of metabolizing these two acids, without acidification of the solution. No
601 growth was observed with the 5-keto-gluconic acid. The ability to produce GA and 2KGA is
602 well conserved among Gram-negative bacteria, but the production of 5KGA or 2,5KGA is
603 more specific to acetic bacteria such as *Acetobacter* or *Gluconobacter* (Sützl *et al.* 2019).
604 Bacteria producing GA and 2KGA have been described as being more effective at
605 solubilizing calcium phosphate than other bacteria (Babu-Khan *et al.* 1995).

606 A focus on the genomic region flanking the plasposon insertion in mutant 25(10F) allowed
607 a better understanding of the link between the 3 weathering null mutants, though they were
608 geographically distant and apparently unrelated. We identified a PQQ-independent enzyme
609 presenting high homology with Glucose/Methanol/Choline (GMC) oxidoreductases (Sützl *et*
610 *al.* 2019; Cavener, 1992). The GMC superfamily encompasses a wide variety of FAD-
611 dependent oxidoreductases (*e.g.*, glucose oxidase (GO), glucose dehydrogenase (GDH),
612 alcohol oxidase (AO), cellobiose dehydrogenase (CDH)) that are present in both prokaryotic

613 and eukaryotic domains (Sützl *et al.* 2019; Zamocký *et al.* 2004). They are composed of three
614 subunits encoding a small subunit, a large subunit, and a cytochrome c subunit (Sützl *et al.*
615 2019; Cavener, 1992), corresponding exactly in term of organization (small subunit: *gdhS*;
616 large subunit: *gdhL* and cytochrome subunit: *gdhC*) and sequence homology to the genomic
617 region impacted in mutant 25(10F). The large subunit of strain PMB3(1) is characterized by
618 several domains according to pfam analysis (a FAD domain and two GMC oxidoreductase
619 domains (N and C)) and the presence of a cysteine rich region. This cysteine rich region was
620 shown to be conserved in the catalytic large subunit of FAD-dependent dehydrogenases
621 (Shiota *et al.* 2016). While we were able to demonstrate the GDH activity and the production
622 of gluconic acid by strain PMB3(1), testing experimentally the FAD dependence of the GDH
623 activity would require the purification of the three subunits in a functional form, which was
624 not possible in our labs. A FAD dependence test performed on lyzed cells was inconclusive.
625 The known FAD-dependent GDHs are involved in the oxidation of different substrates
626 including sugars, alcohols, cholesterol and choline and their conversion to the corresponding
627 lactone. In presence of glucose, these enzymes allow the production of D-glucono-1,5-
628 lactone, which is then converted to gluconic acid through the action of a gluconolactonase or
629 spontaneous hydrolysis (Ferr*i et a*l. 2011; Zamocký *et al*. 2004). Strain PMB3(1) possesses
630 several putative gluconolactonases (NKI70323.1, NKI67822.1, NKI72143.1) in its genome.
631 Although additional experiments are required to support the different steps, we proposed here
632 a hypothetical functional model of the main mechanisms used by strain PMB3(1) to weather
633 minerals and linking the different mutants obtained in this study (Fig. 6).

634 To be functional GMC oxidoreductases also require a mature cytochrome c subunit and a
635 small subunit. The maturation of the cytochrome c subunit is performed through the ResB and
636 ResC system (Crow *et al.* 2005), which allows the translocation of the heme during
637 cytochrome maturation (Kranz *et al.* 2009). Without the heme, the cytochrome c subunit is

638 not functional and does not permit electron transfer. Strain PMB3(1) possesses both ResB and
639 ResC system and very interestingly the *resB* gene is interrupted in our mutant 31(11B), likely
640 interrupting the maturation of the cytochrome c subunit of the GDH identified. Notably, its
641 activity was restored when a complementation in *trans* was done with the gene *resB*. The last
642 small subunit of the GMC oxidoreductases is a protein required for the proper folding and
643 secretion of the large subunit (Shiota *et al.* 2016; Yamaoka *et al.* 2008). Indeed, without a
644 functional small subunit, the GDH of *Burkholderia cepacia* strain SM4 failed to perform the
645 oxidation of its substrate (Yamaoka *et al.* 2008). The small subunit of strain PMB3(1) is also
646 characterized by a TAT-signal. This particularity was also observed in various
647 dehydrogenases organised in three subunits. This strongly suggests that in absence of
648 translocation of the small subunit, the related FAD-dependent enzyme is not functional.
649 Congruent with this hypothesis, Toyama *et al*. (2007) showed that the assembly of the
650 different subunits of the FAD-GADH of *Gluconobacter dioxyacetonicus* was performed in
651 the cytoplasm before the translocation through the TAT system and using the TAT signal of
652 the small subunit. In addition, experimentations showed that TAT system allowed the
653 translocation of folded proteins associated with their co-factor (Allen *et al.* 2002; Palmer and
654 Berks, 2012). In this context, the TAT-signature of the PMB3(1) GdhS (*i.e*., the small subunit
655 of the GDH) represents the link between the mutant 34(11H) and the MW null phenotype.
656 Besides the small subunit of the Glucose dehydrogenase (GdhS) of strain PMB3(1), several
657 other TAT-exported proteins could also play a role in MW, such as those involved in
658 cytochrome biogenesis, oxidoreduction or more especially in the hydrolysis of organic
659 phosphate (Table S4) (Chhabra *et al.* 2013 ; Lidbury *et al.* 2016). However, the complete loss
660 of MW activity in the other two mutants (*i.e.,* 25(10F), 31(11B)) would not support this
661 hypothesis.

662 The conservation of the MW ability among collimonads prompted us to determine the
663 phylogenetic distribution of the GDH type identified in strain PMB3(1) in collimonads and
664 closely related genera. The comparison done on the sequenced genomes of *Collimonas* (*C.*
665 *pratensis* , *C. arenae*, *C. fungivorans*, *C.* sp.) revealed a good conservation of this GDH type
666 as well as of the TAT system, and the absence of PQQ-dependent GDHs or the related *pqq*
667 biogenesis genes in collimonads. Such conservation of the PQQ-independent GDH identified
668 from strain PMB3(1) was expected, as *Collimonas* are known to harbour important gene
669 conservation among the different species described (Mela *et al.* 2012). The comparison done
670 between the large subunit of the GDH of strain PMB3(1) and other dehydrogenases FAD-
671 dependent in other taxa revealed that it was more strongly related to *Glaciimonas*,
672 *Janthinobacterium* and *Burkholderia* than *Herbaspirillum* or *Gluconobacter*. The GDH of
673 PMB3(1) presented higher homology with other alcohol dehydrogenases experimentally
674 identified as sorbitol (FW02570.1), glucose (AAN39686.1) or fructose (BAM93252.1)
675 dehydrogenases, all FAD-dependent. This relative proximity between different type of FAD-
676 dependent dehydrogenases may be due to the low substrate specificity of some of these
677 enzymes, which are capable of oxidizing different substrates, including glucose (Yamaoka *et*
678 *al.* 2008). In contrast, the most distantly related FAD-dependent dehydrogenase of our
679 analysis was the GADH (*i.e.*, Gluconate DH; BAH80545.1) of *Gluconobacter frateurii* that
680 converts gluconate to 2 ketogluconate. Beside the large subunit, the small and cytochrome
681 subunits were also well conserved between collimonads and related taxa. The conservation of
682 an homologous of the GDH of strain PMB3(1) in *Glaciimonas* and the presence of this taxa in
683 nutrient-poor and rocky environments (*i.e*. the forest soil, alpine glaciers) (Udovičić *et al.*
684 2015) suggest that members of this genus are well adapted to oligotrophic conditions and
685 capable of weathering minerals as are collimonads. Their MW ability remains to be tested. In
686 contrast, important differences were found in the sequence of the GMC_C domain and the

687 Cys-rich region of the large subunit of *Herbaspirillum*, a genus reported as non- or poorly-
688 effective at weathering minerals (Uroz *et al.* 2009b). Beside the presence of a GMC
689 oxidoreductase homologue, our genomic analyses revealed that some representative strains of
690 the genus *Burkholderia*, *Glaciimonas*, *Gluconobacter* or *Pantoea* were also characterized by
691 the presence of *pqq*-encoding genes as well as PQQ-dependent enzymes in their genome, but
692 their ability to solubilize minerals is unknown. Such results reveal that bacteria can possess
693 both PQQ-dependent and -independent GDH. Interestingly, a dual transcriptomic and
694 proteomic approach done on another effective MW bacterial strain (i.e. strain PML1(12) of
695 *Caballeronia mineralivorans*) harbouring both systems highlighted that a homologous GMC
696 oxidoreductase was up-regulated in presence of biotite, while the *pqq*-related genes appeared
697 poorly expressed and not regulated (Uroz *et al*. 2020). This GMC oxidoreductase presented
698 high sequence homology with the FAD-dependent GDH of *Burkholderia cepacia* strain SM4
699 and with the GDH identified in strain PMB3(1) in our study, suggesting that such enzyme
700 may be broadly conserved in Burkholderiales.
701

702 **CONCLUSION**

703 This study elucidated the main molecular mechanism used by collimonads to weather
704 mineral. Our results show that their effective MW ability is likely due to an enzymatic activity
705 based on a PQQ-independent GMC-DH capable of oxidizing glucose to gluconic acid (*i.e.*, a
706 PQQ-independent glucose dehydrogenase). This crucial step of the direct oxidative pathway
707 leads to a high production of protons jointly with the oxidation of glucose and an important
708 acidification of the nearby environment and consequently the dissolution of the minerals. This
709 enzyme appears conserved among collimonads explaining their conserved effectiveness. The
710 variations of sequence observed in the GMC_C domain of the large subunit of this enzyme
711 may explain why some of the related taxa (*i.e.*, *Herbaspirillum*, *Janthinobacterium*) harbour a

712 lower effectiveness at weathering minerals than collimonads. Our results represent the first
713 demonstration that MW in bacteria is not exclusively explained by the PQQ-dependent GDH
714 production of gluconate and that an alternative pathway based on a PQQ-independent GDH
715 also exists in MW bacteria lacking the PQQ system. As such enzyme is largely distributed in
716 Bacteria, Archaea and in some Eukaryotes (Sützl *et al.* 2019), it may open new perspectives
717 in the understanding of the MW process by microorganisms and plants. Noticeably, our
718 genomic analyses revealed that some bacteria harbour both the PQQ-dependent and PQQ-
719 independent GDH in their genome as stated in some representative strains of the genus
720 *Burkholderia*, *Glaciimonas*, *Gluconobacter* or *Pantoea* (this study; Uroz *et al.* in press). The
721 next step for future researches will be to determine whether i) the GDH activity is regulated
722 by the availability of inorganic nutrients (P, K, Mg…) as is the case for the PQQ-dependent
723 GDH (Sasnow *et al.* 2016) or the presence/absence of a mineral (Uroz *et al.* 2020) and ii) the
724 GDH genes can be used as biomarkers of the presence of the mineral weathering function into
725 the soil.

726

727

728 **ACKNOWLEDGMENTS**

729 This work was supported by grants from the EC2CO program and the Labex ARBRE to S.U.
730 L.P. was also supported by a fellowship from the French Ministère de l'Enseignement
731 Supérieur, de la Recherche et de l'Innovation. The UMR1136 and UR1138 are supported by
732 the ANR through the Laboratory of Excellence Arbre (ANR-11-LABX-0002-01). The authors
733 thank Dr. P. Frey-Klett for helpful discussions, E. Morin for bioinformatics support and the
734 MICROSCOPE team (Genoscope) for the support provided on the genome of strain PMB31.

735

736 **REFERENCES**

737 Alder NN, Theg SM. Energetics of protein transport across biological membranes: A study of
738 the thylakoid delta pH-dependent/cpTat pathway. *Cell Press*. 2003;112:231–242.
739 Allen SCH, Barrett CML, Ray N, Robinson C. Essential cytoplasmic domains in the
740 *Escherichia coli* TatC protein. *J Biol Chem.* 2002;277:10362–10366.
741 Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman, DJ. Gapped
742 BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic*
743 *Acids Res.* 1997; 25:3389–3402.
744 Babu-Khan S, Yeo TC, Martin WL, et al. Cloning of a mineral phosphate-solubilizing gene
745 from *Pseudomonas cepacia*. *Appl Environ Microbiol* 1995;61:972–978.
746 Balland C, Poszwa A, Leyval C, Mustin C. Dissolution rates of phyllosilicates as a function
747 of bacterial metabolic diversity. *Geochim Cosmochim Acta* 2010; 74:5478–5493.
748 Bendtsen J, Nielsen H, Widdick D, Palmer T, Brunak S. Prediction of twin-arginine signal
749 peptides. *BMC Bioinformatics* 2005; 6:167.
750 Bonneville S, Smits MM, Brown A, *et al.* Plant-driven fungal weathering: Early stages of
751 mineral alteration at the nanometer scale. *Geology*. 2009; 37:615-618.
752 Brantley SL. Kinetics of Mineral Dissolution, p. 151–210. In Brantley, SL, Kubicki, JD,
753 White, AF (eds.), Kinetics of Water-Rock Interaction. 2008; Springer New York, New York,
754 NY.
755 Calvaruso C, Mareschal L, Turpault M-P, Leclerc E. Rapid clay weathering in the rhizosphere
756 of norway spruce and oak in an acid forest ecosystem. *Soil Sci Soc Am J.* 2009; 73:331–338.
757 Calvaruso C, Turpault M-P, Frey-Klett P. Root-associated bacteria contribute to mineral
758 weathering and to mineral nutrition in trees: a budgeting analysis. *Appl Environ Microbiol*.
759 2006; 72:1258–1266.
760 Calvaruso, C., Turpault, M. P., Leclerc, E., Ranger, J., Garbaye, J., Uroz, S., & Frey-Klett, P.
761 Influence of forest trees on the distribution of mineral weathering-associated bacterial

762 communities of the *Scleroderma citrinum* mycorrhizosphere. *Appl Environ Microbiol.*
763 2010;76:4780-4787.
764 Cavener DR. GMC oxidoreductases: A newly defined family of homologous proteins with
765 diverse catalytic activities. *J Mol Biol.* 1992; 223:811-814.
766 Chhabra S, Brazil D, Morrissey J, Burke JI, *et al.* Characterization of mineral phosphate
767 solubilization traits from a barley rhizosphere soil functional metagenome. *Microbiol Open*
768 2013; 2:717–724.
769 Chilton M-D, Currier TC, Farrand SK, Bendich AJ, *et al. Agrobacterium tumefaciens* DNA
770 and PS8 bacteriophage DNA not detected in crown gall tumors. *Proc Natl Acad Sci*
771 1974;71:3672–3676.
772 Colin Y, Nicolitch O, Turpault M-P, Uroz S. Mineral types and tree species determine the
773 functional and taxonomic structures of forest soil bacterial communities. *Appl Environ*
774 *Microbiol*. 2017; 83:e02684-16.
775 Collignon, C., Uroz, S., Turpault, M. P., & Frey-Klett, P. (2011). Seasons differently impact
776 the structure of mineral weathering bacterial communities in beech and spruce stands. *Soil*
777 *Biol and Biochem.* 2011;43:2012-2022.
778 Crow A, Le Brun NE, Oubrie A. The role of ResA in type II cytochrome c maturation.
779 *Biochem Soc Trans.* 2005; 33:149–151.
780 Dennis JJ, Zylstra GJ. Plasposons: modular self-cloning minitransposon derivatives for rapid
781 genetic analysis of Gram-negative bacterial genomes. *Appl Environ Microbiol.* 1998;
782 64:2710–2715.
783 Derrien D, Marol C, Balesdent J. The dynamics of neutral sugars in the rhizosphere of wheat.
784 An approach by 13 C pulse-labelling and GC/C/IRMS. *Plant and Soil*. 2004; 267: 243-253.

785 DeShazer D, Brett PJ, Carlyon R, Woods DE. Mutagenesis of *Burkholderia pseudomallei*
786 with Tn5-OT182: isolation of motility mutants and molecular characterization of the flagellin
787 structural gene. *J Bacteriol.* 1997; 179:2116–2125.
788 Dilks K, Rose RW, Hartmann E, Pohlschroder M. Prokaryotic utilization of the Twin-
789 Arginine Translocation pathway: a genomic survey. *J Bacteriol.* 2003; 185:1478-1483.
790 Driessen AJ, Wickner W. Proton transfer is rate-limiting for translocation of precursor
791 proteins by the Escherichia coli translocase. *Proc Natl Acad Sci*. 1991; 88:2471-2475.
792 Ebert BE, Kurth F, Grund M, Blank LM, *et al.* Response of *Pseudomonas putida* KT2440 to
793 increased NADH and ATP demand. *Appl Environ Microbiol*. 2011; 77:6597-6605.
794 El-Gebali S, Mistry J, Bateman A, Eddy SR, *et al.* The Pfam protein families database in
795 *Nucleic Acids Res.* 2019; 47:D427–D432.
796 Ferri S, Kojima K, Sode K. Review of glucose oxidases and glucose dehydrogenases: A
797 bird's eye view of glucose sensing enzymes. *J Diabetes Sci Technol*. 2011;5:1068–1076.
798 Gadd GM. Geomycology: biogeochemical transformations of rocks, minerals, metals and
799 radionuclides by fungi, bioweathering and bioremediation. *Mycol Res.* 2007; 111:3–49.
800 Goldstein AH. Recent progress in understanding the molecular genetics and biochemistry of
801 calcium phosphate solubilization by Gram negative bacteria. *Biol Agric Hortic.* 1995;12:185–
802 193.
803 Goldstein A, Lester T, Brown J. Research on the metabolic engineering of the direct oxidation
804 pathway for extraction of phosphate from ore has generated preliminary evidence for PQQ
805 biosynthesis in *Escherichia coli* as well as a possible role for the highly conserved region of
806 quinoprotein dehydrogenases. Biochim Biophys Acta BBA - *Proteins Proteomics*. 2003;
807 1647:266–271.

808 Gouy M, Guindon S, Gascuel O. SeaView Version 4: A Multiplatform Graphical User
809 Interface for Sequence Alignment and Phylogenetic Tree Building. *Mol Biol Evol.* 2010;
810 27:221–224.
811 Grayston SJ, Campbell CD. Functional biodiversity of microbial communities in the
812 rhizospheres of hybrid larch (*Larix eurolepis*) and Sitka spruce (*Picea sitchensis*). Tree
813 Physiol. 1996;16:1031-1038.
814 Hinsinger P, Plassard C, Tang C, Jaillard B. Origins of root-mediated pH changes in the
815 rhizosphere and their responses to environmental constraints: A review. *Plant Soil* 2003;
816 248:43–59.
817 Jolivet C, Angers DA, Chantigny MH, Andreux F, Arrouays D. Carbohydrate dynamics in
818 particle-size fractions of sandy spodosols following forest conversion to maize cropping. Soil
819 Biology and Biochemistry. 2006; 38:2834-2842.
820 Jongmans AG, van Breemen N, Lundström U, van Hees PAW, *et al.* Rock-eating fungi.
821 Nature 1997; 389:682–683.
822 Kawai S, Goda-Tsutsumi M, Yakushi T, Kano K, *et al.* Heterologous overexpression and
823 characterization of a flavoprotein-cytochrome c complex fructose dehydrogenase of
824 *Gluconobacter japonicus* NBRC3260. *Appl Environ Microbiol.* 2013; 79:1654–1660.
825 Kovacs N. Identification of *Pseudomonas pyocyanea* by the oxidase reaction. Nature 1956;
826 178:703–703.
827 Kovach ME, Elzer PH, Steven Hill D, Robertson GT, *et al.* Four new derivatives of the
828 broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance
829 cassettes. Gene 1995; 166:175–176.
830 Koele N, Turpault M-P, Hildebrand EE, Uroz S, *et al.* Interactions between mycorrhizal fungi
831 and mycorrhizosphere bacteria during mineral weathering: Budget analysis and bacterial
832 quantification. *Soil Biol Biochem*. 2009; 41:1935–1942.

833 Kranz RG, Richard-Fogal C, Taylor J-S, Frawley ER. Cytochrome c biogenesis: mechanisms
834 for covalent modifications and trafficking of heme and for heme-iron redox control.
835 *Microbiol Mol Biol Rev.* 2009; 73:510–528.
836 Landeweert R, Hoffland E, Finlay RD, Kuyper TW, *et al.* Linking plants to rocks:
837 ectomycorrhizal fungi mobilize nutrients from minerals. *Trends Ecol Evol.* 2001; 16:248–254.
838 Lepleux C, Turpault MP, Oger P, Frey-Klett P, Uroz S. 2012. Correlation of the abundance of
839 betaproteobacteria on mineral surfaces with mineral weathering in forest soils. *Appl Environ*
840 *Microbiol*. 2012; 78:7114-7119.
841 Leveau JHJ, Uroz S, de Boer W. The bacterial genus *Collimonas* : mycophagy, weathering
842 and other adaptive solutions to life in oligotrophic soil environments. *Environ Microbiol*.
843 2010; 12:281–292.
844 Lidbury IDEA, Murphy ARJ, Scanlan DJ, Bending GD, *et al.* Comparative genomic,
845 proteomic and exoproteomic analyses of three *Pseudomonas* strains reveals novel insights
846 into the phosphorus scavenging capabilities of soil bacteria: Comparative genomic, proteomic
847 and exoproteomic analyses of three Pseudomonas strains. *Environ Microbiol*. 2016;18:3535–
848 3549.
849 Liu ST, Lee LY, Tai CY, Hung CH, *et al.* Cloning of an *Erwinia herbicola* gene necessary for
850 gluconic acid production and enhanced mineral phosphate solubilization in *Escherichia coli*
851 HB101: nucleotide sequence and probable involvement in biosynthesis of the coenzyme
852 pyrroloquinoline quinone. *J Bacteriol.* 1992;174:5814–5819.
853 Martin, F., D. Canet, Marchal JP. *In vivo* abundance 13C NMR studies of the carbohydrate
854 storage in ectomycorrhizal fungi. *Physiol Veg*. 1984; 22:733–743.
855 Matsushita, K., Ameyama, M. [24] d-Glucose dehydrogenase from *Pseudomonas fluorescens*,
856 membrane-bound. In Methods in enzymology (Vol. 89, pp. 149-154). 1982; Academic Press.

857 Medeiros PM, Fernandes MF, Dick RP, Simoneit BR. Seasonal variations in sugar contents
858 and microbial community in a ryegrass soil. *Chemosphere*. 2006; 65:832-839.
859 Mela F, Fritsche K, de Boer W, van den Berg M *et al.* Comparative genomics of bacteria
860 from the genus *Collimonas*: linking (dis)similarities in gene content to phenotypic variation
861 and conservation. *Environ Microbiol Rep*. 2012; 4:424–432.
862 Palmer T, Berks BC. The twin-arginine translocation (Tat) protein export pathway. *Nat Rev*
863 *Microbiol.* 2012 10:483–496.
864 Picard L, Oger, P, Turpault, M-P, Uroz, S. Draft genome of *Collimonas pratensis* strain
865 PMB3(1), an effective mineral weathering and chitin hydrolyzing bacterial strain. Microbial
866 Genom. Res. DOI: 10.1128/MRA.00601-20
867 Pospiech AA versatile quick-prep of genomic DNA from Gram-positive bacteria. *Trends*
868 *Genet* . 199511:217–218.
869 Puente ME, Bashan Y, Li CY, Lebsky VK. Microbial populations and activities in the
870 rhizoplane of rock-weathering desert plants. II. Growth promotion of cactus seedlings. *Plant*
871 *Biol*. 2004; 6:643–649.
872 Ranger J, Turpault M-P. Input–output nutrient budgets as a diagnostic tool for sustainable
873 forest management. *For Ecol Manag.* 1999; 122:139–154.
874 Rodríguez H, Fraga R. Phosphate solubilizing bacteria and their role in plant growth
875 promotion. *Biotechnol Adv* . 1999; 17:319–339.
876 Saichana I, Moonmangmee D, Adachi O, Matsushita K, *et al.* Screening of thermotolerant
877 *Gluconobacter* strains for production of 5-Keto-D-gluconic acid and disruption of Flavin
878 Adenine Dinucleotide-containing D-gluconate dehydrogenase. *Appl Environ Microbiol.* 2009;
879 75:4240–4247.
880 Sasnow SS, Wei H, Aristilde L. Bypasses in intracellular glucose metabolism in iron-limited
881 Pseudomonas putida. MicrobiologyOpen 2016; 5:3–20.

882 Shiota M, Yamazaki T, Yoshimatsu K, Kojima K, *et al.* An Fe–S cluster in the conserved
883 Cys-rich region in the catalytic subunit of FAD-dependent dehydrogenase complexes.
884 Bioelectrochem. 2016; 112:178–183.
885 Sützl L, Foley G, Gillam EMJ, Bodén M, *et al.* The GMC superfamily of oxidoreductases
886 revisited: analysis and evolution of fungal GMC oxidoreductases. Biotechnol Biofuels 2019;
887 12:1–18.
888 Toyama H, Furuya N, Saichana I, Ano Y, *et al.* Membrane-bound, 2-keto-D-gluconate-
889 yielding D-gluconate dehydrogenase from Gluconobacter dioxyacetonicus IFO 3271:
890 molecular properties and gene disruption. *Appl Environ Microbiol.* 2007; 73:6551–6556.
891 Toyama H, Soemphol W, Moonmangmee D, Adachi O, *et al.* Molecular properties of
892 membrane-bound FAD-containing D-Sorbitol dehydrogenase from thermotolerant
893 *Gluconobacter frateurii* isolated from Thailand. *Biosci Biotechnol Biochem.* 2005; 69:1120–
894 1129.
895 Turpault MP, Nys C, Calvaruso C. Rhizosphere impact on the dissolution of test minerals in a
896 forest ecosystem. Geoderma. 2009; 153:147–154.
897 Udovičić M, Sievers M, Zhang D-C, Margesin R, *et al. Glaciimonas alpina* sp. nov. isolated
898 from alpine glaciers and reclassification of Glaciimonas immobilis Cr9-12 as the type strain
899 of *Glaciimonas alpina* sp. nov. *Int J Syst Evol Microbiol*. 2015; 65:1779–1785.
900 Uroz S, Calvaruso C, Turpault MP, Pierrat JC, *et al.* Effect of the mycorrhizosphere on the
901 genotypic and metabolic diversity of the bacterial communities involved in mineral
902 weathering in a forest soil. *Appl Environ Microbiol.* 2007;73:3019–3027.
903 Uroz S, Calvaruso C, Turpault M-P, Frey-Klett P. Mineral weathering by bacteria: ecology,
904 actors and mechanisms. *Trends Microbiol.* 2009; 17:378–387.

905 Uroz S, Calvaruso C, Turpault M-P, Sarniguet A, *et al.* 2009b. Efficient mineral weathering is
906 a distinctive functional trait of the bacterial genus Collimonas. *Soil Biol Biochem.* 2009;
907 41:2178–2186.
908 Uroz S, Turpault M-P, Van Scholl L, Palin B, *et al.* Long term impact of mineral amendment
909 on the distribution of the mineral weathering associated bacterial communities from the beech
910 *Scleroderma citrinum* ectomycorrhizosphere. *Soil Biol Biochem.* 2011; 43:2275–2282.
911 Uroz S, Picard L, Turpault M-P, Auer L, *et al.* Dual transcriptomic and proteomic analyses of
912 the early stage of interaction between *Caballeronia mineralivorans* PML1(12) and mineral. In
913 press. Environmental Microbiol. doi:10.1111/1462-2920.15159
914 Vallenet D, Engelen S, Mornico D, Cruveiller S, MicroScope: a platform for microbial
915 genome annotation and comparative genomics. Database 2009.
916 Wagh J, Shah S, Bhandari P, Archana G,*et al.* Heterologous expression of pyrroloquinoline
917 quinone (pqq) gene cluster confers mineral phosphate solubilization ability to *Herbaspirillum*
918 *seropedicae* Z67. Appl Microbiol Biotechnol. 2014; 98:5117–5129.
919 Wang Y-L, Sun L-J, Xian C-M, Kou F-L, *et al.* Interactions between biotite and the mineral-
920 weathering bacterium *Pseudomonas azotoformans* F77. *Appl Environ Microbiol.* 2020
921 86:e02568-19.
922 Yamaoka H, Yamashita Y, Ferri S, Sode K. Site directed mutagenesis studies of FAD-
923 dependent glucose dehydrogenase catalytic subunit of *Burkholderia cepacia*. *Biotechnol Lett.*
924 2008; 30:1967–1972.
925 Zámocký M, Hallberg M, Ludwig R, Divne C, *et al.* Ancestral gene fusion in cellobiose
926 dehydrogenases reflects a specific evolution of GMC oxidoreductases in Fungi. Gene 2004;
927 338:1–14.

Zhang, Z., Huang, J., He, L., Sheng, X. Distinct weathering ability and populations of culturable mineral-weathering bacteria in the rhizosphere and bulk soils of Morus alba. Geomicrobiology J. 2016; 33:39-45.

958 **FIGURE LEGENDS:**

959

960 **Figure 1: Thin layer chromatography analysis of gluconic acid production by the wild**
961 **type strain PMB3(1) of *Collimonas pratensis***

962 A volume of 8 µl of supernatant of ABm culture was spotted on the TLC. Cultures were done
963 in ABm medium supplemented with 10 g/l of glucose inoculated or not with the WT strain
964 PMB3(1). The supernatant was recovered after 4 days of incubation at 25°C. A volume of 4
965 µl of standards was spotted on the TLC. Standards correspond to 100 mM solutions of
966 glucose, gluconate, 2KGA and 5KGA. After revelation, glucose appeared in dark blue spots,
967 gluconate in pink, 2KGA in brown and 5KGA in dark purple spot. The pink arrow indicates
968 the pink spot of the gluconate standard.

969

970 **Figure 2: Biotite dissolution assay by the WT and mutant strains**

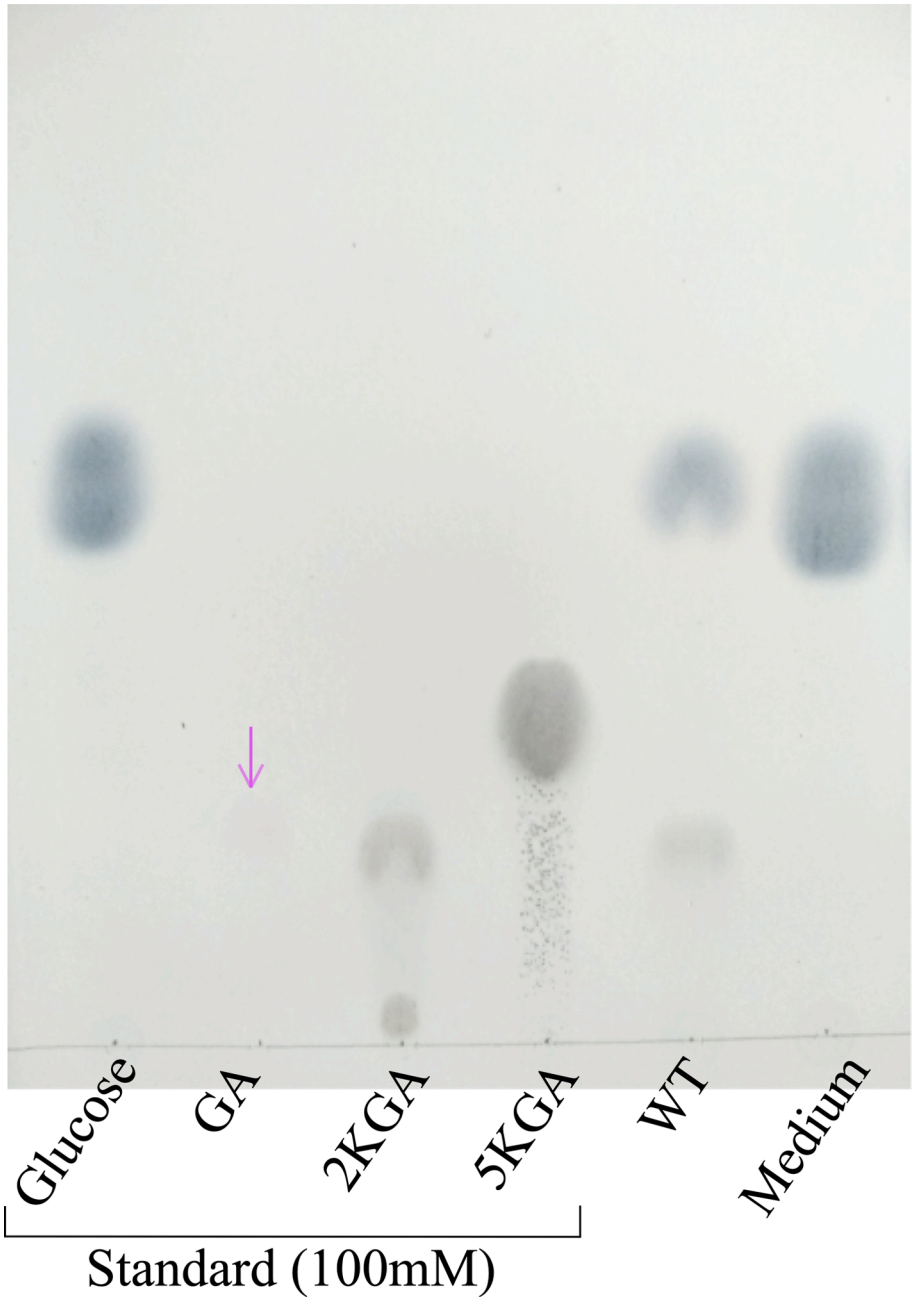
971 **(A**) Measure of iron released from biotite after 7 days of incubation at 25°C in BHm
972 supplemented with (2 g/l) glucose ; **(B)** Measure of pH measurement with bromocresol green.
973 The control treatment termed 'Biotite' corresponds to the non-inoculated treatment. It allows
974 the measurement of the spontaneous release of iron from biotite in the medium in our
975 experimental conditions. The second control termed 'medium' corresponds to the non-
976 inoculated medium without biotite. Each value is the mean of three independent replicates ±
977 the standard error of the mean. Measures with different letters are significantly different
978 according to an ANOVA test ($P<0.05$) ; **(C)** Individual mineral-weathering potentials (pH
979 versus Fe released) of the WT strain and related MW mutants. White and grey squares
980 correspond to the controls of the experiment (white, medium without biotite and non
981 inoculated; grey, medium with biotite and non inoculated). The two curves indicate the

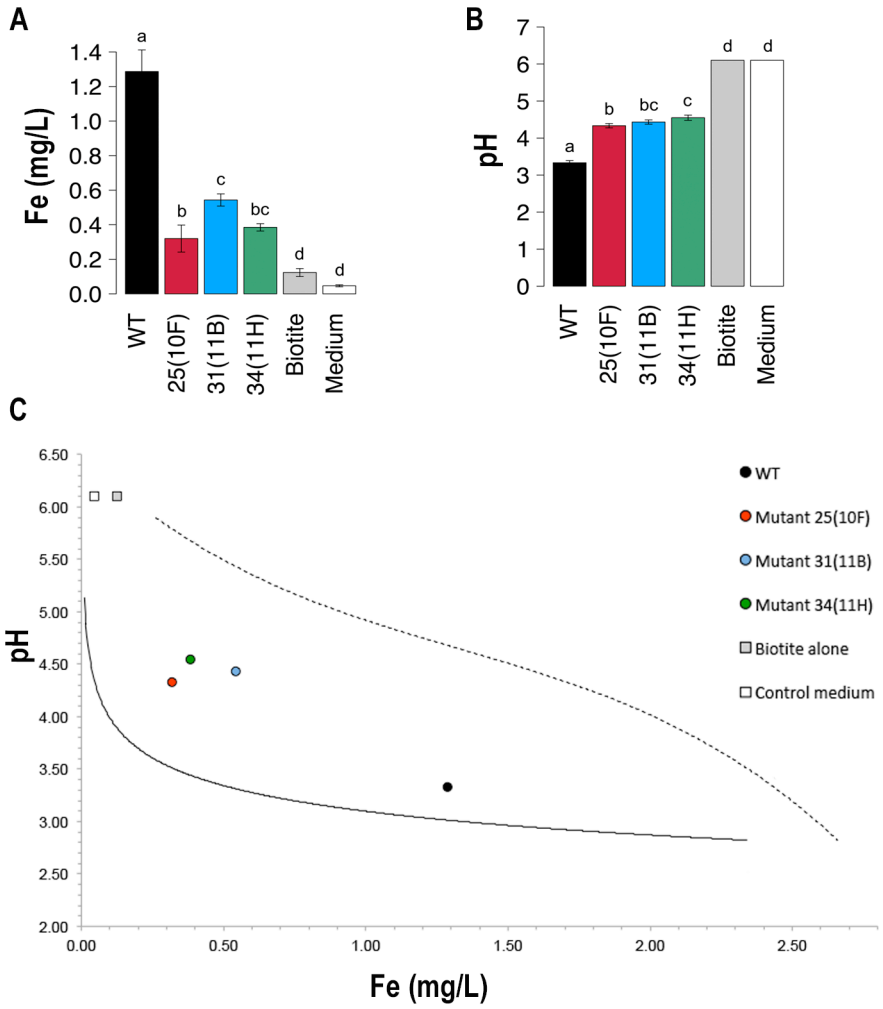
982 mineral-weathering effect of a complexing agent (citric acid) (dashed line) and a strong acid
983 (hydrochloric acid) (solid line).
984
985 **Figure 3: Genomic region of the plasposon insertion in the different mutants**
986 For each mutant, the genes have been represented with large horizontal arrows and their
987 length was scaled according to their nucleotidic size. The genes or group of genes directly
988 impacted by the Tn5-OT182 insertion are filled with colour. The small black arrow indicates
989 the location of the insertion of the Tn5-OT182 in the genome of PMB3(1) as identified by
990 BlastN for the mutants 25(10F) **(A)**, 31(11B) **(B)**, 34(11H) **(C)** and 59(7B) **(D)**. Gene
991 annotation is presented under each gene and the complete description is provided in Table 2
992 and in the supplementary Table S1, S2 and S3.
993
994 **Figure 4: Growth test on WT, mutants and complemented mutants in LB broth**
995 Growth monitoring of the wild type strain and of the mutants 25(10F), 31(11B) and 34(11H)
996 complemented or not was done in LB medium over a period of 120 h. Growth was performed
997 under orbital shaking and the measures correspond to absorbance measurement at λ600 nm.
998 Each point is the mean of 3 independent replicates. For legibility reason, the standard error
999 was not presented.
1000
1001 **Figure 5: Acidification ability and growth on different carbon sources for the WT and**
1002 **mutants strains**
1003 The acidification (right axis) after 5 days of incubation at 25°C of the BHm medium with
1004 different carbon source (glucose, gluconate, 2keto gluconate and 5-keto gluconate) is
1005 presented by coloured circles according to the strain considered. The absorbance at 595 nm
1006 (left axis) representing the growth of the different strains is presented by coloured bars. The

1007 non-inoculated medium was used as control. Each value is the mean of three independent
1008 replicates ± the standard error of the mean.
1009

1010 **Figure 6: Molecular mechanisms of PQQ-independent GDH in mineral weathering**
1011 **process of *Collimonas pratensis* PMB3(1)**

1012 The hypothetical functional model presented in here was generated based on the data
1013 extracted in this study and from the literature. The GDH of the strain *Collimonas pratensis*
1014 PMB3(1) is composed of three subunits GdhS, GdhL and GdhC (blue). The GdhL subunit is
1015 able to bind the cofactor FAD (yellow) in the cytoplasm. The complex GdhS-GdhL is formed
1016 in the cytoplasm and then transported in the periplasm. This transfer is done through the Twin
1017 Arginine Translocation system (green), due to the presence of a TAT signal of the GdhS
1018 subunit. The functioning of the TAT system is based on the proton motive force. The
1019 cytochrome subunit (GdhC) is maturated by the ResABC system (orange) and translocated in
1020 the periplasm where it fixes heme. The GDH converts glucose in gluconate (GA) and protons
1021 (pink). Gluconate is then converted to 2-Keto-gluconate (2KGA) by the gluconate
1022 dehydrogenase (grey). The conversion of glucose to gluconic acid and then to 2KGA
1023 corresponds to the direct oxidative pathway (dotted red box). The 2KGA can go to the
1024 cytoplasm where it is then converted into 6-P-gluconate by a 2-keto gluconate kinase and a 2-
1025 keto gluconate 6-phosphate reductase prior to enter in the Pentose or the Entner-Doudoroff
1026 pathways. Both GA and 2KGA support the growth of strain PMB3(1). Protons produced by
1027 GDH activity lead to an acidification of the extracellular environment and permit the
1028 dissolution of minerals through acidolysis. The mineral weathering makes nutrients available
1029 to bacteria.
1030
1031
1032


Glucose
GA
2KGA
5KGA
WT
Medium
Standard (100mM)


A
Fe (mg/L)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
a
b
c
bc
d
d
WT
25(10F)
31(11B)
34(11H)
Biotite
Medium
B
pH
7
6
5
4
3
2
1
0
a
b
bc
c
d
d
WT
25(10F)
31(11B)
34(11H)
Biotite
Medium
C
pH
6.50
6.00
5.50
5.00
4.50
4.00
3.50
3.00
2.50
2.00
0.00
0.50
1.00
1.50
2.00
2.50
Fe (mg/L)
WT
Mutant 25(10F)
Mutant 31(11B)
Mutant 34(11H)
Biotite alone
Control medium

1057

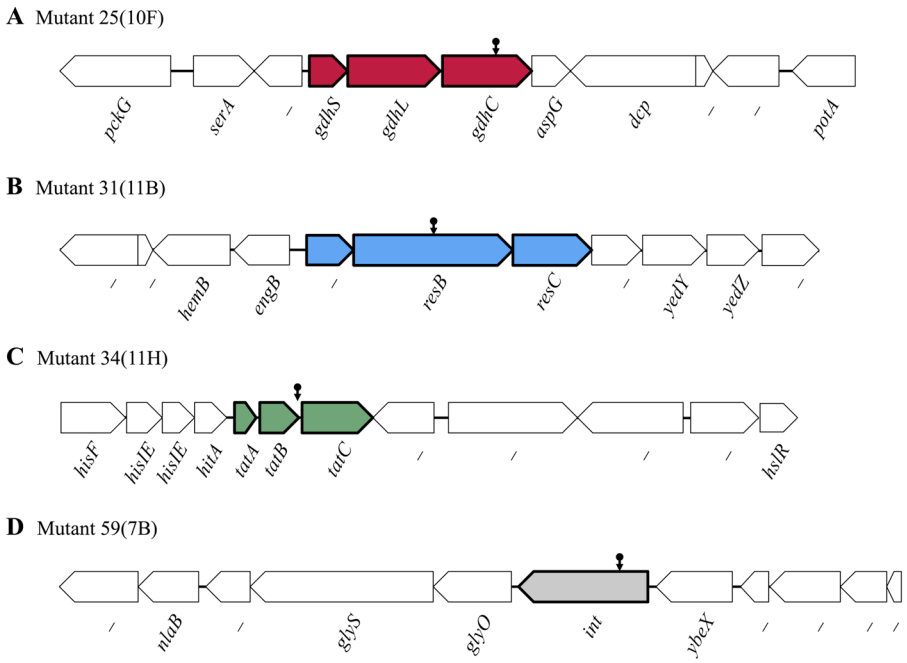
1058

1059

1060

1061

1062



1063

1064

1065

1066

1067

1068

1069

1070

1071

1072

1073

1074

1075

1076

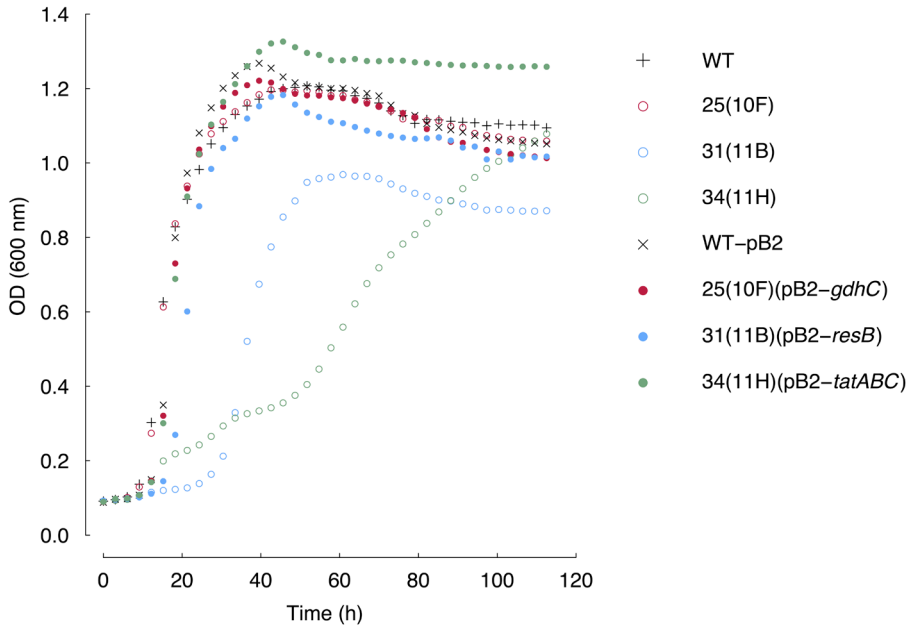
1077

1078

1079

1080

1081



1082

1083

1084

1085

1086

1087

1088

1089

1090

1091

1092

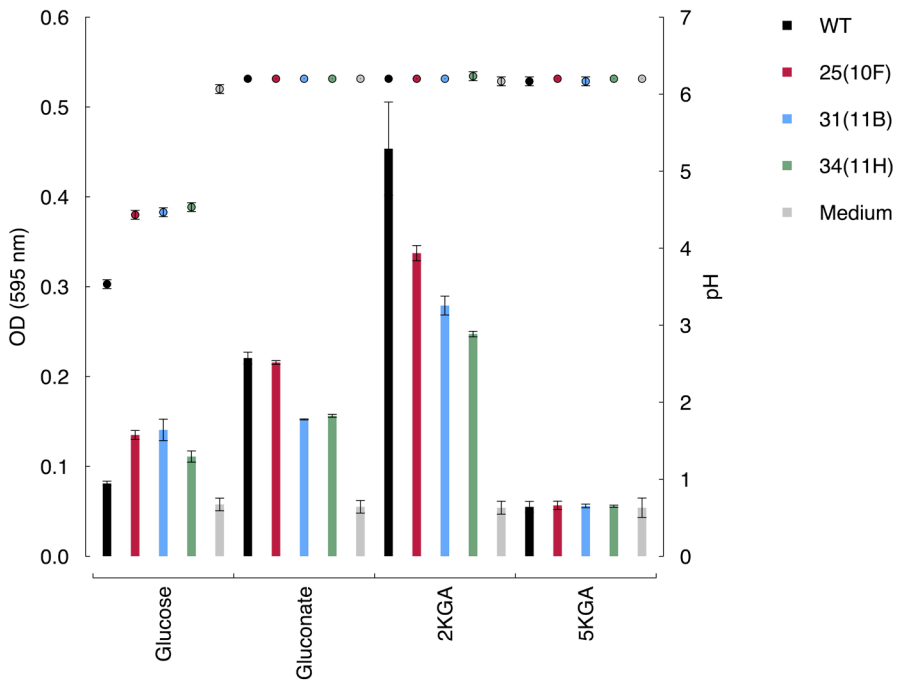
1093

1094

1095

1096

1097



1098

1099

1100

1101

1102

1103

1104

1105

1106

1107

1108

1109

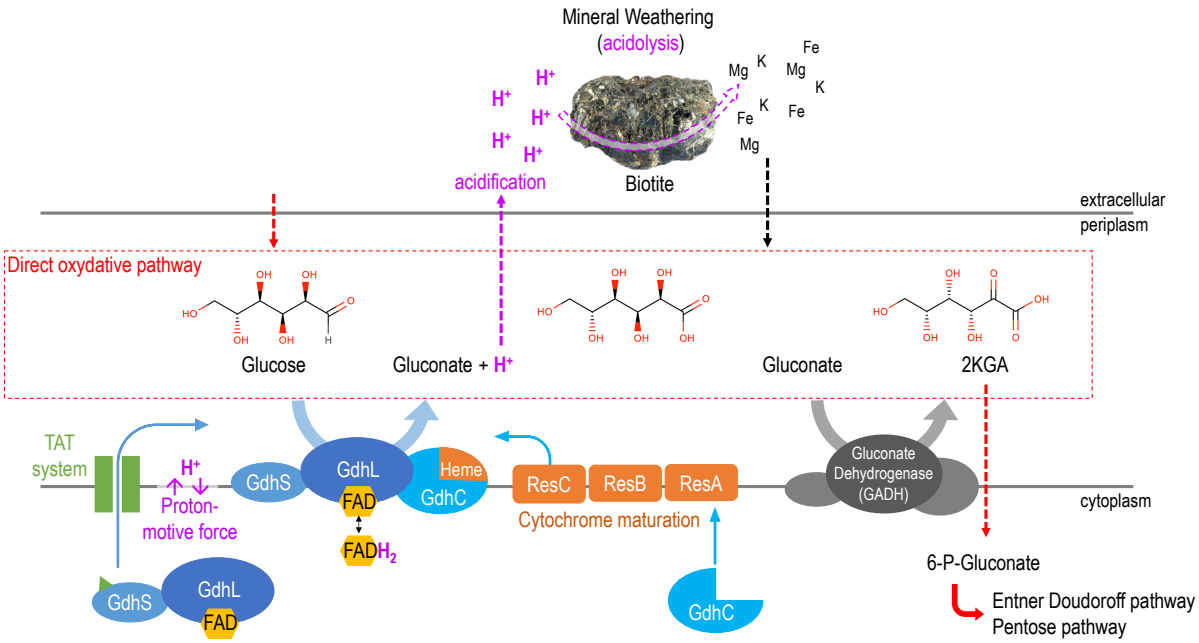
1110

1111

1112

1113

1114



SUPLEMENTAL MATERIAL

Identification of a novel type of glucose dehydrogenase involved in the mineral weathering ability of *Collimonas pratensis* strain PMB3(1)

Picard, L.[1,2], Turpault, M-P.[2], Oger, P.M.[3], Uroz, S.[1,2#]

1133

1134

1135 **Table S1 : Description of the genomic region of the plasposon insertion in mutant 31(11B)**

Accession number	Putative function	Gene annotation	Length (nt)	BLASTP (% Id against NCBI)		
NKI71669.1	Mg2+ transporter protein, cora-like protein	_	1149	93%	Magnesium transporter cora family protein [WP_098495086.1]	Collimonas sp. PA-H2
NI*	Protein of unknown function	_	195	97%	Hypothetical protein cpter291_0433 [AMP12718.1]	Collimonas pratensis
NKI71670.1	Delta-aminolevulinic acid dehydratase	hemB	1068	99%	Porphobilinogen synthase [PFH10198.1]	Collimonas sp. PA-H2
NKI71671.1	Putative GTP-binding protein engb	engB	732	95%	Yiha family ribosome biogenesis GTP-binding protein [WP_098495089.1]	Collimonas sp. PA-H2
NKI71672.1	Cytochrome c4	_	666	99%	Cytochrome c4 [WP_098495090.1]	Collimonas sp. PA-H2
NKI71673.1	**Cytochrome c-type biogenesis protein Ccs1/ResB**	**resB**	**2160**	**94%**	**Cytochrome c biogenesis protein resb [WP_098495091.1]**	**Collimonas sp. PA-H2**
NKI71674.1	Cytochrome c-type biogenesis protein ccsa/resc	resC	1146	99%	C-type cytochrome biogenesis protein ccsb [WP_098497480.1]	Collimonas sp. PA-H2

NKI71675.1	Putative phosphatase yieh	_	684	92%	HAD family hydrolase [WP_098495092.1]	Collimonas sp. PA-H2
NKI71676.1	Exported heme-molybdoenzyme molybdopterin-containing subunit yedy; TAT export	yedY	981	94%	Protein-methionine-sulfoxide reductase catalytic subunit msrp [WP_098495093.1]	Collimonas sp. PA-H2
NKI71677.1	Sulfoxide reductase heme-binding subunit yedz	yedZ	669	89%	Sulfoxide reductase heme-binding subunit yedz [WP_098495094.1]	Collimonas sp. PA-H2
NKI71678.1	Conserved exported protein of unknown function	_	825	87%	Transporter [WP_092414082.1]	Collimonas sp. OK307

1136 * Some of the proteins listed have not been identified (NI) according to the prediction tools used by NCBI, but were predicted in the MICROSCOPE platform.

1137

1138

1139 **Table S2 : Description of the genomic region of the plasposon insertion in mutant 34(11H)**

Accession number	Putative function	Gene annotation	Length (nt)	BLASTP (%Id against NCBI)		
NKI71753.1	Imidazole glycerol phosphate synthase, catalytic subunit with hish	hisF	765	99%	Imidazole glycerol phosphate synthase subunit hisf [WP_092356093.1]	Collimonas sp. OK242
NKI71754.1	Fragment of bifunctional protein [Includes: phosphoribosyl-AMP cyclohydrolase; phosphoribosyl-ATP pyrophosphatase] (part 1)	hisIE	396	97%	Phosphoribosyl-AMP cyclohydrolase [WP_092434873.1]	Collimonas sp. OK607
NKI71755.1	Fragment of bifunctional protein [Includes: phosphoribosyl-AMP cyclohydrolase; phosphoribosyl-ATP pyrophosphatase] (part 2)	hisIE	378	94%	Phosphoribosyl-ATP diphosphatase [WP_098495230.1]	Collimonas sp. PA-H2
NKI71756.1	Protein hita	hitA	372	93%	Histidine triad nucleotide-binding protein [WP_098495231.1]	Collimonas sp. PA-H2
NKI71757.1	Sec-independent protein translocase protein	tatA	231	95%	Twin-arginine translocation protein	Collimonas fungivorans

	tata				tata [AEK60361.1]	Ter331
NKI71758.1	Sec-independent protein translocase protein tatb	tatB	495	93%	Sec-independent protein translocase subunit tatb [WP_098495232.1]	Collimonas sp. PA-H2
NKI71759.1	**Tatabce protein translocation system subunit**	**tatC**	**768**	**95%**	**Twin-arginine translocase subunit tatc [WP_098495233.1]**	**Collimonas sp. PA-H2**
NKI71760.1	Conserved protein of unknown function	_	690	97%	DUF2461 domain-containing protein [WP_098495234.1]	Collimonas sp. PA-H2
NKI71761.1	Emrb/qaca subfamily drug resistance transporter	_	1452	94%	DHA2 family efflux MFS transporter permease subunit [WP_098495235.1]	Collimonas sp. PA-H2
NKI71762.1	Putative htra-like serine protease	_	1167	98%	Do family serine endopeptidase [WP_098495236.1]	Collimonas sp. PA-H2
NKI71763.1	Putative GTP cyclohydrolase 1 type 2	_	768	93%	Dinuclear metal center ybgi/SA1388 family protein [PFH10368.1]	Collimonas sp. PA-H2
NKI71764.1	Heat shock protein 15 homolog	hslR	387	96%	RNA-binding S4 domain-containing protein [WP_098495238.1]	Collimonas sp. PA-H2

1140

1141

1142 **Table S3 : Description of the genomic region of the plasposon insertion in mutant 59(7B)**

Accession number	Putative function	Gene annotation	Length (nt)	Blastp (%id against ncbi)		
NKI72634.1	Zinc metalloprotease	_	984	93%	Hypothetical protein BCF11_1207 [PFH08832.1]	Collimonas sp. PA-H2
NKI72635.1	1-acyl-sn-glycerol-3-phosphate acyltransferase	nlaB	744	98%	1-acyl-sn-glycerol-3-phosphate acyltransferase [WP_098493940.1]	Collimonas sp. PA-H2
NKI72636.1	D-glycero-beta-D-manno-heptose-1,7-bisphosphate 7-phosphatase	_	558	97%	D-glycero-beta-D-manno-heptose 1,7-bisphosphate 7-phosphatase [WP_098493939.1]	Collimonas sp. PA-H2
NKI72637.1	Glycine trna synthetase, beta subunit	glyS	2133	96%	Glycine--trna ligase subunit beta [WP_098493938.1]	Collimonas sp. PA-H2
NKI72638.1	Glycyl-trna synthetase, alpha chain	glyQ	981	99%	Glycine--trna ligase subunit alpha [WP_098493937.1]	Collimonas sp. PA-H2
NKI72639.1	**Apolipoprotein N-acyltransferase**	**lnt**	**1554**	**92%**	**Apolipoprotein N-acyltransferase**	**Collimonas sp. PA-**

					[WP_098493936.1]	**H2**
NKI72640.1	Putative protein involved in divalent ion export	ybeX	882	99%	CBS domain-containing protein [WP_098493935.1]	Collimonas sp. PA-H2
NKI72641.1	Conserved exported protein of unknown function	_	267	83%	Pentapeptide MXKDX repeat protein [WP_098497346.1]	Collimonas sp. PA-H2
NKI72642.1	Oxidoreductase, molybdopterin binding protein	_	780	97%	Molybdopterin-dependent oxidoreductase [WP_098493934.1]	Collimonas sp. PA-H2
NKI72643.1	Thiosulfate reductase cytochrome B subunit (Membrane anchoring protein)	_	615	94%	Cytochrome b/b6 domain-containing protein [WP_098493933.1]	Collimonas sp. PA-H2
NI*	Protein of unknown function	_	114	97%	Hypothetical protein cpter91_4422 [AMP06731.1]	Collimonas pratensis

1143 * Some of the proteins listed have not been identified (NI) according to the prediction tools used by NCBI, but were predicted in the MICROSCOPE platform.

1144

1145

1146

1147

1148 **Table S4: List of proteins carrying a TAT-signal in the strain *Collimonas pratensis* PMB3(1)**

1149 The COG class of each protein is indicated with one-letter abbreviation as followed: P, inorganic ion transport and metabolism; Q, secondary
1150 metabolites biosynthesis, transport and catabolism; C, energy production and conversion; E, amino acid metabolism and transport; F, nucleotide
1151 metabolism and transport; G, carbohydrate metabolism and transport; H, coenzyme and metabolism; J, translation, ribosomal structure and
1152 biogenesis; K, transcription; L, replication, recombination and repair; M, cell wall, membrane, envelope biogenesis; N, cell motility; O,
1153 posttranslational modification, protein turnover, chaperones; S, function unknown; U, intracellular trafficking, secretion, and vesicular transport;
1154 W, extracellular structures. The bold rows indicated the proteins mentioned in the text and could be involved in MW. Some of the proteins listed
1155 have not been identified (NI) according to the prediction tools used by NCBI, but were predicted in the MICROSCOPE platform.

1156

Accession number (MAGE)	Gene annotation	Putative function	COG class
NKI68675.1	yagR	Aerobic-type carbon monoxide dehydrogenase, large subunit	C
NKI69300.1	iorB	Isoquinoline 1-oxidoreductase beta subunit	C
NKI69617.1	soxH	putative cytochrome C oxidase polypeptide II oxidoreductase protein	C
NKI71676.1	yedY	exported heme-molybdoenzyme molybdopterin-containing subunit YedY; TAT export	C
NKI72252.1	mqo	malate dehydrogenase, FAD	C
NKI72324.1	FAD	FMN-dependent dehydrogenase	C

NKI67910.1		exported protein of unknown function	C
NKI69616.1		conserved protein of unknown function	C
NKI68636.1		putative branched-chain amino acid transport protein (ABC superfamily, peri_bind)	E
NKI68806.1		ABC transporter substrate-binding protein	E
NKI69719.1	aguA	putative agmatine deiminase	E
NKI69739.1	ggtA	Gamma-glutamyltranspeptidase	E
NKI72278.1		Monoamine oxidase	E
NKI71314.1		conserved protein of unknown function	E
NKI68557.1	rihA	Pyrimidine-specific ribonucleoside hydrolase RihA	F
NKI71053.1		putative nucleoside hydrolase	F
NKI67981.1	ykgB	putative hemagglutinin-like protein	G
NKI68770.1	bga	Beta-galactosidase	G
NKI69257.1	ytfQ	putative sugar transporter subunit: periplasmic-binding component of ABC superfamily	G
NKI68420.1		conserved membrane protein of unknown function	G
NKI69773.1		conserved exported protein of unknown function	G
NKI72549.1		conserved exported protein of unknown function	H
NKI71161.1		Amidase, Asp-tRNAAsn/Glu-tRNAGln amidotransferase A subunit	J
NKI71645.1	rplP	50S ribosomal protein L16	J
NI		protein of unknown function	J

NKI68641.1		LysR family transcriptional regulator	K
NKI72806.1		transposase	L
NKI68262.1	plcN	Non-hemolytic phospholipase C	M
NKI68666.1	mrcA	Penicillin-binding protein 1A	M
NKI69363.1	acpA	Acid phosphatase	M
NKI69405.1	plcN	Non-hemolytic phospholipase C	M
NKI71095.1	plcN	Non-hemolytic phospholipase C	M
NKI72507.1	amiC	N-acetylmuramoyl-L-alanine amidase	M
NKI68496.1		putative Peptide-N(4)-(N-acetyl-beta-glucosaminyl)asparagine amidase	N
NKI69248.1	ppiA	peptidyl-prolyl cis-trans isomerase A (rotamase A)	O
NKI69912.1	msrB	Methionine sulfoxide reductase	O
NKI72023.1		CcsA-like protein	O
NKI72219.1	lifO	Lipase chaperone	O
NKI72562.1		Methionine-R-sulfoxide reductase	O
NKI70121.1	mdoD	glucan biosynthesis protein, periplasmic	P
NKI71510.1	**phoD**	**Alkaline phosphatase**	P
NKI72113.1	nasF	Nitrate ABC transporter, nitrate-binding protein	P
NKI72201.1		ABC-type nitrate/sulfonate	P
NKI72274.1	Nitrate	sulfonate	P

NKI72605.1		**Alkaline phosphatase**	P
NKI72716.1	metQ	DL-methionine transporter subunit ; periplasmic-binding component of ABC superfamily	P
NKI67804.1	ysgA	putative hydrolase	Q
NKI68502.1		putative multicopper oxidase	Q
NI		protein of unknown function	Q
NKI71458.1		conserved exported protein of unknown function	Q
NKI72652.1		Type 4 fimbrial biogenesis protein PilY1	U
NKI72179.1		exported protein of unknown function	U
NKI70646.1		YadA-like protein	W
NKI68516.1	ars	Hydrolase	S
NKI68701.1		putative transmembrane protein	S
NKI68877.1		Sulfite dehydrogenase (Cytochrome) subunit SorA apoprotein	S
NKI69325.1	yvnB	LPXTG-motif cell wall anchor domain protein	S
NKI70794.1	**gdhS**	**Sorbitol/Glucose/2-Keto-D-gluconate dehydrogenase, membrane-bound, gamma subunit**	S
NKI71018.1		Phosphoserine phosphatase	S
NKI71108.1		putative secreted protein	S
NKI71438.1		Lipoprotein	S
NKI71603.1	ophA	ABC peptide	S
NKI68405.1		conserved exported protein of unknown function	S

NKI68510.1	conserved exported protein of unknown function	S
NKI68544.1	conserved protein of unknown function	S
NKI68550.1	protein of unknown function	S
NKI69712.1	conserved protein of unknown function	S
NKI70227.1	conserved protein of unknown function	S
NKI71034.1	conserved exported protein of unknown function	S
NKI71518.1	conserved exported protein of unknown function	S
NKI71778.1	exported protein of unknown function	S
NKI71790.1	conserved exported protein of unknown function	S
NI	protein of unknown function	S
NI	protein of unknown function	S

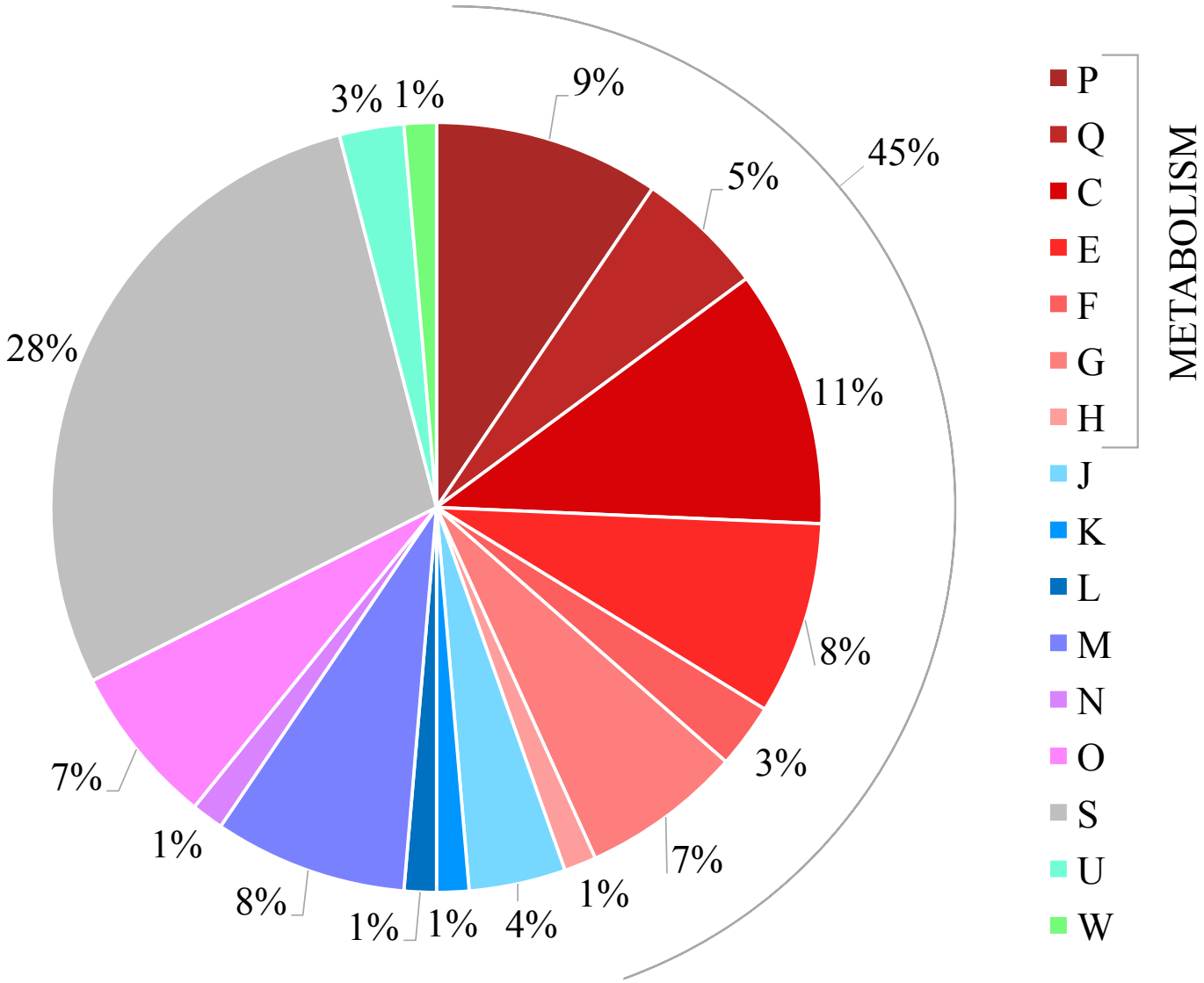


Figure S1 : COG class distribution of the TAT-proteins

The proteins highlighted in red colors light to dark are those involved in metabolism, which encompass ca. 45 % of the total TAT-proteins.

```
1163  Domain 1 FAD          1 --------------------IVGSGVAGALVAHQLASSGASVLVLEAGPRLK--------
1164  Col_PMB31             1 MA--D--N------LSADVVIVGSGVAGALVAHQLASSGASVLVLEAGPRLERWRIVENY
1165  Col.prat_Ter291       1 MA--D--N------LSADVVIVGSGVAGALVAHQLASSGASVLVLEAGPRLERWRIVENY
1166  Col.prat_Ter91        1 MA--D--N------LSADVVIVGSGVAGALVAHQLASSGASVLVLEAGPRLERWRIVENY
1167  Col.arenae_Cal35      1 MA--D--N------LSADVVIVGSGVAGALVAHQLASSGASVLVLEAGPRLERWRIVENY
1168  Col_sp_PA-H2          1 MA--D--N------LSADVVIVGSGVAGALVAHQLASSGASVLVLEAGPRLERWRIVENY
1169  Col.fungi_Ter6        1 MA--D--N------LSADVVIVGSGVAGALVAHQLALSGASVLVLEAGPRLERWRIVENY
1170  Col.fungi_Ter331      1 MA--D--N------LSADVVIVGSGVAGALVAHQLALSGASVLVLEAGPRLERWRIVENY
1171  Col.arenae_Ter282     1 MA--D--N------LSADVVIVGSGVAGALVAHQLALSGASVLVLEAGPRIERWRIVENY
1172  Col.arenae_Ter10      1 MA--D--N------LSADVVIVGSGVAGALVAHQLALSGASVLVLEAGPRIERWRIVENY
1173  Col_sp_OK307          1 MA--D--N------LSADVVIVGSGVAGALVAHQLALSGASVLVLEAGPRLERWRIVENY
1174  Col_sp_OK607          1 MA--D--N------LSADVVIVGSGVAGALVAHQLALSGASVLVLEAGPRLERWRIVENY
1175  Col_sp_OK242          1 MA--E--N------LSADVVIVGSGVAGALVAHQLALSGASVLVLEAGPRLERWRIVENY
1176  Glacii_sp_PCH181      1 MA--D--N------LSADVVIVGSGVAGALVAHQLAKAGVSVLVLEAGPRMERWRIVENY
1177  Glacii_sp_GS1         1 MA--E--N------LSADVVIVGSGVAGAIVAHQLAKSGVSVLVLEAGPRLERWRIVENY
1178  Janthin_UBA1134       1 MN--E--T------LSADVVIVGSGVAGALVAHQMALAGASVLVLEAGPRLERWRLVENF
1179  Herba_seropedic       1 MS--N--P------NSADVVVIGAGVTGALLATRLAKTGKSVLLLESGPRMGRDQLVERF
1180  Herba_rubrisuba       1 MS--N--V------NSADVVVIGAGVTGALLATRLARSGKSVLLLESGPKMGRDQLVERF
1181  Burkholderia          1 MA--DTDT------QKADVVVVGSGVAGAIVAHQLAMAGKAVILLEAGPRMPRWEIVERF
1182  Gluc_japonicus        1 MS--N-ET------LSADVVIIGAGICGSLLAHKLVRNGLSVLLLDAGPRRDRSQIVENW
1183  Pantoea agglo.        1 MKKP---VFTAQGDASADIVIVGSGIVGGMMANELVSQGYSVLVLEAGLRIDRAQAVENW
1184  Gluc_oxydans          1 MSSSN--S------LSADVVIVGSGVAGASIANELARAGLSVIVLEAGPRIDRQHILENF
1185  Glucon_frateuri       1 MSSSN--S------FSADVVIVGSGVAGASIANELARAGLSVIVLEAGPRIDRQHILENF
1186  Glucon_fra (GADH)     1 -M--T--K------KHVDAIVVGAGWAGSILAKELTEAGLSVVMLERGS--ERATAVEGA
1187
1188
1189  domain                  ------------------------------------------------------------
1190  Col_PMB31            51 RNTP--S--KDDFMVPYPSTKYAPHP-EYTP-ENNYLILKGTH--KYNSQY-I-RAVGGT
1191  Col.prat_Ter291      51 RNTP--S--KDDFMVPYPSTKYAPHP-EYTP-ENNYLILKGTH--KYNSQY-I-RAVGGT
1192  C.pratensis_Ter      51 RNTP--S--KDDFMVPYPSTKYAPHP-EYTP-ENNYLILKGTH--KYNSQY-I-RAVGGT
1193  C.arenae_Cal35       51 RNTP--S--KDDFMVPYPSTKYAPHP-EYTP-ENNYLILKGTH--KYNSQY-I-RAVGGT
1194  Collimonas_sp_P      51 RNTP--S--KDDFMVPYPSTKYAPHP-EYTP-ENNYLILKGTH--KYNSQY-I-RAVGGT
1195  C.fungivorans_T      51 RNTP--S--KDDFMVPYPSTKYAPHP-EYTP-ENNYLILKGSH--KYNSQY-I-RAVGGT
1196  C.fungivorans_T      51 RNTP--S--KDDFMVPYPSTKYAPHP-EYTP-ENNYLILKGSH--KYNSQY-I-RAVGGT
1197  C.arenae_Ter282      51 RNTP--S--KDDFMVPYPSTKYAPHP-EYTP-ENNYLILKGTH--KYNSQY-I-RAVGGT
1198  C.arenae_Ter10       51 RNTP--S--KDDFMVPYPSTKYAPHP-EYTP-ENNYLILKGTH--KYNSQY-I-RAVGGT
1199  Collimonas_sp_O      51 RNTP--S--KDDFMVPYPSTKYAPHP-EYTP-ENNYLILKGTH--KYNSQY-I-RAVGGT
1200  Collimonas_sp_O      51 RNTP--S--KDDFMVPYPSTKYAPHP-EYTP-ENNYLILKGTH--KYNSQY-I-RAVGGT
1201  Collimonas_sp_O      51 RNTP--S--KDDFMAPYPSTKYAPHP-EYTP-ENNYLILKGTH--KYNSQY-I-RAVGGT
1202  Glaciimonas_sp_      51 RNTA--S--KDDFMTPYPSTKHAPHP-EYTP-ENNYLIQKGSQ--KYNSQY-I-RAVGGT
1203  Glaciimonas_sp_      51 RNTP--S--KDDFMTPYPSTKYAPHP-EYTP-ENNYLIFKGEH--KYNSQY-I-RAVGGT
1204  Janthin_UBA1134      51 RNTP--A--KDDFMAPYPSTRHAPHP-QYTP-ENNYLIFKGTH--KYNSQY-I-RAVGGT
1205  Herba_seropedic      51 RSSAF----KSDFMAPYPYSPLAPHP-RYTPEPNGYLQQKGPH--AFNAQY-I-RMFGGT
1206  Herba_rubrisuba      51 RSSAF----KSDFMAPYPYSPLAPHP-RYTPEPNGYLQQKGPH--AFNAQY-I-RMFGGT
1207  AAN39686.1Burkh      53 RNQP--D--KMDFMAPYPSSPWAPHP-EYGP-PNDYLILKGEH--KFNSQY-I-RAVGGT
1208  BAM93252.1Gluco      52 RNMPPDNKSQYDYATPYPSVPWAPHT-NYFP-DNNYLIVKGPDRTAYKQGI-I-KGVGGT
1209  AAF21261.1Panto      58 RNMPFANRAGSDFQGLYPQSKFAPAP-LYFP-RNNYVNVTGPNADSFQQGY-L-RTVGGT
1210  AFW02570.1Gluco      53 RTTEN----KGAYQLPYPPVPWAMHP-PDQGSPNGYLHTTGPDGAAYQQGY-L-RVVGGT
1211  BAD60913.1           53 RTTEN----KGAYQLPYPPVPWAMHP-PDQGSPNGYLHTTGPDGAAYQQGY-L-RVVGGT
1212  Glucon_frateuri      48 YPESI----D-ELRGAIRKRLFQNLSQTTVTVRNTIDQTALPY--RQLAAFLPGEGVGGA
1213
```

```
1214  domain                 ------------------------------------------------------------
1215  Col_PMB31          101 TWHWAAAAWRFLPNDFRMKTLYGV------------GRDWPISYDELEPYYYRAEVELGV
1216  Col.prat_Ter291    101 TWHWAAAAWRFLPNDFRMKTLYGV------------GRDWPISYDELEPYYYRAEVELGV
1217  C.pratensis_Ter    101 TWHWAAAAWRFLPNDFRMKTLYGV------------GRDWPISYDELEPYYYRAEVELGV
1218  C.arenae_Cal35     101 TWHWAAAAWRFLPNDFRMKTLYGV------------GRDWPISYDELEPYYYRAEVELGV
1219  Collimonas_sp_P    101 TWHWAAAAWRFLPNDFRMKTLYGV------------GRDWPISYDELEPYYYGAEVELGV
1220  C.fungivorans_T    101 TWHWAAAAWRFLPNDFKMKTLYGV------------GRDWPISYDELEPYYYRAEVELGV
1221  C.fungivorans_T    101 TWHWAAAAWRFLPNDFKMKTLYGV------------GRDWPISYDELEPYYYRAEVELGV
1222  C.arenae_Ter282    101 TWHWAAAAWRFLPNDFKMKTLYGV------------GRDWPISYDELEPYYYRAEVELGV
1223  C.arenae_Ter10     101 TWHWAAAAWRFLPNDFKMKTLYGV------------GRDWPISYDELEPYYYRAEVELGV
1224  Collimonas_sp_O    101 TWHWAAAAWRFLPNDFKMKTLYGV------------GRDWPISYDELEPYYYRAEVELGV
1225  Collimonas_sp_O    101 TWHWAAAAWRFLPNDFKMKTLYGV------------GRDWPISYDELEPYYYRAEVELGV
1226  Collimonas_sp_O    101 TWHWAAAAWRFLPNDFKMKTLYGV------------GRDWPIGYDELEPYYYRAEVELGV
1227  Glaciimonas_sp_    101 TWHWAAAAWRFLPNDFKLKDVYGV------------GRNWPISYDELEPYYYRAEVELGV
1228  Glaciimonas_sp_    101 TWHWAAAAWRFLPSDFEMKTRYGV------------GRDWPINYDELEPYYYRAEVELGV
1229  Janthin_UBA1134    101 TWHWAASTWRLLPNDFRMRSRYGV------------GRDWPLAYEDLEASYYRAELELGV
1230  Herba_seropedic    102 SWHWAAQLWRYVPNDFRQHSLYGV------------GKDWPISYDDLEDYYYQAEVIAGV
1231  Herba_rubrisuba    102 SWHWAAQLWRYVPNDFRQHSLYGV------------GRDWPISYDDLEDYYYQAEVIAGV
1232  AAN39686.1Burkh    103 TWHWAASAWRFIPNDFKMKSVYGV------------GRDWPIQYDDLEPYYQRAEEELGV
1233  BAM93252.1Gluco    108 TWHWAASSWRYLPNDFKLHSTYGV------------GRDYAMSYDELEPYYYEAECEMGV
1234  AAF21261.1Panto    114 TWHWAASCWRHHPSDFVMQSKYGV------------GRDWPIGYDELEPWYCKAENEIGV
1235  AFW02570.1Gluco    106 TWHWAGCAWRYLPSDFELHSRYGV------------GRDWAIKYDDLEPFYYQAEVMMGV
1236  BAD60913.1         106 TWHWAGCAWRYLPSDFELHSRYGV------------GRDWAIKYDDLEPFYYQAEVMMGV
1237  Glucon_frateuri    101 GLHWSGVHERVSPEDLRLRSSVVERYGEKFIPEGMNLQDYGVTYSELEPFFDKAEKVFGT
1238
1239
1240  domain                 ------------------------------------------------------------
1241  Col_PMB31          149 SGPN-D--------GTDLGSPRKQPYPMDHLPLSWNDQRFSAVVNAN-G--HKVVSEPVA
1242  Col.prat_Ter291    149 SGPN-D--------GTDLGSPRKQPYPMDHLPLSWNDQRFSAVVNAN-G--HKVVSEPVA
1243  C.pratensis_Ter    149 SGPN-D--------GTDLGSPRKQPYPMDHLPLSWNDQRFSAVVNAN-G--HKVVSEPVA
1244  C.arenae_Cal35     149 SGPN-D--------GTDLGSPRKQPYPMEHLPLSWNDQRFSAVVNAN-G--HKVVSEPVA
1245  Collimonas_sp_P    149 SGPN-D--------GTDLGSPRKQPYPMEHLPLSWNDQRFSAVVNAN-G--HKVVSEPVA
1246  C.fungivorans_T    149 SGPN-D--------GTDLGSPRKQPYPMDHLPLSWNDQRFSAVVNGA-G--HKVVSEPVA
1247  C.fungivorans_T    149 SGPN-D--------GTDLGSPRKQPYPMDHLPLSWNDQRFSAVVNGA-G--HKVVSEPVA
1248  C.arenae_Ter282    149 SGPN-D--------GTDLGSPRKQPYPMDHLPLSWNDQRFSAVVNAA-G--HKVVSEPVA
1249  C.arenae_Ter10     149 SGPN-D--------GTDLGSPRKQPYPMDHLPLSWNDQRFSAVVNAA-G--HKVVSEPVA
1250  Collimonas_sp_O    149 SGPN-D--------GTDLGSPRKQPYPMDHLPLSWNDQRFSAVVNGN-G--HKVVSEPVA
1251  Collimonas_sp_O    149 SGPN-D--------GTDLGSPRKQPYPMDHLPLSWNDQRFSAVVNSN-G--HKVVSEPVA
1252  Collimonas_sp_O    149 SGPN-D--------GTDLGSPRKQPYPMDHLPLSWNDQRFSAVVNGN-G--HKVVSEPVA
1253  Glaciimonas_sp_    149 SGPN-D--------GTDLGSPRKQPYPMDHLPLSWNDQRFSAVVNGN-G--HKVVSEPVA
1254  Glaciimonas_sp_    149 SGPN-D--------GTDLGSPRKQPYPMDHLPFSWNDQRFSAVVNAH-G--HKVVSEPVA
1255  Janthin_UBA1134    149 SGPN-D--------GTDLGSPRKQPYPMDHLALSHNDQRFFDVLNAK-G--FTVVPEPVA
1256  Herba_seropedic    150 GGSP-E--------S---GSPRKHPFPMERVPASWLQQRVTARLAPE-F---TVLDDSTG
1257  Herba_rubrisuba    150 GGSP-E--------S---GSPRKHPFPMERVPASWLQQRVTARLAPE-F---TVLDDSTG
1258  AAN39686.1Burkh    151 WGPG-P--------EEDLYSPRKQPYPMPPLPLSFNEQTIKTALNNY-DPKFHVVTEPVA
1259  BAM93252.1Gluco    156 MGPNGE--------EITPSAPRQNPWPMTSMPYGYGDRTFTEIVSKL-G--FSNTPVPQA
1260  AAF21261.1Panto    162 AGPN-D-------PARQSPTERSQPYPMDMVPFAHGDNYFASVVNPH-G--YNLVPIPQG
1261  AFW02570.1Gluco    154 AGPN-M-------DVDDLGSPRSHDYPMKEVPLSYGADQFRKLIHEKTN--YRVVHEPQA
1262  BAD60913.1         154 AGPN-M-------DVDDLGSPRSHDYPMKEVPLSYGADQFRKLIHEKTN--YRVVHEPQA
1263  Glucon_frateuri    161 SGEAYKVNGKIVGNGNIFDGDRSDNFPLPAMKDTFTASTFRAAATEAG---YHPYSLPAA
1264
```

```
1265  Domain 2 GMC_N          -----------CCGNNNCMPICP-----------------------AEKAGARVLPDAVV
1266  Col_PMB31          197  RNSRAYDERPNCCGNNNCMPICPIAAM----------YSGIIHVEKAEKAGARVLPDAVV
1267  Col.prat_Ter291    197  RNSRAYDERPNCCGNNNCMPICPIAAM----------YSGIIHVEKAEKAGARVLPDAVV
1268  C.pratensis_Ter    197  RNSRAYDERPNCCGNNNCMPICPIAAM----------YSGIIHVEKAEKAGARVLPDAVV
1269  C.arenae_Cal35     197  RNSRAYDERPNCCGNNNCMPICPIAAM----------YSGIIHVEKAEKAGAKVLADAVV
1270  Collimonas_sp_P    197  RNSRAYDERPNCCGNNNCMPICPIAAM----------YSGIIHVEKAEKAGARVLPDAVV
1271  C.fungivorans_T    197  RNSRAYDERPNCCGNNNCMPICPIAAM----------YSGIIHVEKAEKAGAKVLPDAVV
1272  C.fungivorans_T    197  RNSRAYDERPNCCGNNNCMPICPIAAM----------YSGIIHVEKAEKAGAKVLPDAVV
1273  C.arenae_Ter282    197  RNSRAYDQRPNCCGNNNCMPICPIAAM----------YSGIIHVEKAEKAGAKVLADAVV
1274  C.arenae_Ter10     197  RNSRAYDQRPNCCGNNNCMPICPIAAM----------YSGIIHVEKAEKAGAKVLADAVV
1275  Collimonas_sp_O    197  RNSRAYDQRPNCCGNNNCMPICPIAAM----------YSGIIHVEKAEKAGAKVLADAVV
1276  Collimonas_sp_O    197  RNSRAYDQRPNCCGNNNCMPICPIAAM----------YSGIIHVEKAEKAGAKVLADAVV
1277  Collimonas_sp_O    197  RNSRAYDQRPNCCGNNNCMPICPIAAM----------YSGIIHVEKAEKAGARVLADAVV
1278  Glaciimonas_sp_    197  RNSRAYDARPNCCGNNNCMPICPIAAM----------YSGIIHVEKAEKAGAKVLPNAVV
1279  Glaciimonas_sp_    197  RNSRAYDERPNCCGNNNCMPICPISAM----------YSGIVHVEKAEKAGAKILPEAVV
1280  Janthin_UBA1134    197  RNSRAYDARPNCCGNNNCMPICPIGAM----------YGGIVHVEKAEKAGARVLANAVV
1281  Herba_seropedic    194  RNSRNYDGRPACCGNNNCMPLCPIDAQ----------YHGGLAAQQAEDSGVQVIPEAVV
1282  Herba_rubrisuba    194  RNSRNYDGRPACCGNNNCMPLCPIDAQ----------YHGGLAAQQAQDSGVQVITEAVV
1283  AAN39686.1Burkh    201  RNSRPYDGRPTCCGNNNCMPICPIGAM----------YNGIVHVEKAERAGAKLIENAVV
1284  BAM93252.1Gluco    205  RNSRPYDGRPQCCGNNNCMPICPIGAM----------YNGVYAAIKAEKLGAKIIPNAVV
1285  AAF21261.1Panto    211  RSTRPWEGRPTCCGNNNCQPICPIGAM----------YNGIHHVERAERNGAVVLAEAVV
1286  AFW02570.1Gluco    204  RNTRPYDKRPTCEGNNNCMPICPIGAM----------YNGIHSVNHAEAAGARIIPNAVV
1287  BAD60913.1         204  RNTRPYDKRPTCEGNNNCMPICPIGAM----------YNGIHSVNHAEAAGARIIPNAVV
1288  Glucon_frateuri    218  NASRQYTN-PYGCQMGPCN-FCGYCSGYDCYLYSKASPNVNVLPALRQDPGFTLISDAHV
1289
1290
1291  Domain 2 GMC_N          YRIEVDAK-E-RISAVHYKDPDGAS-HRVTGKYFVLAANGIEIPKLMLISTDDKHKNGVG
1292  Col_PMB31          247  YRIEVDAK-E-RISAVHYKDPDGAS-HRVTGKYFVLAANGIEIPKLMLISTDDKHKNGVG
1293  Col.prat_Ter291    247  YRIEVDAK-E-RISAVHYKDPDGAS-HRVTGKYFVLAANGIEIPKLMLISTDDKHKNGVG
1294  C.pratensis_Ter    247  YRIEVDAK-E-RISAVHYKDPDGAS-HRVTGKYFVLAANGIEIPKLMLISTDDKHKNGVG
1295  C.arenae_Cal35     247  YRIEVDAK-E-RISAVHYKDPEGVT-HRVTGKYFVLAANGIEIPKLMLISTDDRHKNGVG
1296  Collimonas_sp_P    247  YRIEVDAK-E-RISAVHYKDPEGVT-HRVTGKYFVLAANGIEIPKLMLISTDDKHKNGVG
1297  C.fungivorans_T    247  YRIEVDAK-E-RISAVHYKDPEGVT-HRVTGKYFVLAANGIEIPKLMLMSTDDKHGKGVG
1298  C.fungivorans_T    247  YRIEVDAK-E-RISAVHYKDPEGVT-HRVTGKYFVLAANGIEIPKLMLMSTDDKHGKGVG
1299  C.arenae_Ter282    247  YRIEVDAK-E-RISAVHYKDPDGVT-HRVTGKYFVLAANGIEIPKLMLISTDDKHPNGVG
1300  C.arenae_Ter10     247  YRIEVDAK-E-RISAVHYKDPDGVT-HRVTGKYFVLAANGIEIPKLMLISTDDKHPNGVG
1301  Collimonas_sp_O    247  YRIEVDAK-E-RISAVHYKDPEGVT-HRVTGKYFVLAANGIEIPKLMLMSTDDKHGNGVG
1302  Collimonas_sp_O    247  YRIEVDAK-E-RISAVHYKDPEGVT-HRVTGKYFVLAANGIEIPKLMLMSTDDKHGNGVG
1303  Collimonas_sp_O    247  YRLEVDAR-E-RISAVHYKDPQGVT-HRVTGKYFVLAANGIEIPKLMLMSTDDKHGNGVG
1304  Glaciimonas_sp_    247  YRIEVDSA-N-RVSAVHYKDPDGVT-HRVTGKYFVLAANGIEIPKLMLISTDDKHPNGVG
1305  Glaciimonas_sp_    247  YKVEIDSK-D-RVSAVHYKDPGGAT-HRVTGKYFVLAANGIEIPKLMLISTDDKHPNGIG
1306  Janthin_UBA1134    247  YRIEADDK-G-RIAAVHYKDPDGNS-HRVTGKRFVLAANGIETPKLMLMSSSARQPNGVG
1307  Herba_seropedic    244  YRLEHDDK-G-RIAAVHYYDWNKVS-HRITAQTFVLAANAIETPKLLLMSVSDKFPQGIA
1308  Herba_rubrisuba    244  YRLEHDEK-G-RIVAAHYYDWNKLS-HRVTAQTFVLAANAIETPKLLLMSVSDKFPQGIA
1309  AAN39686.1Burkh    251  YKLETGPD-K-RIVAALYKDKTGAE-HRVEGKYFVLAANGIETPKILLMSANRDFPNGVA
1310  BAM93252.1Gluco    255  YAMETDAK-N-RITAISFYDPDKQS-HRVVAKTFVIAANGIETPKLLLLAANDRNPHGIA
1311  AAF21261.1Panto    261  YKMDTDSN-N-RITAVHWLDTSGAS-HKATAKAFALACNGIETPRLLLMAANDANPNGIA
1312  AFW02570.1Gluco    254  YRLETDAS-NKKVVAVNYYDPDKNS-HRVTGKFFVVAAHCIESAKLLLLSADDKNPRGIA
1313  BAD60913.1         254  YRLETDAS-NKKVVAVNYYDPDKNS-HRVTGKFFVVAAHCIESAKLLLLSADDKNPRGIA
1314  Glucon_frateuri    276  LRVDLDETRS-KATGVTYLETKTQKEVSLAADLVILSAFQFHNVHLMLLSGIGK-PYDPV
1315
```

```
1316  Domain 2 GMC_N          NSSDQVGRNLMDHP-----------------------------------------------
1317  Col_PMB31          304  NSSDQVGRNLMDHPGTGVTFLA---N-EDLWPGRGPIEMTSMVDMRDGAFRSEYAAKKLH
1318  Col.prat_Ter291    304  NSSDQVGRNLMDHPGTGVTFLA---N-EDLWPGRGPIEMTSMVDMRDGAFRSEYAAKKLH
1319  C.pratensis_Ter    304  NSSDQVGRNLMDHPGTGVTFLA---N-EDLWPGRGPIEMTSMVDMRDGAFRSEYAAKKLH
1320  C.arenae_Cal35     304  NSSDQVGRNLMDHPGTGVTFLA---N-EDLWPGRGPIEMTSMVDMRDGPFRAQYAAKKLH
1321  Collimonas_sp_P    304  NSSDQVGRNLMDHPGTGVTFLA---N-EDLWPGRGPIEMTSMVDMRDGAFRSEYAAKKLH
1322  C.fungivorans_T    304  NSSDQVGRNLMDHPGTGVTFLA---N-EDLWPGRGPIEMTSIVDMRDGAFRSEYAAKKLH
1323  C.fungivorans_T    304  NSSDQVGRNLMDHPGTGVTFLA---N-EDLWPGRGPIEMTSIVDMRDGAFRSEYAAKKLH
1324  C.arenae_Ter282    304  NSSDQVGRNLMDHPGTGVTFLA---N-EDLWPGRGPIEMTSLVDMRDGAFRSEYAAKKLH
1325  C.arenae_Ter10     304  NSSDQVGRNLMDHPGTGVTFLA---N-EDLWPGRGPIEMTSLVDMRDGAFRSEYAAKKLH
1326  Collimonas_sp_O    304  NSSDQVGRNLMDHPGTGVTFLA---N-EDLWPGRGPIEMTSIVDMRDGAFRAEYAAKKLH
1327  Collimonas_sp_O    304  NSSDQVGRNLMDHPGTGVTFLA---N-EDLWPGRGPIEMTSIVDMRDGAFRAEYAAKKLH
1328  Collimonas_sp_O    304  NSSDQVGRNLMDHPGTGVTFLA---N-EDLWPGRGPIEMTSIVDMRDGAFRAEYAAKKLH
1329  Glaciimonas_sp_    304  NSSDQVGRNLMDHPGTSVTFLA---N-EDMWPGRGSIEMTSIVDMRDGPFRAEYAGKKLH
1330  Glaciimonas_sp_    304  NSSDQVGRNLMDHPGTSVSFLA---N-EDMWPGRGSIEMSSIVDMRDGAFRSEYAAKKLH
1331  Janthin_UBA1134    304  NSSDQVGRNLMDHPGTGVSFLA---D-EDLWPGRGPMELSSIVDMRDGAFRSEYAGKKLH
1332  Herba_seropedic    301  NARDNVGRNLCDHPAIGVTFDV---D-EEIWPGRGPVSPCSIGQFRDGDFRAEHAPFRID
1333  Herba_rubrisuba    301  NARDNVGRNLCDHPAIGVTFDV---D-EEIWPGRGPVSPCSIGQFRDGAFRADHAPFRID
1334  AAN39686.1Burkh    308  NSSDMVGRNLMDHPGTGVSFYA---S-EKLWPGRGPQEMTSLIGFRDGPFRATEAAKKIH
1335  BAM93252.1Gluco    312  NSSDLVGRNMMDHPGIGMSFQS---A-EPIWAGGGSVQMSSITNFRDGDFRSEYAATQIG
1336  AAF21261.1Panto    318  NASDMVGRNMMDHSGFHCSFLT---K-EPVWLGKGPAQSSCMVGYRDGDFRRDYSANKVI
1337  AFW02570.1Gluco    312  NSSDQVGRNMMDHTGVQLSFMS---GNDSLWPGRGPLLTSIIDSFRDGPWRSERGAYLVH
1338  BAD60913.1         312  NSSDQVGRNMMDHTGVQLSFMS---GNDSLWPGRGPLLTSIIDSFRDGPWRSERGAYLVH
1339  Glucon_frateuri    334  SNTGTIGRNFVYQTITSSRIWLPPSKYTNQFIGAGGVGVA-IDDFNSMNF--DHGPLGFV
1340
1341
1342  Domain 3 GMC_C          ------------------------------------------------------PNPE
1343  Col_PMB31          360  LNNMAQTNHAAQSALKD-------G--LVGL--KLNQEIRRRAARTVNINSFHDILPNPE
1344  Col.prat_Ter291    360  LNNMAQTNHAAQSALKD-------G--LVGL--KLNQEIRRRAARTVNINSFHDILPNPE
1345  C.pratensis_Ter    360  LNNMAQTNHAAQSALKD-------G--LVGL--KLNQEIRRRAARTVNINSFHDILPNPE
1346  C.arenae_Cal35     360  LNNMAQTNHAAQSALKE-------G--LVGL--KLNQEIRRRAARTVNINSFHDILPNPE
1347  Collimonas_sp_P    360  LNNMAQTNHAAQSALKD-------G--LVGL--KLNQEIRRRAARTVNINSFHDILPNPE
1348  C.fungivorans_T    360  LNNMAQTNHAAQTALKD-------G--LVGV--KLNQEIRRRAARTVNINSFHDILPDPE
1349  C.fungivorans_T    360  LNNMAQTNHAAQTALKD-------G--LVGL--KLNQEIRRRAARTVNINSFHDILPDPE
1350  C.arenae_Ter282    360  LNNMAQTNHAAQTALKN-------G--LVGV--KLNQEIRRRAARTVNINSFHDILPNPE
1351  C.arenae_Ter10     360  LNNMAQTNHAAQTALKN-------G--LVGV--KLNQEIRRRAARTVNINSFHDILPNPE
1352  Collimonas_sp_O    360  LNNMAQTNHAAQTALKD-------G--LVGI--KLNQEIRRRAARTVNINSFHDILPDPE
1353  Collimonas_sp_O    360  LNNMAQTNHAAQTALKD-------G--LVGV--KLNQEIRRRAARTVNINSFHDILPDPQ
1354  Collimonas_sp_O    360  LNNMAQTNHAGQTALKD-------G--LVGL--KLNQEIRRRAARTVNINSFHDILPDPE
1355  Glaciimonas_sp_    360  LNNMAQTNHAAQTALKQ-------G--LVGT--KLDAEIRHRAARTVNINSFHDILANPE
1356  Glaciimonas_sp_    360  LNNMAQTNHAAQAALKM-------G--LVGS--KLDKEIRHRAARMVNINSFHDILPNPE
1357  Janthin_UBA1134    360  LNNMAQTRAATTKAIQL-------G--LVGA--RLNQEIRHRAARMVNFNSFHDILPDPA
1358  Herba_seropedic    357  ISNASQVASVTKEVIAE-------G--YFGK--RLEEQILFRAARRLSIKNALEQLPDRN
1359  Herba_rubrisuba    357  ISNASQVAAVTKEVVAE-------G--YFGK--RLEEQILFRAARRLSIKNALEQLPDRN
1360  AAN39686.1Burkh    364  LSNLSRIDQETQKIFKA-------G--KLMKPDELDAQIRDRSARYVQFDCFHEILPQPE
1361  BAM93252.1Gluco    368  YNNTAQNSRACMKALSM-------G--LVGK--KLDEEIRRRTAHGVDIYANHEVLPDPN
1362  AAF21261.1Panto    374  LNNISRVVTATQQAMKK-------G--LVGK--ALDEEIRYRAVHSVDLSISLEPLPDPE
1363  AFW02570.1Gluco    369  MVDDNQVDFATGLAIAK-------G--YVGK--ELEEQIRYGSSHAVRLFSHNEGIADPD
1364  BAD60913.1         369  MVDDNQVDFATGLAIAK-------G--YVGK--ELEEQIRYGSSHAVRLFSHNEGIADPD
1365  Glucon_frateuri    391  GGSPVWVNQAGLKPIAAAIGTGGGGTPRWGS--AYKGALIDTYKHSIGVDAHGSNMAYRD
1366
```

```
1367  Domain 3 GMC_C          NRIRPSSEKRDALGIPQPEITYSIDDYVKKSALLTHDTYAS------------------D
1368  Col_PMB31          409  NRIRPSSEKRDALGIPQPEITYSIDDYVKKSALLTHDTYASIAAMFGGTEVVF---A--D
1369  Col.prat_Ter291    409  NRIRPSSEKRDALGIPQPEITYSIDDYVKKSALLTHDTYASIAAMFGGTEVVF---A--D
1370  C.pratensis_Ter    409  NRIRPSSEKRDALGIPQPEITYSIDDYVKKSALLTHDTYASIAAMFGGTEVVF---A--D
1371  C.arenae_Cal35     409  NRIRPSSDKRDALGIPQPEITYSIDDYVKKSAVLTHDTYASIAAMFGGTEVKF---D--D
1372  Collimonas_sp_P    409  NRIRPSSEKRDALGIPQPEITYSIGDYVKKSAVLTHETYASIAAMFGGTEVKF---D--D
1373  C.fungivorans_T    409  NRIRPSTEKRDALGIPQPEITYSINDYVQKSAVLTHQAYASIAAMFGGTEIKF---D--D
1374  C.fungivorans_T    409  NRIRPSAEKRDALGIPQPEITYSINDYVQKSAVLTHQAYASIAAMFGGTEIKF---D--D
1375  C.arenae_Ter282    409  NRIRPSADKRDALGIPQPEITYSIDDYVKKSAVLTHQTYASIAAMFGGTEIQF---A--D
1376  C.arenae_Ter10     409  NRIRPSADKRDALGIPQPEITYSIDDYVKKSAVLTHQTYASIAAMFGGTEIQF---A--D
1377  Collimonas_sp_O    409  NRIRPSTEKRDALGIPQPEITYSINDYVKKSAVLTHQAYASIAAMFGGTEIKF---D--D
1378  Collimonas_sp_O    409  NRIRPSTEKRDALGIPQPEITYSINDYVKKSAVLTHQAYASIAAMFGGTDIKF---D--D
1379  Collimonas_sp_O    409  NRIRPSMEKRDALGIPQPEITYSINDYVKKSAVLTHQAYASIAAMFGGTEIKF---D--D
1380  Glaciimonas_sp_    409  NRIRASTEKRDALGIPQPEITYAIDDYVKKSAVLTHETYASIAAMFGGTEVVF---D--D
1381  Glaciimonas_sp_    409  NRVLPSADHKDALGIPQPEITYAIDDYVKKSAIKTHEVYADIAALFGGTEVTF---D--D
1382  Janthin_UBA1134    409  NRIRPSTEHKDALGIARPEITYSIGDYVKKSAVHTRAAYARIARLFGGTEISF---A--D
1383  Herba_seropedic    406  NRVTLG-DKKDELGLPVPALYWNVDDYVVKGTEKTRAAYDAIAARLGATNIKH---SKSG
1384  Herba_rubrisuba    406  NRVTLG-DKKDELGLPVPALYWNMDDYVIKGTQKTRECYDAIASKLGATNIKH---SKSG
1385  AAN39686.1Burkh    415  NRIVPSKTATDAIGIPRPEITYAIDDYVKRGAAHTREVYATAAKVLGGTDVVF---N--D
1386  BAM93252.1Gluco    417  NRLVLSKDYKDALGIPHPEVTYDVGEYVRKSAAISRQRLMDIAKAMGGTEIEM---T--P
1387  AAF21261.1Panto    423  NRLTLSKTRKDPHGLPCPDIYYDVGDYVRKGAEASHAQLEHIGQLFDAKEFTI---S--Q
1388  AFW02570.1Gluco    418  NRLTLSKTHKDVLGIPHPEVYYKLPEYTVKSCDHTKELFKELMALMSGTDPQW---T--K
1389  BAD60913.1         418  NRLTLSKTHKDVLGIPHPEVYYKLPESTVKSCDHTKELFKELMALMSGTDPQW---T--K
1390  Glucon_frateuri    449  VYLDLDPTWKDAYGQPLLRMTFTWKDNDIRMNQYVVSKIEPIAKAMGATKANLVHLNAGQ
1391
1392
1393  Domain 3 GMC_C          D----FAPNNHIMGAVIMGGDPRDSVVDAHCRSHDHENLFIASSAVMPVAGTVNCTLTIA
1394  Col_PMB31          464  D----FAPNNHIMGAVIMGGDPRDSVVDAHCRSHDHENLFIASSAVMPVAGTVNCTLTIA
1395  Col.prat_Ter291    464  D----FAPNNHIMGAVIMGGDPRDSVVDAHCRSHDHENLFIASSAVMPVAGTVNCTLTIA
1396  C.pratensis_Ter    464  D----FAPNNHIMGAVIMGGDPRDSVVDAHCRSHDHENLFIASSAVMPVAGTVNCTLTIA
1397  C.arenae_Cal35     464  D----FAPNNHIMGAVIMGGDPRDSVVDAHCRSHDHENLFIASSAVMPVAGTVNCTLTIA
1398  Collimonas_sp_P    464  D----FAPNNHIMGAVIMGGDPRDSVVDAHCRSHDHENLFIASSAVMPVAGTVNCTLTIA
1399  C.fungivorans_T    464  D----FAPNNHIMGATIMGGDPRDSVVDAHCRSHDHENLFIASGSVMPVAGTVNCTLTIA
1400  C.fungivorans_T    464  D----FAPNNHIMGATIMGGDPRDSVVDAHCRSHDHENLFIASGSVMPVAGTVNCTLTIA
1401  C.arenae_Ter282    464  D----FAPNNHIMGAVIMGGDPRDSVVDAHCRSHDHENLFIASSAVMPVAGTVNCTLTIA
1402  C.arenae_Ter10     464  D----FAPNNHIMGAVIMGGDPRDSVVDAHCRSHDHENLFIASSAVMPVAGTVNCTLTIA
1403  Collimonas_sp_O    464  D----FAPNNHIMGATIMGGDPRDSVVDAHCRSHDHENLFIASGSVMPVAGTVNCTLTIA
1404  Collimonas_sp_O    464  D----FAPNNHIMGATIMGGDPRDSVVDAHCRSHDHENLFIASGSVMPVAGTVNCTLTIA
1405  Collimonas_sp_O    464  D----FAPNNHIMGATIMGGDPRNSVVDAHCRSHDHENLFIASGSVMPVAGTVNCTLTIA
1406  Glaciimonas_sp_    464  N----FAPNNHIMGATIMGGDPRDSVVDAHCRSHDHENLFIASSAVMPVAGTVNCTLTIA
1407  Glaciimonas_sp_    464  N----FAPNNHIMGSTIMGSDPRDSVVDAHCRSHDHENLFIASSSVMPVAGTVNCTLTIA
1408  Janthin_UBA1134    464  D----FAPNNHIMGSAIMGADSRDSVVDAHCRTHDHENLYIAGSAVMPVAGTVNCTLTIA
1409  Herba_seropedic    462  A----FSNRQHITGTLSMGLDPATSVTDAWGRAHDHENLFMVGTGVMPTVGTCNVTLTAM
1410  Herba_rubrisuba    462  D----FSNRQHIIGTLSMGLDPATSVTDAWGRAHDHENLFMVGTGVMPTVGTCNVTLTAM
1411  AAN39686.1Burkh    470  E----FAPNNHITGSTIMGADARDSVVDKDCRTFDHPNLFISSSATMPTVGTVNVTLTIA
1412  BAM93252.1Gluco    472  Y----FTPNNHITGGTIMGHDPRDSVVDKWLRTHDHSNLFLATGATMAASGTVNSTLTMA
1413  AAF21261.1Panto    478  G----LNANNHIMGGVIMGKNAKEAVVDGNCRAFDHENLWLPGGGAIPSASVVNSTLTMA
1414  AFW02570.1Gluco    473  G----YFPQCHPSGSTIMGTDPTNSVVDGECRTHDHENLFVASSAVFSSVGTGNITLTIA
1415  BAD60913.1         473  G----YFPQCHPSGSTIMGTDPTNSVVDGECRTHDHENLFVASSAVFSSVGTGNITLTIA
1416  Glucon_frateuri    509  PFDTRHYQTTHLGGGAIMGTDPQTSALNRYLQSWDVSNVFVIGSNAFPQGMGYNPTGMVA
1417
```

```
Domain 3 GMC_C          ALSLR--------------------
Col_PMB31          520  ALSLRIADTLRTV-----------L
Col.prat_Ter291    520  ALSLRIADTLRTV-----------L
C.pratensis_Ter    520  ALSLRIADTLRTV-----------L
C.arenae_Cal35     520  ALSLRIADTLRTV-----------L
Collimonas_sp_P    520  ALSLRIADTLRTV-----------L
C.fungivorans_T    520  ALSLRIADTLRTV-----------L
C.fungivorans_T    520  ALSLRIADTLRTV-----------L
C.arenae_Ter282    520  ALSLRIADTLRTV-----------L
C.arenae_Ter10     520  ALSLRIADTLRTV-----------L
Collimonas_sp_O    520  ALSLRIADTLRTV-----------L
Collimonas_sp_O    520  ALSLRIADTLRTV-----------L
Collimonas_sp_O    520  ALSLRIADTLRTV-----------L
Glaciimonas_sp_    520  ALSLRIADTLRTV-----------L
Glaciimonas_sp_    520  ALSLRIADTLRTV-----------L
Janthin_UBA1134    520  ALALRLSDTLR--------------
Herba_seropedic    518  ALALRTADKIIEES-------RH-A
Herba_rubrisuba    518  ALSLRTADKIIEES-------RH-A
AAN39686.1Burkh    526  ALALRMSDTLKKE-----------V
BAM93252.1Gluco    528  ALSLRAADAILNDL-------KQ-G
AAF21261.1Panto    534  ALGLKAAHDISLRM-------KGDA
AFW02570.1Gluco    529  ALALRVAASLKKEM-------LH-A
BAD60913.1         529  ALALRVAASLKKEM-------LH-A
Glucon_frateuri    569  ALAYWAAYHIRTTYLKSPGPLVQ-A
```

Figure S2: Highlight of the active domains detected in the large subunit of the GMC enzyme and of the cysteine rich region

Domain 1 : FAD ; Domain 2 and 3 correspond to GMC sites. Each line corresponds to a protein sequence and its localisation. The abbreviations correspond to those presented in the phylogenetic tree in Figure 6. Gluc_BAH805545.1, gluconate dehydrogenase; Gluc_BAD60913.1, sorbitol dehydrogenases; Burkholderia, glucose dehydrogenase; Gluc_japonicus, fructose dehydrogenase; Pantoea_agglo., choline dehydrogenase; Gluc_oxydans, sorbitol dehydrogenases.

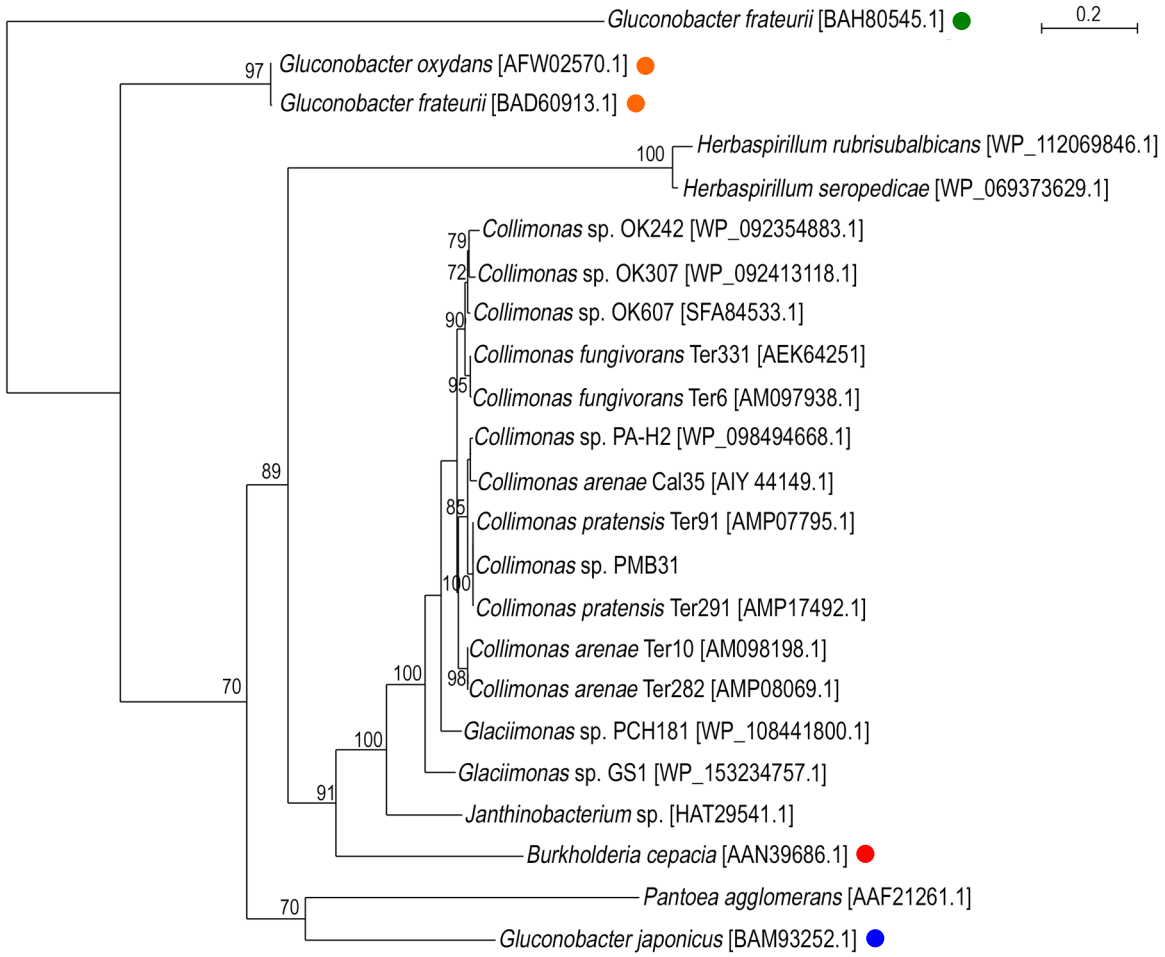


Figure S3: Phylogenetic tree based on protein sequence of the large subunit of different GMC-DH

Comparison of the GdhL sequence from PMB3(1) with homologous protein sequences from a set of 22 different genera or species related to *Collimonas* and available on international databases were analysed among them, 11 were assigned to the *Collimonas* genus. Other genera taxonomically close to *Collimonas* were chosen (*Glaciimonas*, *Janthinobacterium, Herbaspirillum* and *Burkholderia)*. Enzymes belonging to the GMC family and with a demonstrated function were included in the analysis such as Gluconate DH (GADH ; green circle), Sorbitol DH (SDH ; orange circle), Frcutose DH (FDH ; blue circle) and Glucose DH (GDH ; red circle). For these protein sequences, the accession number from NCBI is indicated between brackets.