

1 **Isolation of a *Bacillus safensis* from mine tailings in Peru, genomic**
2 **characterization and characterization of its cyanide- degrading enzyme CynD**

3 Santiago Justo Arevalo^{*1,2}, Daniela Zapata Sifuentes^{1,2}, Andrea Cuba Portocarrero¹,
4 Michella Brescia Reategui¹, Claudia Monge Pimentel¹, Layla Farage Martins², Paulo
5 Marques Pierry², Carlos Morais Piroupo², Alcides Guerra Santa Cruz¹, Mauro Quiñones
6 Aguilar¹, Chuck Shaker Farah², João Carlos Setubal², Aline Maria da Silva²

7 1.- Facultad de Ciencias Biológicas, Universidad Ricardo Palma, Lima, Peru.

8 2.- Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, São
9 Paulo, Brazil.

10 * Correspondence: santiago.jus.are@usp.br

11

12 **ABSTRACT**

13 Cyanide is widely used in industry as a potent lixiviant due to its capacity to tightly bind
14 metals. This property imparts cyanide enormous toxicity to all known organisms. Thus,
15 industries that utilize this compound must reduce its concentration in recycled or waste
16 waters. Physical, chemical, and biological treatments have been used for cyanide
17 remediation; however, none of them meet all the desired characteristics: efficiency, low
18 cost and low environmental impact. A better understanding of metabolic pathways and
19 biochemistry of enzymes involved in cyanide degradation is a necessary step to improve
20 cyanide bioremediation efficacy to satisfy the industry requirements. Here, we used
21 several approaches to explore this topic. We have isolated three cyanide-degrading
22 *Bacillus* strains from water in contact with mine tailings from Lima, Peru, and classified
23 them as *Bacillus safensis* PER-URP-08, *Bacillus licheniformis* PER-URP-12, and *Bacillus*
24 *subtilis* PER-URP-17 based on 16S rRNA gene sequencing and core genome analyses.
25 Additionally, core genome analyses of 132 publicly available genomes of *Bacillus pumilus*
26 group including *B. safensis* and *B. altitudinis* allowed us to reclassify some strains and
27 identify two strains that did not match with any known species of the *Bacillus pumilus*
28 group. We searched for possible routes of cyanide-degradation in the genomes of these
29 three strains and identified putative *B. licheniformis* PER-URP-12 and *B. subtilis* PER-
30 URP-17 rhodanases and *B. safensis* PER-URP-08 cyanide dihydratase (CynD) sequences
31 possibly involved cyanide degradation. We identified characteristic C-terminal residues
32 that differentiate CynD from *B. pumilus* and *B. safensis*, and showed that, differently from
33 CynD from *B. pumilus* C1, recombinant CynD from the *Bacillus safensis* PER-URP-08
34 strain remains active up to pH 9 and presents a distinct oligomerization pattern at pH 8
35 and 9. Moreover, transcripts of *B. safensis* PER-URP-08 CynD (CynD_{PER-URP-08}) are
36 strongly induced in the presence of cyanide. Our results warrant further investigation of *B.*
37 *safensis* PER-URP-08 and CynD_{PER-URP-08} as potential tools for cyanide-bioremediation.

38

39 **INTRODUCTION**

40 Cyanide is a highly toxic compound used in several industrial processes (Mudder et al.,
41 2004) given its capacity to form tight complexes with different metals (Dash et al., 2009)
42 (Hendry-Hofer et al., 2019; Leavesley et al., 2008). Industries that generate cyanide-
43 containing wastes must reduce its concentration before discarding them to the
44 environment, and as such proper strategies have to be implemented for cyanide

45 remediation (Kuyucak & Akcil, 2013). Cyanide bioremediation by bacteria that express
46 nitrilases is one possible low-cost and environmental-friendly approach (Dash et al., 2009).

47 Nitrilases are a superfamily of proteins characterized by a tertiary structure consisting of
48 an alpha-beta-beta-alpha fold and a dimer as a basic unit. This superfamily has been
49 divided into thirteen branches with branch one corresponding to enzymes that cleave the
50 nitrile group into ammonia and its respective carboxylic acid. The other twelve branches
51 are structurally similar but their catalytic activity does not involve cleavage of nitriles (Pace
52 & Brenner, 2001).

53 Two types of nitrilases can degrade cyanide through a hydrolytic pathway: cyanide
54 hydratases (CHTs) and cyanide dihydratases (CynDs). CHTs convert cyanide into
55 formamide using one water molecule in the reaction and are present in fungal genomes.
56 The first enzyme described with this activity was from *Stemphylium loti* (Fry & Millar,
57 1972). Subsequently, CHTs from other fungal species have been studied such as:
58 *Fusarium solani*, *Fusarium oxysporum*, *Gloeocercospora sorghi*, *Leptosphaeria maculans*,
59 and *Aspergillus niger* (Akinpelu et al., 2018; Dumestre et al., 1997; Ping Wang; Hans D.
60 VanEtten, 1992; Rinágelová et al., 2014; Sexton & Howlett, 2000). On the other hand, only
61 the CynDs from *B. pumilus*, *P. stutzeri* and *Alcaligenes xylosoxidans* (Ingvorsen et al.,
62 1991; Meyers et al., 1993; Watanabe et al., 1998) have been experimentally tested. The
63 reaction catalyzed by CynDs generates formic acid and ammonia using two molecules of
64 water. Both, CynDs and CHTs, typically form large helical aggregates of several subunits
65 (Thuku et al., 2009). For example, the CynD of *Bacillus pumilus* is reported to form an 18-
66 subunit oligomer (Jandhyala et al., 2003) whereas the homolog in *Pseudomonas stutzeri*
67 forms a 14-subunit oligomer (Sewell et al., 2003)

68 Several *Bacillus* species have been shown to be capable of metabolizing cyanide using
69 different routes, for instance: B-cyanoalanine synthase in *Bacillus megaterium* (Castric &
70 Strobel, 1969), gamma-cyano-alpha-aminobutyric acid synthase in *B. stearothersophilus*
71 (Omura et al., 2003), rhodanase in *Bacillus cereus* (Itakorode et al., 2019), and CynD in
72 *Bacillus pumilus* (Meyers et al., 1993). On the other hand, some other cyanide-degrading
73 *Bacillus* have still unknown metabolic routes (Al-Badri et al., 2020; Javaheri Safa et al.,
74 2017; Mekuto et al., 2014).

75 The *Bacillus pumilus* group consists mainly of three species: *Bacillus altitudinis*, *Bacillus*
76 *safensis* and *Bacillus pumilus*. These three species share more than 99 % sequence
77 identity in their 16S rRNA gene (Liu et al., 2013), hampering taxonomical classification
78 based solely on this locus. Studies using multiple phylogenetic markers have
79 demonstrated that ~50 % of the *Bacillus pumilus* group genomes deposited in NCBI
80 database could be misclassified (Espariz et al., 2016).

81 It is plausible to speculate that CynDs isolated from *Bacillus* strains from diverse
82 environments could present different properties that could be better for industrial
83 applications. Therefore, the characterization of CynD from other species can expand our
84 understanding on the functioning and plasticity of this enzyme. Furthermore, some aspects
85 of the biology of this enzyme have not been thoroughly studied. For instance, it is known
86 that the oligomeric state of CynD is strongly pH-dependent (Jandhyala et al., 2003);
87 however, the effect on oligomerization at pHs greater than 9 has not been reported. Also, it
88 is unknown whether CynD is constitutively expressed in basal metabolism or is part of a
89 specific physiological response, for instance, induced by the presence of cyanide.

90 Here, we describe the isolation of three indigenous *Bacillus* strains from mine tailing in
91 Peru and their respective genome sequences. We selected a strain that was most efficient

92 in cyanide degradation and investigated its phylogenetic relationship with other species of
93 the *Bacillus pumilus* group. We identified a gene coding for a cyanide dihydratase (CynD)
94 that is most likely the enzyme responsible for cyanide degradation in this selected strain. A
95 recombinant CynD was expressed and purified, its catalytic parameters were determined,
96 and the quaternary structure was studied at different pHs. We also demonstrated that
97 CynD transcripts are strongly induced in the presence of cyanide.

98 MATERIAL AND METHODS

99 Isolation of cyanide-degrading strains

100 Water in contact with mine tailing was collected from a river near Casapalca and La Oroya
101 mines located in San Mateo de Huanchor (Latitude -11.4067 Longitude -76.3361 at 4221
102 MASL). The sample was collected in 2 L sterile bottles and transported at 4 °C.

103 One hundred mL of the sample was added to an Erlenmeyer flask containing 20 ml of 21
104 g/L sodium carbonate, 9 g/L sodium bicarbonate, 5 g/L sodium chloride and 0.5 g/L
105 potassium nitrate. Cultures were incubated for 12 h at 37 °C and after this time 1 mg/L
106 final concentration of cyanide in the form of sodium cyanide was added. The cultures were
107 incubated for another 24 h at 37 °C. Samples of the medium were streaked in petri dishes
108 with nutrient agar (5 g/L peptone, 5 g/L yeast extract, 5 g/L sodium chloride and 1 % agar)
109 and incubated at 37 °C for 24 h. Single colonies were isolated in nutrient broth (5 g/L
110 peptone, 5 g/L yeast extract, 5 g/L sodium chloride) supplemented with 20 % glycerol and
111 stored at -80 °C.

112 Strains stored at -80 °C were reactivated at 37 °C in nutrient agar by streaking a sample.
113 One colony was inoculated in fresh nutrient broth and incubated at 37 °C overnight at 100
114 g. Next, the optical density at 600 nm (OD_{600nm}) of the culture was adjusted to 0.8 and 1
115 mL was centrifuged at 6000 g for 3 min. The pellet was washed twice with 0.2 M Tris-HCl
116 pH 8 and resuspended in 1 mL of 0.2 M Tris-HCl pH 8 supplemented with 0.2 M NaCN.
117 After 2 h of incubation at 30 °C, the culture was centrifuged at 6000 g and 10 µL of the
118 supernatant was taken and diluted in 90 µL of milliQ water. Then, 200 µL of 0.5 % picric
119 acid in 0.25 M sodium carbonate was added and heated for 6 min at 100 °C (Williams &
120 Edwards, 1980). Finally, absorbance at 520 nm was measured and compared to a
121 standard curve of NaCN.

122 Strain identification by 16S rRNA gene sequencing

123 To determine the bacterial genera and/or species of the isolated strains, we used a fresh
124 culture in nutrient agar. Five colonies from these cultures were transferred to 50 µL of
125 milliQ water and then heated to 100 °C for 3 min in a dry bath. The samples were
126 centrifuged at 10 000 g for 5 min and the supernatant was used as a template to amplify a
127 fragment that includes the V6, V7 and V8 variable regions of 16S rRNA gene. One µL of
128 the template, 25 pmol of F_primer, 5' GCACAAGCGGTGGAGCATGTGG 3', and of the
129 R_primer, 5' GCCCGGGAACGTATTCACCG 3', were mixed with 1x Taq buffer, 1.5 mmol
130 of MgCl₂, 0.2 mmol of each dNTP, and 1 U Taq DNA polymerase (ThermoFisher
131 Scientific) in a final reaction of 25 µL. The amplification program was an initial denaturation
132 at 94 °C for 5 min followed by 30 cycles at 94 °C for 45 sec, 50 °C for 45 sec, and 72 °C
133 for 1 min, with a final extension of 10 min at 72 °C. Five µL of the final reaction was used
134 as a template for the sequencing reaction. Sequencing reaction was done using Big Dye
135 terminator v3.1 cycle sequencing kit (ThermoFisher Scientific) consisting of a 1x
136 sequencing buffer, 25 pmol F_primer or R_primer and 2 µL of Big Dye in a final volume of
137 20 µL. The program used was an initial denaturation at 94 °C for 5 min, followed by 40

138 cycles at 94 °C for 30 sec, 50 °C for 30 sec, and 60 °C for 4 min. After the sequencing
139 reaction, 80 µL of 70 % isopropanol was added and the reaction tube was centrifuged at
140 4000 g at 4 °C for 40 min. Then the supernatant was discarded, and the sample was
141 resuspended in 20 µL of milliQ water and injected in an ABI PRISM 3130XL genetic
142 analyzer (ThermoFisher Scientific). The obtained sequences were used to perform
143 BLASTn (Altschul et al., 1990) searches against the Genbank/NCBI database (Benson et
144 al., 2013) to identify most similar sequences.

145 **Genome sequencing, assembling and annotation**

146 Bacterial cultures were grown in 2xTY broth (tryptone 16 g/L, yeast extract 10 g/L, and
147 NaCl 5 g/L) at 37 °C for 18 h at 200 rpm. Genomic DNA purification was done using the
148 Wizard Genomic DNA Purification Kit (Promega). DNA integrity was evaluated by 1 %
149 agarose gel electrophoresis stained with SYBR Safe (Invitrogen) and by Bioanalyzer 2100
150 using Chips Agilent DNA 12000. DNA concentration and purity were estimated using a
151 NanoDrop One/OneC Microvolume UV-Vis Spectrophotometer (ThermoFisher Scientific).
152 Shotgun genomic library was prepared using the Nextera DNA Library Prep (Illumina) with
153 total DNA input of 20-35 ng. The resulting indexed DNA library was cleaned up with
154 Agencourt AMPure XP beads (Beckman Coulter) and fragment size within the range of
155 200-700 bp were verified by running in the 2100 Bioanalyzer using Agilent High Sensitivity
156 DNA chip. Fragment library quantification was performed with KAPA Library Quantification
157 Kit. Genomic libraries prepared for each strain were pooled and subjected to a run using a
158 an Illumina MiSeq Reagent Kit v2 (2 x 250 cycles) which generated ~38 million raw paired-
159 end reads with >75% of bases with quality score > 30.

160 The genome of strain PER-URP-08 was assembled with Discovar (v. 52488) (Weisenfeld
161 et al., 2014). The genomes of strains PER-URP-12 and PER-URP-17 were assembled
162 with A5 (v. 20160825) (Coil et al., 2015). Both software have adapters trimming and read
163 quality checking as part of their respective assembly processes. The tool Medusa (Bosi et
164 al., 2015) was used to generate final genome scaffolds using three sets of five reference
165 genomes, one for each of the genome assemblies (Table S1). The final genome
166 assemblies were submitted to the IMG/M (Chen et al., 2021) and to the NCBI (Benson et
167 al., 2013; Tatusova et al., 2016) for automatic annotation.

168 **Phylogenetic analyses and identification of nitrilases**

169 Annotated genomes belonging to *Bacillus pumilus*, *Bacillus safensis* or *Bacillus altitudinis*
170 species in the category of “Chromosome”, “Scaffold” or “Complete” were downloaded from
171 the Genbank/NCBI (Benson et al., 2013). Using the software cd-hit (Fu et al., 2012; Li &
172 Godzik, 2006) we identified coding sequences that are not duplicated and present in all the
173 genomes (core genes). A total of 1766 core genes with more than 80 % identity and at
174 least 90 % coverage were used in the analysis. Core genes were aligned using MAFFT
175 with the FFT-NS-2 algorithm (Kato & Standley, 2013). The resulting alignments were
176 concatenated and used to calculate a distance matrix based on identity using Biopython
177 (Cock et al., 2009). Phylogenetic inference by maximum likelihood was done using the
178 concatenated alignments as the input and IQ-TREE2 (Minh et al., 2020) with the evolution
179 model GTR+F+R3, ultrafast bootstrap 1000 (Hoang et al., 2018), and 1000 initial trees.

180 IMG/M tools (Chen et al., 2021) were used to identify nitrilases genes in the annotated
181 genomes. Genes encoding the CN_hydrolase domain (PFAM code PF00795) were
182 selected and checked regarding the genomic context and the related literature.

183 **Analysis of CynD sequences from *Bacillus pumilus* group genomes**

184 Protein sequence annotations from genomes belonging to *Bacillus pumilus*, *Bacillus*
185 *safensis* or *Bacillus altitudinis* in the category of “Chromosome”, “Scaffold” or “Complete”
186 were downloaded from GenBank/NCBI (Benson et al., 2013) and used to construct a local
187 database. We ran a BLASTp search (Altschul et al., 1990) using the query sequence
188 AAN77004.1 against the constructed local database, and sequences with more than 90 %
189 identity and 100 % coverage were identified as CynD orthologs. These sequences were
190 aligned using MAFFT with the L-INS-I algorithm (Kato & Standley, 2013). The resulting
191 alignment was used as an input for the phylogenetic inference by maximum likelihood
192 using IQ-TREE2 (Minh et al., 2020) with the evolution model JTTDCMut+I (Kosiol &
193 Goldman, 2005), ultrafast bootstrap 1000 (Hoang et al., 2018), 1000 initial trees and -allnhi
194 option.

195 **Cloning, expression and purification of CynD**

196 The coding sequence for CynD was amplified from genomic DNA of strain PER-URP-08
197 using the primers (restriction sites appear in uppercase): F_CynD (5'
198 tttCATATGatgacaagattttaccggaagtttc 3'), and R_CynD (5' tttCTCGAGcactttttctcaagcaacc
199 3') and cloned in the NdeI and XhoI sites of pET-28 plasmid. Then, this plasmid was used
200 as a template to amplify the CynD coding sequence with a c-terminal 6x-His tag using the
201 primers F_CynD and R_2_CynD (5' tttGAATTCagtggtggtggtggtggtg 3') and cloned in the
202 NdeI and EcoRI sites of pET-11 plasmid.

203 To express CynD protein, we used the *Escherichia coli* BL21(DE3) pLysS strain, induced
204 by 0.3 mM of Isopropyl β -D-1-thiogalactopyranoside for 23 h at 18°C. The cells were lysed
205 by sonication using a lysis buffer (100 mM Tris-HCl pH 8.0, 100 mM NaCl, 50 mM
206 Imidazole) and the suspension was clarified by centrifugation (13000 g). The supernatant
207 was loaded onto Ni-NTA affinity resin (His-trap chelating 5 mL column), washed with 10
208 volumes of lysis buffer, and eluted with a gradient of 50 - 500 mM Imidazole in 20 mM Tris-
209 HCl pH 8.0, 100 mM NaCl. The eluted fractions were further purified by size exclusion
210 chromatography using a Superdex pg 200 16/600 column and 20 mM Tris-HCl pH 8.0, 100
211 mM NaCl as running buffer. The eluted fractions were examined for purity by SDS-PAGE
212 and fractions containing pure protein were concentrated in Amicon Ultra-15 Centrifugal
213 filter units.

214 **Enzymatic assays of recombinant CynD**

215 For the determination of K_m and V_{max} , enzymatic activity of recombinant CynD was
216 measured at pH 8.0 using the Ammonia Assay Kit (Sigma-Aldrich). A concentration of 500
217 nM of CynD was used in all reactions with the following cyanide concentrations: 0.39,
218 0.625, 0.78, 1.25, 1.56, 2.5, 3.125, 5, 6.25, 12.5, 25 mM, with a final volume of 111 μ L at
219 30 °C. Measurements were taken on a plate reader, at 340 nm every 20 sec. Calculations
220 of ammonia concentrations were carried out according to the manufacturer description.

221 To test in which optimal pH for CynD activity, reagent solutions were prepared (40 mM
222 NaCN, 100 mM NaCl and 200 mM Tris-HCl or N-cyclohexyl-3-aminopropanesulfonic acid
223 (CAPS) at pH 8, 9 or 10, 11, respectively). Then, we added 5 μ L of CynD in 100 mM NaCl,
224 Tris-HCl pH 8 to 45 μ L of the reagent solutions to obtain a final concentration of CynD of 0,
225 5, 10, 15, or 20 μ M. The reactions were incubated for 10 min at 37 °C. After that, 100 μ L of
226 picric acid 5 mg/mL, 0.25 M Na_2CO_3 was added, and the reactions were incubated at 99
227 °C for 6 min. Next, 30 μ L of this reaction was transferred to a 96-well plate and
228 absorbance at 520 nm was recorded. Final cyanide concentration was estimated based on
229 calibration curves with cyanide concentrations between 0 – 40 mM.

230 **Size Exclusion Chromatography coupled to Multi-Angle Light Scattering (SEC-**
231 **MALS)**

232 SEC-MALS analysis was used to determine the oligomeric state of recombinant CynD.
233 Molar mass analysis was done in 100 mM NaCl and 20 mM Tris-HCl or CAPS at pH 8, 9
234 or 10, 11, respectively. Protein samples (100 μ L injection of 3.47 mg/mL (89.36 μ M) CynD)
235 were separated using a Superdex 200 increase 10/300 GL coupled to a MiniDAWN
236 TREOS multi-angle light scattering system and an Optilab rEX refractive index detector.
237 Data analysis was performed using the Astra Software package, version 7.1.1 (Wyatt
238 Technology Corp.).

239 **Transmission electron microscopy (TEM)**

240 Ultra-thin carbon layer on lacey carbon-coated copper grids were negatively charged by a
241 glow discharge of 25 sec at 15 mA. Four microliters of purified recombinant CynD in 20
242 mM Tris-HCl pH 8.0 and 100 mM NaCl in different concentrations (3.25 mg/mL or 1.625
243 mg/mL) were placed in the negative charged carbon-coated copper grid for 1 minute. The
244 grids were washed twice with MilliQ water and then stained with 2 % uranyl acetate for 30
245 secs before air drying. Electron micrographs images were obtained using a JEOL JEM
246 2100 transmission electron microscope equipped with a Gatan ORIUS CCD detector at
247 the Institute of Chemistry of the University of Sao Paulo.

248 **RT-qPCR to evaluate *in vivo* induction of *cynD* by cyanide**

249 *Bacillus* strains were grown in meat broth (meat extract 1 g/L, yeast extract 2 g/L, peptone
250 5 g/L, NaCl 5g/L, MnCl₂ 10 mg/L) during 12 h at 30 °C, 200 rpm. One mL of the culture
251 was centrifuged at 500 xg for 1 min, the supernatant was transferred to a clean tube and
252 this tube was centrifuged at 8 000 xg for 3 min. The pellet was resuspended in 1 mL of
253 NaCN ([CN⁻] 100 ppm) in milliQ water. Controls were resuspended in 1 mL milliQ water
254 without NaCN. The tubes were incubated without agitation at 30 °C for 4 h and 100 μ L
255 were retrieved to measure cyanide concentration by the picric acid method (Williams &
256 Edwards, 1980). Nine hundred μ L was centrifuged, and the bacterial pellet was used
257 immediately for total RNA extraction.

258 Total RNA extraction was done using Trizol-chloroform protocol. Briefly, bacterial pellets
259 were treated with 100 μ L of lysozyme 3 mg/mL at 37 °C for 30 min, and extraction was
260 done using a mixture of 5:1 trizol:chloroform. After the extraction, the phase containing
261 RNA was separated and the RNA was precipitated using isopropanol. RNA pellet was
262 washed twice with 75 % ethanol and finally resuspended in 20 μ L of Tris 20 mM-DEPC.
263 Total RNA concentration and purity were estimated in a NanoDrop™ One/OneC
264 Microvolume UV-Vis Spectrophotometer (ThermoFisher Scientific) and the integrity was
265 evaluated in a 2100 Bioanalyzer using an Agilent RNA 6000 Pico chip. After DNase
266 treatment, the samples were subjected to PCR to verify the absence of DNA
267 contamination. cDNA synthesis was performed with 1 μ g of the RNA and Thermo Scientific
268 H Minus First Strand cDNA Synthesis kit. cDNA synthesis was verified by PCR and
269 electrophoresis.

270 Amplification efficiency of the primers used in the RT-qPCR were verified using 300 nM of
271 each primer and a 2-fold dilution series of the cDNA to generate a standard curve
272 composed of 4 concentrations as follows: 62.5, 31.25, 15.625, and 7.8125 ng/ μ L. Each
273 dilution reaction was performed in triplicate using the Maxima SYBR Green/ROX qPCR
274 Master Mix kit (ThermoFisher Scientific) following the manufacturer instructions in a
275 QuantStudio 3 equipment (ThermoFisher Scientific). Primers for the normalizing gene *rpsJ*

276 (F_rpsJ 5' TGAAACGGCTAAGCGTTCTG 3', R_rpsJ 5' ACGCATCTCGAATTGCTCAC
277 3'), and for the nitrilases *cynD* (F_cynD 5' TGCCCAAATGAGCAGGTAC 3', R_cynD 5'
278 AAATGTCTGTGTGCGATGG 3') and *ykrU* (F_ykrU 5' TTGGTGCATGATTTGCTAT 3',
279 R_ykrU 5' GTGTCTCTGCTTGTGCCTGT 3') were tested for efficiency. The amplification
280 efficiency of the qPCR reaction was calculated through the slope of the cDNA curve
281 obtained for each primer pair.

282 Since primer pairs have showed similar efficiency, (*ykrU* = 119.108 %, *cynD* = 108.385 %,
283 *rpsJ* = 104.55 %), we performed each qPCR assay in technical triplicates using 15.625
284 ng/μL of cDNA and the kit Maxima SYBR Green/ROX qPCR Master Mix (ThermoFisher
285 Scientific) in a QuantStudio 3 equipment (ThermoFisher Scientific). $\Delta\Delta CT$ values were
286 calculated in absence or presence of cyanide for the nitrilase genes *ykrU* and *cynD* using
287 *rpsJ* as the normalizing gene. Three biological replicates were performed.

288

289 **RESULTS AND DISCUSSION**

290 **Three *Bacillus* spp. isolates with capacity of cyanide degradation**

291 Several colonies were obtained after selective enrichment in cyanide containing media of
292 water in contact with mine tailing from a river near Casapalca and La Oroya mines located
293 in San Mateo de Huanchor, Lima - Peru. Twenty colonies were screened for the ability to
294 degrade cyanide (Table S2) and three colonies with the greatest efficiency in cyanide
295 degradation (isolates 8, 12, and 17) were selected for further studies (Table S2).

296 Sequencing of the V6, V7, and V8 variable regions of 16S rRNA gene of the three selected
297 isolates and analysis by BLAST showed that they belong to the genus *Bacillus* (Table S3).
298 Isolates 12 and 17 were identified as *Bacillus licheniformis* and *Bacillus subtilis*,
299 respectively (Table S3) and were named *Bacillus licheniformis* PER-URP-12 and *Bacillus*
300 *subtilis* PER-URP-17. Isolate 8 was classified as a member of the *Bacillus pumilus* group
301 based on the 16S rRNA gene sequence (Table S3). However it was not possible to
302 discriminate among the different species in the *Bacillus pumilus* group (Liu et al., 2013)
303 and as such this isolated was provisionally named *Bacillus* sp. PER-URP-08.

304 The three strains were then sequenced in order to obtain a more accurate taxonomical
305 classification as well as to gain insights about possible routes of cyanide degradation in
306 the three strains under study. Table S4 shows a summary of assembly and annotation
307 metrics of these genomes.

308 ***Bacillus* sp. PER-URP-08 is classified as *Bacillus safensis* based on core-genome 309 comparisons**

310 We performed a genome-wide comparative analysis of *Bacillus* sp. PER-URP-08 with 132
311 genomes of species from the *Bacillus pumilus* group retrieved from the GenBank/NCBI
312 database (Benson et al., 2013) and identified 1766 coding sequences present in all the
313 genomes (core genes). An identity matrix based on an alignment of these core genes
314 showed three well defined branches and two genomes that do not belong to any of these
315 three branches (Fig. 1A).

316 Branch 1 (Fig. 1A, brown names) contains several strains already characterized as
317 *Bacillus altitudinis* by different methods (for instance: BA06, ku-bf1, B-388 (X. Fu et al.,
318 2021)) and also 4 strains (TUAT1, MTCB 6033, SH-B11 and C4) previously annotated as
319 *Bacillus pumilus*. However, our analysis clearly demonstrates that they belong to *Bacillus*
320 *altitudinis* and therefore require reclassification (Table S5) as previously suggested

321 (Espariz et al., 2016; X. Fu et al., 2021). The core genes within the *Bacillus altitudinis*
322 branch share more than 0.98 identity whereas they share less than 0.895 identity with core
323 genomes of the other two branches (Fig. 1B).

324 Identity of core genes in branch 2 is greater than 0.96, and this branch is more related to
325 branch 3 (*Bacillus pumilus*, see below) than to branch 1 (*Bacillus altitudinis*) (Fig. 1C).
326 Branch 2 (Fig. 1A, green names) contains the *Bacillus safensis* type strain FO-36b (Satomi
327 et al., 2006) as well as other strains already classified as *Bacillus safensis* such as B4107,
328 B4134, and B4129 (Espariz et al., 2016). *Bacillus* sp. PER-URP-08 appeared inside this
329 branch very near to the type strain FO-36b (99.2 % identity) (Fig. 1A) and so will be named
330 *Bacillus safensis* PER-URP-08 from here on.

331 Branch 3 (Fig. 1A, blue names) contains the SAFR-032 strain that was the first completely
332 sequenced genome of *Bacillus pumilus* (Gioia et al., 2007; Stepanov et al., 2016). This
333 branch 3 appears to be more heterogeneous than the other two branches (*Bacillus*
334 *altitudinis* and *Bacillus safensis*) with more than 0.95 identity of the core genes of this
335 branch (Fig. 1D). Additionally, two genomes isolated from Mexico (CH144a_4T and 145)
336 share less than 0.95 identity with the branch 3 and even less with branches 1 and 2 (Fig.
337 1E). The fact that these two genomes share less than 0.95 identity with all the other
338 genomes in the analysis (Fig. 1E) indicates that CH144a_4T and 145 strains should be
339 classified as different species outside the *Bacillus pumilus* group.

340 **A cyanide dihydratase is likely the responsible for cyanide degradation in *B.*** 341 ***safensis* PER-URP-08**

342 To gain insight regarding the enzymes responsible for cyanide metabolism in the strains *B.*
343 *safensis* PER-URP-08, *B. licheniformis* PER-URP-12, and *B. subtilis* PER-URP-17, we
344 first searched for genes coding for proteins related to nitrilases. The PFAM database
345 annotates homologs of nitrilases as CN_hydrolases under the PFAM code PF00795.
346 Using IMG/M system tools (Chen et al., 2021), we determined the presence of three, two,
347 and two proteins containing CN_hydrolase domains in *B. safensis* PER-URP-08, *B.*
348 *licheniformis* PER-URP-12, and *B. subtilis* PER-URP-17, respectively (Fig. S1). Both *B.*
349 *licheniformis* PER-URP-12 and *B. subtilis* PER-URP-17 present the genes *yhcX* (NCBI
350 locus tags: EGI08_RS06285 and EGI09_16505, respectively) and *mtnU* (EGI08_RS08970
351 and EGI09_01680, respectively). *YhcX* is probably involved in the degradation of indole-3-
352 acetonitrile, a sub product of tryptophan metabolism (Idris et al., 2007) (Fig. S1). On the
353 other hand, MtnU has been described as a possible enzyme catalyzing the conversion of
354 alpha-ketoglutaramate to alpha-ketoglutarate involved in the metabolism of methionine
355 (Ellens et al., 2015; Sekowska & Danchin, 2002) (Fig. S1). None of the enzymes with a
356 CN_hydrolase domain in *B. licheniformis* PER-URP-12 and *B. subtilis* PER-URP-17
357 appears to be responsible for cyanide degradation. However, apart from these proteins,
358 *Bacillus* and other genera present proteins with rhodanese domains (PFAM codes
359 PF12368 and PF00581) (Table S6) that are able to convert cyanide to thiocyanate
360 (Cipollone et al., 2006; Itakorode et al., 2019). Thus, it is likely that those rhodanese
361 enzymes are responsible for the degradation of cyanide by *B. licheniformis* PER-URP-12
362 and *B. subtilis* PER-URP-17. Further studies are necessary to test this hypothesis.

363 *B. safensis* PER-URP-08 presents *yhcX* (EGI07_01665) but not *mtnU*. In addition, this
364 strain carries two other proteins containing a CN_hydrolase domain, EGI07_17510 and
365 CynD (EGI07_08135). EGI07_17510 is a protein of unknown function whereas CynD
366 homologs (Fig. S2) hydrolyzes cyanide to produce ammonia and formic acid (Dash et al.,
367 2009; Ibrahim et al., 2015). We therefore carried out a series of experiments to test the

368 hypothesis that CynD is the enzyme responsible for cyanide degradation in *B. safensis*
369 PER-URP-08.

370 **C-terminal residues differentiate CynD from *B. pumilus* and *B. safensis***

371 We first constructed a maximum likelihood (ML) phylogenetic tree based on the 132 core
372 genomes of strains from *Bacillus pumilus* group (Fig. 2A) and searched for orthologs of
373 CynD in the strains present in the ML tree (see Methods for details of the search). The ML
374 tree confirmed the three branches identified above (Fig. 1A) and that two genomes
375 (CH144a_4T and 145) do not belong to any of these branches (Fig. 2A). Intriguingly,
376 CynD-encoding sequences were found in some representatives of *B. pumilus* (44 out of
377 56) and *B. safensis* (19 out of 23) but not in *B. altitudinis*. Three monophyletic *B. pumilus*
378 and one monophyletic *B. safensis* clades lack CynD (Fig. 2A). This could be due to
379 processes of gene gain and/or loss in the strains, and further studies are necessary to
380 distinguish between these or other possibilities. It is also possible that some *cynD* genes
381 were not sequenced in some incomplete genomes.

382 Next, we identified twenty-three different sequences of CynD in the 132 genomes (Table
383 S7) and a ML phylogenetic tree based on amino acid sequences was constructed,
384 including the sequences of the CynD from strain C1 (CynD_{C1}) (accession id: AAN77004.1)
385 and of the CynD from *B. safensis* PER-URP-08 (CynD_{PER-URP-08}). A clear separation
386 between CynD from *B. safensis* and from *B. pumilus* could be observed in the ML tree
387 (Fig. 2B). Interestingly, CynD_{C1} appear more related to the *B. safensis* group (Fig. 2B).
388 Due to the several taxonomic misclassifications of strains belonging to the *Bacillus pumilus*
389 group (as reported here and by others (Espariz et al., 2016; X. Fu et al., 2021; Liu et al.,
390 2013)), it is likely that strain C1 truly belongs to a *B. safensis* species; however, complete
391 genome of C1 is not available to confirm this hypothesis.

392 The most variable region in the nitrilase protein family is the C-terminal (Benedik & Sewell,
393 2018; Thuku et al., 2009). Thus, we associated a phylogenetic tree obtained from the full-
394 length sequences of identified CynDs homologs to an alignment of the C-terminal region
395 (residues 296 to 330) (Fig. 2B). Residues F314, D318, H323 in *B. safensis* CynD are
396 L314, A318, and N323 in the *B. pumilus* protein. Other residues can vary in one of the
397 species but are strictly conserved in the other, for instance, residues Q309 and I325 in *B.*
398 *safensis* are T309 or N309 and M325 or L325 in *B. pumilus*. Residue 308 can be P or M in
399 *B. safensis* but is strictly D in *B. pumilus* (Fig. 2B). CynD_{C1} has the amino acids strictly
400 conserved in *B. safensis* supporting the conclusion that C1 belongs to *B. safensis* species.
401 Furthermore, residue 27, outside the C-terminal, is E in *B. safensis* and strain C1 but Q in
402 *B. pumilus*.

403 **CynD from *B. safensis* PER-URP-08 is still active at pH 9.**

404 We then went on to characterize some biochemical properties of CynD_{PER-URP-08}. First, we
405 cloned and expressed recombinant CynD_{PER-URP-08} in *E. coli* and determined the basic
406 kinetic constants of the purified recombinant enzyme. Although CynDs are known to be
407 able to adopt different oligomeric states, no evidence of cooperativity was observed in our
408 enzymatic assays (Fig. 3A). Instead, a simple Michaelis-Menten model fit the experimental
409 data adequately. K_m and k_{cat} estimated using this model were 1.93 mM and 6.85 s⁻¹ (Fig.
410 3A, S3). The K_m value is similar to that previously reported for CynD_{C1} and from other
411 species (2.56 mM to 7.3 mM) similar but k_{cat} is lower (433 s⁻¹ to 61600 s⁻¹) (Crum et al.,
412 2015; Crum et al., 2016; Jandhyala et al., 2005; Vargas-Serna et al., 2020).
413

414 Due to the volatility of hydrogen cyanide in its protonated HCN state and its pKa of 9.2
415 (Brüger et al., 2018), bioremediation processes should preferably be carried out at or
416 above pH 9. To test if CynD_{PER-URP-08} is active at pHs greater than 8, we tested its activity
417 at pH 9, 10, and 11. Figure 3B shows that recombinant CynD_{PER-URP-08} carrying a C-
418 terminal 6x-His tag is active up to pH 9 and inactive at pH 10 and 11. Other wild-type
419 CynDs have been shown to be active only up to pH 8 (Crum et al., 2016; Jandhyala et al.,
420 2005) and CynD_{C1} with C-terminal 6x-His tag had its activity compromised at pH 9
421 (Vargas-Serna et al., 2020). The CynD_{C1} and CynD_{PER-URP-08} sequences only differ at five
422 positions: are I18V, S25T, E155D, H305Q and N307Y (first letter correspond to CynD_{C1})
423 with the last two substitutions H305Q and N307Y near the C-terminus.

424

425 Other studies were able to generate active versions of CynD active at pH 9 by introducing
426 mutations in some conserved positions (K93R; Q86R, E96G, D254G) or by replacing the
427 C-terminal from CynD_{C1} with the C-terminal from CynD from *Pseudomonas stutzeri* (Crum
428 et al., 2015; Wang et al., 2012) (note that wild-type CynD from *P. stutzeri* has not been
429 tested at pH 9).

430

431 **Alkaline pH reduces the degree of oligomerization of CynD_{PER-URP-08}**

432 The oligomerization state of nitrilases have been associated with enzyme activity and
433 stability (Crum et al., 2015; Crum et al., 2015; Crum et al., 2016; Martínková et al., 2015;
434 Park et al., 2016; Wang et al., 2012). In the case of CynDs of CynD_{C1} and CynD from *P.*
435 *stutzeri*, mutations in the C-terminal region decrease oligomerization (M. Crum et al., 2016;
436 M. A. N. Crum et al., 2015; Wang et al., 2012). The C-terminal of nitrilases stabilizes the
437 spiral structure through crisscrossed beta sheets in the center of the oligomer (Mulelu et
438 al., 2019; Thuku et al., 2009). Also, pH has been shown to promote higher order
439 oligomerization states of CynDs (D. Jandhyala et al., 2003; Wang et al., 2012); however,
440 the effects of pH greater than 9 have not been reported.

441 Since, CynD_{PER-URP-08} has differences in C-terminal with respect to other CynDs we used
442 SEC-MALS to compare the oligomerization states of CynD_{PER-URP-08} at different pHs. As
443 expected, pHs higher than 8 results in smaller sized oligomers. At pH 11 the monomer
444 (38.5 kDa) is the predominant species (Fig. 4A), whereas pH 10 and 9 presented
445 oligomeric states ranging from ~3-mer to ~5-mer (pH 10, 100.85 to 176.34 kDa) and ~4-
446 mer to ~6-mer (pH 9, 133.19 to 226.99 kDa) (Fig. 4B-C). Furthermore, CynD_{PER-URP-08}
447 presented oligomers ranging from ~24-mer to ~48-mer (918.31 to 1851.39 kDa) at pH 8
448 (Fig. 4D) in contrast to what was reported for CynD_{C1} at pH 8 which forms an 18-mer spiral
449 (D. Jandhyala et al., 2003). These differences could be a result of the differences in
450 aminoacid sequence between CynD_{PER-URP-08} and CynD_{C1} or due to the presence of the C-
451 terminal 6x-His tag in CynD_{PER-URP-08}. Experiments with CynD_{C1} were carried out with
452 untagged protein or with protein carrying an N-terminal 6x-His tag (Crum et al., 2015;
453 Jandhyala et al., 2003; Park et al., 2016; Wang et al., 2012). Electron micrographs of
454 negatively stained CynD_{PER-URP-08} at pH 8 showed spirals of different sizes supporting the
455 conclusion that CynD_{PER-URP-08} at this pH is adopts a range of different oligomerization
456 states (Fig. 4E).

457 **Expression of CynD_{PER-URP-08} from *B. safensis* PER-URP-08 is induced in the** 458 **presence of cyanide**

459 Some previous studies have considered the possibility that CynD gene expression is
460 regulated by cyanide, but this point remains unclear (D. Jandhyala et al., 2003). To
461 address this question, we exposed *B. safensis* PER-URP-08 to 100 ppm CN⁻ (in the form

462 of 38.5 mM NaCN) at 30 °C for 4 h without agitation and the mRNA levels of *cynD* were
463 measured and compared with the levels observed in cells grown in the absence of CN⁻.

464 We observed a 6.7-fold increase in expression of *cynD* in the presence of cyanide (Fig. 5).
465 To evaluate if this overexpression is specific for *cynD* nitrilase and not to other nitrilases of
466 *B. safensis* PER-URP-08, we also measured the mRNA levels of *ykrU* that also possesses
467 a CN₂ hydrolase domain. We did not observe differences in *ykrU* expression in the
468 presence and absence of cyanide. To our knowledge, this is the first report showing
469 induction in the expression of *cynD* in the presence of cyanide. This could possibly be a
470 physiological response of the bacteria in order to protect itself from the toxic effects of the
471 compound, but further studies are necessary to more fully understand the molecular
472 mechanisms behind this response.

473 CONCLUSIONS

474 Here we report the isolation and the genome sequences of three cyanide-degrading
475 *Bacillus* strains obtained from water in contact with mine tailings in Lima – Peru. They
476 were phylogenetically classified and named *Bacillus licheniformis* PER-URP-12, *Bacillus*
477 *subtilis* PER-URP-17 and *Bacillus safensis* PER-URP-08. Comparative genomic analyses
478 indicate that some strains currently classified as *B. pumilus* with publicly available
479 genomes should be reclassified as *Bacillus altitudinis* (strains TUAT1, MTCB 6033, SH-
480 B11, and C4). Furthermore, we propose that strains CH144a_4T and 145 should be
481 classified belonging a new species distinct from *B. pumilus*, *B. safensis*, or *B. altitudinis*.

482 We propose that in *B. licheniformis* PER-URP-12 and *B. subtilis* PER-URP-17 rhodanases
483 (table S6) are possibly the enzymes that confer cyanide degradation capabilities to these
484 strains. In the case of *B. safensis* PER-URP-08, we suggest that EGI07_08135 codes for
485 an ortholog of cyanide dihydratase CynD that imparts the cyanide-degradation ability to
486 this strain.

487 We found that while no *B. altitudinis* strains code for CynD orthologs, some *B. pumilus* and
488 *B. safensis* strains present CynD orthologous sequences. CynD from *B. pumilus* and *B.*
489 *safensis* have high identity (> 97%), however conserved differences in the C-terminus
490 allow us to differentiate between CynD from *B. safensis* or *B. pumilus* (at least in the
491 analyzed genomes). Additionally, sequence analysis of the previously described CynD
492 from strain C1 (CynD_{C1}), named as *B. pumilus* CynD in the literature, is more closely
493 related to CynDs from *B. safensis* than from *B. pumilus*. Thus, indicating that CynD_{C1} is a
494 representative of *B. pumilus* CynDs.

495 We characterized some aspects of CynD from *B. safensis* PER-URP-08 (CynD_{PER-URP-08})
496 corroborating what was described for CynDs from other species and adding new
497 knowledge about these enzymes. First, enzymatic assays with CynD_{PER-URP-08} found no
498 evidence of cooperativity despite the known oligomerization patterns of these enzymes.
499 Second, K_m and K_{cat} of CynD_{PER-URP-08} were 1.93 mM and 6.65 s⁻¹, respectively. Third,
500 despite that CynD_{PER-URP-08} and CynD_{C1} only differ in five positions, CynD_{PER-URP-08} retain
501 almost the same activity in pH 9 whereas CynD_{C1} has been reported as almost inactive at
502 this pH. Fourth, as pH is known to influence the oligomerization of CynDs, we reported that
503 in pH 8, CynD_{PER-URP-08} forms spirals made up of an estimated ~24 to ~48 subunits
504 showing that several oligomeric states are present in this pH. This is different compared
505 with CynD_{C1} that was reported to forms just oligomers of 18 subunits at this pH. Moreover,
506 at pH 11, the CynD_{PER-URP-08} monomer was observed. Finally, we showed for the first time,
507 that the abundance of CynD_{PER-URP-08} transcripts increases 6-fold when bacterial cultures
508 are exposed to CN⁻.

509 Altogether, the results we reported here warrant further investigation to explore the
510 potential of *B. safensis* PER-URP-08 and CynD_{PER-URP-08} for cyanide bioremediation.

511

512 DATA AVAILABILITY

513 The final genomes assemblies are available in IMG/M (Chen et al., 2021) and
514 GenBank/NCBI (Benson et al., 2013) databases under the accessions numbers:
515 2818991268, 2818991267, 2818991266 and RSEW00000000.1, RSEY00000000.1,
516 RSEX00000000.1, respectively for *Bacillus safensis* PER-URP-08, *Bacillus licheniformis*
517 PER-URP-12, *Bacillus subtilis* PER-URP-17.

518

519 REFERENCES

- 520 Akinpelu, E. A., Adetunji, A. T., Ntwampe, S. K. O., Nchu, F., & Mekuto, L. (2018).
521 Performance of fusarium oxysporum EKT01/02 isolate in cyanide biodegradation
522 system. *Environmental Engineering Research*, 23(2), 223–227.
523 <https://doi.org/10.4491/eer.2017.154>
- 524 Al-Badri, B. A. S., Al-Maawali, S. S., Al-Balushi, Z. M., Al-Mahmooli, I. H., Al-Sadi, A. M., &
525 Velazhahan, R. (2020). Cyanide degradation and antagonistic potential of endophytic
526 *Bacillus subtilis* strain BEB1 from *Bougainvillea spectabilis* Willd. *All Life*, 13(1), 92–
527 98. <https://doi.org/10.1080/26895293.2020.1728393>
- 528 Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local
529 alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410.
530 [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- 531 Benedik, M. J., & Sewell, B. T. (2018). Cyanide-degrading nitrilases in nature. *Journal of*
532 *General and Applied Microbiology*, 64(2), 90–93.
533 <https://doi.org/10.2323/jgam.2017.06.002>
- 534 Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., &
535 Sayers, E. W. (2013). GenBank. *Nucleic Acids Research*, 41(D1), 36–42.
536 <https://doi.org/10.1093/nar/gks1195>
- 537 Bosi, E., Donati, B., Galardini, M., Brunetti, S., Sagot, M. F., Lió, P., Crescenzi, P., Fani,
538 R., & Fondi, M. (2015). MeDuSa: A multi-draft based scaffold. *Bioinformatics*,
539 31(15), 2443–2451. <https://doi.org/10.1093/bioinformatics/btv171>
- 540 Brüger, A., Fafilek, G., Restrepo B., O. J., & Rojas-Mendoza, L. (2018). On the
541 volatilisation and decomposition of cyanide contaminations from gold mining. *Science*
542 *of the Total Environment*, 627, 1167–1173.
543 <https://doi.org/10.1016/j.scitotenv.2018.01.320>
- 544 Castric, P. A., & Strobel, G. A. (1969). Cyanide metabolism by *Bacillus megaterium*.
545 *Journal of Biological Chemistry*, 244(15), 4089–4094. [https://doi.org/10.1016/s0021-9258\(17\)36388-3](https://doi.org/10.1016/s0021-9258(17)36388-3)
- 547 Chen, I. M. A., Chu, K., Palaniappan, K., Ratner, A., Huang, J., Huntemann, M., Hajek, P.,
548 Ritter, S., Varghese, N., Seshadri, R., Roux, S., Woyke, T., Eloë-Fadrosch, E. A.,
549 Ivanova, N. N., & Kyrpides, N. C. (2021). The IMG/M data management and analysis
550 system v.6.0: New tools and advanced capabilities. *Nucleic Acids Research*, 49(D1),

- 551 D751–D763. <https://doi.org/10.1093/nar/gkaa939>
- 552 Cipollone, R., Ascenzi, P., Frangipani, E., & Visca, P. (2006). Cyanide detoxification by
553 recombinant bacterial rhodanese. *Chemosphere*, 63(6), 942–949.
554 <https://doi.org/10.1016/j.chemosphere.2005.09.048>
- 555 Cock, P. J. A., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I.,
556 Hamelryck, T., Kauff, F., Wilczynski, B., & De Hoon, M. J. L. (2009). Biopython:
557 Freely available Python tools for computational molecular biology and bioinformatics.
558 *Bioinformatics*, 25(11), 1422–1423. <https://doi.org/10.1093/bioinformatics/btp163>
- 559 Coil, D., Jospin, G., & Darling, A. E. (2015). A5-miseq: An updated pipeline to assemble
560 microbial genomes from Illumina MiSeq data. *Bioinformatics*, 31(4), 587–589.
561 <https://doi.org/10.1093/bioinformatics/btu661>
- 562 Crum, M. A., Park, J. M., Mulelu, A. E., Sewell, B. T., & Benedik, M. J. (2015). Probing C-
563 terminal interactions of the *Pseudomonas stutzeri* cyanide-degrading CynD protein.
564 *Applied Microbiology and Biotechnology*, 99(7), 3093–3102.
565 <https://doi.org/10.1007/s00253-014-6335-x>
- 566 Crum, M. A., Park, J. M., Sewell, B. T., & Benedik, M. J. (2015). C-terminal hybrid mutant
567 of *Bacillus pumilus* cyanide dihydratase dramatically enhances thermal stability and
568 pH tolerance by reinforcing oligomerization. *Journal of Applied Microbiology*, 118(4),
569 881–889. <https://doi.org/10.1111/jam.12754>
- 570 Crum, M. A., Trevor, B., & Benedik, M. (2016). *Bacillus pumilus* cyanide dihydratase
571 mutants with higher catalytic activity. *Frontiers in Microbiology*, 7(AUG), 1–10.
572 <https://doi.org/10.3389/fmicb.2016.01264>
- 573 Dash, R. R., Gaur, A., & Balomajumder, C. (2009). Cyanide in industrial wastewaters and
574 its removal: A review on biotreatment. *Journal of Hazardous Materials*, 163(1), 1–11.
575 <https://doi.org/10.1016/j.jhazmat.2008.06.051>
- 576 Dumestre, A., Chone, T., Portal, J. M., Gerard, M., & Berthelin, J. (1997). Cyanide
577 degradation under alkaline conditions by a strain of *Fusarium solani* isolated from
578 contaminated soils. *Applied and Environmental Microbiology*, 63(7), 2729–2734.
579 <https://doi.org/10.1128/aem.63.7.2729-2734.1997>
- 580 Ellens, K. W., Richardson, L. G. L., Frelin, O., Collins, J., Ribeiro, C. L., Hsieh, Y. F.,
581 Mullen, R. T., & Hanson, A. D. (2015). Evidence that glutamine transaminase and
582 omega-amidase potentially act in tandem to close the methionine salvage cycle in
583 bacteria and plants. *Phytochemistry*, 113, 160–169.
584 <https://doi.org/10.1016/j.phytochem.2014.04.012>
- 585 Espariz, M., Zuljan, F. A., Esteban, L., & Magni, C. (2016). Taxonomic identity resolution of
586 highly phylogenetically related strains and selection of phylogenetic markers by using
587 genome-scale methods: The *Bacillus pumilus* group case. *PLoS ONE*, 11(9), 1–17.
588 <https://doi.org/10.1371/journal.pone.0163098>
- 589 Fry, W. E., & Millar, R. L. (1972). Cyanide degradation by an enzyme from *Stemphylium loti*.
590 *Archives of Biochemistry and Biophysics*, 151(2), 468–474.
591 [https://doi.org/10.1016/0003-9861\(72\)90523-1](https://doi.org/10.1016/0003-9861(72)90523-1)
- 592 Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: Accelerated for clustering the
593 next-generation sequencing data. *Bioinformatics*, 28(23), 3150–3152.

- 594 <https://doi.org/10.1093/bioinformatics/bts565>
- 595 Fu, X., Gong, L., Liu, Y., Lai, Q., Li, G., & Shao, Z. (2021). *Bacillus pumilus* Group
596 Comparative Genomics: Toward Pangenome Features, Diversity, and Marine
597 Environmental Adaptation. *Frontiers in Microbiology*, 12(May), 1–16.
598 <https://doi.org/10.3389/fmicb.2021.571212>
- 599 Gioia, J., Yerrapragada, S., Qin, X., Jiang, H., Igboeli, O. C., Muzny, D., Dugan-Rocha, S.,
600 Ding, Y., Hawes, A., Liu, W., Perez, L., Kovar, C., Dinh, H., Lee, S., Nazareth, L.,
601 Blyth, P., Holder, M., Buhay, C., Tirumalai, M. R., ... Weinstock, G. M. (2007).
602 Paradoxical DNA repair and peroxide resistance gene conservation in *Bacillus*
603 *pumilus* SAFR-032. *PLoS ONE*, 2(9). <https://doi.org/10.1371/journal.pone.0000928>
- 604 Hendry-Hofer, T. B., Ng, P. C., Witeof, A. E., Mahon, S. B., Brenner, M., Boss, G. R., &
605 Beberta, V. S. (2019). A Review on Ingested Cyanide: Risks, Clinical Presentation,
606 Diagnostics, and Treatment Challenges. *Journal of Medical Toxicology*, 15(2), 128–
607 133. <https://doi.org/10.1007/s13181-018-0688-y>
- 608 Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018).
609 UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and*
610 *Evolution*, 35(2), 518–522. <https://doi.org/10.1093/molbev/msx281>
- 611 Ibrahim, K. K., Syed, M. A., Shukor, M. Y., & Ahmad, S. A. (2015). Biological remediation
612 of cyanide: A review. *Biotropia*, 22(2), 151–163.
613 <https://doi.org/10.11598/btb.2015.22.2.393>
- 614 Idris, E. S. E., Iglesias, D. J., Talon, M., & Borriss, R. (2007). Tryptophan-dependent
615 production of Indole-3-Acetic Acid (IAA) affects level of plant growth promotion by
616 *Bacillus amyloliquefaciens* FZB42. *Molecular Plant-Microbe Interactions*, 20(6), 619–
617 626. <https://doi.org/10.1094/MPMI-20-6-0619>
- 618 Ingvorsen, K., Hojer-Pedersen, B., & Godtfredsen, S. E. (1991). Novel cyanide-hydrolyzing
619 enzyme from *Alcaligenes xylosoxidans* subsp. *denitrificans*. *Applied and*
620 *Environmental Microbiology*, 57(6), 1783–1789.
621 <https://doi.org/10.1128/aem.57.6.1783-1789.1991>
- 622 Itakorode, B., Okonji, R., Adedeji, O., Torimiro, O., Famakinwa, T., & Chukwuejim, C.
623 (2019). Isolation, screening and optimization of *Bacillus cereus* for a thiosulphate
624 sulphur transferase production. *Journal of Chemical and Pharmaceutical Sciences*,
625 12(03), 79–84. <https://doi.org/10.30558/jchps.20191203003>
- 626 Jandhyala, D., Berman, M., Meyers, P. R., Sewell, B. T., Willson, R. C., & Benedik, M. J.
627 (2003). CynD, the cyanide dihydratase from *Bacillus pumilus*: Gene cloning and
628 structural studies. *Applied and Environmental Microbiology*, 69(8), 4794–4805.
629 <https://doi.org/10.1128/AEM.69.8.4794-4805.2003>
- 630 Jandhyala, D. M., Willson, R. C., Sewell, B. T., & Benedik, M. J. (2005). Comparison of
631 cyanide-degrading nitrilases. *Applied Microbiology and Biotechnology*, 68(3), 327–
632 335. <https://doi.org/10.1007/s00253-005-1903-8>
- 633 Javaheri Safa, Z., Aminzadeh, S., Zamani, M., & Motallebi, M. (2017). Significant increase
634 in cyanide degradation by *Bacillus* sp. M01 PTCC 1908 with response surface
635 methodology optimization. *AMB Express*, 7(1). [https://doi.org/10.1186/s13568-017-](https://doi.org/10.1186/s13568-017-0502-2)
636 0502-2

- 637 Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version
638 7: Improvements in performance and usability. *Molecular Biology and Evolution*,
639 30(4), 772–780. <https://doi.org/10.1093/molbev/mst010>
- 640 Kosiol, C., & Goldman, N. (2005). Different versions of the dayhoff rate matrix. *Molecular*
641 *Biology and Evolution*, 22(2), 193–199. <https://doi.org/10.1093/molbev/msi005>
- 642 Kuyucak, N., & Akcil, A. (2013). Cyanide and removal options from effluents in gold mining
643 and metallurgical processes. *Minerals Engineering*, 50–51, 13–29.
644 <https://doi.org/10.1016/j.mineng.2013.05.027>
- 645 Leavesley, H. B., Li, L., Prabhakaran, K., Borowitz, J. L., & Isom, G. E. (2008). Interaction
646 of cyanide and nitric oxide with cytochrome c oxidase: Implications for acute cyanide
647 toxicity. *Toxicological Sciences*, 101(1), 101–111.
648 <https://doi.org/10.1093/toxsci/kfm254>
- 649 Li, W., & Godzik, A. (2006). Cd-hit: A fast program for clustering and comparing large sets
650 of protein or nucleotide sequences. *Bioinformatics*, 22(13), 1658–1659.
651 <https://doi.org/10.1093/bioinformatics/btl158>
- 652 Liu, Y., Lai, Q., Dong, C., Sun, F., Wang, L., Li, G., & Shao, Z. (2013). Phylogenetic
653 diversity of the *Bacillus pumilus* group and the marine ecotype revealed by multilocus
654 sequence analysis. *PLoS ONE*, 8(11), 1–11.
655 <https://doi.org/10.1371/journal.pone.0080097>
- 656 Martínková, L., Veselá, A. B., Rinágelová, A., & Chmátal, M. (2015). Cyanide hydratases
657 and cyanide dihydratases: emerging tools in the biodegradation and biodetection of
658 cyanide. *Applied Microbiology and Biotechnology*, 99(21), 8875–8882.
659 <https://doi.org/10.1007/s00253-015-6899-0>
- 660 Mekuto, L., Jackson, V. A., & Obed Ntwampe, S. K. (2014). Biodegradation of Free
661 Cyanide Using *Bacillus* Sp. Consortium Dominated by *Bacillus Safensis*, *Lichenformis*
662 and *Tequilensis* Strains: A Bioprocess Supported Solely with Whey. *Journal of*
663 *Bioremediation & Biodegradation*, 05(02). <https://doi.org/10.4172/2155-6199.s18-004>
- 664 Meyers, P. R., Rawlings, D. E., Woods, D. R., & Lindsey, G. G. (1993). Isolation and
665 characterization of a cyanide dihydratase from *Bacillus pumilus* C1. *Journal of*
666 *Bacteriology*, 175(19), 6105–6112. <https://doi.org/10.1128/jb.175.19.6105-6112.1993>
- 667 Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von
668 Haeseler, A., Lanfear, R., & Teeling, E. (2020). IQ-TREE 2: New Models and Efficient
669 Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and*
670 *Evolution*, 37(5), 1530–1534. <https://doi.org/10.1093/molbev/msaa015>
- 671 Mudder, T. I., Botz, M. M., & Akçil, A. (2004). Cyanide and society: A critical review. *The*
672 *European Journal of Mineral Processing and Environmental Protection*, 4(1), 62–74.
- 673 Mulelu, A. E., Kirykowicz, A. M., & Woodward, J. D. (2019). Cryo-EM and directed
674 evolution reveal how *Arabidopsis* nitrilase specificity is influenced by its quaternary
675 structure. *Communications Biology*, 2(1), 1–11. <https://doi.org/10.1038/s42003-019-0505-4>
676
- 677 Omura, H., Ikemoto, M., Kobayashi, M., Shimizu, S., Yoshida, T., & Nagasawa, T. (2003).
678 Purification, characterization and gene cloning of thermostable O-acetyl-L-
679 homoserine sulfhydrylase forming γ -cyano- α -aminobutyric acid. *Journal of Bioscience*

- 680 *and Bioengineering*, 96(1), 53–58. [https://doi.org/10.1016/S1389-1723\(03\)90096-X](https://doi.org/10.1016/S1389-1723(03)90096-X)
- 681 Pace, H. C., & Brenner, C. (2001). The nitrilase superfamily: Classification, structure and
682 function. *Genome Biology*, 2(1), 1–9. [https://doi.org/10.1186/gb-2001-2-1-](https://doi.org/10.1186/gb-2001-2-1-reviews0001)
683 reviews0001
- 684 Park, J. M., Ponder, C. M., Sewell, B. T., & Benedik, M. J. (2016). Residue Y70 of the
685 nitrilase cyanide dihydratase from *Bacillus pumilus* is critical for formation and activity
686 of the spiral oligomer. *Journal of Microbiology and Biotechnology*, 26(12), 2179–2183.
687 <https://doi.org/10.4014/jmb.1606.06035>
- 688 Ping Wang; Hans D. VanEtten. (1992). Cloning and properties of a cyanide hydratase
689 gene from the phytopathogenic fungus *Gloeocercospora sorghi*. *Biochemical and*
690 *Biophysical Research Communications*, 187(2), 1048–1054.
691 <https://www.sciencedirect.com/science/article/abs/pii/0006291X92913038>
- 692 Rinágelová, A., Kaplan, O., Veselá, A. B., Chmátal, M., Křenková, A., Plíhal, O.,
693 Pasquarelli, F., Cantarella, M., & Martinková, L. (2014). Cyanide hydratase from
694 *Aspergillus niger* K10: Overproduction in *Escherichia coli*, purification,
695 characterization and use in continuous cyanide degradation. *Process Biochemistry*,
696 49(3), 445–450. <https://doi.org/10.1016/j.procbio.2013.12.008>
- 697 Satomi, M., La Duc, M. T., & Venkateswaran, K. (2006). *Bacillus safensis* sp.nov., isolated
698 from spacecraft and assembly-facility surfaces. *International Journal of Systematic*
699 *and Evolutionary Microbiology*, 56(8), 1735–1740. [https://doi.org/10.1099/ijs.0.64189-](https://doi.org/10.1099/ijs.0.64189-0)
700 0
- 701 Sekowska, A., & Danchin, A. (2002). The methionine salvage pathway in *Bacillus subtilis*.
702 *BMC Microbiology*, 2, 1–14. <https://doi.org/10.1186/1471-2180-2-8>
- 703 Sexton, A. C., & Howlett, B. J. (2000). Characterisation of a cyanide hydratase gene in the
704 phytopathogenic fungus *Leptosphaeria maculans*. *Molecular and General Genetics*,
705 263(3), 463–470. <https://doi.org/10.1007/s004380051190>
- 706 Stepanov, V. G., Tirumalai, M. R., Montazari, S., Checinska, A., Venkateswaran, K., &
707 Fox, G. E. (2016). *Bacillus pumilus* SAFR-032 genome revisited: Sequence update
708 and re-annotation. *PLoS ONE*, 11(6), 1–11.
709 <https://doi.org/10.1371/journal.pone.0157331>
- 710 Tatusova, T., Dicuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L.,
711 Lomsadze, A., Pruitt, K. D., Borodovsky, M., & Ostell, J. (2016). NCBI prokaryotic
712 genome annotation pipeline. *Nucleic Acids Research*, 44(14), 6614–6624.
713 <https://doi.org/10.1093/nar/gkw569>
- 714 Thuku, R. N., Brady, D., Benedik, M. J., & Sewell, B. T. (2009). Microbial nitrilases:
715 Versatile, spiral forming, industrial enzymes. *Journal of Applied Microbiology*, 106(3),
716 703–727. <https://doi.org/10.1111/j.1365-2672.2008.03941.x>
- 717 Vargas-Serna, C. L., Carmona-Orozco, M. L., & Panay, A. J. (2020). Biodegradation of
718 cyanide using recombinant *Escherichia coli* expressing *Bacillus pumilus* cyanide
719 dihydratase. *Revista Colombiana de Biotecnología*, 22(1), 27–35.
720 <https://doi.org/10.15446/rev.colomb.biote.v22n1.79559>
- 721 Wang, L., Watermeyer, J. M., Mulelu, A. E., Sewell, B. T., & Benedik, M. J. (2012).
722 Engineering pH-tolerant mutants of a cyanide dihydratase. *Applied Microbiology and*

723 *Biotechnology*, 94(1), 131–140. <https://doi.org/10.1007/s00253-011-3620-9>

724 Watanabe, A., Yano, K., Ikebukuro, K., & Karube, I. (1998). Cloning and expression of a
725 gene encoding cyanidase from *Pseudomonas stutzeri* AK61. *Applied Microbiology*
726 *and Biotechnology*, 50(1), 93–97. <https://doi.org/10.1007/s002530051261>

727 Weisenfeld, N. I., Yin, S., Sharpe, T., Lau, B., Hegarty, R., Holmes, L., Sogoloff, B.,
728 Tabbaa, D., Williams, L., Russ, C., Nusbaum, C., Lander, E. S., Maccallum, I., &
729 Jaffe, D. B. (2014). Comprehensive variation discovery in single human genomes.
730 *Nature Genetics*, 46(12), 1350–1355. <https://doi.org/10.1038/ng.3121>

731 Williams, H. J., & Edwards, T. G. (1980). Estimation of cyanide with alkaline picrate.
732 *Journal of the Science of Food and Agriculture*, 31(1), 15–22.
733 <https://doi.org/10.1002/jsfa.2740310104>

734

735 **Acknowledgements**

736 We are very grateful to the Ricardo Palma University High-Performance Computational
737 Cluster (URPHPC) managers Gustavo Adolfo Abarca Valdiviezo and Roxana Paola Mier
738 Hermoza at the Ricardo Palma Informatic Department (OFICIC) for their contribution in
739 programs and remote use configuration of URPHPC. We also thank Germán Sgro for help
740 with negative staining.

741 **Funding Information**

742 Funding for this work was provided by São Paulo Research Foundation (FAPESP) student
743 fellowship 2015/13318-4 to S.J.A and research grant 2017/17303-7 to C.S.F. The work
744 was also supported by Coordination for the Improvement of Higher Education Personnel
745 (CAPES) research grant 3385/2013 to A.M.D.S. P.M.P. received fellowships from CAPES
746 (88882.160114/2017-1). C.S.F., and J.C.S., A.M.D.S. received research fellowship awards
747 from the National Council for Scientific and Technological Development (CNPq). The
748 funders had no role in study design, data collection and analysis, decision to publish, or
749 preparation of the manuscript.

750 **Authors Contributions**

751 Conceptualization: S.J.A., A.G.S., A.M.D.S., Methodology: S.J.A., D.Z.S., A.C.P., M.B.R.,
752 C.M.P., L.F.M., P.M.P., C.M. Computing resources: M.Q.A., J.C.S., Data curation: S.J.A.,
753 D.Z.S., A.C.P., C.M.P. Formal analysis: S.J.A., D.S.Z., A.C.P., C.M.P., C.S.F.
754 Visualization: S.J.A., D.Z.S. Writing – original draft preparation: S.J.A., Writing – review
755 and editing: S.J.A., J.C.S., C.S.F., A.M.D.S. Supervision: S.J.A., A.G.S., M.Q.A., C.S.F.,
756 A.M.D.S. Funding acquisition: A.G.S., M.Q.A., C.S.F., J.C.S., A.M.D.S. All authors read,
757 provided critical review, and approved the final manuscript.

758 **Conflicts of interest**

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

760 **FIGURES:**

761 **Figure 1. Core genome identity matrix to classified genomes of *Bacillus pumilus***
762 **group genomes.** A) An identity matrix of 132 core genomes of *Bacillus pumilus* group

763 showing delimitations between three species: *Bacillus altitudinis* (brown names), *Bacillus*
764 *safensis* (green names), *Bacillus pumilus* (blue names). Two core genomes (red names)
765 appear outside of these three species. B – E) Plots showing the range of identity when
766 compare *B. altitudinis* (B), *B. safensis* (C), *B. pumilus* (D) or *B. sp* (E) with itself or with
767 other groups.

768 **Figure 2. CynD is present in some genomes of *B. pumilus* and *B. safensis* and they**
769 **are mainly differentiated by C-terminal residues.** A) Maximum likelihood tree of core
770 genomes of 132 *Bacillus pumilus* group strains showing separation between three species.
771 Color of the circles represent absence (green) or presence (blue) of CynD homologue in
772 the genome. Circles with black and red borders represent complete genomes
773 (“chromosome” or “complete” sequencing status in NCBI) and possibly not complete
774 genomes (“scaffold” sequencing status in NCBI). B) Maximum likelihood tree of full-length
775 CynD sequences associated to and alignment of their C-terminal region (residues 296 to
776 330). Shown in number blue or green are the positions that are completely conserved in
777 *Bacillus safensis* or *B. pumilus*, respectively.

778 **Figure 3. CynD_{PER-URP-08} have similar kinetic constants to other CynD homologues**
779 **and is still active up to pH 9.** A) Plot of CynD_{PER-URP-08} Initial velocity (Vo) versus initial
780 concentration of cyanide adjusted to the Michaelis Menten equation. Km and Kcat
781 constants calculated assuming this model are shown in the graphic. B) Percentage of
782 cyanide removal in different pHs using different CynD_{PER-URP-08} concentrations. CynD_{PER-}
783 _{URP-08} showed considerable activity in pH 8 and 9 but not in 10 and 11.

784 **Figure 4. SEC-MALS of CynD_{PER-URP-08} showed that higher pHs reduced its**
785 **oligomerization states and TEM showed that CynD_{PER-URP-08} presents a helical**
786 **structure.** A-D) Plot of UV intensity/molar mass for CynD_{PER-URP-08} in different pHs. A
787 pattern of decrease the oligomeric state while increasing the pH was observed. E) TEM
788 micrographs at pH 8 in two different magnifications (right and left) showed helical
789 structures of CynD_{PER-URP-08}.

790 **Figure 5. cynD_{PER-URP-08} but not ykrU is induced in the presence of cyanide.** Relative
791 expression measured by RT-qPCR showed that when *Bacillus safensis* is in presence of
792 cyanide the RNA levels of *cynD* are 6.67-fold greater than when in absence of cyanide. In
793 contrast, other nitrilase gene (*ykrU*) have the same RNA levels in presence or absence of
794 cyanide.

795

796

797

798

799

800

801

802 **SUPPLEMENTARY FIGURES**

803 **Figure S1. Proteins containing CN_hydrolase domain in the three genomes studied.**
804 Four CN_hydrolases containing-proteins were identified in the analyzed genomes. YkrU
805 and CynD are present only in *B. safensis* PER-URP-08. MtnU was found in *B.*
806 *licheniformis* PER-URP-12 and *Bacillus subtilis* PER-URP-17. YhcX was found in the three
807 genomes.

808 **Figure S2. Alignments of identical protein group CynDs with CynD_{PER-URP-08} and**
809 **CynD_{C1}.** Homology of CynD from *Bacillus safensis* PER-URP-08 is clearly showed in the
810 protein sequence alignments of several CynD homologs including those with tested
811 enzymatic activity as CynD from *B. pumilus* strain C1.

812 **Figure S3. CynD_{PER-URP-08} production of NH₄ by time.** Linear adjust of the product
813 formation (NH₄) by CynD in the first 40 seconds of reaction using different initial
814 concentrations of cyanide.

815

816 **SUPPLEMENTARY TABLES:**

817 **Table S1. Accession numbers of reference genomes used in the assembly process.**

818 **Table S2. Cyanide removal percentage of the twenty isolates from mine tailings in**
819 **Peru.**

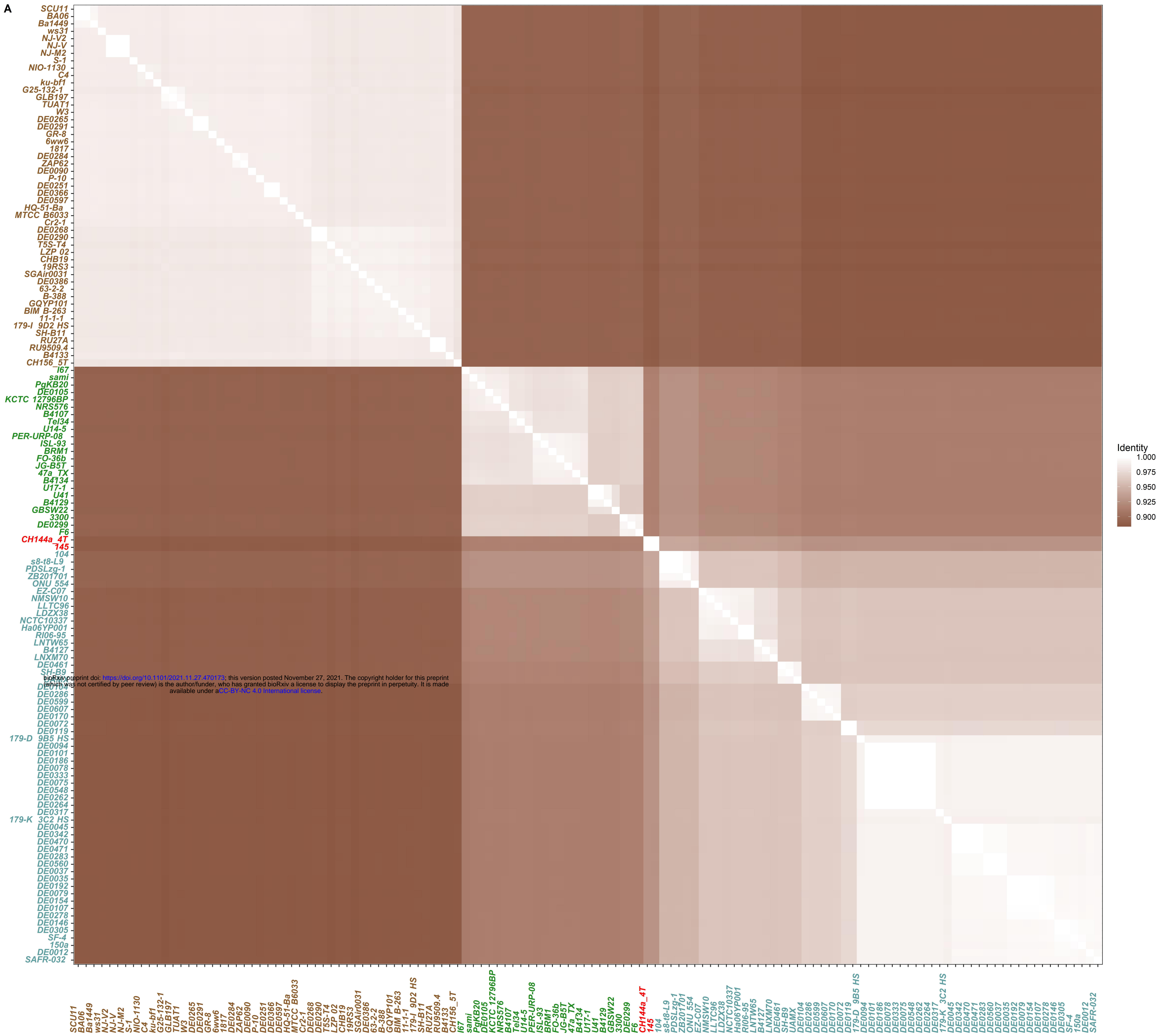
820 **Table S3. BLAST best-hits of the partial 16S rRNA gene for each tested strain.**

821 **Table S4. Summary of IMG/M annotations of the three *Bacillus* genomes.**

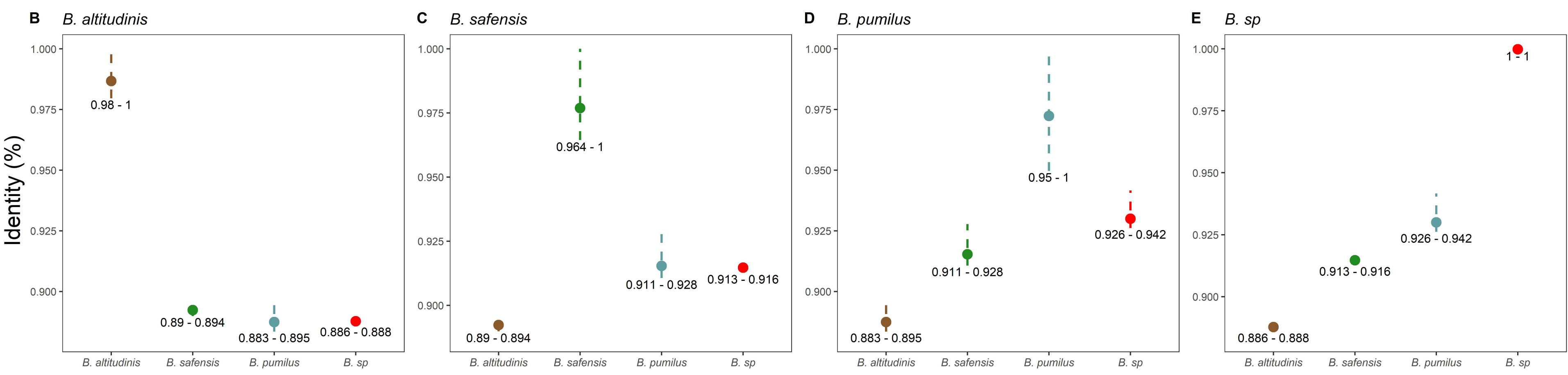
822 **Table S5. Summary information of the 132 genomes used in the core genomes**
823 **analysis.**

824 **Table S6. Rhodanese domain coding ORFs in the three sequenced genomes.**

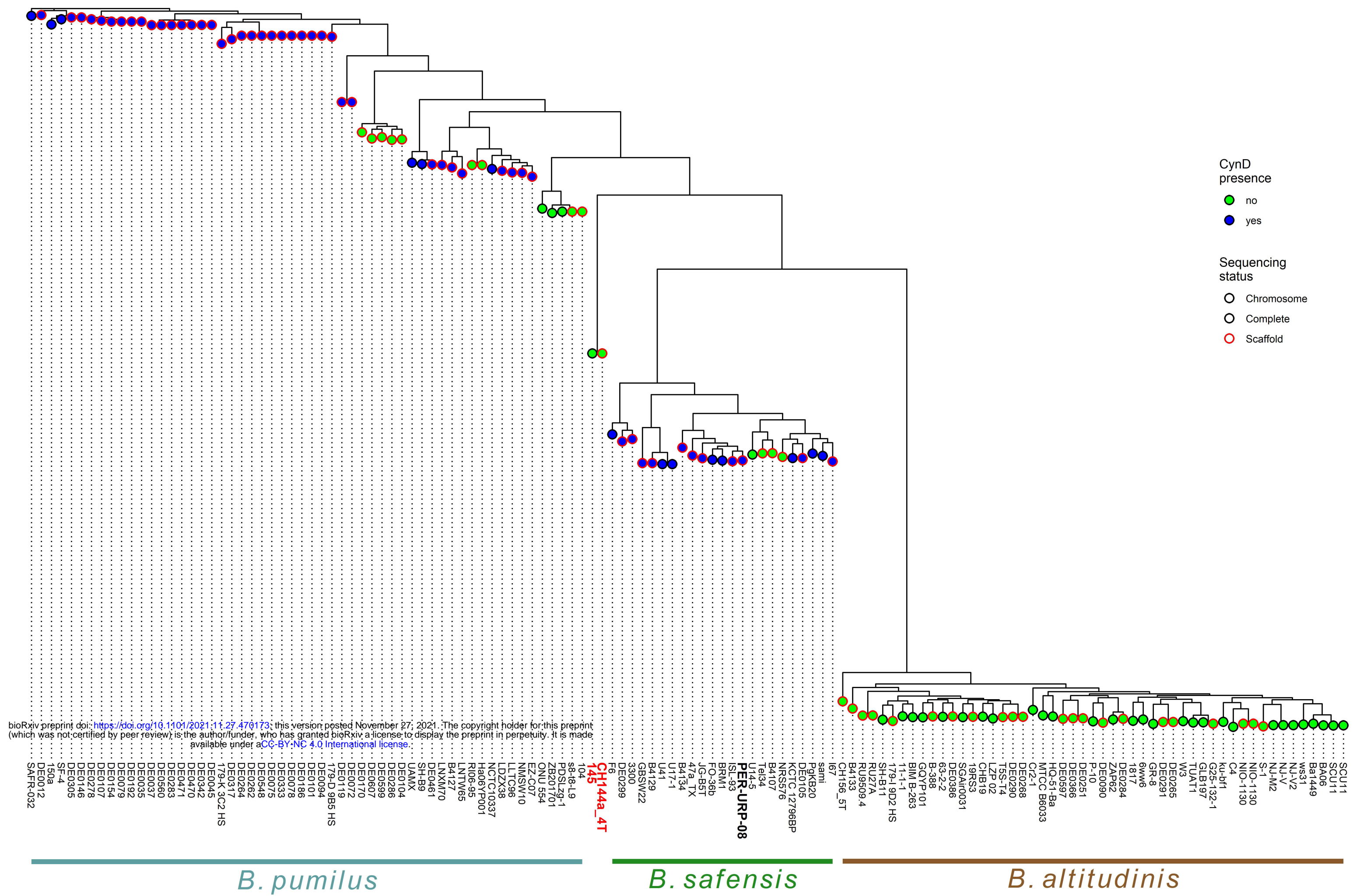
825 **Table S7. Identical protein groups (IPG) NCBI accession IDs by strain and species.**



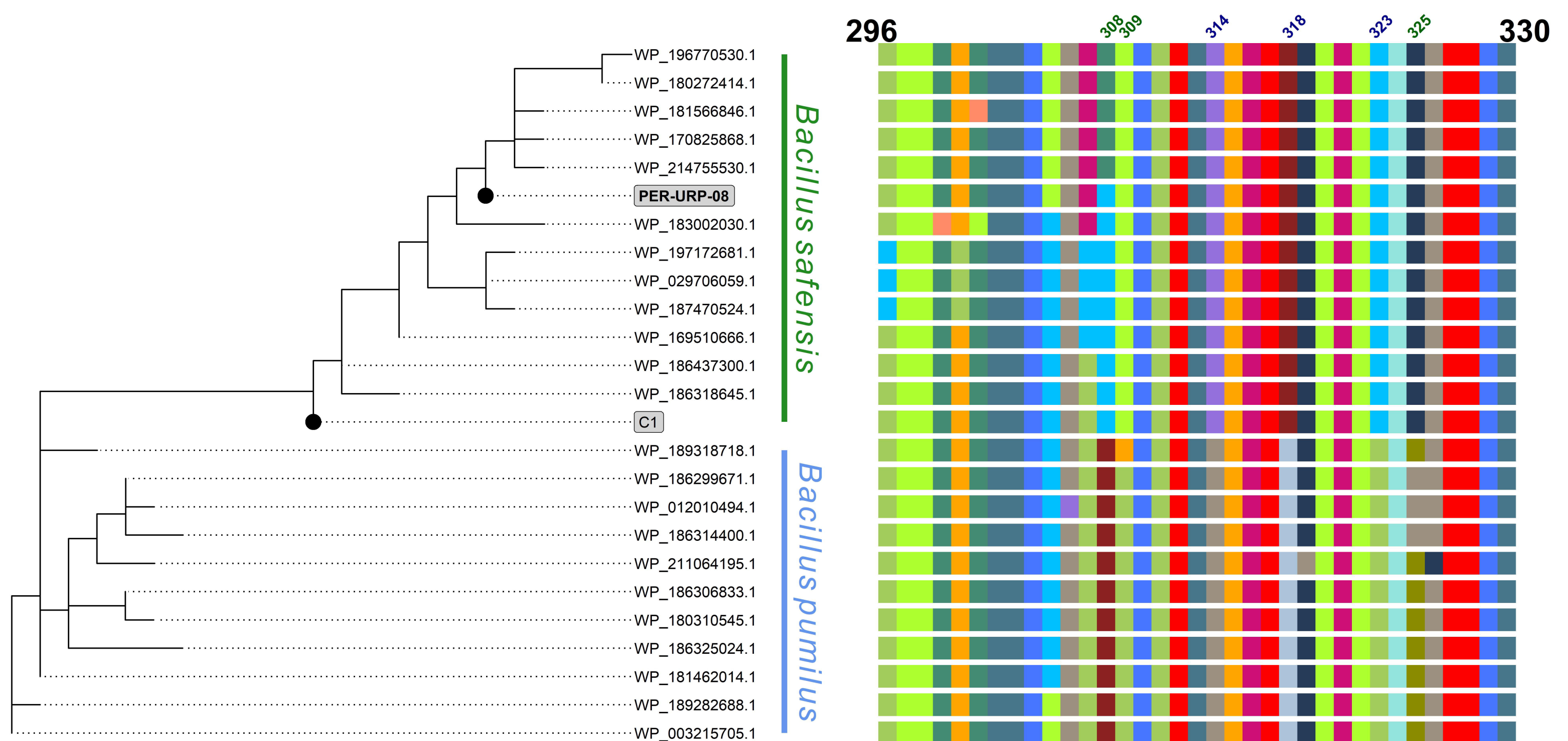
bioRxiv preprint doi: <https://doi.org/10.1101/2021.11.27.470173>; this version posted November 27, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

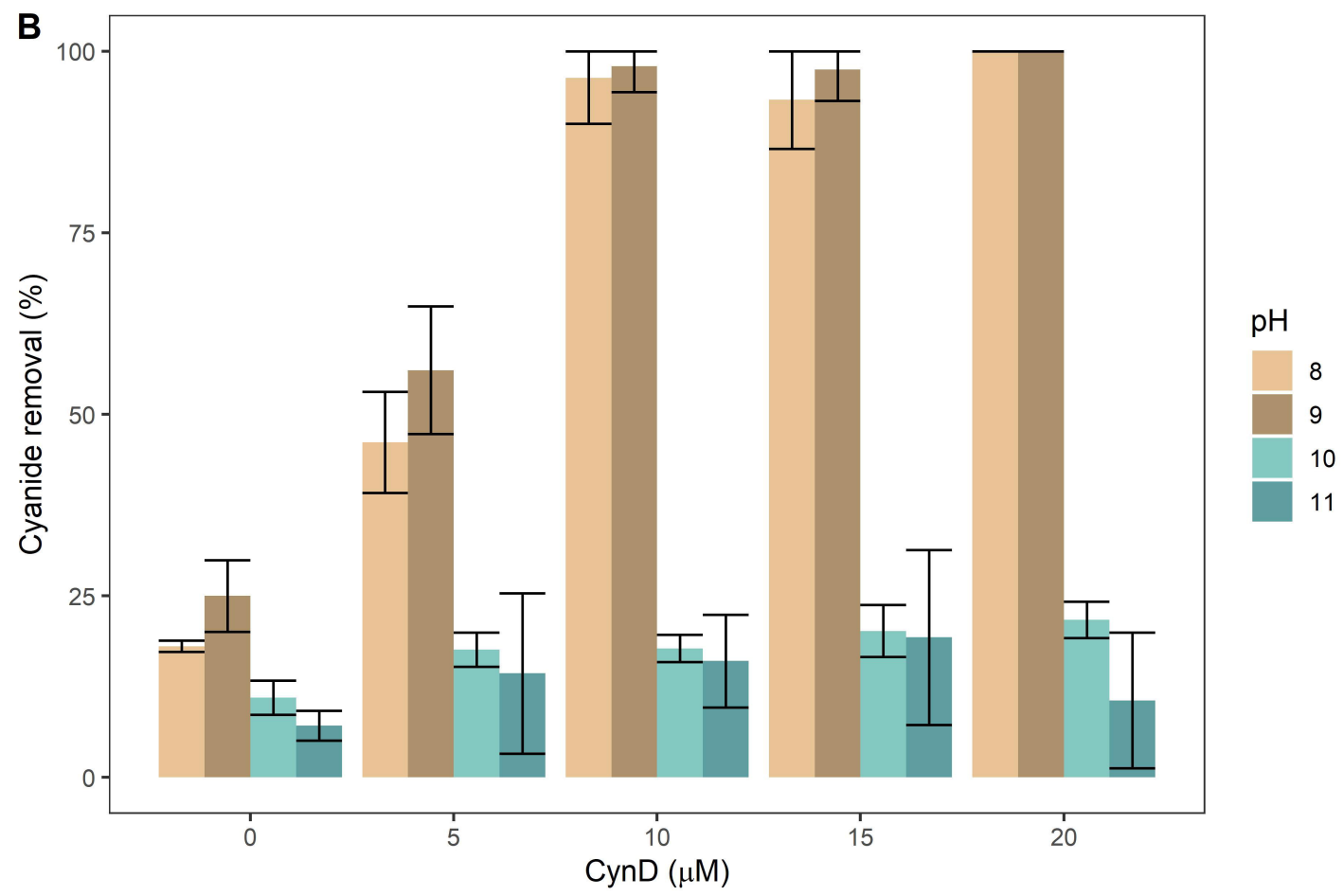
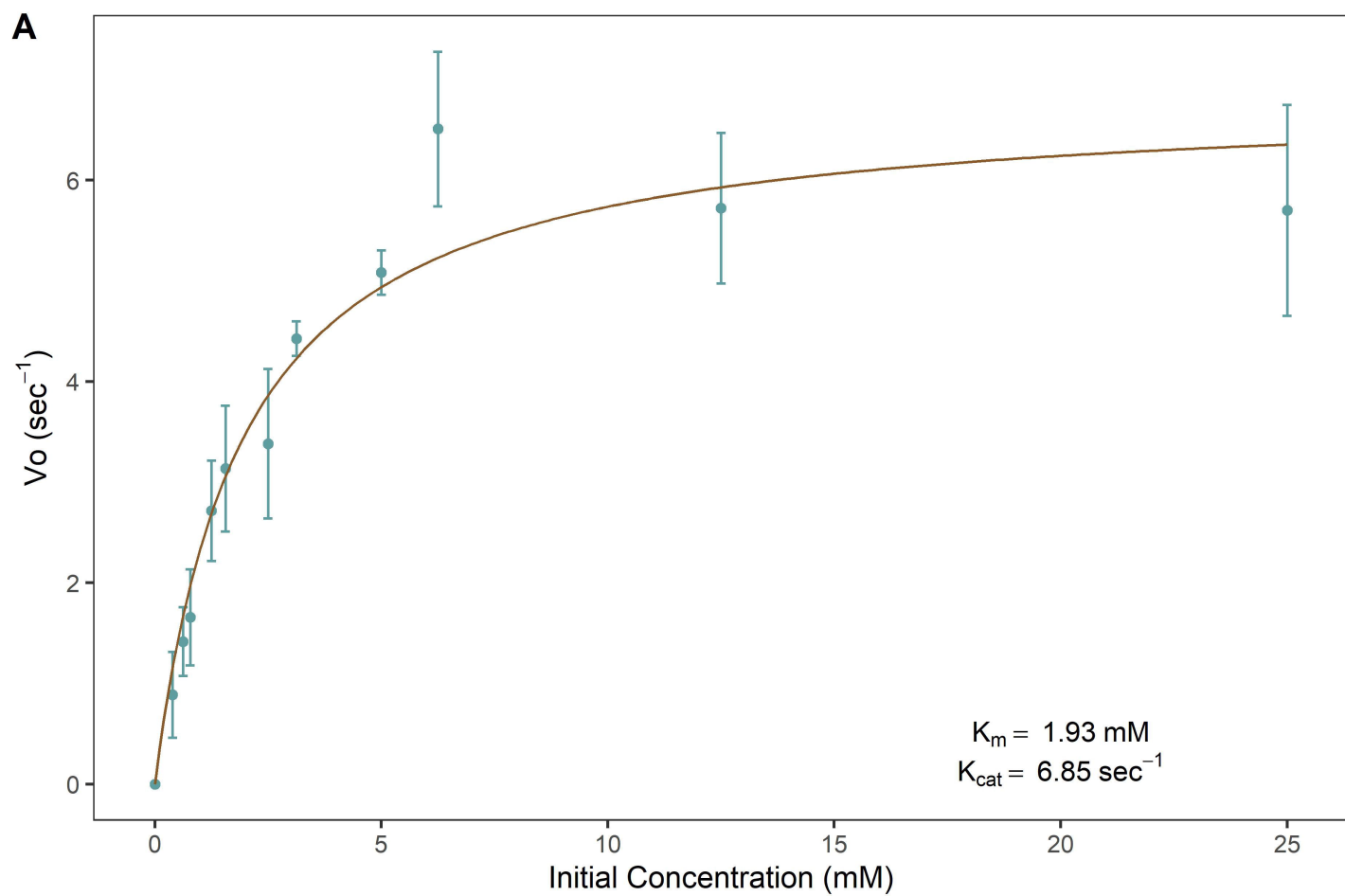


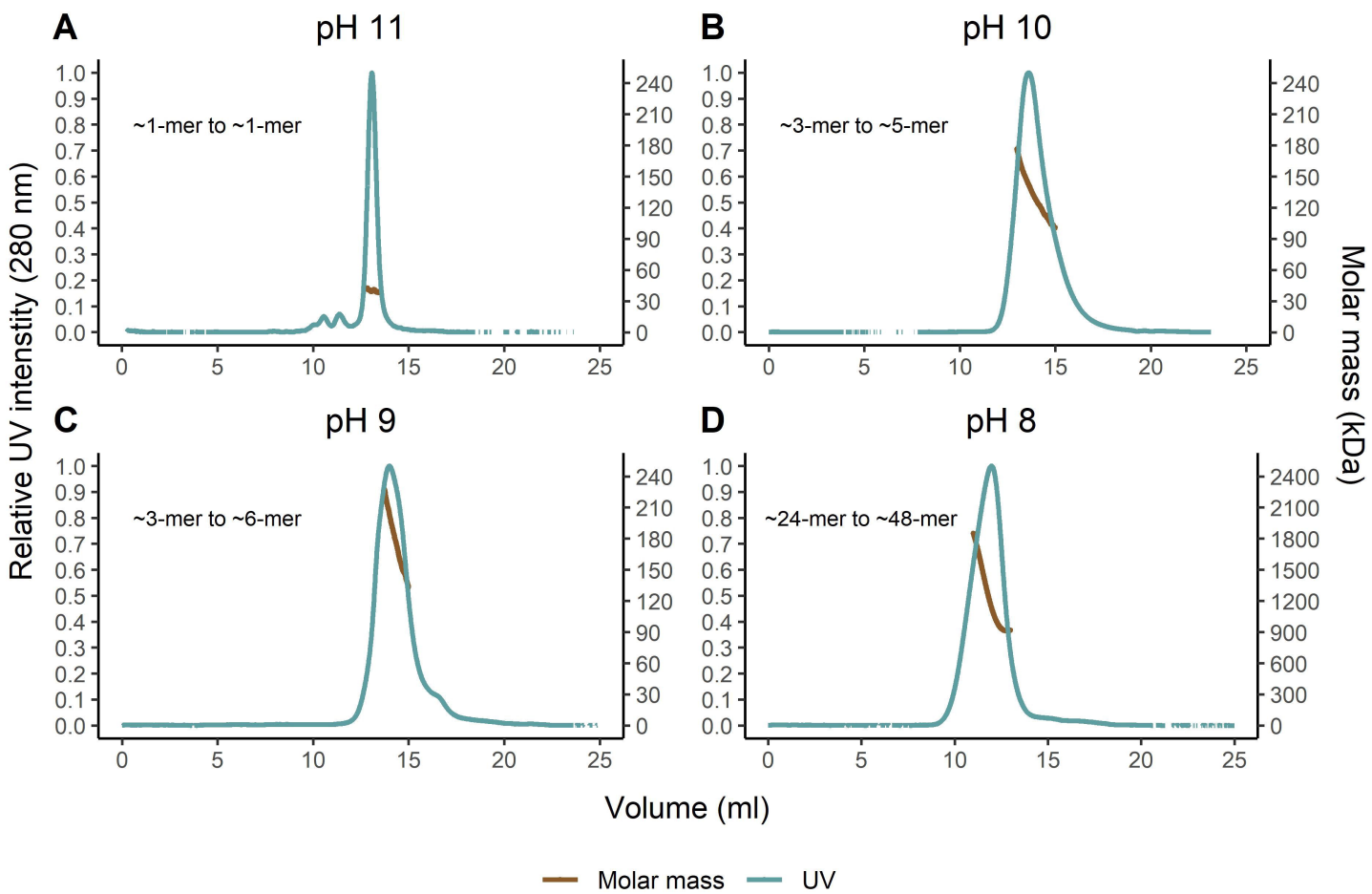
A



B







E

