Linking Genomic and Physiological Characteristics of Psychrophilic Arthrobacter to Metagenomic Data to Explain Global Environmental Distribution

Liang Shen

Anhui Normal University   https://orcid.org/0000-0003-0374-4198

Yongqin Liu (✉ yqliu@itpcas.ac.cn)

Institute of Tibetan Plateau Research Chinese Academy of Sciences and CAS Center for Excellence in Tibetan Plateau Earth Sciences   https://orcid.org/0000-0003-2876-7484

Michelle A. Allen

University of New South Wales - Kensington Campus: University of New South Wales

Baiqing Xu

Institute of Tibetan Plateau Research Chinese Academy of Sciences

Ninglian Wang

Northwest University

Timothy J. Williams

University of New South Wales - Kensington Campus: University of New South Wales

Feng Wang

Institute of Tibetan Plateau Research Chinese Academy of Sciences

Yuguang Zhou

Institute of Microbiology Chinese Academy of Sciences

Qing Liu

Institute of Microbiology Chinese Academy of Sciences

Ricardo Cavicchioli

University of New South Wales - Kensington Campus: University of New South Wales

Research

1 **Linking genomic and physiological characteristics of psychrophilic *Arthrobacter* to**

2 **metagenomic data to explain global environmental distribution**

3

4 **Running title: Polar and alpine *Arthrobacter***

5

6 **Liang Shen[1, 2], Yongqin Liu[1, 3*], Michelle A. Allen[4], Baiqing Xu[1, 3], Ninglian Wang[3, 5],**

7 **Timothy J. Williams[4], Feng Wang[1], Yuguang Zhou[6], Qing Liu[6], Ricardo Cavicchioli[4*]**

8

9 [1]Key Laboratory of Tibetan Environment Changes and Land Surface Processes, Institute of

10 Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, 100085, China

11 [2]College of Life Sciences, Anhui Normal University, Wuhu, 241000, China

12 [3]CAS Center for Excellence in Tibetan Plateau Earth Sciences, Beijing, 100085, China

13 [4]School of Biotechnology and Biomolecular Sciences, University of New South Wales,

14 Sydney, NSW 2052, Australia

15 [5]College of Urban and Environmental Science, Northwest University, Xian, 710069, China

16 [6]China General Microbiological Culture Collection Center (CGMCC), Institute of

17 Microbiology, Chinese Academy of Sciences, Beijing 100101, China

18

19 *Correspondence: yqliu@itpcas.ac.cn

20 [1]Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, 100101, China

21 r.cavicchioli@unsw.edu.au

22 [4]School of Biotechnology and Biomolecular Sciences, University of New South Wales,

23 Sydney, NSW 2052, Australia

Abstract

Background: Microorganisms drive critical global biogeochemical cycles and dominate the biomass in Earth's expansive cold biosphere. Determining the genomic traits that enable psychrophiles to grow in cold environments informs about their physiology and adaptive responses. However, defining important genomic traits of psychrophiles has proven difficult, with the ability to extrapolate genomic knowledge to environmental relevance proving even more difficult.

Results: Here we examined the bacterial genus *Arthrobacter*, and assisted by genome sequences of new Tibetan Plateau isolates, defined a clade, Group C, that represents isolates from polar and alpine environments. Group C had a superior ability to grow at -1ºC, and possessed genome G+C content, amino acid composition, predicted protein stability and functional capacities (e.g., sulfur metabolism and mycothiol biosynthesis) that distinguished it from non-polar or alpine Group A-*Arthrobacter*. Interrogation of more than 1,000 metagenomes identified an over-representation of Group C in Canadian permafrost communities from a simulated spring-thaw experiment, indicative of niche adaptation, and an under-representation of Group A in all polar and alpine samples, indicative of a general response to environmental temperature.

Conclusion: The findings illustrate a capacity to define genomic markers of specific taxa that potentially have value for environmental monitoring of cold environments, including environmental change arising from anthropogenic impact. More broadly, the study illustrates the challenges involved in extrapolating from genomic and physiological data to an environmental setting.

Keywords: Genomics, metagenomics, psychrophiles, polar environment, alpine environment, microbial adaptation

Background

Many biotic and abiotic factors influence the ability of microorganisms to become indigenous members of environmental communities. Certain environmental factors can limit or prevent the growth of microorganisms, while enhancing, or being essential for others, resulting in ecological niches that support specific microbiome structures [1]. This phenomenon is well illustrated by a Winogradsky column where light and oxygen can be seen to exert major influences on the diversity and the dynamic of microorganisms throughout its length. In more recent times, particularly through technological advancement (e.g., metagenomics), the understanding of microbial ecology and the contributions that microorganisms make to the natural world has grown considerably. Appreciation for microorganisms has accrued from discoveries of new biomes capable of supporting microbial colonization, such as the deep subsurface [2]; new examples of life hidden within microbial 'dark matter' (e.g. Asgard archaea [3]), and; the dynamic nature of microbial responses, particularly those that provide surprises, such as the major societal upheaval caused by the COVID19 coronavirus. There is a growing realization that microorganisms constitute the life support system of the biosphere and must be properly accounted for when devising strategies to mitigate the impacts of human activity on the natural world [4]. In essence, we are living in a period in history when the need for society to learn about microbial responses to natural and anthropogenic influences is of unprecedented relevance [4-6].

Metagenomic methods have provided a level of insight into microbial communities [7, 8] that could possibly be equated to the advances made by Carl Woese and colleagues when using rRNA sequencing to discover Archaea as the third domain of life. Applied to the cold biosphere, Earth's single largest biome, metagenomic analyses have catalogued diverse ways in which microbial life has evolved. As an example, Antarctic, marine-derived lake communities have been shown to have evolved independently over their relatively short history of 3,000-5,000 years, adapting not just to low temperature but also to a variety of important environmental factors specific to each lake system (reviewed in REF. [9]). Metagenomic analyses have also begun to be used to uncover the ways in which communities

80 in polar environments respond to changing environmental conditions; for example, the effects

81 of the seasonal polar sunlight cycle on Antarctic marine and marine-derived lake communities

82 [10, 11], and the roles that Arctic bacteria play in melting permafrost acting as a CO_2 source

83 and atmospheric CH_4 sink [12, 13].

84 *Arthrobacter* (Actinobacteria; Micrococcales; Micrococcaceae) are a globally distributed

85 genus of bacteria commonly found in soil, but also in a broad range of environments

86 including water, human skin, and sewage [14-17]. *Arthrobacter* are reported to play important

87 roles in global biogeochemical cycles and decontamination of polluted environments [16, 18].

88 Their responses to temperature, desiccation, ionizing radiation, oxygen radicals, and a range

89 of chemicals have been described [19-21]. Their growth in the laboratory is characterized by

90 nutritional versatility that translates to an ability to grow aerobically in media utilizing a wide

91 range of carbon and nitrogen sources [15]. Some strains closely related to the type species *A.*

92 *globiformis* were isolated from a Lapland glacier region [22], and the identification of

93 psychrophilic species has led to the characterization of a number of *Arthrobacter* enzymes for

94 their biotechnological potential (e.g., REF. [23]). Cold environments that *Arthrobacter* have

95 been isolated from include permafrost and glaciers [24-26].

96 Due to the large scale of the Earth's cold biosphere and its relevance to global

97 biogeochemical cycles, and the biotechnological potential of psychrophiles and their

98 products, numerous studies have been performed to attempt to define the critical traits of

99 psychrophiles (discussed in REFs [9, 23, 27-33]). In the current study, we sequenced the

100 genomes of 13 *Arthrobacter* strains isolated from the Tibetan Plateau and utilized the

101 existence of more than 100 *Arthrobacter* genomes to assess traits that may explain the

102 presence of the genus in naturally cold environments. After identifying a clade characteristic

103 of polar and alpine environments and determining that representatives had a superior ability to

104 grow at low temperature, we used available metagenome data to assess the environmental

105 relevance of the findings. What we learned illustrated the complexities involved in attempting

106 to extrapolate from genomic and physiological data to an environmental setting. It also

107 revealed possible avenues for utilizing *Arthrobacter* as biomarkers of environmental

108   warming.

109

110   **Results**

111   **Phylogenomics**

112   To increase the number of *Arthrobacter* genomes from polar and alpine (PA) environments,

113   isolates from the Tibetan Plateau were sequenced resulting in high-quality genomes (see

114   Materials and methods) for three glacier and 10 lake isolates (Table S1). The phylogenomic

115   relationships of a total of 210 non-redundant high-quality Micrococcaceae genomes were

116   analyzed by constructing ML and Bayesian trees (Fig. S1). The two trees were congruent and

117   most tree-nodes (194/208) were supported by high bootstrap values (>70 %) (Fig. S1). The

118   *Arthrobacter* lineage formed a cluster with 106 representatives that was clearly separated

119   from other Micrococcaceae genera (Fig. S1). The *Arthrobacter* genomes represented PA

120   isolates (total 31, including the 13 new genomes), with the remainder from a broad range of

121   non-polar or alpine (NPA) environments (Table S1).

122   The *Arthrobacter* lineage separated into three clades that branched from the root of the

123   *Arthrobacter* tree (Fig. 1). The 31 PA strains were distributed across the tree, although 11 PA

124   strains formed a cluster with three NPA strains in the central clade (Fig. 1a). Within the

125   central clade, 10 strains grouped together (Fig. 1b, blue font) with an *F* measure of 0.95,

126   defining them as an operationally monophyletic lineage [34]. This was supported by NMDS

127   analysis of amino acid composition (Fig. 1c) and both ANI and AAI distributions (Fig. 1d, e).

128   The grouping of 10 strains (Fig. 1b, blue font) consisted of nine PA strains, plus *A.*

129   *psychrolactophilus* B7 which was isolated from Pennsylvania soil following snow melt; the

130   strain was isolated as a source of cold-active enzymes and was capable of growth at 0 °C [35].

131   The monophyletic lineage of 10 strains was defined as Group C (Fig. 1b, blue font), with all

132   other PA strains as Group B (Fig. 1b, green font), and all strains from NPA environments as

133   Group A (Fig. 1b, red font).

134

135   **Low temperature growth capacity of Group C**

136   To evaluate the growth temperature response of Group A, B and C, three strains from each

137   group were grown at 25, 5 and -1 °C and growth monitored (OD_{600}) over time (Fig. 2). The

138   three Group C strains exhibited a markedly enhanced rate of growth at -1 °C (Fig. 2c)

139   particularly compared to Group A strains, and had a reduced rate of growth at 25 °C compared

140   to some of the Group A and B strains (Fig. 2a, b).

141

142   **Genomic characteristics**

143   The size of the 106 *Arthrobacter* genomes ranged from 3.24 – 5.89 Mbp (Table S1). Between

144   Group A and C, no significant differences occurred in genome size, 16S rRNA and tRNA

145   gene copy number, or coding density (Fig. S2). However, a significant difference was

146   observed in amino acid composition and G+C content (Fig. 3, Fig. S2, Table S1). In Group C,

147   the content of N, K, M, I, S, T, F, Q, W and H was significantly higher, while A, E, G, P, D

148   and R was significantly lower (one-way ANOVA, $p < 0.05$, Fig. S3). The correlation between

149   amino acid composition and G+C content was significantly negative for N, I, M, S, F, K, Q

150   and G+C content (R^2 ranged from 0.46 – 0.81; $p < 0.01$), and significantly positive for W, G,

151   D, P, R and A (R^2 ranged from 0.39 – 0.77; $p < 0.01$) (Fig. S4).

152   To evaluate the potential structural relevance of the amino acid compositional

153   differences, temperature-dependent protein stability predictions were made using SCooP,

154   which predicts stability assuming proteins are monomeric and follow a two-state folding

155   transition [36]. From 180 proteins targeted for evaluation (mostly single-copy genes; see

156   Materials and Methods), 86 produced robust stability curves (Fig. 3, Fig. S5, Table S2).

157   Stability was calculated at -1 °C to match the growth temperature at which Group C showed a

158   marked difference in growth ability (Fig. 2). The 86 Group C proteins had significantly higher

159   ΔG values (Group A, -4.27; Group B, -4.29; Group C, -3.60; $p < 0.01$), with 32 proteins being

160   responsible for the reduced predicted stability (Fig. S5, Table S2). These 32 proteins

161   contained a particularly high representation of the amino acids that were most over-

162   represented in Group C (i.e., N, K and R; Fig. S3). The 32 proteins represented 12 functional

163   categories, primarily metabolism (28 proteins; 9 categories), with four involved in respiration,

164 stress response, or cell division and cell cycle (Table S2). The marked amino acid

165 compositional differences, broad representation of functional categories, and high proportion

166 of proteins with predicted decreases in stability (~ 1/3rd of those tested), demonstrates that

167 Group C *Arthrobacter* possess broad genomic differences to Group A *Arthrobacter*. If the

168 decreases in predicted protein stability translate to an increased capacity to perform catalysis

169 at low temperature, this may contribute to the higher growth rates of Group C at -1°C (Fig. 2).

170   To further explore the influence of amino acid composition on functional potential, Bray-

171 Curtis distances of genome-wide amino acid composition were evaluated for proteins

172 representing 26 functional categories (Fig. 3). The greatest distance was for the category

173 'phages, prophages, transposable elements, plasmids', consistent with previous studies

174 associating transposable elements with cold adapted microorganisms [37-39]. The functional

175 potential of Group C was also compared to Group A using enrichment analysis [40]

176 performed on proteins representing the 26 functional categories. Group C was over-

177 represented in sulfur metabolism; cofactors, vitamins, prosthetic groups, pigments; protein

178 metabolism; stress response; cell division and cell cycle (Fig. 3, Fig. S6, Table S5), whereas

179 Group A, was over-represented in metabolism of aromatic compounds; nitrogen metabolism;

180 amino acids and derivatives; regulation and cell signaling (Fig. S6, Table S5). The category

181 'sulfur metabolism' also exhibited signatures of amino acid bias (Bray-Curtis distance; Fig.

182 3), suggesting selection for this functional capacity occurred at the levels of both gene

183 complement and amino acid composition.

184   Functional assessments were extended to attempt to identify specific genes unique to

185 Group C. A number of genes involved in the synthesis of amino acids, vitamins, and

186 nucleosides were present in all Group C genomes (Fig. 4). The specific genes also tended to

187 be present in the other four, non-Group C members (Group A and Group B) of the central

188 clade (Fig. 4), but had low representation in other *Arthrobacter* genomes (Fig. 4). The most

189 marked feature was a complete mycothiol (MSH) biosynthesis pathway that was present in all

190 Group C genomes (Fig. 4); MSH is a redox-active thiol, functionally analogous to glutathione

191 (which is typically absent from Actinobacteria), that maintains intracellular redox balance and

192 can therefore protect against oxidative damage [41]. Further, MSH potentially serves as a

193 stable reservoir of carbon and sulfur in bacteria [42]. The ability to respond effectively to

194 oxidative damage may be an important trait of microorganisms from cold environments,

195 particularly for facilitating growth at low temperature limits [43-47]. The MSH pathway was

196 also present in the other four, non-Group C members of the central clade, plus one other

197 Group B member (Fig. 4). Therefore, the MSH pathway plus the individual genes involved in

198 the synthesis of amino acids, vitamins, and nucleosides, are characteristic, but not unique to

199 Group C *Arthrobacter*. If MSH or the other individual genes fulfill roles in facilitating growth

200 at low temperature, the genes may be under stronger positive selection in Group C, while also

201 being retained within Group A and Group B populations (pan genome), but at a significantly

202 lower level.

203   A total of 48 Group C gene families had significantly higher, and 66 had significantly

204 lower average gene copy number compared to Group A, including four which were absent in

205 Group C genomes (Table S6, $p < 0.05$). The absence of two specific genes is noteworthy:

206 adenosylhomocysteinase, which hydrolyzes S-adenosyl-L-homocysteine (a product of methyl

207 transfer reactions that involve S-adenosyl-L-methionine) to homocysteine and adenosine [48];

208 and formate--tetrahydrofolate (THF) ligase, which catalyzes the initial recruitment of single

209 carbon units for THF-mediated one-carbon metabolism [49]. The absence of both genes

210 would be expected to disrupt the synthesis of methionine from homocysteine, and instead

211 favor the alternative pathway of synthesizing methionine from cysteine; the latter pathway

212 may be connected to MSH metabolism, in that accumulation of cysteine (a precursor of MSH

213 synthesis) is toxic to cells [42], so surplus cysteine not required for MSH synthesis, or

214 resulting from MSH degradation, could be directed to methionine synthesis (Table S6).

215   Some of the gene families had particularly high copy numbers per genome (~30 in Group

216 A) with large reductions (~6) in Group C (Table S6); this trend was observed for 3-oxoacyl-

217 [acyl-carrier protein] reductase (FabG), glycerate kinase (GlxK), and alcohol dehydrogenase

218 (Adh). For FabG, this likely reflects a reduced capacity of Group C to catalyze the formation

219 of long-chain fatty acids (Table S6). GlxK is an important catabolic enzyme, in that diverse

220 substrates are degraded to glycerate, and GlxK links these degradation pathways to central

221 carbon metabolism [42]. Decreased copy numbers of Adh likely indicates decreased

222 capacities to utilize alcohols. Thus, decreases in GlxK and Adh might reflect decreased

223 substrate preferences by these *Arthrobacter* strains. It was noteworthy that the copy number

224 of cold shock protein (*csp*) genes was lower in Group C. While *csp* genes are sometimes

225 equated with an ability to grow in the cold or survive cold shock, these nucleic acid binding

226 proteins can perform diverse roles in cellular function (reviewed in REF. [33]); the findings

227 here reinforce the notion that *csp* and other 'stress' genes are not good molecular markers for

228 identifying psychrophiles [4, 32].

229

230 **Ecology of Group C *Arthrobacter***

231 We hypothesized that if the laboratory-generated growth data (Fig. 2) and genomic traits (Fig.

232 3) translated to competitiveness in cold environments, Group C *Arthrobacter* would be over-

233 represented in metagenome data from PA vs NPA environments. The relative abundance of

234 *Arthrobacter* in environmental samples (publicly available metagenome data) tends to be low,

235 with no metagenome assembled genomes (MAGs) present in the ~8,000 that were constructed

236 from ~1,500 metagenomes [50], and a total of 12 (> 90% completeness) present in 76,831

237 IMG MAGs (December 2019). To facilitate metagenome analyses, group-specific genes

238 (Table S3) were examined in 797 metagenomes representing PA, temperate and tropical

239 environments (Table S4), with representation shown relative to the *Arthrobacter* pan genome

240 (Fig. 5a) (see **Materials and methods** for description of analytical approach).

241   Group C-specific genes were more highly represented in 11 permafrost metagenomes

242 (Fig. 5b). All of the 11 metagenomes came from a single site: Axel Heiberg Island, Nunavut,

243 Canada [51]. The Axel Heiberg Island study reported 76 metagenomes derived from 1 m

244 cores that were used during a controlled thawing experiment [51]. Most Group C-specific

245 genes were enriched in the 65 cm depth active-layer (7 metagenomes), with one from the 35

246 cm active-layer and three from the 80 cm permafrost-layer (Table S7). A total of 94% of the

247 variability that exists in hits to Group C for the 797 metagenomes (Fig 5b) was traced to pre-

248   existing variability in hits to the *Arthrobacter* pan genome, and when this covariance was

249   removed by ANCOVA analysis, a statistically significant difference in the y-intercepts for the

250   regression lines ($p < 0.0001$) remained; this confirms the over-representation of Group C-

251   specific genes in the 11 metagenomes compared to the remaining 786 metagenomes.

252   To assess whether Group C *Arthrobacter* were generally enriched in permafrost regions,

253   all other publicly available permafrost metagenomes (361 datasets) were analyzed (Fig. 5c). A

254   number appeared somewhat enriched in Group C-specific genes (e.g., a metagenome from

255   Stordalen Mire near Abisko, Sweden; marked by an arrow in Fig. 5c), but as the slopes of the

256   two regression lines were not parallel, it was not valid to compare the y-intercepts, and hence

257   the significance of the difference between them could not be evaluated [52].

258   For Group B-specific genes, no obvious trends separated the PA from NPA metagenomes

259   (Fig. 5d). However, the distribution of Group A-specific genes clustered according to climate,

260   with PA metagenomes showing lower Group A content ($y=0.0161x + 0.1398$, $R^2 = 0.96773$)

261   compared to temperate and tropical metagenomes ($y=0.0244x - 1.8835$, $R^2 = 0.96432$) (Fig.

262   5e). The 11 Axel Heiberg Island metagenomes had a statistically significant under-

263   representation of Group A-specific genes compared to all other metagenomes (ANCOVA, $p <$

264   0.0001). This pattern indicates there is selection against Group A *Arthrobacter* in PA

265   environments, and/or selection for Group A in NPA environments.

266   To attempt to define variables that may explain the niche adaptation of Group C in the

267   Axel Heiberg Island permafrost, available abiotic and biotic data were used from the

268   permafrost study [12, 13]. A range of physicochemical data were available for each of the four

269   depths (5, 35, 65 and 80 cm) but as the timing of sampling for physicochemical data (0, 4, 6,

270   8, 11 and 12 weeks) did not align with the timing of sampling for the metagenomes (0, 0.25,

271   6, 12 and 18 months), the physicochemical data were ultimately not useful for interpreting

272   Group C distribution. Depth, treatment group and sample core did not explain the variation in

273   species composition across the sites, and although incubation time had some explanatory

274   power for the distribution of the entire permafrost study dataset (data not shown), the

275   metagenomes that were enriched in Group C *Arthrobacter* were widely distributed and did

276   not cluster together, suggesting the importance of specific microniches in enrichment of these

277   species. Assessment of functional potential of the microbial communities in each of the 76

278   metagenomes using presence/absence of KO groups also did not identify any significant

279   functional differences (PERMANOVA, $p > 0.05$; data not shown).

280   In contrast, strong taxonomic associations were identified with many members of the

281   microbial community. Analyses were performed to assess taxa that correlated with Group A

282   and Group C, just Group A, and just Group C (Table S8). A total of 107 OTUs positively

283   correlated with Group C *Arthrobacter*, and 63 negatively correlated (Table S8). Of the

284   107 OTUs positively correlated to Group C, 72 were also positively correlated to Group A

285   above the threshold of 0.5 (the remainder were positively correlated with values 0.359 -

286   0.499), and no OTUs were positively correlated to Group A that did not also correlate to

287   Group C. The positively correlating OTUs were dominated by both spore- and non-spore-

288   forming members of Actinobacteria and Firmicutes, as well as members of Proteobacteria; the

289   majority of these OTUs were isolated from soil. Negatively correlating bacterial OTUs

290   mainly belonged to marine or lacustrine members of Bacteroidetes, Cyanobacteria, and

291   Proteobacteria, as well as certain eukaryotes (fungi, plants, marine annelid worm).

292

293   **Discussion**

294   Numerous studies have been performed to attempt to define the critical traits of a

295   psychrophile (discussed in REF. [32, 33, 53]). The current study explored genomic

296   characteristics of a lineage with less than 3.5 % difference in 16S rRNA gene identity. The

297   analyses revealed that the *Arthrobacter* lineage contains a clade with members (Group C)

298   possessing a clear capacity to grow faster than their relatives (Group A and B) under

299   laboratory growth conditions at -1 °C (Fig. 2). Genomic characteristics that potentially

300   explain the physiological capacity include an amino acid composition that is predicted to

301   reduce the stability of a large proportion of proteins thereby enhancing enzyme activity at low

302   temperature [52]. Group C genomes are enriched in sulfur metabolism genes, and sulfur is

303   required for the cysteine component of mycothiol. The synthesis of mycothiol may potentially

protect Group C *Arthrobacter* against oxidative damage that may otherwise accumulate as cell division decreases towards the lower temperature limit of growth [43]. Group C also exhibits a relatively high proportion of mobile elements, which is a trait shared with some other cold-adapted microorganisms [37-39]. Collectively, the physiological and genomic traits appear compelling for denoting Group C, a cold-adapted clade of *Arthrobacter*.

However, from assessing available metagenome data, we infer that these traits do not translate to a generally enhanced ability to compete in low temperature environments. Other than the 11 specific Axel Heiberg Island permafrost metagenomes, Group C *Arthrobacter* were not highly represented in the other metagenomes from cold environments, including 144 from Arctic peat soil, 22 associated with glaciers, 42 from polar deserts, and importantly, 365 from other permafrost environments. Even at the Axel Heiberg Island site, Group C-specific genes were not highly abundant at 5 and 20 cm depths. Instead, the pattern of abundance of Group C appears to derive not just from low temperature, but from niche-specific conditions.

Attempting to identify specific niche conditions is not trivial. For the Axel Heiberg Island study, the permafrost microbial community was reported to be dominated by Actinobacteria and Proteobacteria, with significant increases at depth for Firmicutes and Actinobacteria and significant decreases for Acidobacteria, Proteobacteria and Verrucomicrobia [13]. However, despite these taxonomic differences, we did not identify significant predicted functional differences by depth. When we turned to specifically correlating the abundance of Group C to OTUs from the metagenome data, a large number of OTUs with positive or negative correlations were identified (Table S8). At a broad level, the environmental data of the positively correlating taxa are consistent with Group C associating with other soil bacteria. While this provides scope for investigating specific taxa that may help shape the niche that Group C occupy, determining which taxa are important and the nature of their interactions will require a dedicated effort.

For the negatively correlating cohort, they tend to represent isolates from non-soil environments (Table S8) and may therefore represent non-indigenous microorganisms that have been introduced. The permafrost samples were obtained from an "upland polygonal

332   terrain in proximity to the McGill Arctic Research Station at Expedition Fjord (79°24'57"N,

333   90°45'46"W)" [12]. The prevalence of negatively correlating OTUs matching to

334   Proteobacteria isolated from sea water may reflect aeolian carriage from Expedition Fjord,

335   which is located ~8 km from the Research Station. As the samples were obtained for a

336   simulated permafrost-thaw experiment [51], the negatively correlating OTUs may also reflect

337   environmental disturbance.

338

339   **Conclusions**

340   Our study commenced with the analysis of genome sequences of new Group C *Arthrobacter*

341   isolated from the Tibetan Plateau and progressed through to a rationalization of Group C

342   abundance in global metagenomes. Group C was clearly distinguished from Group A-

343   *Arthrobacter* by possessing genomic signatures consistent with its representation in PA

344   environments and an ability to grow faster when cultivated at -1ºC. Assessment of available

345   metagenome data point to the Group C traits as being more relevant to cold niches rather than

346   competitiveness across global permafrost or cold soil environments. The challenge in being

347   able to define the specific niche parameters enabling Group C *Arthrobacter* to be relatively

348   competitive illustrates the inherent difficulties associated with defining 'cause and effect' for

349   explaining 'why' microorganisms reside in the environments in which they are found; that is,

350   the characteristics of the ecological niches that define microbiome structure [1]. Without

351   knowing the specific effectors, the ability to understand and predict responses to

352   environmental changes, is greatly compromised [4, 6, 54, 55]. Establishing long-term data

353   records that include comprehensive metadata associated with monitoring sites, including

354   metadata for each biological sample, will be essential for learning how to link environmental

355   parameters to microbial processes. In a study of sulfate reduction in Arctic marine sediments,

356   growth yield was reasoned to be the most relevant factor for determining competitiveness of

357   sulfate-reducing bacteria in permanently cold marine sediments [53]. These findings illustrate

358   that for cold environments, linking genomic and metagenomic data to measurements of

359   metabolic rates, growth rates and growth yields will undoubtedly help to clarify how specific

360   microbial processes and associated taxa are influenced by environmental temperature.

361   While the characteristics that define the Group C niche are still to be defined, at sites where

362   Group C *Arthrobacter* are relatively abundant they may have value as a biomarker for

363   monitoring the stability of those locations. Moreover, Group A *Arthrobacter* may serve as a

364   more broadly useful biomarker of soil microbial communities. Group A exhibited high

365   relative abundance across NPA metagenomes and relatively low abundance across PA

366   metagenomes. As the data indicate environmental temperature exerts a broad, strong influence

367   on Group A *Arthrobacter*, we predict that environmental warming will generally increase the

368   relative abundance of Group A. Similar influences of environmental temperature have been

369   described for the marine SAR11 clade, including the predicted displacement of polar

370   specialists by phylotypes from warmer latitudes [56]. Depending on how strongly the

371   environmental factors other than temperature select for Group C in permafrost, the apparent

372   broad influence of temperature on Group A suggests it will displace Group C from the niches

373   in which it is currently relatively competitive.

374

375   **Methods**

376   *Arthrobacter* **isolation and genome sequencing**

377   Isolation of *Arthrobacter* strains from glaciers and lakes on the Tibetan Plateau was

378   performed as previously described [25, 26] (Table S1). All strains were deposited in the China

379   General Microbiological Culture Collection Center (CGMCC) with accession numbers:

380   CGMCC 1.16187-1.16198, 1.16223 and 1.16312. Genomic DNA for 16 isolates was

381   extracted using a TIANamp Bacteria DNA Kit (Tiangen, Beijing) following manufacturer's

382   instructions. Paired-end libraries with an insert size of 500 bp were constructed and

383   sequenced using Illumina Hiseq 2000 platform. Prior to *de novo* sequence assembly, low-

384   quality reads were filtered out using Fastp with default options [57]. Filtered sequencing reads

385   were subjected to assembly using SPAdes v3.11.1 with default options [58]. The assembled

386   genome sequences were deposited in DDBJ/ENA/GenBank under the BioProject

387   PRJNA421662.

388

Growth temperature response

Arthrobacter strains were grown in R2A broth at 25, 5 and −1°C for up to 10 days with

optical density measured at 600 nm (OD_{600}) using a Microplate Reader (MD, SpectraMax

M5). For cultivation at −1°C, flasks were placed in ice produced by an ice maker (TKKY,

FM40) with flasks placed in a ~ 4°C refrigerator and ice replaced every 12 h. Growth at 5 and

25°C was performed using a constant-temperature incubator as described previously [59]. All

cultures were grown statically, with flasks swirled to resuspend biomass prior to recording

OD_{600}. Strains used for growth temperature profiles were: Group A: *A. luteolus*, *A.

globiformis* and *A. subterraneus*; Group B: *Arthrobacter* sp. 4R501, *Arthrobacter* sp. 9E14

and *Arthrobacter* sp. 08Y14; Group C: *A. alpinus*, *Arthrobacter* sp. A3 and *Arthrobacter* sp.

N199823.

Preparation of *Arthrobacter* genomes for analysis

As the taxonomic assignment of the genus *Arthrobacter* is not consistent, in August 2018 all

genome sequences with the taxonomy identifier 'Arthrobacter' or 'Micrococcaceae' were

retrieved from GenBank, providing a total of 427 genomes including the 16 new Tibetan

Plateau genomes. The completeness of each genome was calculated using CheckM v1.0.7

with default options [60]. Genomes composed of > 300 contigs, with an N50 of < 20 kb,

completeness of < 95%, and contamination > 5% were removed. Genomes were dereplicated

to remove genomes with an average amino acid identity (AAI) ≥ 99.5%. AAI values were

calculated using CompareM with default options (https://github.com/dparks1134/CompareM).

Average nucleotide identity (ANI) was calculated using the ANI calculator (http://enve-

omics.ce.gatech.edu/ani/). A total of 210 genomes met quality requirements, which included

13 of the 16 new Tibetan Plateau genomes (Table S1). Gene families were clustered using

FastOrtho software (--pv_cutoff 1-e5 --pi_cutoff 70 --pmatch_cutoff 70)

(http://enews.patricbrc.org/fastortho/) with the cutoff values set according to Parks *et al* [34].

A gene family matrix was produced using custom PERL scripts, and non-functional based

416　group specific genes were calculated based on this matrix. The annotation of genes was

417　standardized by annotating all genomes using RAST (Rapid Annotation using Subsystem

418　Technology) [61] and PROKKA [62].

419

420　**Phylogenetic and genomic analyses**

421　For phylogenomic clustering, *Cellulomonas carbonis* T26 and *C. fimi* ATCC 484 were chosen

422　as the outgroup as they are close relatives of Micrococcaceae [63], and species that are

423　closely related to the in-group are more suitable for phylogenetic reconstruction than distantly

424　related species [64]. A Maximum Likelihood (ML) phylogenomic tree was constructed using

425　PhyloPhlAn2 with default options [65]. A Bayesian tree was constructed using MPI Mrbayes

426　v3.2 (prset aamodelpr = mixed, mcmc nchains = 16, ngen = 300,000, and leaving other

427　parameter values as default) [66]. The F measure (harmonic mean of precision) provides a

428　metric for determining if taxa are operationally monophyletic (F measure $\geq$ 0.95) [34]; and

429　was calculated as F = 2*((precision * recall) / (precision + recall)). Genome wide amino acid

430　composition was calculated using CompareM with the function aa_usage. The stability curves

431　of proteins were predicted by SCooP [36] using the PDB (Protein Data Bank) files modeled

432　by SWISS-MODEL [67]. The stability equations of the same protein from different hosts

433　were visualized and smoothed using ggplot2 v3.2.1 [68]. The stability curves were analyses

434　for 180 single-copy genes that were shared by most genomes; a small number of genomes had

435　multiple copies of genes, and up to three genomes were permitted to have the absence of the

436　gene in order to account for the use of unclosed genomes (99 of the 106). After retrieval from

437　SWISS-MODEL of all possible PDB files matching the candidate genes, a total of 17,339

438　stability equations were generated (Table S2). Ordination and statistical analyses, including

439　three-dimensional nonmetric multidimensional scaling (3D-NMDS) and gene enrichment

440　analyses were performed with R v3.3.3 and Origin v9.5. For comparisons between Group A,

441　B and C *Arthrobacter*, group-specific genes or functions were defined as being present in

442　95% of the target group (e.g., Group C) genomes and absent in 95% of each of the other

443　group(s) (e.g., Group A) genomes. Group-specific genes were identified (Group A, 74

444   genomes, 16,149 specific genes; Group B, 22 genomes, 4,675 specific genes; Group C, 10

445   genomes, 969 specific genes; Table S3) and normalized to account for the different number of

446   genomes used for each group. To account for differences in gene content between strains,

447   comparisons were calculated relative to the total *Arthrobacter* gene complement from all 106

448   *Arthrobacter* genomes (referred to as the *Arthrobacter* pan genome). Gene copy number was

449   calculated as the average number of the gene for each genome in a group (e.g., Group C),

450   with gene loss or gain calculated from the average copy number for the groups (e.g., Group C

451   vs Group A). To assess the bias of amino acid composition of different functional classes of

452   proteins, genes were assigned to functional categories (assigned by RAST) and total amino

453   acid composition for all proteins from the functional category were compared between groups

454   (e.g., Group C vs Group A). Similarity was measured by Bray-Curtis distance with larger

455   Bray-Curtis distances denoting stronger bias, possibly indicative of selection pressure [40].

456   The functional potential of groups were also compared using enrichment analysis [40].

457   Briefly, the presence or absence of KEGG Ortholog (KO) groups in genomes and

458   metagenomes (see **Collection and analysis of metagenomes**) was assessed [40]and non-

459   parametric one-way ANOVA was used to identify differentially abundant categories using R

460   [69].

461

462   **Collection and analysis of metagenomes**

463   Assembled metagenomes were downloaded from IMG (https://genome.jgi.doe.gov/portal/).

464   Classification into polar and alpine (PA) or non-polar or alpine (NPA) environments were

465   made using metadata associated with metagenomes, supplemented by Köppen-Geiger climate

466   classifications (to define temperate and tropical regions) using ArcGIS location data [70].

467   Analyses were initially performed using 797 metagenomes from environments with mean

468   ambient annual temperature (MAT) ranging from -24 °C to 28 °C , representing PA (n = 286,

469   gene count = 244 million), temperate (n = 294, gene count = 291 million), and tropical (n =

470   217, gene count = 981 million) zones (Table S4). Subsequently, all additional (361) available

471   (May 2020) assembled permafrost metagenomes were analyzed. Analyses assessed the

472 relative abundance of each *Arthrobacter* group (A, B and C) using group-specific genes (see

473 **Phylogenetic and genomic analyses**) by performing a local alignment search against the

474 metagenomes using DIAMOND v0.9.24 with the arguments --outfmt 6, --query-cover 70, --id

475 70, --evalue 1e-5, and leaving others as default [71]. One-way ANCOVA was used to assess

476 statistical differences between regression lines for groups of metagenomes [72] using the data

477 import webform for k=2 at http://vassarstats.net/vsancova.html. Correlation analyses were

478 performed between *Arthrobacter* groups and other members of the microbial community from

479 76 metagenomes derived from a simulated permafrost thaw experiment [13]. Operational

480 taxonomic units (OTUs) were assigned from IMG phylodist matches which are based on the

481 top taxon in the IMG isolate database; < 4% of OTUs had < 35% identity. The raw abundance

482 of all OTUs was determined, with 956 meeting the criteria of average abundance ≥2, and

483 detection in at least 56 of the 76 metagenomes. The 956 OTUs were used to construct a

484 correlation matrix using SparCC [73] implemented in python3 with default parameters (20

485 iterations). One hundred simulated datasets were created by random shuffling of the original

486 input with replacement, and their correlation matrices constructed in the same way. The

487 simulated datasets were used to calculate the one- and two-sided pseudo *p*-values. The

488 selected threshold for strong correlations was > | 0.5 |. To assess if depth, incubation time,

489 treatment group, sample core, or physicochemical data explained the variation in species

490 composition across permafrost sites, a generalized linear latent variable model was employed

491 as implemented in the R package gllvm [74]. All custom scripts are available at

492 (https://github.com/environmental-genomes/Arthro).

493

499

Availability of data and materials

The 16 newly isolated *Arthrobacter* strains were deposited in the China General Microbiological Culture Collection Center (CGMCC) with accession numbers: CGMCC 1.16187-1.16198, 1.16223 and 1.16312. The assembled genome sequences for newly isolated *Arthrobacter* strains were deposited in DDBJ/ENA/GenBank under the BioProject PRJNA421662. All custom scripts are available at (https://github.com/environmental-genomes/Arthro).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

528   The authors declare that they have no conflict of interest.

529   **References**

530   1. Berg G, Rybakova D, Fischer D, Cernava T, Vergès M-CC, Charles T, et al. Microbiome
531   definition re-visited: old concepts and new challenges. Microbiome. 2020;8.

532   2. Colman DR, Poudel S, Stamps BW, Boyd ES, Spear JR. The deep, hot biosphere: Twenty-
533   five years of retrospection. Proc Natl Acad Sci USA. 2017;114:6895-903.

534   3. Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, Backstrom D, Juzokaite L, Vancaester E,
535   et al. Asgard archaea illuminate the origin of eukaryotic cellular complexity. Nature.
536   2017;541:353-8.

537   4. Cavicchioli R, Ripple WJ, Timmis KN, Azam F, Bakken LR, Baylis M, et al. Scientists'
538   warning to humanity: microorganisms and climate change. Nat Rev Microbiol.
539   2019;17:569-86.

540   5. Timmis K, Cavicchioli R, Garcia JL, Nogales B, Chavarria M, Stein L, et al. The urgent
541   need for microbiology literacy in society. Environ Microbiol. 2019;21:1513-28.

542   6. Edwards A, Cameron KA, Cook JM, Debbonaire AR, Furness E, Hay MC, et al. Microbial
543   genomics amidst the Arctic crisis. Microb Genom. 2020;6.

544   7. Chen LX, Anantharaman K, Shaiber A, Eren AM, Banfield JF. Accurate and complete
545   genomes from metagenomes. Genome Res. 2020;30:315-33.

546   8. Nkongolo KK, Narendrula-Kotha R. Advances in monitoring soil microbial community
547   dynamic and function. J Appl Genet. 2020;61:249-63.

548   9. Cavicchioli R. Microbial ecology of Antarctic aquatic systems. Nat Rev Microbiol.
549   2015;13:691-706.

550   10. Grzymski JJ, Riesenfeld CS, Williams TJ, Dussaq AM, Ducklow H, Erickson M, et al. A
551   metagenomic assessment of winter and summer bacterioplankton from Antarctica
552   Peninsula coastal surface waters. ISME J. 2012;6:1901-15.

553   11. Panwar P, Allen MA, Williams TJ, Hancock AM, Brazendale S, Bevington J, et al. Influence
554   of the polar light cycle on seasonal dynamics of an Antarctic lake microbial community.
555   Microbiome. 2020;8:116.

556 12. Lau MCY, Stackhouse BT, Layton AC, Chauhan A, Vishnivetskaya TA, Chourey K, et al.
557   An active atmospheric methane sink in high Arctic mineral cryosols. ISME J. 2015;9:1880-
558   91.

559 13. Stackhouse BT, Vishnivetskaya TA, Layton A, Chauhan A, Pfiffner S, Mykytczuk NC, et
560   al. Effects of simulated spring thaw of permafrost from mineral cryosol on CO_2 emissions
561   and atmospheric CH_4 uptake. J Geophys Res Biogeo. 2015;120:1764-84.

562 14. Conn H, Dimmick I. Soil bacteria similar in morphology to *Mycobacterium* and
563   *Corynebacterium*. J Bacteriol. 1947;54:291.

564 15. Cacciari I, Lippi D. *Arthrobacters*: Successful arid soil bacteria: A review. Arid Soil Res
565   Rehab. 1987;1:1-30.

566 16. Niewerth H, Schuldes J, Parschat K, Kiefer P, Vorholt JA, Daniel R, et al. Complete genome
567   sequence and metabolic potential of the quinaldine-degrading bacterium *Arthrobacter* sp.
568   Rue61a. BMC Genomics. 2012;13:534.

569 17. Busse HJ, Wieser M. *Arthrobacter*. In: Trujillo ME, Dedysh S, DeVos P, Hedlund B,
570   Kämpfer P, Rainey FA, et al. (eds). Bergey's Manual of Systematics of Archaea and Bacteria.
571   (Wiley, Hoboken, NJ, 2018) pp 1-43.

572 18. Unell M, Nordin K, Jernberg C, Stenstrom J, Jansson JK. Degradation of mixtures of
573   phenolic compounds by *Arthrobacter chlorophenolicus* A6. Biodegradation. 2008;19:495-
574   505.

575 19. Boylen CW. Survival of *Arthrobacter crystallopoietes* during prolonged periods of extreme
576   desiccation. J Bacteriol. 1973;113:33-7.

577 20. Hayashi T, Mukouyama M, Sakano K, Tani Y. Degradation of a sodium acrylate oligomer
578   by an *Arthrobacter* sp. Appl Environ Microbiol. 1993;59:1555-9.

579 21. Dsouza M, Taylor MW, Turner SJ, Aislabie J. Genomic and phenotypic insights into the
580   ecology of *Arthrobacter* from Antarctic soils. BMC Genomics. 2015;16:36.

581 22. Gounot A. Effects of temperature on the growth of psychrophilic bacteria from glaciers.
582   Can J Microbiol. 1976;22:839-46.

583 23. Feller G. Psychrophilic enzymes: from folding to function and biotechnology. Scientifica

584 (Cairo). 2013;2013:512840.

585 24. Chen X-M, Jiang Y, Li Y-T, Zhang H-H, Li J, Chen X, et al. Regulation of expression of
586   trehalose-6-phosphate synthase during cold shock in *Arthrobacter* strain A3. Extremophiles.
587   2011;15:499-508.

588 25. Shen L, Liu YQ, Wang N, Jiao NZ, Xu BQ, Liu XB. Variation with depth of the abundance,
589   diversity and pigmentation of culturable bacteria in a deep ice core from the Yuzhufeng
590   Glacier, Tibetan Plateau. Extremophiles. 2018;22:29-38.

591 26. Liu YQ, Priscu JC, Yao TD, Vick-Majors TJ, Michaud AB, Sheng L. Culturable bacteria
592   isolated from seven high-altitude ice cores on the Tibetan Plateau. J Glaciol. 2019;65:29-
593   38.

594 27. Margesin R, Feller G. Biotechnological applications of psychrophiles. Environ Technol.
595   2010;31:835-44.

596 28. Cavicchioli R, Charlton T, Ertan H, Mohd Omar S, Siddiqui KS, Williams TJ.
597   Biotechnological uses of enzymes from psychrophiles. Microb Biotechnol. 2011;4:449-60.

598 29. Anesio AM, Laybourn-Parry J. Glaciers and ice sheets as a biome. Trends Ecol Evol.
599   2012;27:219-25.

600 30. Siddiqui KS, Williams TJ, Wilkins D, Yau S, Allen MA, Brown MV, et al. Psychrophiles.
601   Annu Rev Earth Planet Sci. 2013;41:87-115.

602 31. Boetius A, Anesio MA, Deming WJ, Mikucki AJ, Rapp ZJ. Microbial ecology of the
603   cryosphere: sea ice and glacial habitats. Nat Rev Microbiol. 2015;13:677–90.

604 32. Cavicchioli R. On the concept of a psychrophile. ISME J. 2016;10:793-5.

605 33. Collins T, Margesin R. Psychrophilic lifestyles: mechanisms of adaptation and
606   biotechnological tools. Appl Microbiol Biotechnol. 2019;103:2857-71.

607 34. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, et al. A
608   standardized bacterial taxonomy based on genome phylogeny substantially revises the tree
609   of life. Nat Biotechnol. 2018;36:996-1004.

610 35. Loveland J, Gutshall K, Kasmir J, Prema P, Brenchley JE. Characterization of
611   psychrotrophic microorganisms producing beta-galactosidase activities. Appl Environ

612   Microbiol. 1994;60:12-8.

613   36. Pucci F, Kwasigroch JM, Rooman M. SCooP: an accurate and fast predictor of protein

614   stability curves as a function of temperature. Bioinformatics. 2017;33:3415-22.

615   37. DeLong EF, Preston CM, Mincer T, Rich V, Hallam SJ, Frigaard N-U, et al. Community

616   genomics among stratified microbial assemblages in the ocean's interior. Science.

617   2006;311:496-503.

618   38. Lauro FM, Tran K, Vezzi A, Vitulo N, Valle G, Bartlett DH. Large-scale transposon

619   mutagenesis of *Photobacterium profundum* SS9 reveals new genetic loci important for

620   growth at low temperature and high pressure. J Bacteriol. 2008;190:1699-709.

621   39. Allen MA, Lauro FM, Williams TJ, Burg D, Siddiqui KS, De Francisci D, et al. The genome

622   sequence of the psychrophilic archaeon, *Methanococcoides burtonii*: the role of genome

623   evolution in cold adaptation. ISME J. 2009;3:1012-35.

624   40. Bai Y, Muller DB, Srinivas G, Garrido-Oter R, Potthoff E, Rott M, et al. Functional overlap

625   of the *Arabidopsis* leaf and root microbiota. Nature. 2015;528:364-9.

626   41. Newton GL, Buchmeier N, Fahey RC. Biosynthesis and functions of mycothiol, the unique

627   protective thiol of Actinobacteria. Microbiol Mol Biol Rev. 2008;72:471-94.

628   42. Bzymek KP, Newton GL, Ta P, Fahey RC. Mycothiol import by *Mycobacterium smegmatis*

629   and function as a resource for metabolic precursors and energy production. J Bacteriol.

630   2007;189:6796-805.

631   43. Methé BA, Nelson KE, Deming JW, Momen B, Melamud E, Zhang XJ, et al. The

632   psychrophilic lifestyle as revealed by the genome sequence of *Colwellia psychrerythraea*

633   34H through genomic and proteomic analyses. Proc Natl Acad Sci USA. 2005;102:10913-

634   8.

635   44. Williams TJ, Lauro FM, Ertan H, Burg DW, Poljak A, Raftery MJ, et al. Defining the

636   response of a microorganism to temperatures that span its complete growth temperature

637   range (-2 °C to 28 °C) using multiplex quantitative proteomics. Environ Microbiol.

638   2011;13:2186-203.

639   45. Dsouza M, Taylor MW, Turner SJ, Aislabie J. Genome-based comparative analyses of

640   Antarctic and temperate species of *Paenibacillus*. PLoS One. 2014;9:e108009.

641   46. Goordial J, Raymond-Bouchard I, Zolotarov Y, de Bethencourt L, Ronholm J, Shapiro N,

642   et al. Cold adaptive traits revealed by comparative genomic analysis of the

643   eurypsychrophile *Rhodococcus* sp. JG3 isolated from high elevation McMurdo Dry Valley

644   permafrost, Antarctica. FEMS Microbiol Ecol. 2016;92:fiv154.

645   47. Mackelprang R, Burkert A, Haw M, Mahendrarajah T, Conaway CH, Douglas TA, et al.

646   Microbial survival strategies in ancient permafrost: insights from metagenomics. ISME J.

647   2017;11:2305-18.

648   48. Palmer JL, Abeles RH. The mechanism of action of S-adenosylhomocysteinase. J Biol

649   Chem. 1979;254:1217-26.

650   49. Sah S, Aluri S, Rex K, Varshney U. One-carbon metabolic pathway rewiring in *Escherichia*

651   *coli* reveals an evolutionary advantage of 10-formyltetrahydrofolate synthetase (Fhs) in

652   survival under hypoxia. J Bacteriol. 2015;197:717-26.

653   50. Parks DH, Rinke C, Chuvochina M, Chaumeil PA, Woodcroft BJ, Evans PN, et al. Recovery

654   of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. Nat

655   Microbiol. 2017;2:1533-42.

656   51. Chauhan A, Layton AC, Vishnivetskaya TA, Williams D, Pfiffner SM, Rekepalli B, et al.

657   Metagenomes from thawing low-soil-organic-carbon mineral cryosols and permafrost of

658   the canadian high arctic. Genome Announc. 2014;2:e01217-14.

659   52. Siddiqui KS, Cavicchioli R. Cold-adapted enzymes. Annu Rev Biochem. 2006;75:403-33.

660   53. Scholze C, Jørgensen BB, Røy H. Psychrophilic properties of sulfate-reducing bacteria in

661   Arctic marine sediments. Limnol Oceanogr. 2020:10.1002/lno.11586.

662   54. Webster NS, Wagner M, Negri AP. Microbial conservation in the Anthropocene. Environ

663   Microbiol. 2018;20:1925-8.

664   55. Cavicchioli R. A vision for a 'microbcentric' future. Microb Biotechnol. 2019;12:26-9.

665   56. Brown MV, Lauro FM, DeMaere MZ, Muir L, Wilkins D, Thomas T, et al. Global

666   biogeography of SAR11 marine bacteria. Mol Syst Biol. 2012;8:595.

667   57. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor.

668   Bioinformatics. 2018;34:i884-90.

669   58. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. Spades: A

670   new genome assembly algorithm and its applications to single-cell sequencing. J Comput

671   Biol. 2012;19:455-77.

672   59. Shen L, Liu YQ, Xu BQ, Wang NL, Zhao HB, Liu XB, et al. Comparative genomic analysis

673   reveals the environmental impacts on two *Arcticibacter* strains including sixteen

674   Sphingobacteriaceae species. Sci Rep. 2017;7:2055.

675   60. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the

676   quality of microbial genomes recovered from isolates, single cells, and metagenomes.

677   Genome Res. 2015;25:1043.

678   61. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al. The SEED and the rapid

679   annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res.

680   2014;42:D206-14.

681   62. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068-

682   9.

683   63. Yarza P, Richter M, Peplies J, Euzeby J, Amann R, Schleifer KH, et al. The All-Species

684   Living Tree project: A 16S rRNA-based phylogenetic tree of all sequenced type strains.

685   Syst Appl Microbiol. 2008;31:241-50.

686   64. Yang ZH. Computational molecular evolution. Great Britain: Oxford University Press;

687   2006.

688   65. Segata N, Bornigen D, Morgan XC, Huttenhower C. PhyloPhlAn is a new method for

689   improved phylogenetic and taxonomic placement of microbes. Nat Commun. 2013;4:2304.

690   66. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, et al. MrBayes

691   3.2: efficient Bayesian phylogenetic inference and model choice across a large model space.

692   Syst Biol. 2012;61:539-42.

693   67. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: An automated protein

694   homology-modeling server. Nucleic Acids Res. 2003;31:3381-5.

695   68. Wickham H. ggplot2: Elegant graphics for data analysis. New York: Springer; 2016.

696   69. Ihaka R, Gentleman R. R: A language for data analysis and graphics. J Comput Graph Statist.

697   1996;5:299-314.

698   70. Peel MC, Finlayson BL, Mcmahon TA. Updated world map of the Köppen-Geiger climate

699   classification. Hydrol Earth Syst Sci. 2007;11:259-63.

700   71. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat

701   Methods. 2015;12:59-60.

702   72. McDonald JH. Handbook of Biological Statistics (3rd ed.). Baltimore, Maryland: Sparky

703   House Publishing; 2014.

704   73. Friedman J, Alm EJ. Inferring correlation networks from genomic survey data. PLoS

705   Comput Biol. 2012;8:e1002687.

706   74. Niku J, Hui FKC, Taskinen S, Warton DI. gllvm: Fast analysis of multivariate abundance

707   data with generalized linear latent variable models in R. Methods Ecol Evol. 2019;10:2173–

708   82.

709

710

711   **Fig. 1 *Arthrobacter* phylogeny and genome compositional profiling. (a)** Maximum

712   Likelihood *Arthrobacter* phylogenomic tree. The *Arthrobacter* portion of ML Micrococcaceae

713   phylogenomic tree (Fig. S1a) is reproduced with each leaf marked as polar and alpine (PA,

714   grey highlight) or non-polar and alpine (NPA). The tree has three major clades with the

715   central clade highlighted (purple box). **(b)** As for (a) except strain names denoted and font

716   color used to depict Group A (red font; NPA environments), Group B (green font; PA

717   environments clustering with sequences from NPA environments) and Group C (blue font; PA

718   environments that formed an operationally monophyletic lineage with an *F* measure of 0.95).

719   The specific types of cold environments from where Group C *Arthrobacter* were isolated are

720   shown to the right of tree. **(c)** 3D-NMDS plot of genome-wide amino acid composition. **(d)**

721   Distribution of pairwise ANI. **(e)** Distribution of pairwise AAI.

722

723   **Fig. 2 Growth temperature profiles of Group A, B and C *Arthrobacter*.** OD_{600} growth
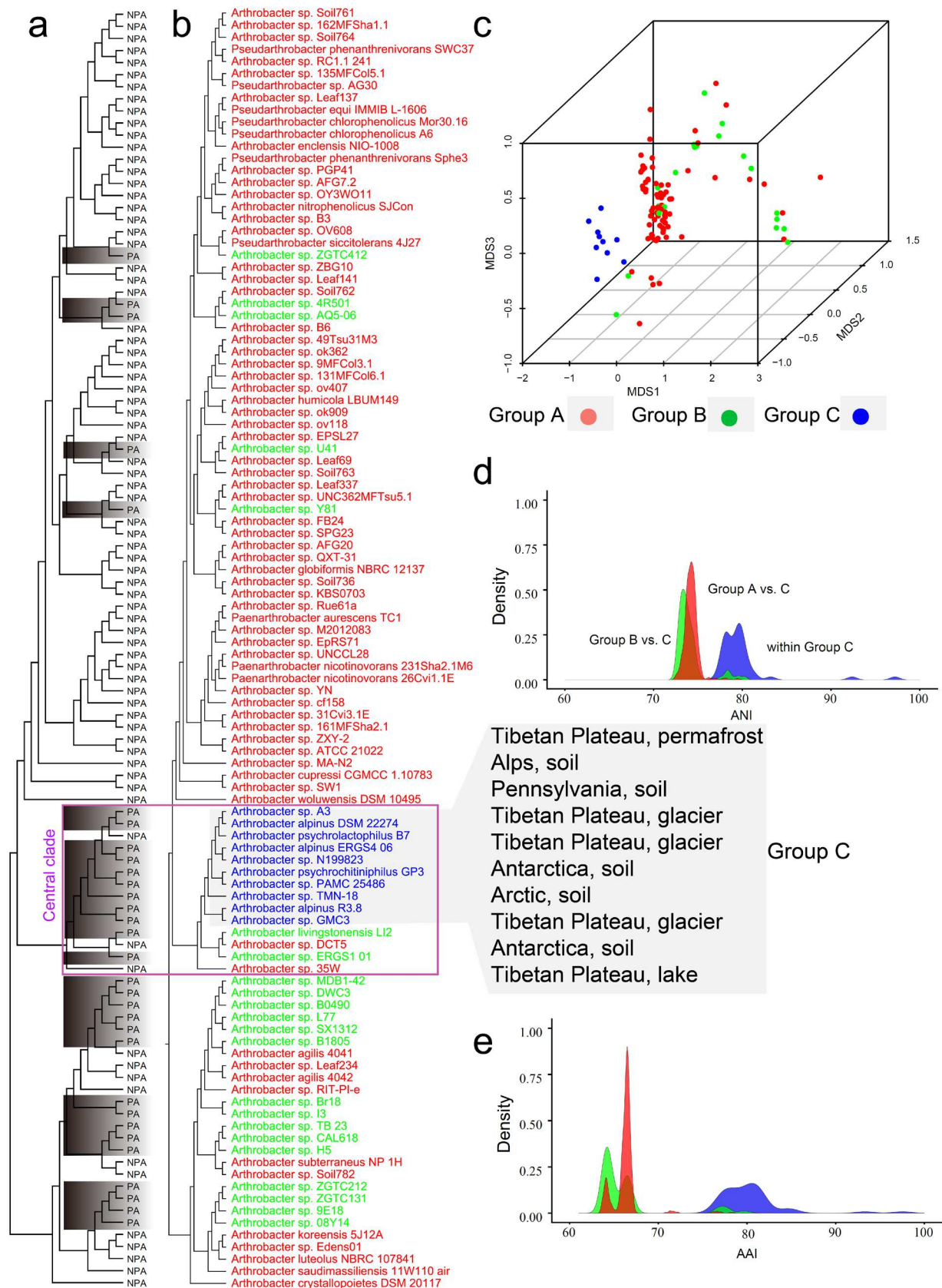
724 curves for representative *Arthrobacter* of Group A (red symbols and line; *A. luteolus*, *A.*

725 *globiformis* and *A. subterraneus*), Group B (green symbols and line; *Arthrobacter* sp. 4R501,

726 *Arthrobacter* sp. 9E14 and *Arthrobacter* sp. 08Y14) and Group C (blue symbols and line; *A.*

727 *alpinus*, *Arthrobacter* sp. A3 and *Arthrobacter* sp. N199823) at **(a)** 25 °C, **(b)** 5 °C and **(c)** -

728 1 °C.

729

730 **Fig. 3 Overview of genomic characteristics of Group C *Arthrobacter*. (a)** Box plot of G+C

731 content. Group A (red); Group C (blue); boxes represent the interquartile range with

732 horizonatal lines showing maximum and minimum values, excluding outliers. Group C had

733 significantly lower G+C content. **(b)** Scatter plot of amino acid composition. Group A (light

734 red circles); Group C (blue circles); ***, $P < 0.005$; *, $P = 0.05 – 0.01$; ns, not significant.

735 The composition of numerous amino acids varied significantly between Group C and Group A

736 *Arthrobacter*. **(c)** Protein stability predictions calculated using SCooP. Group A (red line);

737 Group C (blue line). The curve is for coenzyme A biosynthesis bifunctional protein, CoaC,

738 and is representative of one of the 32 Group C proteins from a total of 86 which had reduced

739 predicted stability (Table S2, Fig. S5) **(d)** Box plot of amino acid bias for functional

740 categories. Boxes represent the interquartile range of the Bray-Cruise distances; lines

741 extending from boxes show the maximum and minimum Bray-Cruise distances; dots beyond

742 the lines represent outliers. Biases in amino acid composition (b) were reflected in specific

743 functional categories. **(e)** Representation of functional categories. Specific functional

744 categories were over- or under-represented in Group C; arrows indicate relative increases (up

745 arrow) or decreases (down arrow) in functional categories in Group C. **(f)** Representation of

746 specific functions. Specific functional processes defined by genes or pathways were

747 characteristic of Group C (up arrow) or had a restricted capacity in Group C (down arrow)

748 compared to Group A (also see Fig. 4).

749

750 **Fig. 4 *Arthrobacter* genes typifying the functional potential of Group C. (a)** ML

751 *Arthrobacter* phylogenomic tree as for Fig. 1. **(b)** Heat map of the representation of specific
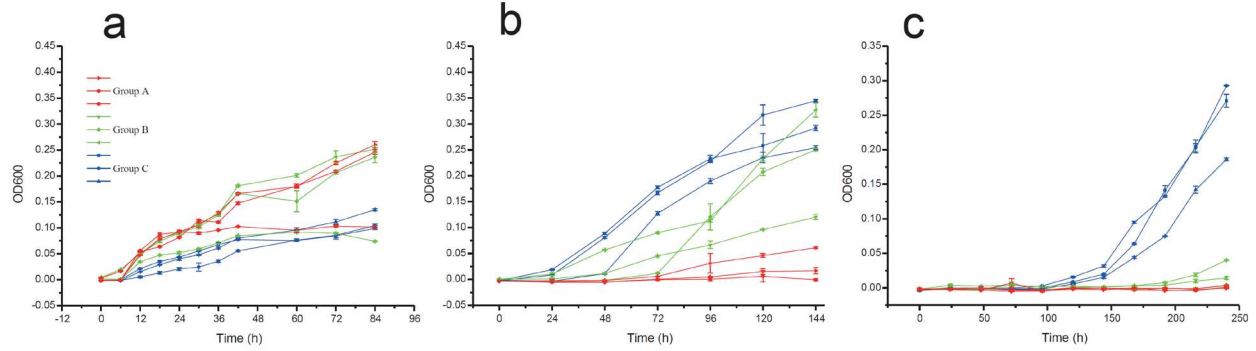
genes in *Arthrobacter* genomes, highlighting those present in Group C and the central clade. *i*,

branched-chain acyl-CoA dehydrogenase; *ii*, enoyl-CoA hydratase; *iii*, biotin repressor; *iv*,

hydrolase in cluster with formaldehyde/S-nitrosomycothiol reductase; *v*, mycothiol-dependent

formaldehyde dehydrogenase.

Fig. 5 Metagenome analysis of Group C *Arthrobacter*. (a) Depiction of the (mean annual

temperature (MAT) of surface air at a height of 2 m ([European Centre for Medium-Range

Weather Forecasts](#)) relative to latitude. The 797 metagenomes are divided into thermal

categories: PA (black squares, 286 metagenome), temperate (grey squares, 294 metagenomes)

and tropical (purple squares, 217 metagenomes). **(b)** Linear regression showing the

correlation of the abundance of Group C-specific genes within each of the 797 metagenomes

(see panel (a)) relative to the abundance of Group C-specific genes within the *Arthrobacter*

pan genome. The 95% prediction interval (dark red band) and 95% confidence interval (light

red band) is shown for each regression line (panels b, c and e). The upper cluster contains 11

Axel Heiberg Island permafrost metagenomes. **(c)** As for panel (b), except with the addition

of 361 permafrost metagenomes (total 1158 metagenomes). The Stordalen Mire (Abisko,

Sweden) metagenome is shown by an arrow. **(d)** As for panel (b), except showing Group B-

specific genes. **(e)** As for panel (b), except showing Group A-specific genes present in PA

genomes (lower line) and NPA genomes (upper line). The regression line for the 11 Axel

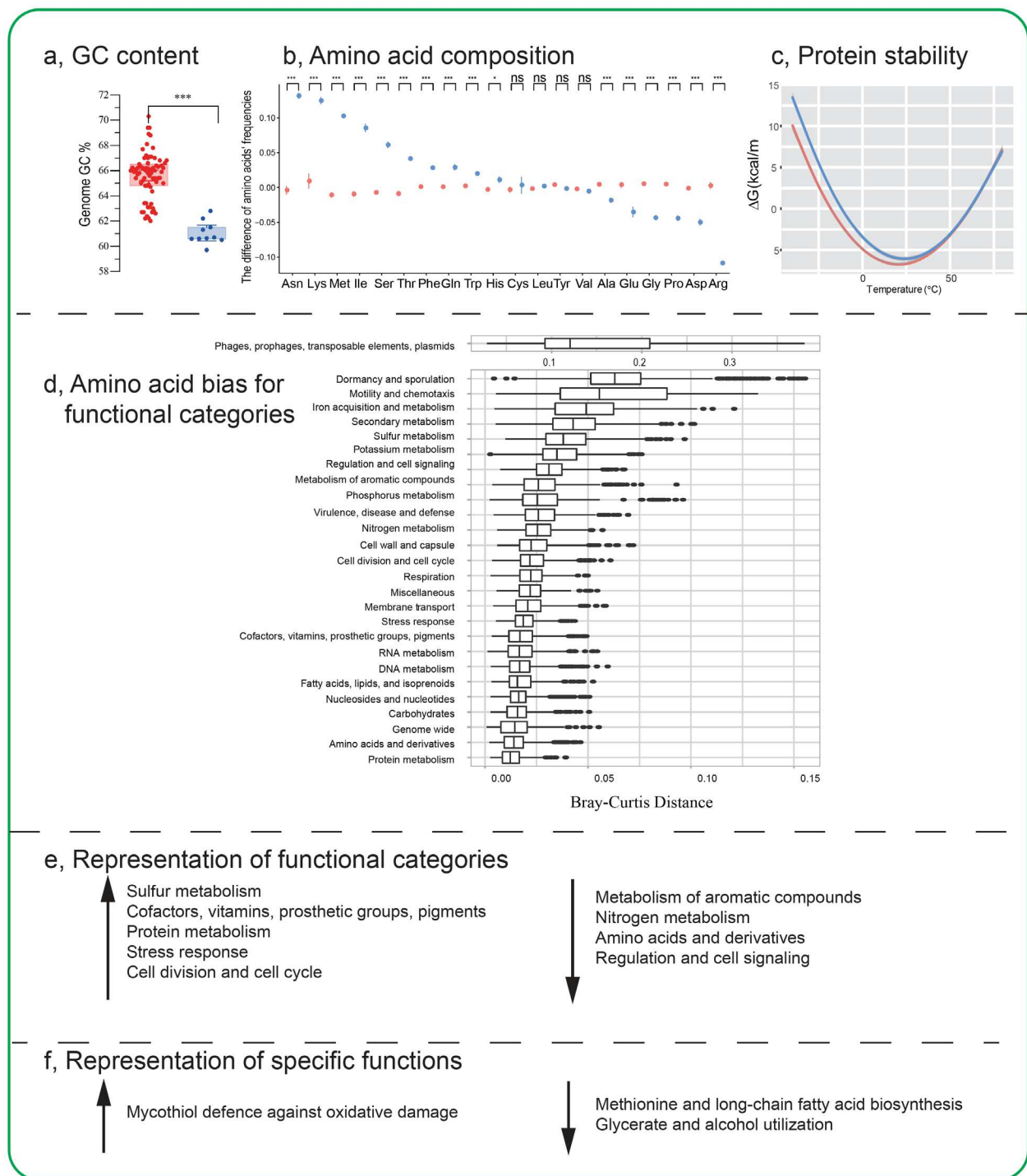Heiberg Island permafrost metagenomes is not shown.

a

b
Arthrobacter sp. Soil761
Arthrobacter sp. 162MFSha1.1
Arthrobacter sp. Soil764
Pseudarthrobacter phenanthrenivorans SWC37
Arthrobacter sp. RC1.1 241
Arthrobacter sp. 135MFCol5.1
Pseudarthrobacter sp. AG30
Arthrobacter sp. Leaf137
Pseudarthrobacter equi IMMIB L-1606
Pseudarthrobacter chlorophenolicus Mor30.16
Pseudarthrobacter chlorophenolicus A6
Arthrobacter enclensis NIO-1008
Pseudarthrobacter phenanthrenivorans Sphe3
Arthrobacter sp. PGP41
Arthrobacter sp. AFG7.2
Arthrobacter sp. OY3WO11
Arthrobacter nitrophenolicus SJCon
Arthrobacter sp. B3
Arthrobacter sp. OV608
Pseudarthrobacter siccitolerans 4J27
Arthrobacter sp. ZGTC412
Arthrobacter sp. ZBG10
Arthrobacter sp. Leaf141
Arthrobacter sp. Soil762
Arthrobacter sp. 4R501
Arthrobacter sp. AQ5-06
Arthrobacter sp. B6
Arthrobacter sp. 49Tsu31M3
Arthrobacter sp. ok362
Arthrobacter sp. 9MFCol3.1
Arthrobacter sp. 131MFCol6.1
Arthrobacter sp. ov407
Arthrobacter humicola LBUM149
Arthrobacter sp. ok909
Arthrobacter sp. ov118
Arthrobacter sp. EPSL27
Arthrobacter sp. U41
Arthrobacter sp. Leaf69
Arthrobacter sp. Soil763
Arthrobacter sp. Leaf337
Arthrobacter sp. UNC362MFTsu5.1
Arthrobacter sp. Y81
Arthrobacter sp. FB24
Arthrobacter sp. SPG23
Arthrobacter sp. AFG20
Arthrobacter sp. QXT-31
Arthrobacter globiformis NBRC 12137
Arthrobacter sp. Soil736
Arthrobacter sp. KBS0703
Arthrobacter sp. Rue61a
Paenarthrobacter aurescens TC1
Arthrobacter sp. M2012083
Arthrobacter sp. EpRS71
Arthrobacter sp. UNCCL28
Paenarthrobacter nicotinovorans 231Sha2.1M6
Paenarthrobacter nicotinovorans 26Cvi1.1E
Arthrobacter sp. YN
Arthrobacter sp. cf158
Arthrobacter sp. 31Cvi3.1E
Arthrobacter sp. 161MFSha2.1
Arthrobacter sp. ZXY-2
Arthrobacter sp. ATCC 21022
Arthrobacter sp. MA-N2
Arthrobacter cupressi CGMCC 1.10783
Arthrobacter sp. SW1
Arthrobacter woluwensis DSM 10495
Arthrobacter sp. A3
Arthrobacter alpinus DSM 22274
Arthrobacter psychrolactophilus B7
Arthrobacter alpinus ERGS4 06
Arthrobacter sp. N199823
Arthrobacter psychrochitiniphilus GP3
Arthrobacter sp. PAMC 25486
Arthrobacter sp. TMN-18
Arthrobacter alpinus R3.8
Arthrobacter sp. GMC3
Arthrobacter livingstonensis LI2
Arthrobacter sp. DCT5
Arthrobacter sp. ERGS1 01
Arthrobacter sp. 35W
Arthrobacter sp. MDB1-42
Arthrobacter sp. DWC3
Arthrobacter sp. B0490
Arthrobacter sp. L77
Arthrobacter sp. SX1312
Arthrobacter sp. B1805
Arthrobacter agilis 4041
Arthrobacter sp. Leaf234
Arthrobacter agilis 4042
Arthrobacter sp. RIT-PI-e
Arthrobacter sp. Br18
Arthrobacter sp. I3
Arthrobacter sp. TB 23
Arthrobacter sp. CAL618
Arthrobacter sp. H5
Arthrobacter subterraneus NP 1H
Arthrobacter sp. Soil782
Arthrobacter sp. ZGTC212
Arthrobacter sp. ZGTC131
Arthrobacter sp. 9E18
Arthrobacter sp. 08Y14
Arthrobacter koreensis 5J12A
Arthrobacter sp. Edens01
Arthrobacter luteolus NBRC 107841
Arthrobacter saudimassiliensis 11W110 air
Arthrobacter crystallopoietes DSM 20117

Central clade

c

MDS3

MDS1

MDS2

Group A   Group B   Group C

d

Density

ANI

Group B vs. C

Group A vs. C

within Group C

Tibetan Plateau, permafrost
Alps, soil
Pennsylvania, soil
Tibetan Plateau, glacier
Tibetan Plateau, glacier
Antarctica, soil
Arctic, soil
Tibetan Plateau, glacier
Antarctica, soil
Tibetan Plateau, lake

Group C

e

Density

AAI

1

Fig. 1 *Arthrobacter* phylogeny and genome compositional profiling. (a) Maximum Likelihood *Arthrobacter* phylogenomic tree. The *Arthrobacter* portion of ML Micrococcaceae phylogenomic tree (Fig. S1a) is reproduced with each leaf marked as polar and alpine (PA, grey highlight) or non-polar and alpine (NPA). The tree has three major clades with the central clade highlighted (purple box). **(b)** As for (a) except strain names denoted and font color used to depict Group A (red font; NPA environments), Group B (green font; PA environments clustering with sequences from NPA environments) and Group C (blue font; PA environments that formed an operationally monophyletic lineage with an *F* measure of 0.95). The specific types of cold environments from where Group C *Arthrobacter* were isolated are shown to the right of tree. **(c)** 3D-NMDS plot of genome-wide amino acid composition. **(d)** Distribution of pairwise ANI. **(e)** Distribution of pairwise AAI.

1
2
3 **Fig. 2 Growth temperature profiles of Group A, B and C *Arthrobacter*.** OD_{600} growth curves for
4 representative *Arthrobacter* of Group A (red symbols and line; *A. luteolus*, *A. globiformis* and *A.*
5 *subterraneus*), Group B (green symbols and line; *Arthrobacter* sp. 4R501, *Arthrobacter* sp. 9E14 and
6 *Arthrobacter* sp. 08Y14) and Group C (blue symbols and line; *A. alpinus*, *Arthrobacter* sp. A3 and
7 *Arthrobacter* sp. N199823) at **(a)** 25 °C, **(b)** 5 °C and **(c)** -1 °C.

Fig. 3 Overview of genomic characteristics of Group C *Arthrobacter*. (a) Box plot of G+C content. Group A (red); Group C (blue); boxes represent the interquartile range with horizonatal lines showing maximum and minimum values, excluding outliers. Group C had significantly lower G+C content. **(b)** Scatter plot of amino acid composition. Group A (light red circles); Group C (blue circles); ***, *P* < 0.005; *, *P* = 0.05 – 0.01; ns, not significant. The composition of numerous amino acids varied significantly between Group C and Group A *Arthrobacter.* **(c)** Protein stability predictions calculated using

9   SCooP. Group A (red line); Group C (blue line). The curve is for coenzyme A biosynthesis bifunctional

10   protein, CoaC, and is representative of one of the 32 Group C proteins from a total of 86 which had

11   reduced predicted stability (Table S2, Fig. S5) **(d)** Box plot of amino acid bias for functional categories.

12   Boxes represent the interquartile range of the Bray-Cruise distances; lines extending from boxes show the

13   maximum and minimum Bray-Cruise distances; dots beyond the lines represent outliers. Biases in amino

14   acid composition (b) were reflected in specific functional categories. **(e)** Representation of functional

15   categories. Specific functional categories were over- or under-represented in Group C; arrows indicate

16   relative increases (up arrow) or decreases (down arrow) in functional categories in Group C. **(f)**

17   Representation of specific functions. Specific functional processes defined by genes or pathways were

18   characteristic of Group C (up arrow) or had a restricted capacity in Group C (down arrow) compared to

19   Group A (also see Fig. 4).
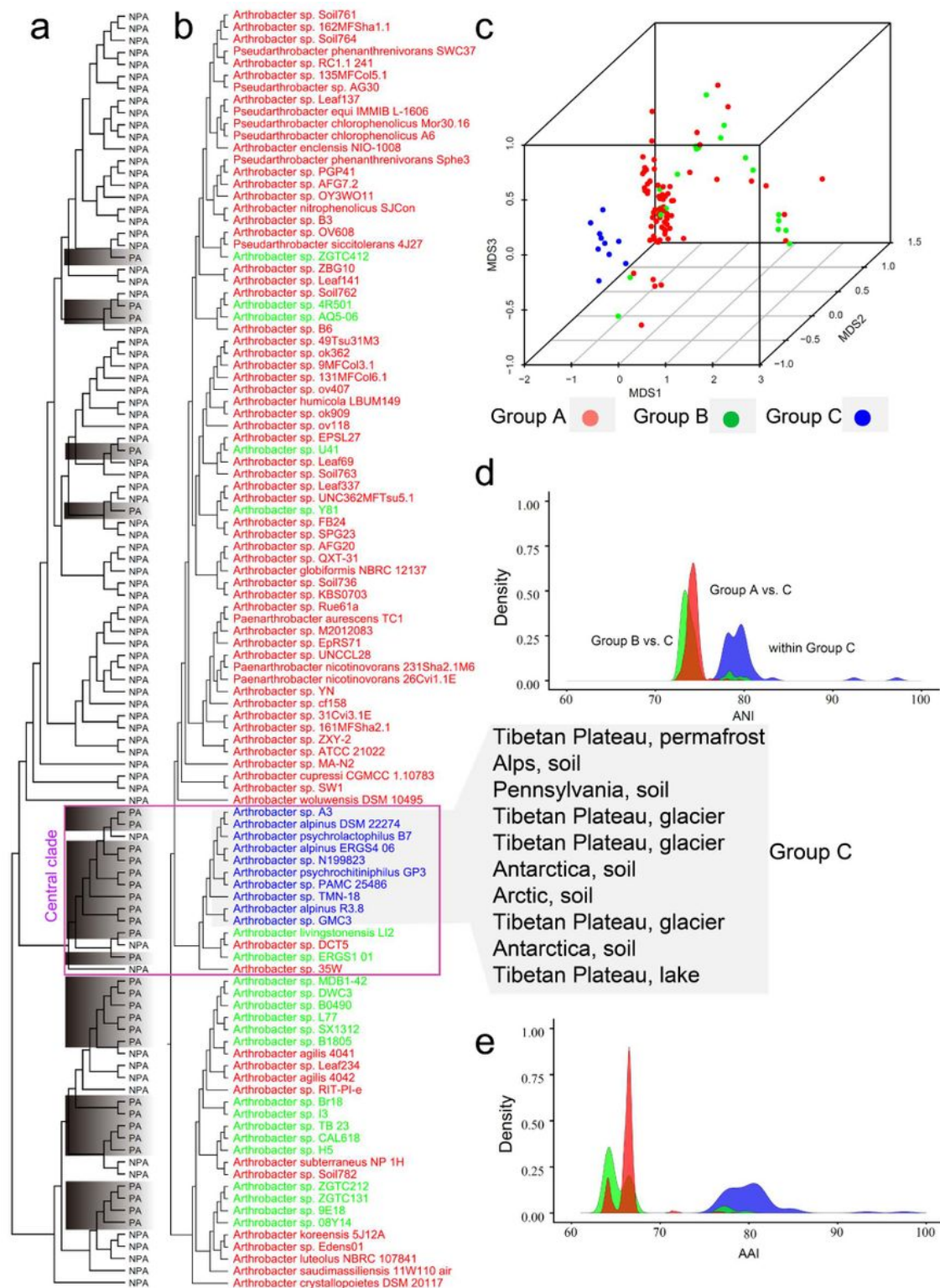
3   **Fig. 4 *Arthrobacter* genes typifying the functional potential of Group C. (a)** ML *Arthrobacter*

4   phylogenomic tree as for Fig. 1. **(b)** Heat map of the representation of specific genes in *Arthrobacter*

5   genomes, highlighting those present in Group C and the central clade. *i*, branched-chain acyl-CoA

6   dehydrogenase; *ii*, enoyl-CoA hydratase; *iii*, biotin repressor; *iv*, hydrolase in cluster with formaldehyde/S-

7   nitrosomycothiol reductase; *v*, mycothiol-dependent formaldehyde dehydrogenase.

Fig. 5 Metagenome analysis of Group C *Arthrobacter*. (a) Depiction of the (mean annual temperature (MAT) of surface air at a height of 2 m (European Centre for Medium-Range Weather Forecasts) relative to latitude. The 797 metagenomes are divided into thermal categories: PA (black squares, 286 metagenome), temperate (grey squares, 294 metagenomes) and tropical (purple squares, 217 metagenomes). **(b)** Linear regression showing the correlation of the abundance of Group C-specific genes within each of the 797 metagenomes (see panel (a)) relative to the abundance of Group C-specific genes within the *Arthrobacter* pan genome. The 95% prediction interval (dark red band) and 95% confidence interval (light red band) is shown for each regression line (panels b, c and e). The upper cluster contains 11 Axel Heiberg Island permafrost metagenomes. **(c)** As for panel (b), except with the addition of 361 permafrost metagenomes (total 1158 metagenomes). The Stordalen Mire (Abisko, Sweden) metagenome is shown by an arrow. **(d)** As for panel (b), except showing Group B-specific genes. **(e)** As for panel (b), except showing Group A-specific genes present in PA genomes (lower line) and NPA genomes (upper line). The regression line for the 11 Axel Heiberg Island permafrost metagenomes is not shown.
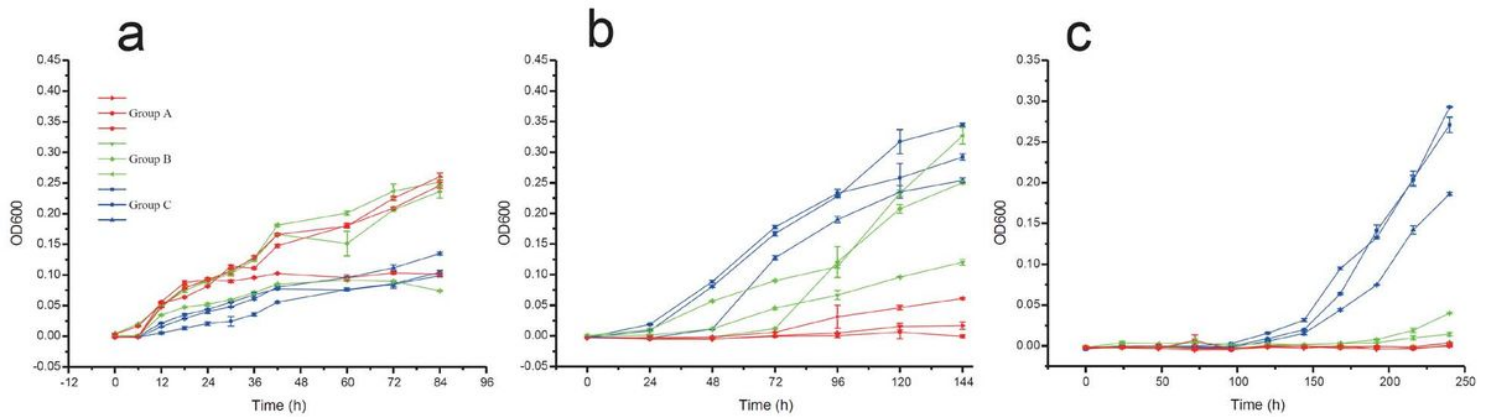
Figures

Figure 1

Arthrobacter phylogeny and genome compositional profiling. (a) Maximum Likelihood Arthrobacter phylogenomic tree. The Arthrobacter portion of ML Micrococcaceae phylogenomic tree (Fig. S1a) is reproduced with each leaf marked as polar and alpine (PA, grey highlight) or non-polar and alpine (NPA).
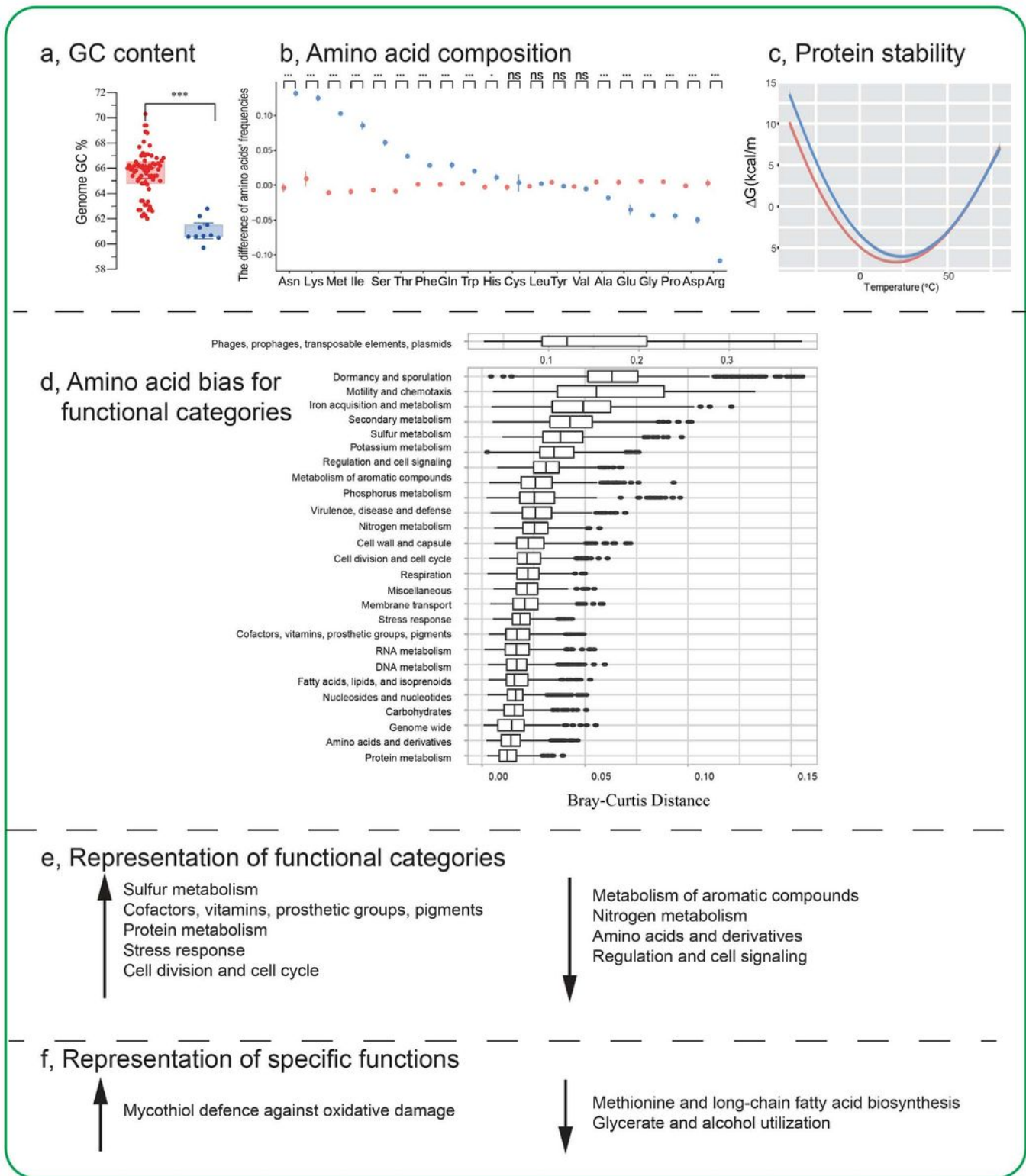
The tree has three major clades with the central clade highlighted (purple box). (b) As for (a) except strain names denoted and font color used to depict Group A (red font; NPA environments), Group B (green font; PA environments clustering with sequences from NPA environments) and Group C (blue font; PA environments that formed an operationally monophyletic lineage with an F measure of 0.95). The specific types of cold environments from where Group C Arthrobacter were isolated are shown to the right of tree. (c) 3D-NMDS plot of genome-wide amino acid composition. (d) Distribution of pairwise ANI. (e) Distribution of pairwise AAI.

Figure 2

Growth temperature profiles of Group A, B and C Arthrobacter. OD600 growth curves for representative Arthrobacter of Group A (red symbols and line; A. luteolus, A. globiformis and A. subterraneus), Group B (green symbols and line; Arthrobacter sp. 4R501, Arthrobacter sp. 9E14 and Arthrobacter sp. 08Y14) and Group C (blue symbols and line; A. alpinus, Arthrobacter sp. A3 and Arthrobacter sp. N199823) at (a) 25 °C, (b) 5 °C and (c) -1 °C.

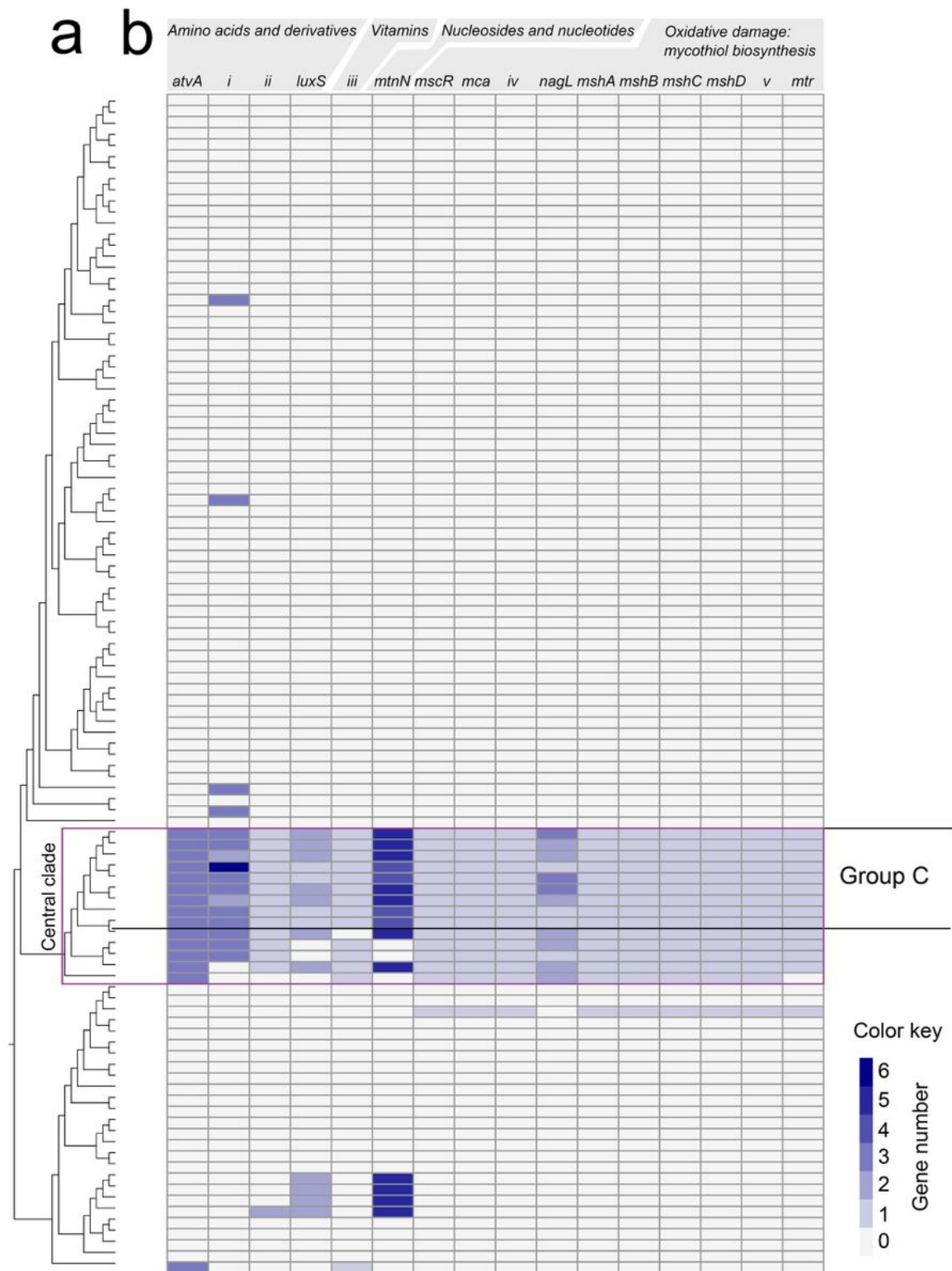**a, GC content**

b, Amino acid composition

c, Protein stability

d, Amino acid bias for functional categories

Bray-Curtis Distance

e, Representation of functional categories

Sulfur metabolism
Cofactors, vitamins, prosthetic groups, pigments
Protein metabolism
Stress response
Cell division and cell cycle

Metabolism of aromatic compounds
Nitrogen metabolism
Amino acids and derivatives
Regulation and cell signaling

f, Representation of specific functions

Mycothiol defence against oxidative damage

Methionine and long-chain fatty acid biosynthesis
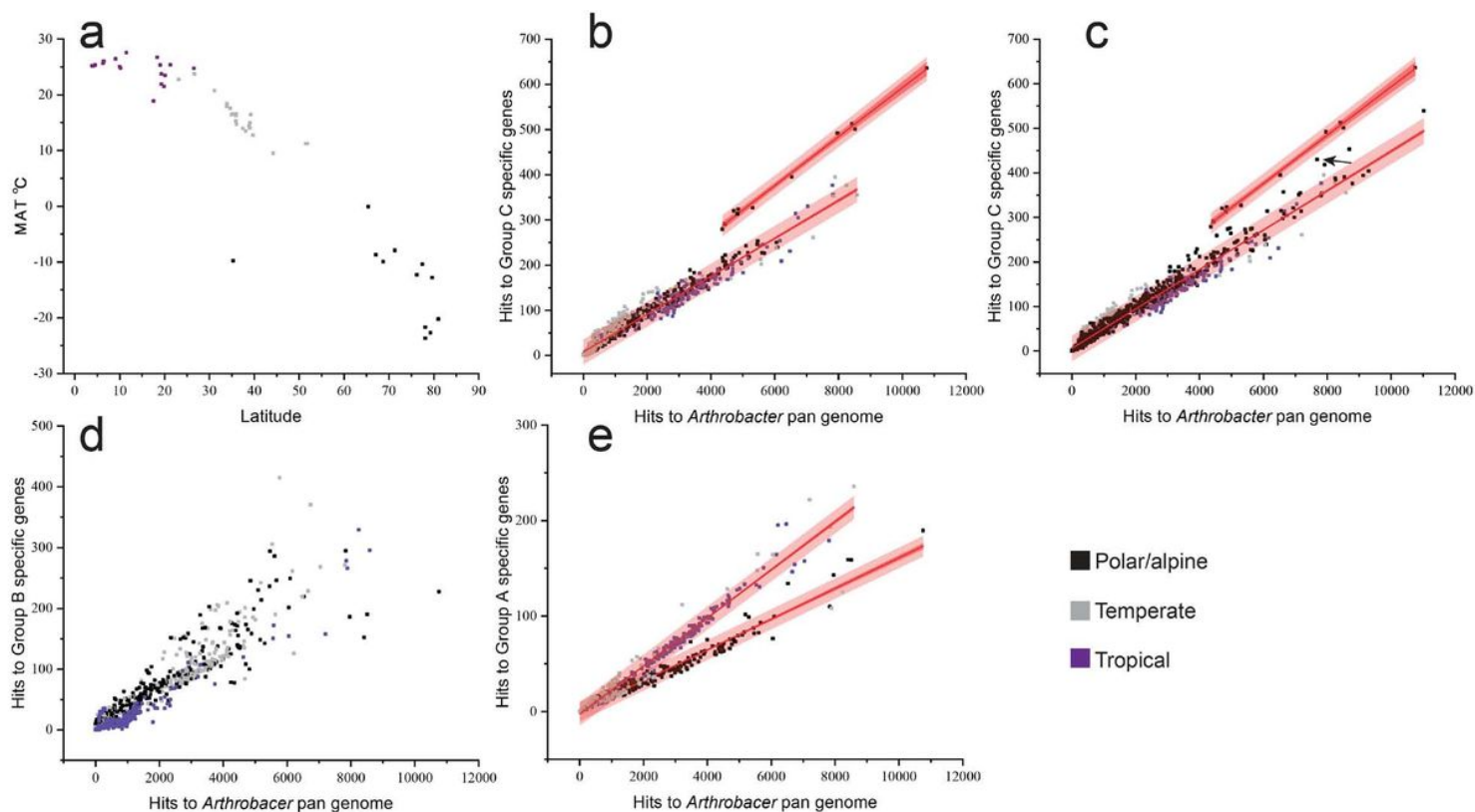Glycerate and alcohol utilization

Figure 3

Overview of genomic characteristics of Group C Arthrobacter. (a) Box plot of G+C content. Group A (red); Group C (blue); boxes represent the interquartile range with horizonatal lines showing maximum and minimum values, excluding outliers. Group C had significantly lower G+C content. (b) Scatter plot of amino acid composition. Group A (light red circles); Group C (blue circles); ***, P < 0.005; *, P = 0.05 – 0.01; ns, not significant. The composition of numerous amino acids varied significantly between Group C

and Group A *Arthrobacter*. (c) Protein stability predictions calculated using SCooP. Group A (red line); Group C (blue line). The curve is for coenzyme A biosynthesis bifunctional protein, CoaC, and is representative of one of the 32 Group C proteins from a total of 86 which had reduced predicted stability (Table S2, Fig. S5) (d) Box plot of amino acid bias for functional categories. Boxes represent the interquartile range of the Bray-Cruise distances; lines extending from boxes show the maximum and minimum Bray-Cruise distances; dots beyond the lines represent outliers. Biases in amino acid composition (b) were reflected in specific functional categories. (e) Representation of functional categories. Specific functional categories were over- or under-represented in Group C; arrows indicate relative increases (up arrow) or decreases (down arrow) in functional categories in Group C. (f) Representation of specific functions. Specific functional processes defined by genes or pathways were characteristic of Group C (up arrow) or had a restricted capacity in Group C (down arrow) compared to Group A (also see Fig. 4).

Figure 4

Arthrobacter genes typifying the functional potential of Group C. (a) ML Arthrobacter phylogenomic tree as for Fig. 1. (b) Heat map of the representation of specific genes in Arthrobacter genomes, highlighting those present in Group C and the central clade. ⓘ, branched-chain acyl-CoA dehydrogenase; ⓘ, enoyl-CoA hydratase; ⓘ, biotin repressor; ⓘ, hydrolase in cluster with formaldehyde/S nitrosomycothiol reductase; ⓘ, mycothiol-dependent formaldehyde dehydrogenase.

Figure 5

Metagenome analysis of Group C Arthrobacter. (a) Depiction of the (mean annual temperature (MAT) of surface air at a height of 2 m (European Centre for Medium-Range Weather Forecasts) relative to latitude. The 797 metagenomes are divided into thermal categories: PA (black squares, 286 metagenome), temperate (grey squares, 294 metagenomes) and tropical (purple squares, 217 metagenomes). (b) Linear regression showing the correlation of the abundance of Group C-specific genes within each of the 797 metagenomes (see panel (a)) relative to the abundance of Group C-specific genes within the Arthrobacter pan genome. The 95% prediction interval (dark red band) and 95% confidence interval (light red band) is shown for each regression line (panels b, c and e). The upper cluster contains 11 Axel Heiberg Island permafrost metagenomes. (c) As for panel (b), except with the addition of 361 permafrost metagenomes (total 1158 metagenomes). The Stordalen Mire (Abisko, Sweden) metagenome is shown by an arrow. (d) As for panel (b), except showing Group B-specific genes. (e) As for panel (b), except showing Group A-specific genes present in PA genomes (lower line) and NPA genomes (upper line). The regression line for the 11 Axel Heiberg Island permafrost metagenomes is not shown.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- CavicchioliSupplementaryinformation.pdf
- CavicchioliTableS1.xlsx

- CavicchioliTableS2.xlsx
- CavicchioliTableS3.xlsx
- CavicchioliTableS4.xlsx
- CavicchioliTableS5.xlsx
- CavicchioliTableS6.xlsx
- CavicchioliTableS8.xlsx