Janthinobacterium CG23_2: Comparative genome analysis reveals enhanced environmental sensing and transcriptional regulation for adaptation to life in an Antarctic supraglacial stream

Markus Dieser[1,2,*], Heidi J. Smith[1,4], Thiruvarangan Ramaraj[3] and C. M. Foreman[1,2]

[1] Center for Biofilm Engineering, Montana State University, Bozeman, MT 59717, USA

[2] Department of Chemical & Biological Engineering, Montana State University, Bozeman, MT 59715, USA

[3] School of Computing, College of Computing & Digital Media, DePaul University, Chicago, IL 60604, USA

[4] Department of Microbiology & Immunology, Montana State University, Bozeman, MT 59717, USA

* Corresponding Author:

Markus Dieser

Montana State University

Center for Biofilm Engineering

366 Barnard Hall

Bozeman, MT 59717, USA

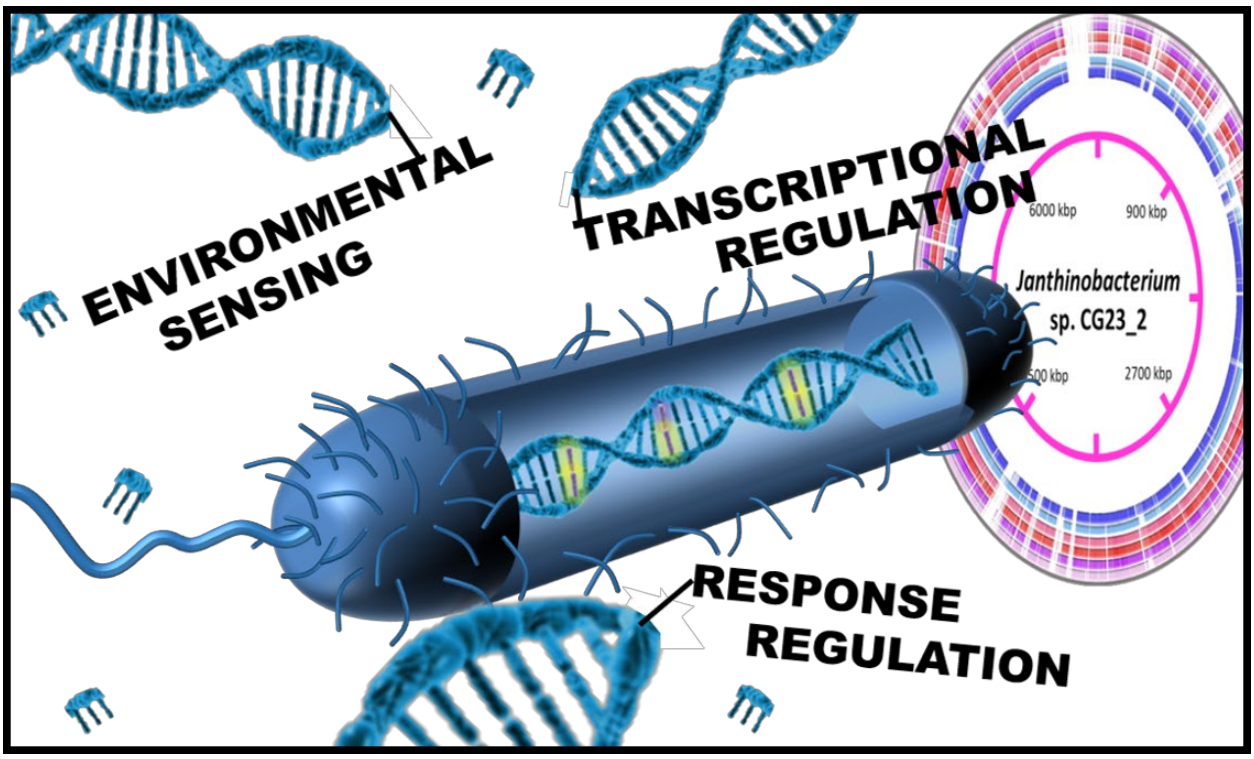
Email: markus.dieser@montana.edu

Draft for submission to Microorganisms-special issue: Microbial Evolution of Extremophiles

Abstract

As many bacteria detected in Antarctic environments are neither true psychrophiles nor endemic species, their proliferation despite environmental extremes gives rise to genome adaptations. *Janthinobacterium* sp. CG23_2 is a bacterial isolate from the Cotton Glacier stream, Antarctica. To understand how *Janthinobacterium* sp. CG23_2 has adapted to its environment, we investigated its genomic traits in comparison to genomes of 35 published Janthinobacteria. While we hypothesized that genome shrinkage and specialization to narrow ecological niches would be energetically favorable for dwelling in an ephemeral Antarctic stream, the genome of *Janthinobacterium* sp. CG23_2 was on average 1.7±0.6 Mb larger and predicted 1,411±499 more coding domain sequences compared to the other *Janthinobacterium* spp. Putatively identified horizontal gene transfer events contributed 0.92 Mb to the genome size expansion of *Janthinobacterium* sp. CG23_2. Genes with high copy numbers in the species-specific accessory genome of *Janthinobacterium* sp. CG23_2 were associated with environmental sensing, locomotion, response and transcriptional regulation, stress response, and mobile elements; functional categories which also showed molecular adaptation to cold. Our data suggest that genome plasticity and the abundant complimentary genes for sensing and responding to the extracellular environment supported adaptation of *Janthinobacterium* sp. CG23_2 to this extreme environment.

Graphical abstract



Keywords: Janthinobacterium, comparative genomics, horizontal gene transfer, cold adaptation, environmental sensing

Environments with temperatures permanently below 5°C dominate the Earth's biosphere (>80%). While much of the cold biosphere is made up of the world's oceans, a combined 37% of the land area consists of permafrost regions, 198,000 glaciers, and two ice sheets [1-3]. Adaptation to cold temperatures and associated stresses (i.e. desiccation, water activity, radiation, pH, ionic strength, and nutrient availability) [4] allows psychrophilic/psychrotolerant microorganisms to inhabit and even thrive in these extreme and often inhospitable environments. Integrating 'omics' technologies with physiological studies on cold adaptation has advanced our understanding of functional and evolutionary biological processes at the molecular level [5-8]. Specifically, genome comparisons are a powerful tool that enables the exploration of molecular level adaptations in microorganisms inhabiting a broad temperature spectrum. Comparative approaches have identified two overarching trends across bacterial genomes. First, genome size appears to correlate with environmental niches and lifestyle, and second, diversification is explained by horizontal gene transfer [9]. Paradoxically, in genomes from cold-adapted bacterial species both diversification [10,11] and streamlining are observed [6,12], resulting in an increased and decreased genome size, respectively.

In the present study, we investigated the genomic properties of *Janthinobacterium* sp. CG23_2 [13], a bacterium isolated from a supraglacial stream on the Cotton Glacier, Antarctica. *Janthinobacterium* spp. are frequently isolated from alpine and polar regions [13-22], while others have been identified as amphibian symbionts [23] or pathogens [24]. The genus *Janthinobacterium* (*Betaproteobacteria*, *Oxalobacteraceae*) includes Gram-negative, motile, aerobic, chemoorganoheterotrophic bacteria, that are often resistant to a wide range of antibiotics and heavy metals [16,25,26]. *Janthinobacterium* spp. commonly produce violacein, a purple, water insoluble, secondary metabolite [27]; however, pigment color can vary [14,16,17,19].

Unlike most supraglacial streams, which form seasonally in the ablation zone by melt water incisions, the Cotton Glacier stream is a perennial feature. Water only flows during summer melt (4-12 weeks), with variations in stream hydrology related to local daily, seasonal, and interannual climate conditions [28]. The stream bed is devoid of significant amounts of sediments but is flanked by large para-fluvial sediment deposits [29]. Due to the dynamic ephemeral nature of the Cotton Glacier stream, its changing velocity, low in-stream nutrient levels, short residence times, temperatures near or below the freezing point, daily freeze-thaw cycles in summer, months of deep-freeze in winter, and extreme transparency to solar UV radiation [28,29], microbial inhabitants must be capable of major metabolic and physiological adjustments year-round. We hypothesized that in order to optimize adaptation to this extremely variable environment, free living microorganisms would form long-lasting metabolic interactions to stabilize their close environment. According to the Black Queen Hypothesis [30], such functional dependencies of sharing genomic resources would select smaller genomes, thereby reducing energetically costly requirements associated with the maintenance of genetic material and metabolic activities; genomic features which are characteristic for oligotrophic environment [31]. Yet, the genome of *Janthinobacterium* sp. CG23_2 was significantly larger than any previously sequenced *Janthinobacterium* genome [13]. To discern the genetic attributes of *Janthinobacterium* sp. CG23_2 that contribute to a larger genome and allow proliferation under these extremely harsh and transient conditions within the Cotton Glacier stream, a comparative genomic approach with 35 *Janthinobacterium* spp. was performed.

Methodology

Comparative genome and phylogenetic analyses

The genomes of 35 *Janthinobacterium* species/strains publicly available at NCBI were used for comparison to the *Janthinobacterium* CG23_2 genome (Table 1). Due to the low percentage match to the other *Janthinobacterium* spp. (Table S1), genomes from *Janthinobacterium* spp. B9-8, Marseille, HH01, CG3, and NBRC.102515 were excluded from core gene analysis. Core genes were identified using the algorithm described in Cleary et al. [32], which constructs a compressed de Bruijn graph (CDBG) of a genome population and identifies core genes using the frequently visited regions in the graph. Subsequently, the coding domain sequences (CDS) of the *Janthinobacterium* CG23_2 genome were mapped to the core genes using GMAP v1.8 [33]. Default setting were applied except for the "maximum number of paths to show" flag, which was set to 1. The GMAP alignment was filtered for ≥ 90% query coverage and ≥80% identity. The same parameters were selected in the GMAP v1.8 software package for pairwise comparisons of CDS identified in the 36 *Janthinobacterium* spp. genomes.

A *Janthinobacterium* spp. genome tree was inferred using CheckM [34]. CheckM identifies, annotates, concatenates, and aligns 43 marker genes/amino acid sequences, which were subsequently placed onto a reference genome tree using the implemented pplacer software package [35]. The Maximum Likelihood tree was computed in MEGA7 [36]. The phylogenetic tree was visualized using iTOL version 4.3 [37]. Whole genome sequences *in silico* DNA-DNA hybridization (DDH) and average nucleotide identity (ANI) between *Janthinobacterium* sp. CG23_2 and *Janthinobacterium* spp. (Table 1) were performed using the default settings for the genome-to-genome distance calculator GGDC 2.1 in combination with the BLAST+ alignment tool [38] and the Average Nucleotide Identity calculator [39,40], respectively. As recommended by the developers [38], results from formula 2 were considered for DDH, as these estimates are more robust against the use of incomplete draft genomes.

121

122 Molecular analysis of cold-adaptation

123 Cold-adaptation of proteins from *Janthinobacterium* sp. CG23_2 was identified using a
124 publicly available python script [6; https://github.com/ColdAdaptationScripts/cold_adaptation].
125 Two custom databases were generated, using *Janthinobacterium* spp. isolated from mesophilic
126 (31 annotated genomes) and polar/glacial environments (4 annotated genomes) (Table 1). The
127 annotated genome from *Janthinobacterium* sp. CG23_2 was compared to these databases using
128 BLASTP [41] with a cutoff score *E*-value $\leq 10^{-15}$. Cold-adaptation scores were assigned to each
129 protein based on the following parameters: arginine to lysine ratio; frequency of acidic residues;
130 proline residues; aromaticity; aliphacity; and grand average of hydropathicity (GRAVY) [6].
131 Cold-adaptation was inferred for each index if the direction of change was significant for fewer
132 proline and acidic residues, and lower R/K (arginine/lysine) ratios, aliphacity, aromaticity, and
133 GRAVY. Proteins with ≥3 cold-adaptation indices were determined to be cold-adapted. [6].
134 Clusters of orthologous groups (COGs) were annotated with WebMGA using default settings
135 [42] and the updated COGs database [43].

136

137 Horizontal gene transfer (HGT)

138 HGTector [44; https://github.com/DittmarLab/HGTector] was used to identify putative
139 horizontally transferred genes. All-against-all BLASTP [41] was performed against the NCBI
140 non-redundant protein sequences database with an *E*-value cutoff $\leq 10^{-10}$, ≥30% identity, and
141 ≥70% sequence coverage. The protein sequence database, taxonomy database, and a protein-to-
142 taxonomy dictionary were downloaded on Aug 15th, 2018 from the NCBI website. Up to 100
143 non-redundant hits per protein were preserved. Hits to more than one organism under the same

species were excluded. *Janthinobacterium* was defined as the *self* group (NCBI taxonomic ID: 29580), and *Burkholderiales* as the *close* group (NCBI taxonomic ID: 1224). The *distal* group was comprised of all other organisms. Cutoffs of 7.04, 5.55, and 0.79 for the *self*, *close*, and *distal* group, respectively, were computed using the kernel density estimation function. The cutoff in the *self* weight distribution was included to predict putatively HGT-derived genes that were acquired by specific organisms within the *self* group. COGs of predicted HGT-derived genes were annotated with WebMGA using default settings [42] and the updated COGs database [43]. Due to their phylogenetic distance and lack of gene similarity between the other *Janthinobacterium* spp. (0.0-3.8%; Table S1), the *Janthinobacterium* spp. Marseille and B9-8 were excluded from this analysis.

Results

Genome statistics for all 36 *Janthinobacterium* spp. are summarized in Table 1. Genome sizes (4.11-7.85 Mb) and number of CDS (3,870-6,859) varied across the analyzed species/strains. All species/strains had a high mole percent GC content (60.5-65.5%), except for *Janthinobacterium* spp. Marseille (54.25) and B9-8 (48.7). *Janthinobacterium* sp. CG23_2 had the largest genome with 7.85 Mb. Its genome was on average 1.7±0.6 Mb larger and predicted 1,411±499 more CDS compared to the other *Janthinobacterium* spp. (Table 1). Calculated DDH and ANI values for *Janthinobacterium* sp. CG23_2 were 21.5±1.8% and 79.3±0.4%, respectively, and thus below the threshold boundaries (DDH: 70% and ANI: 95%) for members of the same species. Phylogenetic reconstruction of the 36 *Janthinobacterium* spp. using 43 marker genes identified by CheckM placed *Janthinobacterium* sp. CG23_2 near the root of the tree (Fig. 1). *Janthinobacterium* spp. isolated in close proximities (i.e. *Janthinobacterium* spp. CG23_2/CG3; 551a/334; GW, and HH) were more closely related. *Janthinobacterium* spp.

Marseille and B9-8, the two species with the smallest genome sizes and considerably lower GC contents, were most distantly related to all other members of the genus *Janthinobacterium* investigated (Fig. 1).

Collectively, *Janthinobacterium* sp. CG23_2 had only 1,282 CDS (18.7%) in common with the other 35 *Janthinobacterium* spp. as determined by pairwise comparison. Overlap between *Janthinobacterium* sp. CG23_2 and each of the 35 *Janthinobacterium* spp. genomes was low and ranged between 7.2-12.1% (Fig. 2). Likewise, *Janthinobacterium* spp. B9-8, Marseille, HH01, CG3, and NBRC.102515 had little genome overlap with the other *Janthinobacterium* spp. (0-25% on average; Table S1). These six *Janthinobacterium* spp. with low percentage match (<25%) to the other *Janthinobacterium* spp. were excluded for core gene analysis. For the remaining 30 *Janthinobacterium* spp. 3,260 core CDS were identified. Only 164 CDS were shared between *Janthinobacterium* sp. CG23_2 and the core genes, indicating the prevalence of *Janthinobacterium* sp. CG23_2 species-specific accessory genome. Core genes were twelve- and four-fold enriched in the COG categories J and C, respectively, relative to those found specific to *Janthinobacterium* sp. CG23_2 (Fig. 3). Conversely, the species-specific accessory genome of *Janthinobacterium* sp. CG23_2 was thirteen-fold enriched in COG category T, and four- to five-fold in categories V, M, G, and Q, relative to those identified as core genes (Fig. 3). COG categories N, W, and X were only identified in the species-specific accessory genome of *Janthinobacterium* sp. CG23_2.

The species-specific accessory genome of *Janthinobacterium* sp. CG23_2 contained 81% of the total number of CDS, which equates to 5,144 amino acid protein sequences, clustering into 3,793 orthologs. COGs with copy numbers $n \geq 10$ are summarized in Fig. 4. The highest percentage of the genes in these COG categories were associated with signal transduction

191 histidine kinases (n=123). Likewise, c-di-GMP synthetases (n=46) and c-di-GMP
192 phosphodiesterases (n=28), CheY-like receivers (n=43), and response regulators containing a
193 CheY-like receiver domain (n=84) were prevalent. Further, prominent COGs of the species-
194 specific accessory genome were transcriptional regulators (n=115) and transposases (n=27).
195 Numerous copies of short-chain alcohol dehydrogenases (n=21) and glutathione S-transferases
196 (n=15) were identified. Intercellular competition was mediated by Rhs family proteins (n=35;
197 COG3209). Motility was supported by methyl-accepting chemotaxis protein (n=29), different
198 chemotaxis proteins/regulators/signal transduction proteins (n=19), and a variety of proteins
199 involved in the assembly and function of pili and flagella (n=112). A total of 42 COGs, including
200 lysozymes (n=6), integrases (n=4), terminases (n=3), and genes encoding phage components
201 (n=13), were related to phage proteins.
202
203 Cold-adaptation
204 Cold-adaptation of the entire *Janthinobacterium* sp. CG23_2 genome was inferred from
205 substitution patterns across amino acids using the amino acid sequences of the other 35
206 *Janthinobacterium* spp. as comparative databases. Across the *Janthinobacterium* sp. CG23_2
207 genome, 27% (n=1,760) and 9% (n=577) of the amino acid sequences indicated cold-adaptation
208 when compared to the mesophilic (i.e. 31 *Janthinobacterium* spp.) and polar/glacial (i.e. 4
209 *Janthinobacterium* spp.) database, respectively. Noteworthy differences were found in the
210 number of amino acid sequences that were classified as neutral between *Janthinobacterium* sp.
211 CG23_2 and the two databases. The number of amino acid sequences with no significant
212 changes in the amino acid content for *Janthinobacterium* sp. CG23_2 and *Janthinobacterium*
213 spp. isolated from other polar/glacial environments was 2.3 times (n=1,687) higher compared to

214 the database built from *Janthinobacterium* spp. isolated from mesophilic environments. Overall,
215 *Janthinobacterium* sp. CG23_2 had significantly more proteins that possessed lower aliphacity,
216 R/K (arginine/lysine) ratios, and aromaticity when compared to their counterparts from
217 mesophilic environments (Bonferonni corrected $P \leq 0.001$; Fig 5A/C). Conversely, the GRAVY
218 index was found to be significantly enriched (Bonferonni corrected $P = 0.001$; Fig. 5A/C). When
219 compared to the four *Janthinobacterium* spp. isolated from polar/glacial environments, the
220 *Janthinobacterium* sp. CG23_2 genome was significantly cold-adapted for aliphacity while
221 enriched in proline and acidic residues (Bonferonni corrected $P \leq 0.005$; Fig. 5B/C).
222 COGs were identified for 90% (n=1583) of the cold adapted amino acid sequences in
223 *Janthinobacterium* sp. CG23_2. Cold-adapted proteins were associated with key processes
224 including transport, environmental sensing, locomotion, defense and stress response,
225 macromolecular syntheses, degradation, and repair, and key enzymes in central pathways and
226 their intermediates (Table 2). COGs in high abundance were ABC transporters (i.e. components
227 of the ATP binding cassette) involved in amino acid (n=19; COG0411, COG0683, COG0834,
228 COG1126, COG0765, COG4177, COG0559), multidrug (n=13; COG1131, COG1132), sugar
229 (n=10; COG1129, COG1653, COG3839, COG0395, COG1175), and
230 nitrate/sulfonate/bicarbonate (n=8; COG0715, COG1116, COG0600) transport. Multiple copies
231 of cold-adapted ABC transporters were found for antimicrobial peptides (n=7; COG0577,
232 COG1136) and organic solvent resistance (n=3; COG1127, COG2854, COG0767). Choline
233 dehydrogenases (n=2; COG2303) and the corresponding choline-glycine betaine ABC-type
234 transport system (COG1732) were found to be cold-adapted. Thirty-nine cold-adapted proteins
235 were characterized as outer membrane receptor proteins, mainly involved in iron transport
236 (n=28; COG1629). Cold-adapted efflux pumps such as arabinose efflux permeases (n=17;

COG2814), cation/multidrug efflux pumps (n=9; COG0841), and multidrug resistance efflux pumps (n=5; COG1566) were indicative of proteins engaged in transmembrane transport. Secretion of proteins into the extracellular space was represented by 28 copies of Type II secretory proteins (Table 2).

A large number of cold-adapted proteins (n=238) was associated with environmental sensing (Table 2). The genome of *Janthinobacterium* sp. CG23_2 contained multiple copies of CheY chemotaxis proteins (n=5; COG0784) and a methyl-accepting chemotaxis protein (n=16; COG0840). Major proteins involved in sensing of and adaptation to environmental signals were signal transduction histidine kinases (n=48; COG0642) and c-di-GMP synthetases (n=27; COG2199). Three proteins involved in the biosynthesis of histidine were identified as cold-adapted [imidazoleglycerol-phosphate dehydratase (COG0131), histidinol phosphatase (COG0241), and imidazolonepropionase (COG1228)]. Cellular response regulators were predominantly CheY-like receivers (n=67; COG3706, COG0745, COG2204, COG3437, COG2197, COG2201). Proteins affecting flagellar activity were associated with the motor (n=3; COG1291, COG1360, COG1536), basal body (n=14; COG2063, COG1766, COG1558, COG1706, COG1261, COG1815, COG1580), and hook (n=9; COG1344, COG1749, COG1256, COG1677) of the flagellar complex. Further, cold-adaptation included several proteins involved in flagellar biosynthesis (Table 2) and flagellar biosynthesis chaperones (COG2882, COG1516). Proteins related to environmental stress responses were primarily identified for detoxification (n=17; COG0625, COG0491, COG0346) and protection against oxidative stress (n=20; COG0753, COG1764, COG1225, COG0494, COG0225, COG1858, COG0386, COG0189, COG0695, COG3118, COG0783). Virulence factors such as RTX toxins (n=2; COG2931), Rhs family proteins regulating intercellular competition (n=5; COG3209), Type IV protein secretion

systems participating in virulence and antibacterial activities (n=5; COG3519, COG3157, COG3455, COG3523), and beta-lactamases providing antibiotic resistance (n=7; COG1680) were identified as potential defense mechanisms.

Janthinobacterium sp. CG23_2 had 24 and 18 cold-adapted COGs crucial for DNA replication and repair, respectively (Table 2). Overall, 113 proteins denoting 27 COGs were associated with transcriptional functions, dominated by transcriptional regulators (n=72; COG1309, COG1522, COG4977, COG3829, COG1167, COG2909, COG4650). Multiple copies of proteins responsible for methylation, a mechanism protecting newly synthesized DNA from endonucleases, were identified for the methylase of chemotaxis methyl-accepting proteins (n=3; COG1352) and polypeptide chain release factors (n=2; COG2890).

Several proteins involved in key metabolic steps of the citric acid cycle (TCA) were cold-adapted. These included pyruvate dehydrogenase (COG2609), citrate synthase (COG0372), isocitrate dehydrogenase (COG0473, COG0538), 2-oxoglutarate dehydrogenase (COG0508), and succinate dehydrogenase (COG1053, COG2009). Glucokinase (COG0837), the first step of glycolysis, and multiple copies of short-chain alcohol dehydrogenases (n=18; COG1028, COG4221) and lactate dehydrogenases (n=2; COG1052) both essential during fermentation were identified. Additionally, proteins supporting carboxydotrophy were cold-adapted (Table 2). Of relevance were proteins efficient in generating precursors or intermediates (e.g. galactose, propionate, glucose, ribulose-5-phosphate, fructose-6-phosphate, pyruvate, acytl-CoA, oxaloacetate) (Table 2) that can be used in catabolic and anabolic pathways to generate ATP or synthesize macromolecular subunits.

A subset of 48 cold-adapted proteins (Table 2) were involved in the biosynthesis of membrane constituents and included phosphoglycerides [e.g. glycerol-3-phosphate

dehydrogenase (n=2; COG0240)], phospholipids [phosphatidylserine synthases (n=4; COG1502), peptidoglycan [e.g. glycosyltransferase (n=6; COG0438), D-alanine-D-alanine ligase (n=3; COG1181)], fatty acids [e.g. acyl-CoA synthetases (n=4; COG0318), 3-oxoacyl-(acyl-carrier-protein) synthase (n=3; COG0304)], and lipopolysaccharides [e.g. sugar transferases involved in lipopolysaccharide synthesis (n=3; COG2148)]. *Janthinobacterium* sp. CG23_2 had cold-adapted phospholipase (COG3240), amidase (COG3023, COG0860), transglycosylase (COG0741), and peptidase (COG2173, COG1686) required for the continuous remodeling of cellular membranes. Thiol-disulfide isomerase (n=5; COG0526) and acetyltransferases (n=7; COG0456, COG1670), which catalyze protein folding and acetylation, respectively were among the cold-adapted proteins. Three out of the four proteins involved in beta-oxidation were cold-adapted and included multiple copies of acyl-CoA dehydrogenases (n=8; COG1960), enoyl-CoA hydratase/carnithine racemases (n=3; COG1024), and 3-hydroxyacyl-CoA dehydrogenases (n=3; COG1250).

Horizontal gene transfer

Putative HGT events across 34 *Janthinobacterium* spp. (excluding Marseille and B9-8) are summarized in Fig. 6. *Janthinobacterium* sp. CG23_2 had the highest number of horizontally transferred genes, with 11.5% of its protein coding genes predicted to be the result of HGT events (Fig. 6A). Notably, this was 8.8±1.6% above the average for the other *Janthinobacterium* strains. The predicted gain by HGT was 0.92 Mb in *Janthinobacterium* sp. CG23_2, which was substantially higher than the average of 0.17±0.11 Mb for the other *Janthinobacterium* strains (Fig. 6A). *Janthinobacterium* spp. that were isolated from the same site such as H100/H103, GW456P,W/GW460P,W, and 551a/344 showed almost identical HGT events (Fig. 6B). Genes

306 linked to HGT in *Janthinobacterium* sp. CG23_2 were predominately derived from
307 *Pseudomonadales* (17.1%), *Xanthomonadales* (9.6%), *Neisseriales* (7.3%), *Rhizobiales* (5.5%),
308 and *Nitrosomonadales* (4.8%), similar to the other *Janthinobacterium* species (Fig. 6B).
309 Specifically, *Pseudomonas* spp. (n=89), *Rugamonas rubra* (n=28), *Hyalangium minutum* (n=19),
310 and *Lysobacter dokdonensis* DS-58 (n=16) were dominant predicted gene donors.
311 Of the 741 putatively identified HGT events in *Janthinobacterium* sp. CG23_2, 292 or
312 4.5% of the whole genome of *Janthinobacterium* sp. CG23_2 clustered into COGs. COG
313 categories of genes involved in environmental sensing included signal transduction histidine
314 kinases (n=6; COG0642), second messengers (n=5; COG2199), and response regulators (n=3;
315 COG2197) (Table 3). Noticeable were functions associated with defense mechanisms,
316 production of compatible solutes, and the mobilome. These genes comprised Rhs family proteins
317 (n=13; COG3209), RTX toxins and related Ca^{2+}-binding proteins (n=7; COG2931), beta-
318 lactamase and other penicillin binding proteins (n=4, COG1680), choline dehydrogenases (n=4;
319 COG2303), phage proteins (n=18), and transposases (n=6) (Table 3). Other abundant HGT-
320 acquired genes included short-chain alcohol dehydrogenases (n=8; COG1028, COG4221) and
321 mannose-6-phosphate isomerases (n=4; COG0662). HTG was predicted for genes encoding for
322 the biosynthesis of peptidoglycan (glycosyltransferase; n=5; COG0438) and fatty acids (3-
323 oxoacyl-(acyl-carrier-protein) synthase; n=4; COG0304). Acetyltransferases (n=6; COG0456,
324 COG1670) with relevance to protein modifications and proteins containing pentapeptide repeats
325 (n=6; COG1357) were of HGT origin.
326
327 **Discussion**

328 Unlike core genomes, which may consist of conserved genes essential to the lifestyle of
329 specific taxonomic groups, the accessory genome is more likely subject to genome evolution and
330 provides selective advantage under specific environmental conditions [45]. Only 18.7% of all
331 CDS identified in the *Janthinobacterium* sp. CG23_2 genome matched protein sequences to one
332 or more of the other 35 *Janthinobacterium* spp. genomes. Further, with merely 164 CDS being
333 identified in both the *Janthinobacterium* sp. CG23_2 genome and the core gene set of 30
334 *Janthinobacterium* spp., these results reinforce the importance of a species-specific accessory
335 genome and the genomic variability of *Janthinobacterium* sp. CG23_2. While it should be noted
336 that gene duplication was not determined, putatively identified HGT events alone increase the
337 genome size of *Janthinobacterium* sp. CG23_2 by 0.92 Mb. HGT events were mainly associated
338 with signal transduction histidine kinases, second messengers, response regulators, and functions
339 linked to defense/stress mechanisms (Table 3); all advantageous traits for survival and adaptation
340 to extreme environments (discussed below). HGT is made possible primarily by the mobilome,
341 including transposons and bacteriophages. Notable was the occurrence of 27 transposases in the
342 species-specific accessory genome of *Janthinobacterium* sp. CG23_2, which catalyze the
343 rearrangement or transfer of mobile genetic elements (i.e. transposons) within or between cells
344 [46]. In a meta-analysis of 384 bacterial genomes, Newton and Bordenstein [47] determined that
345 up to ~6% of bacterial genomes could be the result of bacteriophage genes. These authors also
346 established a correlation between larger genome sizes and an increase in the number of
347 bacteriophage genes. While the *Janthinobacterium* sp. CG23_2 genome is by far the largest
348 genome of the 36 *Janthinobacterium* species investigated, bacteriophage genes account for only
349 0.6% of its gene content. Smith et al. [48] reported virus to bacterium ratios ranging from 0.12-
350 0.44 for the Cotton Glacier stream, 10-1000-fold lower compared to other polar inland waters

[49]. Such low viral abundance in the Cotton Glacier stream may have limited the integration of phage genes into the bacterial chromosome of *Janthinobacterium* sp. CG23_2.

The species-specific accessory genome of *Janthinobacterium* sp. CG23_2 is dominated by functions associated with environmental signaling and transcriptional regulation (Fig. 4). While both functions are predominant in the core genome of the genus *Janthinobacterium* [22], their enrichment in the species-specific accessory genome of *Janthinobacterium* sp. CG23_2 underscores their role in the adaptation to life in an ephemeral supraglacial Antarctic stream. Moreover, the importance of environmental sensing and orchestrating gene expression were firmly established in the cold-adaptation patterns of amino acid sequences (Table 2). Of relevance were gene categories related to signal transduction histidine kinases and response regulators containing CheY-like receivers. Histidine kinases and response regulators are the building blocks of the two-component signal transduction system, enabling an adaptive response to environmental stimuli (e.g. changes in pH and osmolarity levels, thermal and oxidative stress, light, nutrients and metal ions, and antimicrobials) mainly through gene expression [50]. Moreover, histidine kinases play a central role in the signal integration of the bacterial chemotaxis pathway, where auto-phosphorylated substrates transfer the phosphoryl group to CheY (CheY-P) [51]. Subsequently, the diffusible response regulator CheY-P interacts with the flagellar motor and reverses the rotation of flagella [51]. In line with these findings, the species-specific accessory genome of *Janthinobacterium* sp. CG23_2 possesses chemotaxis genes for sensing environmental cues and the movement towards factors that favor survival. Methyl-accepting chemotaxis proteins were the predominant chemoreceptors; proteins that are involved in biofilm formation and exopolysaccharide production, flagellum biosynthesis, degradation of xenobiotic compounds, and the production of toxins [52]. In addition to chemotaxis proteins, the

presence of c-di-GMP phosphodiesterases and c-di-GMP synthetases suggests the possibility of reciprocal interactions between different chemosensory systems. c-di-GMP, a second messenger, inhibits the methyltransferase activity of methyl-accepting chemotaxis proteins. Ultimately, this modulation affects the phosphorylation of the CheY-like proteins and chemotactic responses [53]. As such, the c-di-GMP signaling system regulates the transition between motile-sessile states [54], lifestyle switches that enhance adaptation to fluctuations in the environment [55]. The species-specific accessory genome of *Janthinobacterium* sp. CG23_2 is equipped with gene categories associated with flagellar biosynthesis, basal body, hook, and motor proteins (n=80) as well as pilus assembly proteins (n=32). While the latter aids the adhesion of a bacterial cell to surfaces, flagella permit chemotaxis-navigated motility systems that allow for active locomotion. Temperature, osmolarity, pH, and nutrient concentration can trigger the expression of the flagellar master operon, which facilitates switching between a motile and sessile state [56]. With their involvement in detecting wetness [57], flagella participate collectively in the sensing of environmental conditions crucial for successful propagation in a supraglacial stream.

The genome composition of *Janthinobacterium* sp. CG23_2 revealed temperature and oxidative stress as major environmental challenges associated with a supraglacial stream environment. Overall, 1,760 and 577 amino acid sequences in the *Janthinobacterium* sp. CG23_2 genome were predicted to be cold-adapted when compared to the *Janthinobacterium* species isolated from mesophilic and polar/glacial habitats, respectively. Both the increased levels of UV radiation above the Antarctic Ice Sheet and low temperatures can lead to the formation of reactive oxygen species, posing a lethal threat to bacterial cells. Protection against oxidative damage in the genome of *Janthinobacterium* sp. CG23_2 included genes such as catalases, hydroperoxide reductases, peroxiredoxins, cytochrome c peroxidases, glutathione

peroxidases, glutaredoxins, thioredoxin reductases, many of which were cold-adapted (Table 2). The species-specific accessory genome of *Janthinobacterium* sp. CG23_2 also contains 15 copies of cold-adapted glutathione S-transferases. Not only do bacterial glutathione transferases provide protection against oxidative stresses, they also play a key role in cellular detoxification including processes such as the biodegradation of xenobiotics and antimicrobial drug resistance [58]. Further, choline dehydrogenases (n=7) were found in the species-specific accessory genome of *Janthinobacterium* sp. which oxidize the first of the two enzymatic steps in the production of glycine-betaine [59]. This compatible solute is a known cryoprotectant and osmolyte and is believed to prevent cold induced aggregation of proteins and maintain membrane fluidity [59,60]. In addition to genes coping with environmental stresses, Rhs protein families (n=35) are included in the species-specific accessory genome of *Janthinobacterium* sp. CG23_2. Rhs proteins are part of a contact-dependent growth inhibition system. Intercellular competition is mediated by injecting toxins that inhibit the growth of neighboring cells [61]; thereby, providing a competitive advantage in a low-nutrient environment such as the Cotton Glacier stream [29].

For bacteria to sense and adapt to their ever-changing environment, modifications in signaling and gene regulation pathways are essential [62]; and, by implication, genotypic selection would depend on the complexity of the environment. Clearly, the temporal heterogeneity of the supraglacial Cotton Glacier stream, Antarctica, poses challenges for its microbial inhabitants within the time frame of both a single and multiple generations. A major survival advantage in this fluctuating environment would be the ability to anticipate changes in the environment [63,64]. Investigations by Mitchell et al. [65], for instance, showed that by using heat shock as the early stimulus, certain bacterial or yeast cells gained protection against stresses

to come (e.g. oxidative stress, oxygen depletion). Similarly, a specific response to one stress could increase the resistance to another [66]. While experimental evidence for anticipating stressors or the physiological cross-protection to secondary stresses were beyond the scope of the present study, the genome of *Janthinobacterium* sp. CG23_2 is well equipped with ample genes associated with environmental sensing related functions, transcription regulators, and stress response.

In the context of the Black Queen Hypothesis, cells can evolve in two ways, by either loosing gene functions and mutually depending on other members of a community or by retaining large genomes expressing many genes that are not essential to central metabolism, growth, and reproduction [30]. Although the latter would seem energetically unfavorable in an extreme environment such as the Cotton Glacier stream, *Janthinobacterium* sp. CG23_2 has evolved through genome plasticity (i.e. horizontal gene transfer and transposase activity); features that have been suggested to enable adaptation to life in cold environments [67,68]. Whether this gene acquisition, however, represents a common trend in the adaptation to the Cotton Glacier stream environment or *Janthinobacterium* sp. CG23_2 acquired a key status as a function-performing helper within the microbial community according to the Black Queen Hypothesis invite further studies on the network of interactions between co-occurring organisms and their genome evolution. Both whole genome *in silico* DDH (21.5±1.8%) and ANI (79.3±0.4%) qualified well below the cut-off value for species boundaries [69]. These results agreed with the distant branching of *Janthinobacterium* sp. CG23_2 within the Maximum Likelihood tree (Fig. 1). Based on these molecular and phylogenetic indicators, *Janthinobacterium* sp. CG23_2 appears sufficiently different to constitute a separate species. The

new species *Janthinobacterium cottonii* is proposed. Additional taxonomic studies will help in placing *Janthinobacterium cottonii* within the genus *Janthinobacterium*.

Supplementary Materials:

Supplemental Table 1: Pairwise comparisons of coding domain sequences (CDS) identified in 36 *Janthinobacterium* spp. genomes

Supplemental Table 2: List of clusters of orthologous groups (COGs)

Acknowledgments:

We thank R. Mueller at Montana State University for guidance on bioinformatics analyses. This work was supported by the National Science Foundation under grant ANT-0838970 to CMF. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Author Contributions:

H.J.S. and C.M.F. conceived the study. M.D., H.J.S., T.R., and C.M.F. developed the analytical frame work. M.D. and T.R. analyzed the data. M.D., H.J.S., T.R., and C.M.F. wrote the paper.

Conflicts of Interest:

The authors declare no conflict of interest.

Figure Legends

Figure 1: Maximum Likelihood tree of 36 *Janthinobacterium* spp. based on the JTT matrix-based model. The tree with the highest log likelihood (-41655.66) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. Circles indicate bootstrap values ≥ 0.9.

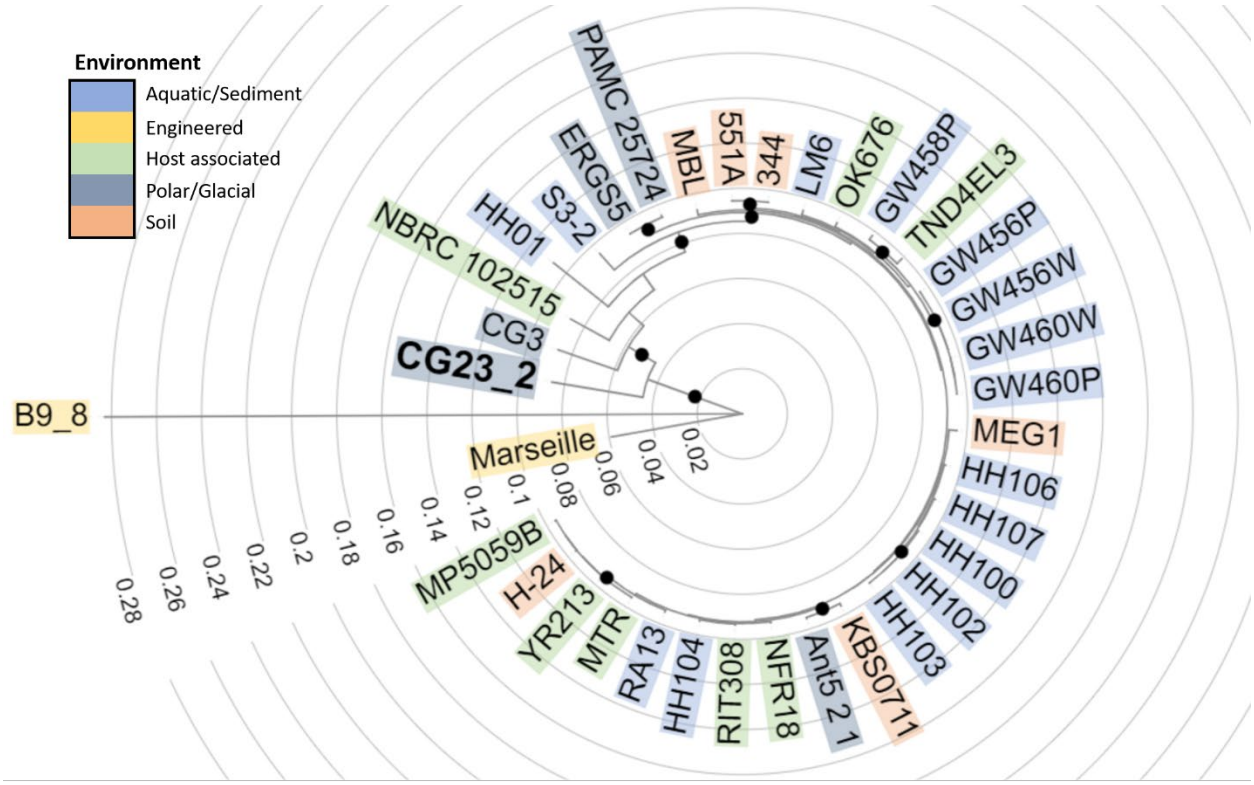
Figure 2: Pairwise gene comparison between *Janthinobacterium* sp. CG23_2 and the 35 *Janthinobacterium* spp. with a 90% query coverage and an 80% identity. The total number of genes from CG23_2 is in the figure center. Each petal depicts the percentage of shared genes (bold) and the total number of predicted protein coding sequences (in parentheses) for each strain.

Figure 3: Relative abundance of functional classification of annotated protein coding genes normalized to the total number of protein coding genes for the core genome, shared between *Janthinobacterium* sp. CG23_2 and n ≥ 2 species/strains, specific to *Janthinobacterium* sp. CG23_2, horizontally transferred genes, and cold adapted proteins. Protein coding genes lacking specific functional assignments were excluded.

Figure 4: *Janthinobacterium* sp. CG23_2 species-specific COG categories. Only COGs with n ≥ 10 copy numbers are shown.

Figure 5: Genome wide molecular cold-adaptation in CG23_2 compared to (A) 31 *Janthinobacterium* spp. isolated from mesophilic environments and (B) four *Janthinobacterium* spp. found in polar /glacial regions. (C) Adaptation ratio decreased:enriched with significant indices indicated with an asterisk (Bonferonni corrected $P \leq 0.005$). Decrease indicates cold-adaptation.

Figure 6: (A) Percentage of putatively horizontally transferred genes identified for the 34 *Janthinobacterium* spp. (B) Relative abundance of predicted HGT donors based on the twenty most abundant orders identified in *Janthinobacterium* sp. CG23_2. Data are presented at the order level Shaded colors indicate the five environments. Arrows mark *Janthinobacterium* sp. CG23_2.



499

500 Figure 1

501

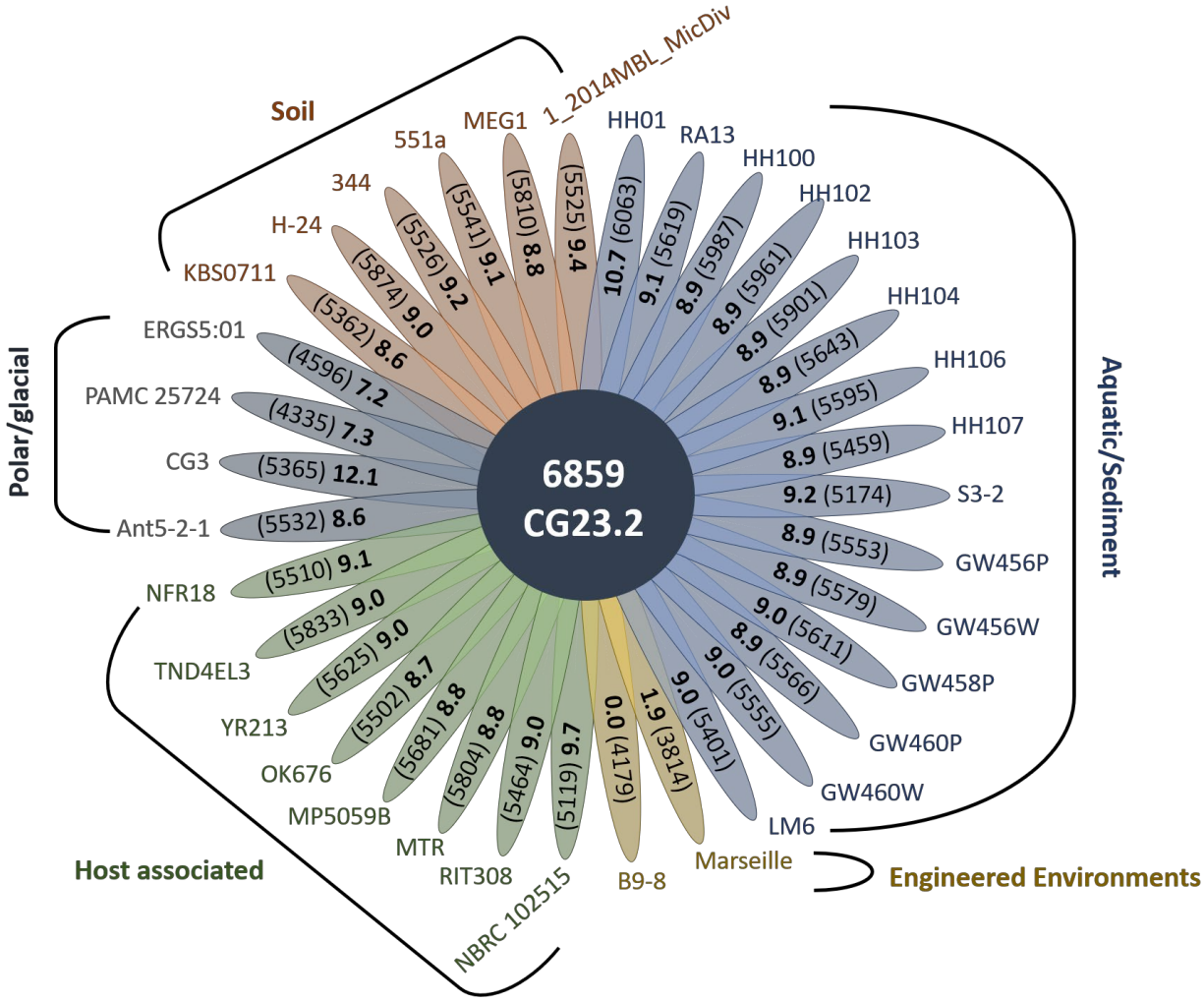

6859
CG23.2
Soil
Polar/glacial
Host associated
Aquatic/Sediment
Engineered Environments
1_2014MBL_MicDiv
MEG1
551a
344
H-24
KBS0711
ERGS5:01
PAMC 25724
CG3
Ant5-2-1
NFR18
TND4EL3
YR213
OK676
MP5059B
MTR
RIT308
NBRC 102515
B9-8
Marseille
LM6
GW460W
GW460P
GW458P
GW456W
GW456P
S3-2
HH107
HH106
HH104
HH103
HH102
HH100
RA13
HH01
(5525) 9.4
(5810) 8.8
(5541) 9.1
(5526) 9.2
(5874) 9.0
(5362) 8.6
(4596) 7.2
(4335) 7.3
(5365) 12.1
(5532) 8.6
(5510) 9.1
(5833) 9.0
(5625) 9.0
(5502) 8.7
(5681) 8.8
(5804) 8.8
(5464) 9.0
(5119) 9.7
0.0 (4179)
1.9 (3814)
9.0 (5401)
9.0 (5555)
8.9 (5566)
9.0 (5611)
8.9 (5579)
8.9 (5553)
9.2 (5174)
8.9 (5459)
9.1 (5595)
8.9 (5643)
8.9 (5901)
8.9 (5961)
8.9 (5987)
9.1 (5619)
10.7 (6063)

Figure 2

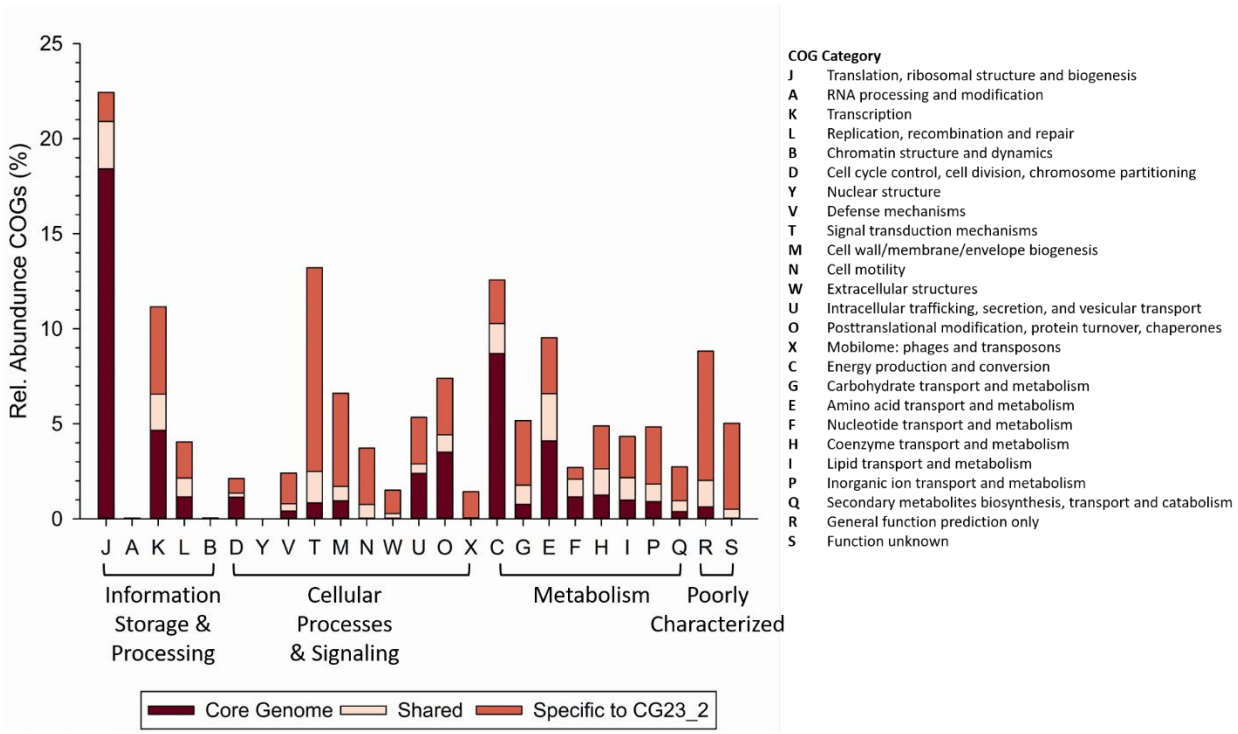
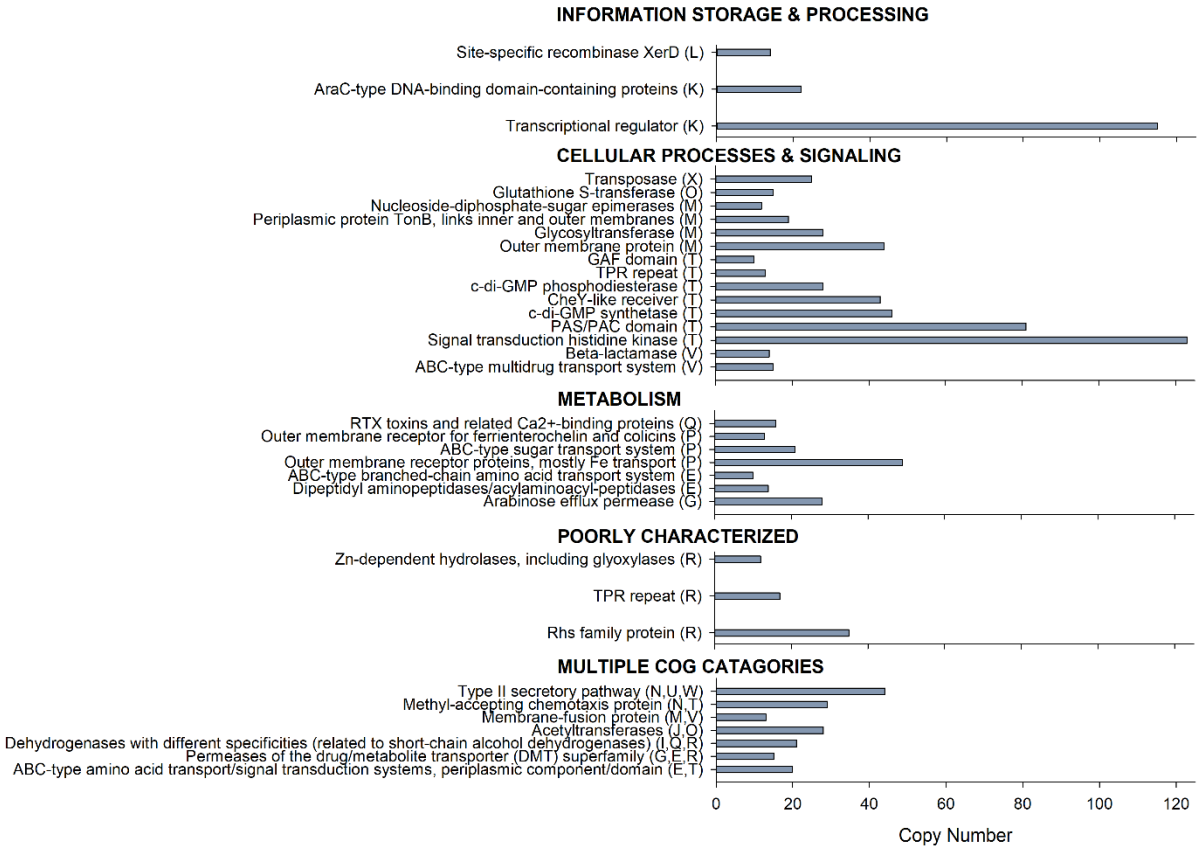

Rel. Abundance COGs (%)
25
20
15
10
5
0
J A K L B D Y V T M N W U O X C G E F H I P Q R S
Information Storage & Processing
Cellular Processes & Signaling
Metabolism
Poorly Characterized
Core Genome
Shared
Specific to CG23_2
COG Category
J Translation, ribosomal structure and biogenesis
A RNA processing and modification
K Transcription
L Replication, recombination and repair
B Chromatin structure and dynamics
D Cell cycle control, cell division, chromosome partitioning
Y Nuclear structure
V Defense mechanisms
T Signal transduction mechanisms
M Cell wall/membrane/envelope biogenesis
N Cell motility
W Extracellular structures
U Intracellular trafficking, secretion, and vesicular transport
O Posttranslational modification, protein turnover, chaperones
X Mobilome: phages and transposons
C Energy production and conversion
G Carbohydrate transport and metabolism
E Amino acid transport and metabolism
F Nucleotide transport and metabolism
H Coenzyme transport and metabolism
I Lipid transport and metabolism
P Inorganic ion transport and metabolism
Q Secondary metabolites biosynthesis, transport and catabolism
R General function prediction only
S Function unknown

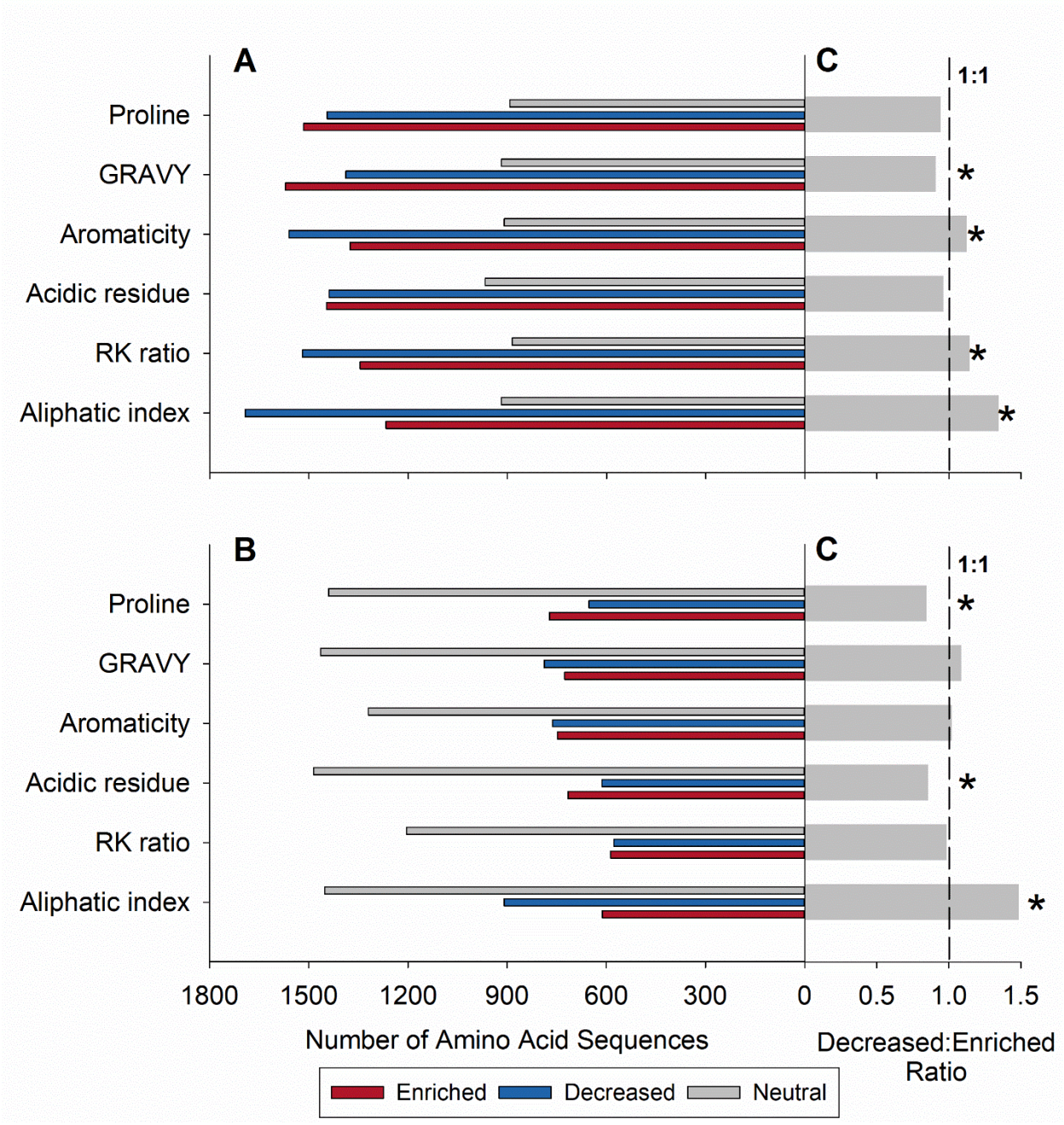
Figure 3



511

512 Figure 4

513



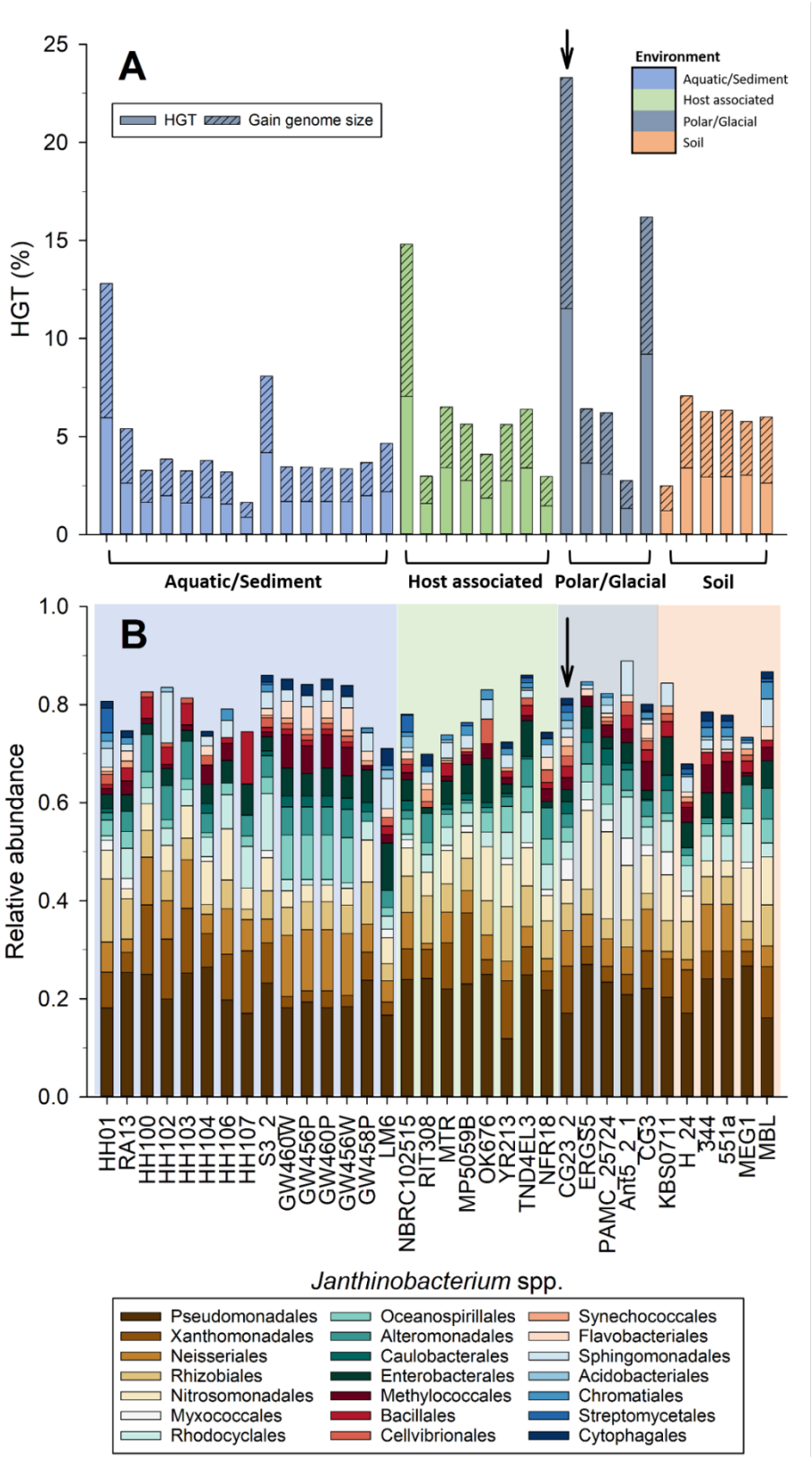
514

515 Figure 5

516

517

518



519

520 Figure 6

Table 1: Summary of the 36 *Janthinobacterium* genome statistics.

Genome name	Isolation source	Size (Mb)	# contigs	N50 (Mb)	Total # Genes	# Protein coding seqs	# 16S rRNA	GC %	NCBI ID (GCA_)
***Janthinobacterium* sp. CG23_2**	**polar, glacial**	**7.85**	**4**	**4.20**	**6,974**	**6,859**	**7**	**63.5**	**001485665.1**
J. sp. HH01	aquatic, sediments	7.11	2	4.21	6,172	6,063	7	64.2	000335815.1
J. sp. RA13	aquatic, sediments	6.42	1	6.42	5,740	5,619	8	62.5	000745325.1
J. sp. HH100	aquatic, sediments	6.70	150	0.09	6,044	5,987	1	62.6	001758685.1
J. sp. HH102	aquatic, sediments	6.65	125	0.31	6,010	5,961	10	62.4	001758625.1
J. sp. HH103	aquatic, sediments	6.58	141	0.10	5,964	5,901	1	62.5	001758705.1
J. sp. HH104	aquatic, sediments	6.39	65	0.23	5,701	5,643	1	62.6	001758715.1
J. sp. HH106	aquatic, sediments	6.31	73	0.19	5,651	5,595	1	62.9	001758725.1
J. sp. HH107	aquatic, sediments	6.16	116	0.11	5,517	5,459	1	63.1	001758765.1
J. psychrotolerans sp. S3-2	aquatic, sediments	5.84	62	0.26	5,265	5,174	8	63.0	001677885.1
J. sp. GW456P	aquatic, sediments	6.27	92	0.13	5,633	5,553	7	62.9	002127655.1
J. sp. GW456W	aquatic, sediments	6.26	149	0.08	5,660	5,579	7	62.9	002127615.1
J. sp. GW458P	aquatic, sediments	6.28	157	0.08	6,584	5,611	4	63.3	002127585.1
J. sp. GW460P	aquatic, sediments	6.27	120	0.09	5,650	5,566	8	62.9	002127625.1
J. sp. GW460W	aquatic, sediments	6.27	104	0.11	5,637	5,555	7	62.9	002127575.1
J. sp. LM6	aquatic, sediments	6.29	1	6.29	5,523	5,401	8	63.0	002002885.1
J. agaricidamnosum NBRC 102515	host associated	5.95	1	5.95	5,204	5,119	1	61.1	000723165.1
J. lividum RIT308	host associated	6.21	44	0.30	5,554	5,464	4	62.8	000632025.1
J. lividum MTR	host associated	6.54	144	0.15	5,874	5,804	3	62.4	000783415.1
J. sp. MP5059B	host associated	6.46	25	0.38	5,775	5,681	6	62.7	001758645.1
J. sp. OK676	host associated	6.27	35	0.35	5,589	5,502	2	62.8	900103595.1
J. sp. YR213	host associated	6.42	16	0.84	5,753	5,625	6	62.7	900099875.1
J. sp. TND4EL3	host associated	6.52	78	0.28	5,915	5,833	6	62.9	900156175.1
J. lividum NFR18	host associated	6.30	26	0.50	5,592	5,510	2	62.5	900119665.1
J. sp. Ant5-2-1	polar, glacial	6.20	1703	0.003	5,597	5,532	8	62.5	001445815.1
J. sp. CG3	polar, glacial	6.27	41	0.78	5,460	5,365	3	65.5	000344615.1
J. lividum PAMC 25724	polar, glacial	4.98	48	0.25	4,428	4,335	9	60.6	000242815.2
J. lividum ERGS5:01	polar, glacial	5.17	16	3.37	4,715	4,596	8	60.5	001678745.2
J. sp. KBS0711	soil	6.07	149	0.09	5,438	5,362	1	62.7	000988085.1
J. lividum H-24	soil	6.71	125	0.12	5,916	5,874	4	62.4	001758635.1
J. sp. 344	soil	6.44	35	0.43	5,612	5,526	2	63.7	900112025.1
J. sp. 551a	soil	6.46	38	0.51	5,626	5,541	1	63.6	900103675.1
J. lividum MEG1	soil	6.60	16	1.11	5,893	5,810	1	62.3	001854915.1
J. sp. 1_2014MBL_MicDiv	soil	6.45	1	6.45	5,648	5,525	8	63.6	001865675.1
J. sp. Marseille	Engineered systems	4.11	1	4.11	3,870	3,814	2	54.2	000013625.1
J. sp. B9-8	Engineered systems	4.73	1	4.73	4,295	4,179	10	48.7	000969645.2

Table 2: Function and COG category for cold adapted amino acid sequences in *Janthinobacterium* sp. CG23_2. (For COG ID description see Table S2).

Function	COG ID and (#) of amino acid sequences
TRANSPORT	
ABC transporters (n=84)	1131 (11), 0411 (7), 1136 (4), 0683 (4), 0834 (3), 0577 (3), 1116 (3), 0600 (3), 1653 (3), 3839 (3), 0488 (3), 0715 (2), 1132 (2), 1126 (2), 3842 (2), 1129 (2), 0555 (2), 0765 (1), 2274 (1), 4177 (1), 1120 (1), 0609 (1), 1135 (1), 1464 (1), 4148 (1), 1668 (1), 4555 (1), 1117 (1), 0581 (1), 3639 (1), 4618 (1), 1176 (1), 1177 (1), 0395 (1), 1175 (1), 5265 (1), 1127 (1), 2854 (1), 0767 (1), 1732 (1), 1172 (1), 0559 (1)
Efflux pumps (n=31)	2814 (16), 0841 (9), 1566 (5), 1230 (1)
Symporters (n=4)	2211 (3), 1301 (1)
Outer membrane proteins (n=30)	2885 (9), 1538 (6), 3203 (5), 3713 (3), 3133 (2), 3047 (2), 3248 (1), 3017 (1), 2913 (1)
Outer membrane receptors (n=45)	1629 (28), 4771 (7), 0810 (6), 4774 (3), 4773 (1)
Secretion (n=27)	2165 (7), 1459 (6), 2804 (4), 3267 (3), 4796 (3), 1450 (2), 4795 (1), 1989 (1)
Membrane proteins (n=21)	0845 (8), 5009 (2), 1477 (2), 4942 (2), 0744 (1), 4953 (1), 0481 (1), 4260 (1), 0330 (1), 1585 (1), 2244 (1)
Electron transfer (n=29)	0604 (5), 0277 (3), 2863 (2), 2146 (2), 2010 (1), 2193 (2), 1271 (1), 1294 (1), 4654 (1), 3175 (1), 2124 (1), 0437 (1), 0633 (1), 0056 (1), 0055 (1), 0355 (1), 0356 (1), 0711 (1), 0655 (1)
Other transporters (n=20)	0697 (5), 0475 (3), 0531 (2), 2217 (2), 0848 (2), 0811 (2), 1638 (1), 3090 (1), 0004 (1), 0735 (1)
ENVIRONMENTAL SENSING	
Chemotaxis/chemosensory (n=26)	0840 (16), 0784 (5), 0643 (3), 1776 (1), 3143 (1)
Signaling (n=89)	0642 (48), 2199 (27), 4191 (4), 2202 (4), 1956 (1), 5000 (1), 0835 (1), 3852 (1), 2206 (1), 5581 (1)
Response regulator (n=72)	0745 (21), 3706 (19), 2204 (12), 2197 (7), 3437 (5), 2201 (3), 3279 (3), 3707 (2)
Locomotion-flagellar (n=26)	1344 (6), 2063 (3), 1766 (3), 1558 (3), 1706 (2), 1261 (1), 1815 (1), 1580 (1), 1749 (1), 1256 (1), 1677 (1), 1291 (1), 1360 (1), 1536 (1)
Pilus (n=25)	5008 (3), 3745 (2), 4965 (2), 2064 (2), 4964 (2), 3063 (2), 4972 (2), 4961 (1), 4970 (1), 4963 (1), 3170 (1), 4968 (1), 3167 (1), 3168 (1), 4726 (1), 2805 (1), 3419 (1)
DEFENSE (n=21)	3209 (5), 1680 (7), 2931 (2), 0515 (2), 3519 (2), 3157 (1), 3455 (1), 3523 (1)
STRESS RESPONSE (n=57)	0625 (7), 0664 (5), 0491 (5), 0346 (5), 0457 (3), 0753 (3), 1764 (3), 1225 (3), 0790 (3), 0494 (3), 0589 (2), 0225 (2), 1278 (1), 1825 (1), 1188 (1), 3109 (1), 1858 (1), 3210 (1), 0386 (1), 0189 (1), 0695 (1), 3118 (1), 0492 (1), 0631 (1), 0783 (1)
DNA/RNA/REPAIR/CELL DEVISION	
Replication (n=24)	0513 (4), 4974 (3), 0305 (2), 0188 (2), 0514 (2), 0749 (1), 0587 (1), 0847 (1), 0358 (1), 4973 (1), 1961 (1), 0187 (1), 0553 (1), 0593 (1), 0272 (1), 0632 (1)
Repair (n=18)	0467 (2), 0708 (2), 0420 (1), 1381 (1), 0210 (1), 1112 (1), 1197 (1), 0692 (1), 1573 (1), 0389 (1), 0249 (1), 3663 (1), 1864 (1), 0178 (1), 2818 (1), 0122 (1)
Transcription/translation (n=110)	1309 (40), 1522 (21), 2207 (7), 4977 (6), 0568 (3), 1943 (3), 2916 (3), 1191 (2), 1595 (2), 3829 (2), 1167 (2), 0480 (2), 0060 (2), 0008 (2), 1741 (2), 2909 (1), 3327 (1), 3901 (1), 0215 (1), 0042 (1), 4650 (1), 0154 (1), 0324 (1), 0430 (1), 0242 (1), 0776 (1)
Post-transcription (n=9)	0564 (2), 1187 (2), 0130 (1), 0030 (1), 0144 (1), 0566 (1), 2501 (1)
Degradation (n=10)	0084 (2), 1530 (2), 0507 (1), 1875 (1), 1295 (1), 0349 (1), 0328 (1), 0164 (1)
Methylation (n=7)	1352 (3), 2890 (2), 0350 (1), 0863 (1)
Cell division (n=13)	3116 (4), 0489 (3), 0424 (2), 1475 (2), 0772 (1), 2891 (1)
CHAPERONS (n=12)	0542 (3), 0443 (3), 0265 (1), 0071 (1), 0576 (1), 1281 (1), 2882 (1), 1516 (1)
METABOLISM	
TCA cycle (n=9)	0508 (2), 0372 (1), 0473 (1), 0538 (1), 1249 (1), 1053 (1), 2009 (1), 2609 (1)
Calvin cycle/PPP (n=1)	0120 (1)
Glycolysis (n=1)	0837 (1)
Fermentation (n=23)	1028 (15), 4221 (3), 1052 (2), 1064 (2), 1062 (1)
Aerobic carboxidotrophy (n=5)	2080 (2), 1529 (2), 1319 (1)
Precursor/intermediates (n=33)	0451 (5), 3386 (3), 3338 (3), 0183 (3), 0663 (2), 2133 (2), 1171 (2), 0662 (1), 0686 (1), 0221 (1), 0836 (1), 0800 (1), 1607 (1), 1472 (1), 1087 (1), 0449 (1), 3265 (1), 2065 (1), 0176 (1), 2301 (1)

Table 2: Continued

Function	COG ID and (#) of amino acid sequences
BIOSYNTHESIS	
Cell wall	
Phosphoglycerides (n=8)	0584 (2), 0240 (2), 0204 (2), 0818 (1), 0554 (1)
Phospholipids (n=8)	1502 (4), 3540 (2), 0575 (1), 0688 (1)
Peptidoglycan (n=13)	0438 (6), 1181 (3), 0812 (1), 0796 (1), 0773 (1), 2348 (1)
Fatty acids (n=15)	0318 (4), 0304 (3), 0331 (2), 0365 (1), 0236 (1), 1043 (1), 2030 (1), 4281 (1), 0427 (1)
Lipopolysaccharides (n=4)	2148 (3), 1109 (1)
Cell wall turnover (n=8)	0741 (2), 1686 (2), 3240 (1), 3023 (1), 0860 (1), 2173 (1)
Proteins	
Proteins (n=6)	1020 (1), 1186 (1), 0682 (1), 1952 (1), 0690 (1), 1862 (1)
Protein folding (n=7)	0526 (5), 1651 (1), 4232 (1)
Post-translational modification (n=10)	0456 (4), 1670 (3), 1182 (2), 1247 (1)
Amino acids (n=20)	0367 (2), 0031 (2), 1246 (2), 1045 (2), 0241 (2), 1228 (2), 0065 (1), 2008 (1), 1794 (1), 0436 (1), 0119 (1), 0135 (1), 2309 (1), 0131 (1)
Nucleosides (n=7)	0213 (1), 0125 (1), 0207 (1), 0104 (1), 0299 (1), 0107 (1), 0134 (1)
Flagellar (n=7)	1298 (2), 1684 (2), 1377 (1), 1987 (1), 1157 (1)
Ribosomes (n=6)	1358 (1), 0244 (1), 0256 (1), 0335 (1), 0052 (1), 0049 (1)
Vitamins (n=7)	0161 (3), 0352 (1), 0028 (1), 0054 (1), 0414 (1)
Glutamate and glutamine (n=7)	0493 (2), 0067 (1), 0347 (1), 0174 (1)
Molybdenum cofactor (n=8)	2896 (3), 0303 (1), 0521 (1), 0314 (1), 1763 (1), 0746 (1)
Ubiquinone (n=10)	0654 (4), 2226 (4), 2227 (2)
DEGRADATION	
Beta oxidation (n=14)	1960 (8), 1024 (3), 1250 (3)
Aldehydes (n=6)	1012 (6)
Proteins (n=9)	1506 (5), 0616 (2), 0265 (2)
Glycine (n=4)	0665 (4)
Organic acids (n=3)	0657 (3)
NUTRIENTS	
Sulphur (n=7)	0607 (4), 2041 (1), 0369 (1), 4117 (1)
Nitrogen (n=3)	1140 (1), 2180 (1), 3256 (1)

Table 3: Function and COG category for abundant putatively identified HGT events in *Janthinobacterium* sp. CG23_2. (For COG ID description see Table S2).

Function	COG ID and (#) of amino acid sequences
TRANSPORT	
ABC transporters (n=5)	ABC-type amino acid transport/signal transduction systems, periplasmic component/domain: COG0834 (5)
Outer membrane proteins (n=4)	Choline dehydrogenase and related flavoproteins: COG2303 (4)
ENVIRONMENTAL SENSING	
Signaling (n=11)	Signal transduction histidine kinase: COG0642 (6)
	c-di-GMP synthetase (diguanylate cyclase, GGDEF domain): COG2199 (5)
Response regulator (n=3)	Response regulator containing a CheY-like receiver domain and an HTH DNA-binding domain: COG2197 (3)
DEFENSE (n=27)	RTX toxins and related Ca^{2+}-binding proteins: COG2931 (7)
	Rhs family protein: COG3209 (13)
	Beta-lactamase class C and other penicillin binding proteins: COG1680 (4)
MOBILOME (n=24)	Phage proteins: COG3497 (3), COG3772 (2), COG4626 (2), COG4695 (2), COG3500 (1), COG3561 (1), COG3628 (1), COG3645 (1), COG3646 (1), COG3740 (1), COG3948 (1), COG4653 (1), COG5362 (1)
	Transposase: COG2963 (2), COG3666 (1), COG3335 (1), COG3415 (1), COG2801 (1)
DNA/RNA/REPAIR (n=27)	Superfamily I DNA and RNA helicases: COG0210 (3)
	Transcriptional regulators: COG1309 (5), COG0583 (5), COG1846 (2), COG1609 (1)
	AraC-type DNA-binding domain-containing proteins: COG2207 (5)
METABOLISM	
Fermentation (n=8)	Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases): COG1028 (6)
	Short-chain alcohol dehydrogenase of unknown specificity: COG4221 (2)
Precursor/intermediates (n=4)	Mannose-6-phosphate isomerase: COG0662 (4)
BIOSYNTHESIS	
Cell wall	
Peptidoglycan (n=5)	Glycosyltransferase: COG0438 (5)
Fatty acids (n=4)	3-oxoacyl-(acyl-carrier-protein) synthase: COG0304 (4)
Proteins (n=12)	Pentapeptide repeats containing protein: COG1357 (6)
	Acetyltransferases: COG0456 (3)
	Acetyltransferases, including N-acetylases of ribosomal proteins: COG1670 (3)

References

1. Benn, D.; Evans, D.J. *Glaciers and glaciation*. Arnold: London, England, 1998.

2. Goordial, J.; Lamarche-Gagnon, G.; Lay, C.Y.; Whyte, L. Left out in the cold: life in cryoenvironments. In *Polyextremophiles, life under multiple forms of stress*; Seckbach, J., Oren, A., Stan-Lotter, H., Eds.; Springer: Dordrecht, Netherlands, 2013, Volume 27, pp 335-363.

3. Pfeffer, W.T.; Arendt, A.A.; Bliss, A.; Bolch, T.; Cogley, J.G.; Gardner, A.S.; Hagen, J.O.; Hock, R.; Kaser, G.; Kienholz, C.; Miles, E.S.; Moholdt, G.; Mölg, N.; Paul, F.; Radić, V.; Rastner, P.; Raup, B.H.; Rich, J.; Sharp, M.J.; The Randolph Consortium. The Randolph Glacier Inventory: a globally complete inventory of glaciers. *J Glaciol* **2014**, *60*, 537-552.

4. De Maayer, P.; Anderson, D.; Cary, C.; Cowan, D.A. Some like it cold: understanding the survival strategies of psychrophiles. *EMBO Rep* **2014**, *15*, 508-517.

5. Casanueva, A.; Tuffin, M.; Cary, C.; Cowan, D.A. Molecular adaptations to psychrophily: the impact of 'omic'technologies. *Trends Microbiol* **2010**, *18*, 374-381.

6. Goordial, J.; Raymond-Bouchard, I.; Zolotarov, Y.; de Bethencourt, L.; Ronholm, J.; Shapiro, N.; Woyke, T.; Stromvik, M.; Greer, C.W.; Bakermans, C.; Whyte, L. Cold adaptive traits revealed by comparative genomic analysis of the eurypsychrophile *Rhodococcus* sp. JG3 isolated from high elevation McMurdo Dry Valley permafrost, Antarctica. *FEMS Microbiol Ecol* **2016**, *92*, fiv154.

7. Bowman, J.P. Genomics of psychrophilic bacteria. In *Psychrophiles: From biodiversity to biotechnology*; Margesin, R., Ed.; Springer: Cham, Switzerland, 2017; pp. 345-387.

8. Kawamoto, J.; Kurihara, T.; Esaki, N. Proteomic insights of psychrophiles. In *Psychrophiles: From biodiversity to biotechnology*; Margesin, R., Ed.; Springer: Cham, Switzerland, 2017; pp. 423-435.

9. Mann, S.; Chen, Y.P. Bacterial genomic G+C composition-eliciting environmental adaptation. *Genomics* **2010**, *95*, 7-15.

10. Math, R.K.; Jin, H.M.; Kim, J.M.; Hahn, Y.; Park, W.; Madsen, E.L.; Jeon, C.O. Comparative genomics reveals adaptation by *Alteromonas* sp. SN2 to marine tidal-flat conditions: cold tolerance and aromatic hydrocarbon metabolism. *PLoS One* **2012**, *7*, e35784.

11. Feng, S.; Powell, S.M.; Wilson, R.; Bowman, J.P. Extensive gene acquisition in the extremely psychrophilic bacterial species *Psychroflexus torquis* and the link to sea-ice ecosystem specialism. *Genome Biol Evol* **2014**, *6*, 133-148.

12. Dsouza, M.; Taylor, M.W.; Turner, S.J.; Aislabie, J. Genome-based comparative analyses of Antarctic and temperate species of *Paenibacillus*. *PloS One* **2014**, *9*, e108009.

13. Smith, H.J.; Foreman, C.M.; Akiyama, T.; Franklin, M.J.; Devitt, N.P.; Ramaraj, T. Genome sequence of *Janthinobacterium* sp. CG23_2, a violacein-producing isolate from an Antarctic supraglacial stream. *Genome Announc* **2016**, *4*, e01468-15.

14. Schloss, P.D.; Allen, H.K.; Klimowicz, A.K.; Mlot, C.; Gross, J.A.; Savengsuksa, S.; McEllin, J.; Clardy, J.; Ruess, R.W.; Handelsman, J. Psychrotrophic strain of *Janthinobacterium lividum* from a cold Alaskan soil produces prodigiosin. *DNA Cell Biol* **2010**, *29*, 533-541.

15. Kim, S.J.; Shin, S.C.; Hong, S.G.; Lee, Y.M.; Lee, H.; Lee, J.; Choi, I.G.; Park, H. Genome sequence of *Janthinobacterium* sp. strain PAMC 25724, isolated from alpine glacier cryoconite. *J Bacteriol* **2012**, *194*, 2096.

16. Avguštin, J.A.; Bertok, D.Ž.; Kostanjšek, R.; Avguštin, G. Isolation and characterization of a novel violacein-like pigment producing psychrotrophic bacterial species *Janthinobacterium svalbardensis* sp. nov. *Anton Leeuw Int J G* **2013**, *103*, 763-769.

17. Smith, H.; Akiyama, T.; Foreman, C.; Franklin, M.; Woyke, T.; Teshima, H.; Davenport, K.; Daligault, H.; Erkkila, T.; Goodwin, L.; Gu, W.; Xu, Y.; Chain, P. Draft genome sequence and description of *Janthinobacterium* sp. strain CG3, a psychrotolerant Antarctic supraglacial stream bacterium. *Genome Announc* **2013**, *1*, e00960-13.

18. Koo, H.; Strope, B.M.; Kim, E.H.; Shabani, A.M.; Kumar, R.; Crowley, M.R.; Andersen, D.T.; Bej, A.K. Draft genome sequence of *Janthinobacterium* sp. Ant5-2-1, isolated from proglacial Lake Podprudnoye in the Schirmacher Oasis of East Antarctica. *Genome Announc* **2016**, *4*, e01600-15.

19. Gong, X.; Skrivergaard, S.; Korsgaard, B.S.; Schreiber, L.; Marshall, I.P.; Finster, K.; Schramm, A. High quality draft genome sequence of *Janthinobacterium psychrotolerans* sp. nov., isolated from a frozen freshwater pond. *Stand Genomic Sci* **2017**, *12*, 8.

20. Rafiq, M.; Hayat, M.; Anesio, A.M.; Jamil, S.U.; Hassan, N.; Shah, A.A.; Hasan, F. Recovery of metallo-tolerant and antibiotic resistant psychrophilic bacteria from Siachen glacier, Pakistan. *PloS One* **2017**, *12*, e0178180.

21. Chiriac, C.; Baricz, A.; Coman, C. Draft genome sequence of *Janthinobacterium* sp. strain ROICE36, a putative secondary metabolite-synthesizing bacterium isolated from Antarctic snow. *Genome Announc* **2018**, *6*, e01553-17.

22. Kumar, R.; Acharya, V.; Singh, D.; Kumar, S. Strategies for high-altitude adaptation revealed from high-quality draft genome of non-violacein producing *Janthinobacterium lividum* ERGS5: 01. *Stand Genomic Sci* **2018**, *13*, 11.

23. Brucker, R.M.; Harris, R.N.; Schwantes, C.R.; Gallaher, T.N.; Flaherty, D.C.; Lam, B.A.; Minbiole, K.P. Amphibian chemical defense: antifungal metabolites of the microsymbiont *Janthinobacterium lividum* on the salamander *Plethodon cinereus*. *J Chem Ecol* **2008**, *34*, 1422-1429.

24. Lincoln, S.P.; Fermor, T.R.; Tindall, B.J. *Janthinobacterium agaricidamnosum* sp. nov., a soft rot pathogen of *Agaricus bisporus*. *Int J Syst Evol Microbiol* **1999**, *49*, 1577-1589.

25. Rossolini, G.M.; Condemi, M.A.; Pantanella, F.; Docquier, J.D.; Amicosante, G.; Thaller, M.C. Metallo-β-lactamase producers in environmental microbiota: new molecular class B enzyme in *Janthinobacterium lividum*. *Antimicrob Agents Ch* **2001**, *45*, 837-844.

26. Hornung, C.; Poehlein, A.; Haack, F.S.; Schmidt, M.; Dierking, K.; Pohlen, A.; Schulenburg, H.; Blokesch, M.; Plener, L.; Jung, K.; Bonge, A.; Krohn-Molt, I.; Utpatel, C.; Timmermann, G.; Spieck, E.; Pommerening-Röser, A.; Bode, E.; Bode, H.B.; Daniel, R.; Schmeisser, C.; Streit, W.R. The *Janthinobacterium* sp. HH01 genome encodes a homologue of the V. cholerae CqsA and *L. pneumophila* LqsA autoinducer synthases. *PLoS One* **2013**, *8*, e55045.

27. Gillis, M.; De Ley, J. The genera *Chromobacterium* and *Janthinobacterium*. In *The Prokaryotes A handbook on the biology of bacteria*, 3rd ed.; Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K-H., Stackebrandt, E., Eds.; Springer: New York, USA, 2006; pp:737-746.

28. SanClements, M.D.; Smith, H.J.; Foreman, C.M.; Tedesco, M.; Chin, Y.P.; Jaros, C.; McKnight, D.M. Biogeophysical properties of an expansive Antarctic supraglacial stream. *Antarct Sci* **2017**, *29*, 33-44.

29. Foreman, C.M.; Cory, R.M.; Morris, C.E.; SanClements, M.D.; Smith, H.J.; Lisle, J.T.; Miller, P.L.; Chin, Y.P.; McKnight, D.M. Microbial growth under humic-free conditions in a supraglacial stream system on the Cotton Glacier, Antarctica. *Environ Res Lett* **2013**, *8*, 035022.

30. Morris, J.J.; Lenski, R.E.; Zinser, E.R. The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. *MBio* **2012**, *3*, e00036-12.

31. Giovannoni, S.J.; Thrash, J.C.; Temperton, B. Implications of streamlining theory for microbial ecology. *ISME J* **2014**, *8*, 1553.

32. Cleary, A.; Ramaraj, T.; Kahanda, I.; Mudge, J.; Mumey, B. Exploring frequented regions in pan-genomic graphs. *IEEE ACM T Comput Bi* **2018**, DOI: 10.1109/TCBB.2018.2864564.

33. Wu, T.D.; Watanabe, C.K. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **2005**, *21*, 1859-1875.

34. Parks, D.H.; Imelfort, M.; Skennerton, C.T.; Hugenholtz, P.; Tyson, G.W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **2015**, *25*, 1043-1055.

35. Matsen, F.A.; Kodner, R.B.; Armbrust, E.V. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* **2010**, *11*, 538.

36. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* **2016**, *33*, 1870-1874.

37. Letunic, I.; Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* **2016**, *44*, W242-245.

38. Meier-Kolthoff, J.P.; Auch, A.F.; Klenk, H.P.; Göker, M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* **2013**, *14*, 60.

39. Goris, J.; Konstantinidis, K.T.; Klappenbach, J.A.; Coenye, T.; Vandamme, P.; Tiedje, J.M. DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* **2007**, *57*, 81-91.

40. Rodriguez-R, L.M.; Konstantinidis, K.T. The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *PeerJ Preprints* **2016**, *4*, e1900v1.

41. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **1997**, *25*, 3389-3402.

42. Wu, S.; Zhu, Z.; Fu, L.; Niu, B.; Li, W. WebMGA: a customizable web server for fast metagenomic sequence analysis. *BMC Genomics* **2011**, *12*, 444.

43. Galperin, M.Y.; Makarova, K.S.; Wolf, Y.I.; Koonin, E.V. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res* **2015**, *43*, D261-D269.

44. Zhu, Q.; Kosoy, M.; Dittmar, K. HGTector: an automated method facilitating genome-wide discovery of putative horizontal gene transfers. *BMC Genomics* **2014**, *15*, 717.

45. González-Torres, P.; Rodríguez-Mateos, F.; Antón, J.; Gabaldón, T. Impact of homologous recombination on the evolution of prokaryotic core genomes. *mBio* **2019**, *10*, e02494-18.

46. Frost, L.S.; Leplae, R.; Summers, A.O.; Toussaint, A. Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* **2005**, *3*, 722.

47. Newton, I.L.; Bordenstein, S.R. Correlations between bacterial ecology and mobile DNA. *Curr Microbiol* **2011**, *62*, 198-208.

48. Smith, H.J.; Foster, R.A.; McKnight, D.M.; Lisle, J.T.; Littmann, S.; Kuypers, M.M.; Foreman, C.M. Microbial formation of labile organic carbon in Antarctic glacial environments. *Nat Geosci* **2017**, *10*, 356-359.

49. Säwström, C.; Lisle, J.; Anesio, A.M.; Priscu, J.C.; Laybourn-Parry, J. Bacteriophage in polar inland waters. *Extremophiles* **2008**, *12*, 167-175.

50. Bhate. M.P.; Molnar. K.S.; Goulian, M.; DeGrado, W.F. Signal transduction in histidine kinases: insights from new structures. *Structure* **2015**, *23*, 981-994.

51. Sarkar, M.K.; Paul, K.; Blair, D. Chemotaxis signaling protein CheY binds to the rotor protein FliN to control the direction of flagellar rotation in *Escherichia coli*. *PNAS* **2010**, *107*, 9370-9375.

52. Ud-Din, A.I.M.S.; Roujeinikova, A. Methyl-accepting chemotaxis proteins: a core sensing element in prokaryotes and archaea. *Cell Mol Life Sci* **2017**, *74*, 3293-3303.

53. Xu, L.; Xin, L.; Zeng, Y.; Yam, J.K.; Ding, Y.; Venkataramani, P.; Cheang, Q.W.; Yang, X.; Tang, X.; Zhang, L.H.; Chiam, K.H. A cyclic di-GMP–binding adaptor protein interacts with a chemotaxis methyltransferase to control flagellar motor switching. *Sci Signal* **2016**, *9*, ra102.

54. Pesavento, C.; Hengge, R. Bacterial nucleotide-based second messengers. *Curr Opin Microbiol* **2009**, *12*, 170-176.

55. Rossi, E.; Paroni, M.; Landini, P. Biofilm and motility in response to environmental and host-related signals in Gram negative opportunistic pathogens. *J Appl Microbiol* **2018**, *125*, 1587-1602.

56. Prüß, B.M.; Besemann, C.; Denton, A.; Wolfe, A.J. A complex transcription network controls the early stages of biofilm development by *Escherichia coli*. *J Bacteriol* **2006**, *188*, 3731-9373.

57. Wang, Q.; Suzuki, A.; Mariconda, S.; Porwollik, S.; Harshey, R.M. Sensing wetness: a new role for the bacterial flagellum. *EMBO J* **2005**, *24*, 2034-2042.

58. Allocati, N.; Federici, L.; Masulli, M.; Di Ilio, C. Glutathione transferases in bacteria. *FEBS J* **2009**, *276*, 58-75.

59. Sleator, R.D.; Hill, C. Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence. *FEMS Microbiol Rev* **2002**, *26*, 49-71.

60. Doyle, S.; Dieser, M.; Broemsen, E.; Christner, B. General characteristics of cold-adapted microorganisms. In *Polar microbiology: Life in a deep freeze*, Miller, R.V., Whyte, L., Eds; ASM Press, Wahington DC, USA, 2012; pp. 103-125.

61. Jamet, A.; Nassif, X. New players in the toxin field: polymorphic toxin systems in bacteria. *MBio* **2015**, *6*, e00285-15.

62. Tetsch, L.; Jung, K. The regulatory interplay between membrane-integrated sensors and transport proteins in bacteria. *Mol Microbiol* **2009**, *73*, 982-991.

63. Bleuven, C.; Landry, C.R. Molecular and cellular bases of adaptation to a changing environment in microorganisms. *P Roy Soc B-Biol Sci* **2016**, *283*, 20161458.

64. Pinto, D.; Mascher, T. (Actino) bacterial “intelligence”: using comparative genomics to unravel the information processing capacities of microbes. *Curr Genet* **2016**, *62*, 487-498.

65. Mitchell, A.; Romano, G.H.; Groisman, B.; Yona, A.; Dekel, E.; Kupiec, M.; Dahan, O.; Pilpel, Y. Adaptive prediction of environmental changes by microorganisms. *Nat*. **2009**, *460*, 220-224.

66. Wang, G.; Doyle, M.P. Heat shock response enhances acid tolerance of *Escherichia coli* O157: H7. *Lett Appl Microbiol* **1998**, *26*, 31-34.

67. Allen, M.A.; Lauro, F.M.; Williams, T.J.; Burg, D.; Siddiqui, K.S.; De Francisci, D.; Chong, K.W.; Pilak, O.; Chew, H.H.; De Maere, M.Z.; Ting, L.; Katrib, M.; Ng, C.; Sowers, K.R.; Galperin, M.Y.; Anderson, I.J.; Ivanova, N.; Dalin, E.; Martinez, M.; Lapidus, A.; Hauser, L.; Land, M.; Thomas, T.; Cavicchioli, R. The genome sequence of the psychrophilic archaeon, *Methanococcoides burtonii*: the role of genome evolution in cold adaptation. *ISME J* **2009**, *3*, 1012.

68. Shen, L.; Yao, T.; Liu, Y.; Jiao, N.; Kang, S.; Xu, B.; Zhang, S.; Liu, X. Downward-shifting temperature range for the growth of snow-bacteria on glaciers of the Tibetan Plateau. *Geomicrobiol J* **2014**, *31*, 779-787.

69. Richter, M.; Rosselló-Móra, R. Shifting the genomic gold standard for the prokaryotic species definition. *PNAS* **2009**, *106*, 19126-19131.