Title: Free-living psychrophilic bacteria of the genus *Psychrobacter* are descendants of pathobionts

Running Title: psychrophilic bacteria descended from pathobionts

Daphne K. Welter[1], Albane Ruaud[1], Zachariah M. Henseler[1], Hannah N. De Jong[1], Peter van Coeverden de Groot[2], Johan Michaux[3,4], Linda Gormezano[5¥], Jillian L. Waters[1], Nicholas D. Youngblut[1], Ruth E. Ley[1]*

1. Department of Microbiome Science, Max Planck Institute for Developmental Biology,
Tübingen, Germany.
2. Department of Biology, Queen's University, Kingston, Ontario, Canada.
3. Conservation Genetics Laboratory, University of Liège, Liège, Belgium.
4. Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD), UMR ASTRE, Montpellier, France.
5. Department of Vertebrate Zoology, American Museum of Natural History, New York, NY, USA.
¥deceased
*Correspondence: rley@tuebingen.mpg.de

Abstract

Host-adapted microbiota are generally thought to have evolved from free-living ancestors. This process is in principle reversible, but examples are few. The genus *Psychrobacter* (family *Moraxellaceae*, phylum *Gamma-Proteobacteria*) includes species inhabiting diverse and mostly polar environments, such as sea ice and marine animals. To probe *Psychrobacter's* evolutionary history, we analyzed 85 *Psychrobacter* strains by

comparative genomics and phenotyping under 24 different growth conditions. Genome-based phylogeny shows *Psychrobacter* are derived from *Moraxella*, which are warm-adapted pathobionts. *Psychrobacter* strains form two ecotypes based on growth temperature: flexible (FE, growth at 4 - 37^{o}C), and restricted (RE, 4 - 25^{o}C). FE strains, which can be either phylogenetically basal or derived, have smaller genomes and higher transposon copy numbers. RE strains have larger genomes, and show genomic adaptations towards a psychrophilic lifestyle and are phylogenetically derived only. We then assessed *Psychrobacter* abundance in 86 mostly wild polar bear stools and tested persistence of select strains in germfree mice. *Psychrobacter* (both FE and RE) was enriched in stool of polar bears feeding on mammals, but only FE strains persisted in germfree mice. Together these results indicate growth at 37^{o}C is ancestral in *Psychrobacter*, lost in many derived species, and likely necessary to colonize the mammalian gut.

Introduction

Whether microbiota are associated with vertebrate hosts or not is the largest factor driving differences in the composition of microbiomes sampled globally [1, 2]. Recent analysis of metagenome-assembled genomes from multiple habitats shows that many of these genomes are either animal host-enriched or environment-enriched, but generally not both [3]. Strikingly, such specialization can also be seen at higher taxonomic levels, indicating that whole lineages may have diverged once animal hosts were first successfully colonized. For instance, within the *Bacteroidetes*, the taxa that are mammal-gut associated are derived from phylogenetically basal clades that include free-living and invertebrate-associated taxa [4]. These patterns of distribution imply that specialization to the warm animal host habitat is mostly incompatible with fitness in other environments.

There are known exceptions: microbiota with complex lifestyles adapted to life both on and off the warm animal host. A few taxa from the phylum *Proteobacteria*, many of which are pathogens and pathobionts, inhabit mammalian bodies and have environmental reservoirs [5, 6]. Genomic adaptations that support fitness across several different environment types have been identified, many of which increase infectivity in mammals. For instance, genes that allow bacteria to avoid predation by protozoa, amoebas, and nematodes, also contribute to virulence in mammalian infections in species such as *Vibrio cholera*, *Burkholderia pseudomallei*, and *Yersinia pestis* [7, 8]. Genes regulating the formation of biofilms have also been implicated in the infectivity of organisms such as *V. cholera [9]*. Type IV pili, organelles that are important for the asymptomatic colonization of plant tissues, are also associated with mammalian tissue invasion [10]. Particular strains of virulent *Escherichia coli* serotype O157H7 show increased or decreased ability to persist in soil depending on mutations in their stress response genes, which impact survival in acid and other selection pressures [11]. In the food-borne pathogen *Listeria monocytogenes*, mutations in genes important for cell invasion also affect cold tolerance, such that strains

with increased persistence in processed foods are more likely to have high virulence [12]. For the majority of nonpathogenic animal-associated microbiota, the adaptations to life in and on the host seem to preclude sustaining populations outside the host.

The evolutionary history of pathogens has been studied in depth in a few cases, and indicate an environmental ancestry. The evolutionary trajectory of commensal microbiota is less well characterized, but many likely follow the same patterns. *Mycobacterium tuberculosis* and *Y. pestis* are both pathogens thought to be derived from non-pathogenic, environmental organisms [13, 14]. Many *Mycobacteria* are soil microbes that can sometimes cause disease in mammals, while *M. tuberculosis* itself is only found in humans and has no known environmental reservoir [14, 15]. *Y. pestis* is closely related to pathogenic species *Y. pseudotuberculosis* and *Y. enterocolitica*, though other *Yersinia* spp. are nonpathogenic soil and water microbes [13, 16]. There may be cases of the inverse: where environmental bacteria are derived from animal host-associated ancestors, but well characterized examples are lacking. Experimental evolution studies in *Pseudomonas aeruginosa* and *Serratia marcescens* have established that a trajectory from host to environment is possible, however, and results in attenuated virulence [17, 18]. The more bacterial genomes become available for comparative studies, the better the understanding will be of how certain lineages may have moved between animal hosts and their environments.

Here, we investigated the evolutionary history of the genus *Psychrobacter*, a group of closely related bacteria with a broad environmental distribution. Species of *Psychrobacter* have been recovered through culture-based and sequenced-based methods from a range of animal microbiomes, including marine mammal skin [19], respiratory blow [20], and guts [21–23]; the gastrointestinal tracts of birds [24, 25] and fish [26]; and many nonhost environments such as sea water [27], sea ice [28], marine sediment [29], glacial ice [30], and permafrost soil [31]. Some *Psychrobacter* species are capable of causing disease in mammalian hosts [32, 33]. However, *Psychrobacter* infections are very rare, and the virulence factors involved are relatively uncharacterized. Intriguingly, a previous comparative genomics analysis of 26 *Psychrobacter* spp. and metadata gleaned from public sources revealed differences in cold-

97 adaptation of protein coding sequences between warm-host-associated strains versus
98 derived marine and terrestrial strains [34]. Furthermore, warm-adapted strains were basal in
99 the *rpoB* gene phylogeny, suggesting that *Psychrobacter* evolved from a mesophilic
100 ancestor. These observations make *Psychrobacter* an interesting candidate to assess how
101 ecotype maps onto phylogeny and source of isolation, and to probe into the evolutionary
102 history of a genus with a wide habitat range.

103 We tripled the collection of *Psychrobacter* genomes, which we use for phylogenomic
104 analysis, and combine these data with extensive phenotyping applied consistently for all
105 strains. We use a large collection of wild polar bear feces collected on ice and land to assess
106 the presence of *Psychrobacter* ecotypes as a function of diet determined by cytochrome b
107 barcode sequencing. Finally, we conducted tests with select strains for colonization of the
108 mammal gut using germ-free mice. Our results confirm a mesophilic ancestry for
109 *Psychrobacter* and a common ancestor with the genus *Moraxella*. Our phenotyping revealed
110 that overall, *Psychrobacter* tolerate a wide range of salinity, but growth at 37 ^{o}C divided the
111 accessions into two ecotypes: those that retained the ability to grow at warm temperatures
112 and colonize mammalian hosts (flexible ecotype, FE), and those that exhibit adaptive
113 evolution towards a psychrophilic lifestyle (restricted ecotype, RE). Genomic analysis of the
114 two ecotypes shows genome reduction in the FE strains with high transposon copy numbers,
115 and adaptation to cold in RE strains. We show that *Psychrobacter* that are basal are FE, but
116 that FE are also interspersed with RE, indicating either re-adaptation to the animal host or
117 retention of the basal traits. Although both FE and RE ecotypes were detected in the feces
118 of wild polar bears, only FE strains tested could colonize the germfree mouse gut. Together
119 our results indicate the evolutionary history of the genus *Psychrobacter* indicates a
120 pathobiont losing its ability to associate with animals in its adaptation to nonhost
121 environments.
122
123

Materials and Methods

All software versions, parameters, and the relevant citations for our analyses are detailed in Table S1.

***Moraxellaceae* family genomics and phenotypic data.** The *Moraxallaceae* family includes three well characterized genera: *Moraxella*, *Acinetobacter* and *Psychrobacter*. To build a family-level phylogeny, we downloaded genomes of 18 species of *Acinetobacter* and 18 species of *Moraxella* from NCBI (June 2020) (Table S2). We also included 15 *Psychrobacter* genomes generated in this study (described below). We determined the phylogenetic relationship between the genomes using whole-genome marker gene analysis software PhyloPhlAn, and determined genome quality and summary characteristics using checkM and Prokka. The *Moraxellaceae* phylogenetic tree was visualized and annotated using the interactive Tree of Life (iTOL) web interface.

We analyzed the *Moraxellaceae* pan-genome using the PanX pipeline. The input genomes used by PanX were initially annotated using Prokka. After their assignment into orthologous clusters using MCL, we re-annotated gene clusters using eggNOG mapper. We explored genome content by calculating a distance matrix using the Jaccard metric through the R package ecodist from a binary gene presence-absence table, followed by dimensional reduction of that distance matrix through principle coordinate decomposition (PCoA) with the cmdscale function from the R package stats. We investigated variables contributing to the separation of the PCoA using the envfit function from the R package vegan; we tested whether genes significantly contributing to separation were “core” (genes present in 90% to 100% of strains), “shell” (genes present in greater than two strains, but in fewer than 90%), or “cloud” (genes present in only one strain), and if general gene function - summarized by Cluster of Orthologous Groups (COG) category - contributed to the separation. We collected

150 growth temperature range data from type strain publications [23, 24, 27, 28, 35–76]. We
151 used the R package phytools [77] to map the temperature ranges onto the phylogeny.
152
153 ***Psychrobacter* strains.** We obtained 92 isolates of *Psychrobacter* from strain
154 catalogues for phenotypic and genotypic characterization. These represent 38 validly
155 published species of *Psychrobacter* as well as unclassified strains, all isolated from a wide
156 variety of geographical locations and diverse environmental and host samples. All strains
157 were purchased and maintained in compliance with the Nagoya Protocol on Access to
158 Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their
159 Utilization to the Convention on Biological Diversity. For a full list of accessions and their
160 catalogue, isolation, and cultivation information, see Table S3. Unless mentioned, we grew
161 accessions as recommended by the strain catalogue (medium and temperature) from which
162 they were purchased. We performed electron microscopy as described in the Supplemental
163 Text.
164
165 ***Psychrobacter* phenotypic screen.** For the 85 *Psychrobacter* accessions that
166 passed genome quality control (see Supplemental Text, Table S4), we tested their ability to
167 grow under 24 different conditions; a growth condition being a combination of a medium
168 (complex or defined), salt concentration (0, 2.5, 5 or 10% NaCl) and incubation temperature
169 (4, 25 or 37 ^{o}C). For more information regarding the media choice, see Supplemental Text.
170 We randomly assigned *Psychrobacter* accessions to blocks of ten strains to be tested
171 simultaneously (several accessions were included in multiple blocks, see Table S2). We first
172 grew strains to saturation, washed, diluted to optical density at 600 nm (OD_{600}) = 0.3 in
173 sterile phosphate buffered saline (PBS), and inoculated in 100 µL of medium with a final
174 ratio of 1:1000. We used 96-well plates, with 10 inoculum in 5 replicates and 10 uninoculated
175 media wells per plate, all periphery wells were filled with water to reduce edge effects. Plates
176 were incubated at each temperature to reflect all 24 growth conditions. We measured the
177 OD_{600} (Spark® plate reader, Tecan, Zürich, Switzerland) every eight hours during the first week of

incubation, then every twenty-four hours, and let cultures grow until stationary phase (3 to 12 weeks). We tested the carbon utilization of a subset of *Psychrobacter* accessions as described in the Supplemental Text.

Growth probabilities. We scored each replicate as either 'growth positive', meaning the accession grew, or else, 'growth negative,' if the replicate never reached a maximum OD_{600} of 0.15 over the course of the experiment. For each strain and condition (medium, salt and temperature), using base R functions we calculated a growth probability, which corresponds to the median value of growth positivity/negativity of all replicates for that condition.

Genome sequencing, assembly and annotation. Genomic DNA was extracted from cultures grown in their preferred conditions using the Gentra Puregene Tissue Kit (Qiagen, Valencia, CA, USA). Samples were sequenced using the MiSeq 2x250 bp and HiSeq 2x150 bp paired-end read technology (Illumina, San Diego, CA, USA) as previously described [78], with some samples undergoing additional long-read sequencing using Ligation Sequencing (Oxford Nanopore, Oxford, UK). For details regarding quality control and assembly, see Supplemental Text, and for a summary of genome quality, see Table S4. Note that of the initial 92 accessions, 85 had genomes that passed all QC measures (Table S4). All further analyses use these 85 accessions.

We annotated genomes with Prokka and eggNOG mapper. A phylogeny of the accessions was generated using PhyloPhlAn, with *Moraxella lincolnii* as an outgroup. Again the phylogeny was visualized and annotated using iTOL. We used PanX to analyze the *Psychrobacter* pan-genome and R to explore gene presence-absence data as described above with the Moraxellaceae family. Pseudogenes were predicted using the DFAST core workflow. We used treeWAS to perform a pan-genome wide association study relating data collected in the phenotypic screen to genomic data.

205

206 **Potential cold-adaptive trait analysis.** For every gene cluster of the *Psychrobacter*
207 pan-genome obtained, we used DIAMOND to map the consensus sequence against the
208 UniRef90 database, excluding results from the family *Moraxellaceae*. We thus obtained
209 homologous sequences while ensuring that the analysis would not be biased by self-
210 comparison. Gene clusters without UniRef90 homologs were removed from further analysis.
211 The putative cold adaptive traits of every protein-coding gene from the *Psychrobacter*
212 set and their UniRef90 homologs were evaluated following the methods of Bakermans [34].
213 Briefly, we used the relative abundance of every amino acid per coding sequence (CDS) to
214 calculate the arginine to lysine ratio, acidity, hydrophobicity and KVYWREP ratios [34]. We
215 additionally calculated the grand average of hydropathicity (GRAVY) [79] and the isoelectric
216 point (using Protein Analysis method class from Biopython). Next, we compared the
217 distributions of these measures from *Psychrobacter* protein sequences to the distributions of
218 the measures from the UniRef sequences. For each protein sequence, if a majority of
219 measures were in the top 25% of the homologs distribution, we defined the protein sequence
220 as “highly cold adaptive”. The top 25% were defined as follows for each trait: highest 25% for
221 glycine relative abundance, acidity, hydrophobicity, and GRAVY; and the lowest 25% for
222 proline relative abundance, arginine relative abundance, arginine-lysine ratio, isoelectric
223 point, and KVYWREP.

224

225 **Microbiome diversity of polar bear feces.** We collected 86 polar bear fecal
226 samples from several regions in Canada, including samples from 10 captive bears, fed
227 varying diets or fasted, and 76 samples from an unknown number of wild bears, whose diets
228 were determined from stool analysis (below). The feces from the captive bears were
229 forwarded from institutions within Canada and did not require permitting for their passage to
230 Queen's University. The captive samples comprised: five fecal samples from a single bear
231 sequentially fed varying diets of Arctic char, harp seal and a “zoo diet” at the Polar Bear
232 Habitat in Cochrane during 2010; two samples from each of two bears held at the Metro

Toronto Zoo, fed a consistent "zoo diet" during 2010; and a single sample from a bear held at the Churchill Polar Bear Holding Facility in Churchill during 2010, where the bears are given only water until release. We collected all wild bear faeces from M'Clintock Channel and Hudson Strait in Nunavut in accordance with permits prior to shipping to Queen's University: we collected 24 samples from M'Clintock Channel under *Wildlife Research* permits issued in 2007, 2008, 2009, 2010 and 2011 to Peter Van Coeverden de Groot, and nine samples from Hudson Straight collected in 2011 under *Wildlife Research* permit to Grant Gilchrest (Environment Canada). We collected 43 wild samples from the Wapusk National Park in Manitoba from 2007-2010 under a Canada Parks permit to Robert Rockwell (American Museum of Natural History).

We confirmed that samples collected from wild bears originated from bears by sequencing a cytochrome b gene fragment using metabarcoding approaches based on Ion Torrent (Ion Torrent Systems Inc., Gilford, NH, USA) and 454 pyrosequencing (454 Life Sciences, Branford, CT, USA) next-generation sequencing technologies [80]. We also used this method to analyze what prey animals the polar bears had been feeding on at the time of sample deposition. We visually inspected the samples for confirmation on the dietary analysis. We extracted DNA from the fecal samples using the DNeasy Blood and Tissue kit (Qiagen), and characterized the gut bacterial community by amplification and sequencing of the V4 region of the 16S rRNA gene as described previously [81]. We used QIIME for sequence processing and Silva to assign taxonomy. As described in Supplemental Text, we isolated a strain of *Psychrobacter* (*P. faecalis PBFP-1)* from polar bear feces that we included in our genomic and phenotypic analyses.

Gnotobiotic mouse colonizations. Based on their source of derivation and phylogenetic breadth, we selected eight accessions of *Psychrobacter - P. cibarius JG-220, P. ciconiae, P. faecalis PBFP-1, P. immobilis S3, P. lutiphocae, P. okhotskensis MD17, P. namhaensis,* and *P. pacificensis* – to monocolonize germfree mice. We grew each accession in its preferred conditions (Table S3) to saturation, spun at 10 °C at 2500 rpm for

25 minutes, washed with sterile PBS, resuspended in 15% (v/v) glycerol in PBS, and flash frozen in liquid N_2. Inoculum samples were stored at -80 ^{o}C until the mouse experiments were performed by animal caretakers as follows. For *P. ciconiae*, *P. faecalis*, *P. namhaensis*, and *P. pacificensis*, experiments were performed at Taconic Biosciences (Rensselaer, NY, USA), while for *P. cibarius*, *P. immobilis*, and *P. okhotskensis*, experiments were performed at the Max Planck Institute for Developmental Biology. *P. lutiphocae* was included in both Taconic and MPIDB experiments.

Five to six-week old germfree male C57BL/6J mice were orally inoculated with approximately 10^7 cfu (n=4 per *Psychrobacter* accession, n = 8 for *P. lutiphocae*). Mice inoculated with the same strain were co-housed (Taconic) or single-housed (MPIDB) in sterile cages (IsoCage P, Tecniplast) and provided autoclaved water and sterile chow (NIH31M) ad libitum. Three weeks post-colonization, mice were euthanized and cecal contents were immediately collected, flash frozen, and stored at -80 ^{o}C. The Taconic experiments were performed in compliance with Taconic's IACUC, and the MPIDB experiments were approved by and performed in accordance with the local animal welfare authority's legal requirements.

To determine the bacterial colonization density in the mouse cecum, we serially diluted two aliquots per mouse of cecal material (50 mg each), incubated them on plates under their preferred conditions (Table S3) for three to five days. If no colonies were observed, we spread an inoculating loop of the undiluted aliquot onto Brain-Heart Infusion Agar and incubated at 37 ^{o}C for two weeks. For samples where we again observed no colonies, we categorized these *Psychrobacter* accessions as "non-persistent." For the samples that did show colony growth, we confirmed the identity of the colonies as *Psychrobacter* using Sanger sequencing of 16S rRNA gene amplicons as described in the Supplemental Text.

Statistical analysis. We performed all data processing and statistical analysis using R or Python. We compared means between groups using a Kruskal-Wallis test, followed by

289 a pairwise Wilcoxon rank sum test to identify which groups differed when more than two
290 groups were compared. We tested differences in frequencies between groups with more
291 than 10 observations with a χ^2-test, repeated a 100 times with down-sampling in order to
292 correct for sampling sizes between groups. We measured phylogenetic signal using a log-
293 likelihood ratio test on Pagel's λ (Pagel's λ fitted using the phylosig function from the
294 phytools R package, and the null hypothesis being λ = 0). When applicable, we tested for the
295 confounding of phylogeny with our groups of interest using the aov.phylo function from the R
296 package geiger. We adjusted all p-values for multiple comparisons with the Benjamini-
297 Hochberg (BH) correction method.
298
299 ## Results
300
301 ***Psychrobacter* forms a clade whose basal members are of the *Moraxella***
302 **genus.** To explore the evolutionary history of *Psychrobacter*, we built a phylogeny based on
303 400 conserved marker genes using publically available whole genomes derived from
304 cultured isolates of 51 species from the Moraxellaceae family. We included 18 species each
305 of *Acinetobacter* and *Moraxella* obtained from NCBI, and 15 *Psychrobacter* genomes that we
306 generated in this study (Table S2). For all genomes, we incorporated phenotypic data
307 collected from previously published type-strain research. Our analysis shows the 51 species
308 formed three distinct clades with robust bootstrap support. The *Acinetobacter* clade consists
309 uniquely of *Acinetobacter* species (labeled A in Fig 1A). The Acinetobacter clade is a sister
310 taxon to the *Moraxella* (M) clade, consisting entirely of *Moraxella* species, and to the
311 *Psychrobacter* (P) clade, which contains all of the *Psychrobacter* species, as well as four
312 *Moraxella* species (*M. boevrei*, *M. atlantae*, *M. osloensis*, and *M. lincolnii*) that are basal in
313 the P clade. This phylogeny indicates that these four *Moraxella* species are more closely
314 related to *Psychrobacter* than to other *Moraxella*. P-clade *Moraxella* have more similar
315 phenotypes compared to M-clade *Moraxella* than to *Psychrobacter* strains. Review of type

316 strain descriptions revealed that *Psychrobacter* are consistently urease positive, nitrate
317 reducing, salt tolerant, and non-fastidious, whereas P-clade *Moraxella* are inconsistent with
318 their urease and nitrate reducing phenotypes, are sensitive to high salt concentrations, and
319 are often nutritionally fastidious - they have complex growth requirements (in particular, often
320 blood or bile for growth) [23, 24, 39, 42, 50, 51, 66, 67].

321 Consistent with the topology of the phylogeny, a principal coordinates analysis
322 (PCoA) of gene presence/absence data shows that the greatest variation within the family
323 (*i.e*., Principle Coordinate [PC] 1) is the separation between the A clade versus the M and P
324 clades, which are grouped (Fig. 1). The P-clade *Moraxella* spp. fall between the P-clade
325 *Psychrobacter* and the M-clade *Moraxella* when visualizing PC1-2 (Fig 1B) as well as PC2-3
326 (Fig 1C). To assess what genes and gene functions may be contributing to the separation
327 between the A-, M- and P-clades' gene presence/absence, we performed an envfit function
328 analysis using the R package vegan. This analysis indicates that the separation along PC1
329 and 2 is due to differences in shell genes (*i.e.,* genes present in greater than two strains, but
330 in fewer than 90%), not to differences in their core genes (*i.e.,* genes present in between
331 90% and 100% of strains in the clade). Genes annotated with very diverse functions, falling
332 under almost every Cluster of Orthologous Groups (COG) category, strongly contribute to
333 the separation between clades (Fig. 1B, Table S5).

334 Despite their close phylogenetic relationship and high similarity in gene
335 presence/absence, *Psychrobacter* and *Moraxella* have different genomic properties. When
336 examined by genus rather than clade, *Psychrobacter* species have an average genome size
337 of 3.12 ± 0.27 Mb, while the average of *Moraxella* is 2.41 ± 0.28 Mb (pairwise Wilcoxon rank
338 sum test, p-value = 8e-07). *Moraxella* species have an average coding density of 86.4 ±
339 1.33%, which is significantly higher than the *Psychrobacter* species average of 82.7 ± 1.33%
340 (pairwise Wilcoxon rank sum test, p-value = 1e-5). Despite being phylogenetically more
341 related to *Psychrobacter* species than to the other *Moraxella* species, the P-clade *Moraxella*
342 had significantly smaller genomes than *Psychrobacter* species and greater coding density

(pairwise Wilcoxon rank sum test, p-values = 0.004) while not significantly different from the M-clade *Moraxella* spp. (pairwise Wilcoxon rank sum test, p-values = 0.9).

***Psychrobacter* spp. ranges of growth temperatures differ from those of *Moraxella*.** To examine the phenotypic behaviors of the *Moraxellaceae* family, we applied onto the previously generated phylogeny continuous trait mapping of the ranges of temperatures at which species from the *Moraxellaceae* family were reported to grow [23, 24, 27, 28, 35–75] (Fig 2). The *Psychrobacter* spp. included here are reported to have a broad range of growth temperatures (0 - 38 ^{o}C), but several strains, such as *P. frigidicola* and *P. glacincola*, are psychrophilic (restricted to growth below 20 ^{o}C), which is a phenotype that is not seen elsewhere in the family. Using the growth temperature information reported in the literature for this comparison, we observed that *Psychrobacter* spp. have lower minimum growth temperatures than *Moraxella* spp. from either the P- or M-clades (pairwise Wilcoxon rank sum test, p-value = 5e-06), which have a narrow range of temperatures at which they can grow (between 22 ^{o}C and 40 ^{o}C). In contrast to the minimum growth temperatures, there is little variation in the maximum growth temperatures, except for the notable exceptions of several *Psychrobacter* spp. that are restricted to growth at low temperatures.

***Psychrobacter* spp. from diverse isolation sources show differences in cultivation temperatures.** To explore further *Psychrobacte's* phenotypic diversity, we established a strain collection of 85 *Psychrobacter* accessions isolated from diverse locations (Fig S1A) and ecological sources (Fig S1B). The optimal growth temperatures for each accession - provided by the catalogues from which the accessions were ordered - vary by isolation source (Kruskal-Wallis χ^2 = 43.4, df = 9, p-value = 2e-06), with mammalian-derived strains reported as having significantly higher cultivation temperatures than strains from fish, invertebrates, sea water, terrestrial water, and soil samples (pairwise Wilcoxon rank sum test, adjusted p-values < 0.05) (Fig. S1C). Latitude of isolation has a confounding effect on reported optimal growth temperature, but explains very little of the variance (r^2 =

371 0.06, F(1,67) = 5.5, p-value = 0.02). Examples of *Psychrobacter* morphology (P*. ciconiae*
372 and *P. immobilis A351*) visualized by scanning and transmission electron microscopy are
373 shown (Fig. S1D-G).
374
375 **Few *Psychrobacter* strains can grow at 37 °C.** For a direct comparison of
376 *Psychrobacter* phenotypes, we assessed the 85 *Psychrobacter* accessions for their ability to
377 grow under 24 different combination of medium, salt concentration and temperatue (Methods
378 & Supplemental Text). We calculated growth probabilities, or the fraction of growth positive
379 conditions out of total conditions tested, for every strain and given variable of the growth
380 curve screen, and compared them across the phylogeny and by isolation source. We
381 generated a robust genus-level phylogeny for *Psychrobacter* using 400 conserved marker
382 genes, with *M. lincolnii* as an outgroup (Fig. 3A). In agreement with single marker gene trees
383 generated using *rpoB* sequences [34], 16S rRNA gene sequences [82], and the P-clade
384 structure of *Moraxellaceae* family tree generated in this study, there is a phylogenetically
385 basal group of strains mostly isolated from mammals, and a phylogenetically derived group
386 isolated from mixed sources. Across the entire phylogeny, closely related strains have
387 similar growth probabilities (Pagel's λ ranging from 0.78 to 0.97, all corrected p-values < 1e-
388 3). We observed that most *Psychrobacter* strains are tolerant of a wide variety of
389 temperatures between 4 and 25 °C and of salt concentrations between 0 and 5%; more than
390 90% of all strains can grow under these conditions. However, only 54% of the tested
391 accessions can grow at 10% added salt, and only 31% at 37 °C.
392 Since 37 °C was the most restrictive condition tested, we divided the strains into two
393 ecotypes: the "flexible ecotype" (FE) corresponds to strains that could grow at 37 °C, and the
394 "restricted ecotype" (RE) corresponds to strains that could not grow at 37 °C. FE strains are
395 psychrotrophic (mesophilic organisms with a low minimum growth temperature but an
396 optimal growth temperature above 15 °C), and RE strains are either psychrotrophs or true
397 psychrophiles (unable to grow at temperatures higher than 20 °C).

Notably, the basal clade of the *Psychrobacter*-only tree (Fig 3A) consists solely of FE strains, while the rest of the phylogeny is made up of a mixture of FE and RE strains. Furthermore, the basal FE strains have higher growth probabilities at 37 ^{o}C compared to other FE strains (Pagel's λ = 0.89, p-value = 2e-5). Frequencies of RE and FE strains vary significantly across sources of isolation: the FE group is significantly enriched in strains derived from mammalian sources, and the RE group is significantly enriched in strains derived from fish, sea water and food sources (χ^2-test, all p-values adjusted for group size < 0.05). Nonetheless, both FE and RE ecotypes contain strains from other environments, including mammalian-derived strains within the RE group. FE strains have higher growth probabilities at low- to mid-salt concentrations (Wilcoxon rank sum test, all p-values < 0.05), though there is no difference between FE and RE strains at 10% salt (Wilcoxon rank sum test, W = 811.5, p = 0.7). After accounting for phylogenetic relatedness, the growth probabilities under different salt concentrations are no longer significantly different between FE and RE strains (F(1,83) < 24.0, p-value < 0.4), indicating that ecotype grossly maps onto phylogeny. Finally, FE strains show higher growth probabilities in complex media compared to defined media, while RE strains showed little difference between the two (Wilcoxon rank sum test, W = 1.13e3, p-value = 0.0005). The difference in FE and RE strains' growth probabilities under rich media remains significant after accounting for phylogenetic relatedness (F(1,83) = 43.4, p-value = 0.002).

FE and RE *Psychrobacter* spp. have differences in genomic content. We looked for genes differentiating the FE and RE ecotypes by performing a microbial pan-genome wide association analysis (pan-GWAS) with the R package treeWAS [83] using gene presence/absence data. While this analysis returned no significant results, strong phylogenetic signals in the gene presence/absence data may have resulted in a loss of power for the pan-GWAS. When exploring gene presence/absence via PCoA, accessions from the basal FE-only subclade cluster closely together, indicating similar gene content. In agreement with the phylogeny, these basal FE accessions are separated from the other

426 accessions on PC1, while the derived FE and RE accessions are more scattered, indicating
427 more diverse gene content (Fig 3B). As with the separation in the *Moraxellaceae* PCoA of gene
428 content, the separation between the basal FE-clade and the rest of the accessions is due largely
429 to presence/absence patterns in shell genes (present in greater than one strain but fewer than
430 90%), not core (present in between 90 to 100% of strains) or cloud genes (present only in one
431 strain), and all contributing genes were not unique to either clade. The separation is most
432 strongly driven by genes from the COG categories T, signal transduction (r^2 = 0.65, p-value
433 = 0.001); U, trafficking and secretion (r^2 = 0.63, p-value = 0.001); P, inorganic ion transport
434 and metabolism (r^2 = 0.56, p-value = 0.001), and X, unassigned or no homologs in the COG
435 database (r^2 = 0.51, p-value = 0.001)(Fig 5B, Table S5).
436 We examined if there were qualitative differences in gene content between the FE
437 and RE ecotypes, especially in functions related to host colonization versus cold adaptation.
438 Both ecotypes carry genes associated with virulence [84–86] as well as genes that are
439 potentially related to psychrophilic lifestyles [87] (Table S6).
440 We next tested whether there was a predicted difference in adaptation to cold
441 environments in FE versus RE strains based on the amino acid properties of the protein-
442 coding genes of their genomes. Amino acid traits have been implicated in psychrophilic
443 lifestyles, including increased abundance of glycine, decreased abundance of proline,
444 decreased arginine-to-lysine ratio, increased acidity, decreased isoelectric point, increased
445 hydrophobicity, and increased GRAVY, when comparing homologs from psychrophiles to
446 mesophiles or thermophiles [88–91]. Using these traits to define a protein sequence as cold-
447 adaptive, REs have a higher proportion of cold-adapted proteins compared to FEs (Wilcoxon
448 rank sum test, W = 407, p-value = 0.0006) (Fig. 3C). However, after including phylogenetic
449 relatedness as a covariate, the difference is no longer significant ($F(1,83) = 16.4$, $p = 0.2$),
450 indicating that cold adaptation in the RE is associated with their derivation from the basal
451 strains. When comparing core proteins (present in 99 - 100% of accessions) RE strains have
452 a higher percentage of cold-adapted core proteins than FE strains (Wilcoxon rank sum test,

453 W = 404, p-value = 0.0004) (Fig. 3D), which is again confounded with phylogenetic
454 relatedness (F(1,83) = 12.4, p = 0.2).
455
456 **FEs exhibit higher transposon copy numbers than RE strains.** To determine
457 whether there are genomic differences between the FE and RE ecotypes at a broader level
458 than individual genes, we next examined the proportions of each genome devoted to each
459 COG category. There are few differences between the ecotypes by COG category (Fig. S2).
460 FEs have a significantly higher proportion of "L" category genes (replication-, recombination-,
461 and repair-related genes) than REs (Wilcoxon rank sum test, W = 1.15e3, p-value = 0.0005;
462 Fig. 4A). In particular, this difference stems from a higher proportion of transposon copies
463 per FE genome than per RE genome (Wilcoxon rank sum test, W = 1.08e3, p-value = 0.003;
464 Fig. 4B). As expected from the distribution of the RE and FE across the phylogeny,
465 phylogenetic relatedness is confounded with ecotype in the effect on these difference
466 (F(1,83) < 21.3, p = 0.1).
467 Increased transposon activity can lead to interruption and decay of functional protein-
468 coding genes, leading to an increase in pseudogenes [92]. Given the higher number of
469 transposons in FE strains, we next examined the number of predicted pseudogenes
470 between the ecotypes. FE genomes are predicted to have a higher number of pseudogenes
471 than RE strains (Wilcoxon rank sum test, W = 997, p-value = 0.03; Fig. 4C). Finally, we
472 compared the average genome size between the ecotypes, as bacterial genomes are known
473 to strongly select against the accumulation of pseudogenes [93]. FE strains have
474 significantly smaller genomes than RE strains (Wilcoxon rank sum test, W = 524, p-value =
475 0.02; Fig. 4D). As with the other genomic properties, phylogeny confounds the comparison
476 between the FE and RE groups' pseudogene proportion and genome size (F(1,83) < 3.4, p-
477 value < 0.6).
478
479 **Polar bear feces collected from the Arctic ice have high abundance of**
480 ***Psychrobacter*.** We surveyed the gut microbial diversity of 86 polar bear fecal samples, 76

481 wild and 10 captive, by 16S rRNA gene amplicon sequencing. *Psychrobacter* was detectable
482 in 76/86 of the samples (Fig 5A). The large majority of *Psychrobacter* sequences (83%) were
483 assigned to unclassified *Psychrobacter* spp.. We detected RE strain *P. immobilis* in 50% of
484 samples with a mean abundance of 3%, and FE strain *P. pulmonis* in 8% of samples with a
485 mean abundance of 0.5%.

486 Polar bear diet significantly impacted the abundance of *Psychrobacter* spp. (Kruskal-
487 Wallis χ^2 = 13.5, df = 3, p-value = 0.004); we found that polar bears feeding on mammalian
488 prey, including seals and reindeer, had significantly higher abundances of unclassified
489 *Psychrobacter spp.* than polar bears feeding on avian prey or mixed diets (pairwise Wilcoxon
490 rank-sum test, adjusted p-values < 0.05) (Fig 5B). Unsurprisingly, diet data is confounded
491 with location (Kruskal-Wallis χ^2 = 8.6, df = 4, p-value = 0.0009) and year (Kruskal-Wallis χ^2 =
492 16.6, df = 5, p-value = 0.005) of sample collection. Captive status did not significantly impact
493 mean abundances, but there is a trend of wild bear samples having higher unclassified
494 *Psychrobacter* spp. mean relative abundance than samples from captive bears (Wilcoxon
495 rank sum test, W = 222, p-value = 0.08).
496
497 **Carbon source utilization patterns.** To elucidate the effect that host dietary
498 nutrition may have on *Psychrobacter* growth, we tested the maximum change in absorbance
499 for 190 different carbon sources by a subset of *Psychrobacter* accessions including 9 FE
500 and 10 RE strains (Fig 5C). All *Psychrobacter* spp. reached significantly higher OD_{600}
501 growing on amino acid carbon sources compared to carbohydrates or sugar alcohols
502 (pairwise Wilcoxon rank-sum test, adjusted p-values < 0.05). *Psychrobacter* spp. reach the
503 highest OD_{600} growing on fatty acids, surfactants, and peptides. There was no significant
504 difference between FE and RE strains' changes in absorbance in these assays (Wilcoxon
505 rank sum test, W = 1.55e6 , p-value = 0.08).
506
507 ***Psychrobacter* strain survival in gnotobiotic mice.** To assess the survivorship of
508 *Psychrobacter* in a mammalian gut, we tested 8 accessions for persistence in the

gastrointestinal tracts of germ-free mice (Fig. 5D). Chosen for phylogenetic breadth, we tested 4 FE strains, *P. ciconiae*, *P. faecalis PBFP-1*, *P. lutiphocae*, and *P. pacificensis*, and 4 RE strains, *P. cibarius JG-220*, *P. immobilis S3*, *P. namhaensis,* and *P. okhotskensis MD17*. Of the four FE strains tested, three were able to persist in the mice, while the FE strain *P. faecalis PBFP-1* and none of the RE strains were detectable after three weeks. Phylogenetic relatedness does not correlate with ability to colonize, as FE strain *P. pacificensis* was able to persist, while closely related RE strains, *P. namhaensis* and *P. okhotskensis*, were not. Phylogenetic placement does correlate with colonization density however, as the two most basal strains tested, *P. lutiphocae* and *P. ciconiae*, colonized at significantly higher densities than the most derived strain that was successful, *P. pacificensis* (pairwise Wilcoxon rank sum test, both adjusted p-values = 0.0007).

Discussion

The phylogenomic and phenotypic characterizations of the *Psychrobacter* genus indicates a common ancestor with *Moraxella*, all of which are restricted to growth at higher temperatures. Furthermore, the most basal members of the *Psychrobacter* clade are *Moraxella* species and species of *Psychrobacter* that can grow at 37 ^{o}C, unlike most of the derived *Psychrobacter* species. Our extensive phenotyping indicated that members of the *Psychrobacter* genus grow at a wide range of salinities and temperatures, but it is the ability to grow at 37 ^{o}C that distinguishes strains the most, and which we used to define the two ecotypes, FE and RE. Our analysis of a large collection of wild polar bear feces shows both RE and FE strains are present, however tests in germfree mice support the notion that only FE may colonize the mammal gut, whereas RE may be allochthonous members or environmental contaminants. Together with previous reports, this work indicates the genus *Psychrobacter* is a lineage of pathobionts, some of which have evolved to inhabit the colder environments of their warm-bodied hosts.

Our results corroborate those of Bakermann, who used the isolation source of *Psychrobacter* as a proxy for temperature adaptation to conclude the genus has a mesophilic ancestor [34]. By assessing growth under the same 24 conditions for 85 strains, we remove any ambiguity that can stem from whether an isolate can indeed grow at the temperature of its source of isolation. This is particularly important in light of our results showing that several strains isolated from mammals proved to be RE, and that many of the FE strains came from sea water or other relatively cold environments.

Psychrobacter's sister taxon *Moraxella* in particular is commonly isolated from host mucosal tissues, and exhibits the reduced genome size and nutritional fastidiousness common to many host-dependent organisms. *Moraxella* contains species that are frequently associated with human respiratory infections, primarily *M. catarrhalis* [84], as well as livestock conjunctivitis, for example, *M. bovis* or *M. equis* [94]. Since they are commonly found in healthy individuals and can cause disease in healthy individuals [95], *Moraxella* are best categorized as pathobionts and not dedicated or opportunistic pathogens. Several species of *Moraxella* appear basally in the P-clade of the *Moraxellaceae* family-level phylogeny, suggesting that *Psychrobacter* evolved from a "*Moraxella*-like" ancestor. This is supported by the fact that both phylogenetically basal and derived *Psychrobacter* strains carry genes related to virulence functions, and that many of the basal *Psychrobacter* strains exhibit growth defects in liquid culture, similar to the fastidiousness of *Moraxella*.

Despite clear phenotypic differentiation, *Psychrobacter* and *Moraxella* have similar genomic content, although *Psychrobacter* genomes are larger. A psychrophile emerging from an apparently mesophilic background through widespread horizontal gene transfer has been suggested before in the genus *Psychroflexus* [96], though the study was limited to comparing two genomes. In fact, it has been suggested before that this is *Psychrobacter's* evolutionary trajectory [34], and although many *Psychrobacter* trees are constructed using a *Moraxella* outgroup, *Psychrobacter's* potential pathogenic origin has not been widely discussed. Horizontal gene transfer would explain *Psychrobacter's* larger genome size compared to *Moraxella* despite lower coding density, as many newly acquired horizontally

564 transferred genes are expected to be inactivated and pruned by the recipient genome [96,
565 97].

566 Our data show that the largest phenotypic divide within the genus is the ability to
567 grow at 37 ^{o}C, which we used to sort strains into FE and RE. FE strains make up the basal
568 clade of the *Psychrobacter* phylogeny. Their smaller genomes show less cold-adaptation in
569 their protein-coding genes than RE strains with proportionally fewer cold-adaptive proteins
570 and more transposons. The three of four FE strains tested were able to colonize germ-free
571 mice, whereas none of the RE strains could, indicating that growth at 37 ^{o}C may be
572 necessary (although not sufficient) to colonize mammals. Opportunistic infections in
573 mammals caused by *Psychrobacter* strains are limited to *P. sanguinis*, *P. phenylpyruvicus*,
574 *P. faecalis*, and *P. pulmonis* [32], which while are all FE strains. Our results suggest that the
575 FE strains are maintaining an ancestral ability to grow at mammalian body temperatures and
576 colonize mammalian host bodies, while RE strains have adopted a psychrophilic lifestyle.

577 Adaptation to psychrophilic lifestyles usually arises conjointly with high tolerance for
578 salt, but at very high salt concentrations, RE strains had similar growth probabilities as FE
579 strains. Both FE and RE strains carry genes relating to both salt tolerance and cold
580 adaptation functions, such as compatible solute accumulation or membrane fluidity control.
581 The similar salt tolerances of FE and RE strains is thus not surprising, given that both
582 ecotypes have enough psychrophilic adaptation to grow well at 4 ^{o}C. There may be a
583 difference between ecotypes in salt tolerance and cold tolerance in more extreme conditions
584 than those we tested.

585 Phylogenetically basal FE strains show stronger growth at 37 ^{o}C than
586 phylogenetically derived FE strains. It may be that genes responsible for ancestral strains'
587 ability to grow at 37 ^{o}C were lost between the differentiation of the basal and derived clades.
588 Subsequently derived *Psychrobacter* strains could have then acquired other genes providing
589 the same FE phenotype as the basal strains, making pinpointing the genes responsible for
590 the FE phenotype difficult. Genes that are unannotated and do not have homologs in
591 databases such as UniRef or COG may also play a role in phenotypic and genomic

592 differentiation of FE and RE strains, so future *in vitro* phenotypic screens may be helpful in
593 assigning function to the many hypothetical genes in the *Psychrobacter* pan-genome.
594 The difference in ecotypes could also be due to gene regulation rather than gene
595 presence and absence. There has been some transcriptional work done in *P. arcticus* [98],
596 but it focused on comparing gene expression at ambient temperatures (around 20 ^{o}C) to
597 ultra-low temperatures (-10 ^{o}C) to study *P. arcticus*'s evolutionary approach to cold-
598 adaptation. Future studies might use the same techniques to identify differences in gene
599 expression responsible for FE strains' phenotypic plasticity by comparing gene expression at
600 low temperatures (for example, 4 ^{o}C) to higher temperatures (37 ^{o}C).
601 Our data reveal clear genomic signatures between the ecotypes. In particular, we
602 observed higher transposon proportions and smaller genomes in FE strains compared to RE
603 strains. Transposon-mediated genome reduction in host-associated bacteria compared to
604 their free-living relatives has been observed in as-of-yet unculturable light-producing
605 symbionts from ceratoid deep-sea anglerfish [99]. While none of the *Psychrobacter* spp.
606 exhibit the fastidiousness of truly host-dependent bacteria [100], it is striking that FE strains
607 were more likely to grow in nutritionally complex media than defined media, compared to the
608 RE strains, for which nutritional complexity had little impact.
609 It is possible that FE strains have enough contact with mammalian hosts that it
610 remains advantageous for them to maintain their ability to grow at higher temperatures in
611 rich nutritional environments. *Psychrobacter* spp. have been reported previously in the skin,
612 respiratory, and gut microbiomes of several marine mammals, including whales, porpoises,
613 seals, and sea lions; it could be argued that *Psychrobacter* presence is due to contamination
614 from sea water. Polar bears hunt seals on sea ice and their diet during the summer includes
615 bird eggs. Our data show that polar bears consuming seal meat have higher *Psychrobacter*
616 abundance than those consuming eggs, which may result from their time on the sea ice,
617 particularly since FE and RE were equally represented. *Psychrobacter* spp. grew to high
618 densities when grown on amino acids, peptide mixtures, and fatty acids: all carbon sources
619 that would be abundant in the gut of a polar bear eating fatty seal meat. This may allow FE

620 *Psychrobacter* strains to thrive in the bear gut. Given that the majority of the *Psychrobacter*
621 diversity detected from the polar bear feces could not be classified, much remains to be
622 learned about the natural history of this genus.
623 The history of the genus *Psychrobacter* is one of an ancestral pathobiont or
624 pathogen, some of the descendants of which attenuated their own pathogenicity to broaden
625 their ecological distribution. The emergence of a psychrotroph - a remarkable generalist -
626 from a background of a more specialized pathobiont or pathogen showcases the adaptability
627 of bacteria, and particularly *Proteobacteria*, to their environments.
628
629 **Data availability.** Raw sequences for the *Psychrobacter* genome sequencing and
630 polar bear feces 16S rRNA gene sequencing, as well as assembled *Psychrobacter*
631 genomes, are available in the European Nucleotide Archive under the accession
632 PRJEB40380. Annotated *Psychrobacter* genomes are available at
633 ftp://ftp.tue.mpg.de/pub/ebio/dwelter. Raw data, R notebooks, and Python scripts for the
634 analyses are available at https://github.com/dkwelter/Welter_et_al_2020.
635
636 **Acknowledgments.** This work was supported by the Max Planck Society. We thank
637 Jacobo de la Cuesta-Zuluaga, Sara Di Rienzi, Hagay Enav, Angela Poole, Jessica Sutter,
638 Taichi Suzuki, and William Walters for discussions regarding project design and analysis,
639 and Andrea Belkacemi, Ilja Bezrukov, Pablo Carbonell, Silke Dauser, Julia Hildebrandt, and
640 Christa Lanz for their advice on and assistance with sequencing. We also thank Jürgen
641 Berger and Katharina Hipp for performing electron microscopy. We thank Markus Dyck and
642 Patricia Morin for providing the Polar Bear Habitat samples, Maria Frank for samples from
643 the Metro Toronto Zoo, Daryll Hedman and Manitoba Conservation for the sample from the
644 Polar Bear Holding Facility in Churchill, and Sam Iverson for samples from Hudson Straight.
645 The collection of Nunavut samples would not have been possible without collaboration of
646 colleagues at the Gjoa Haven Hunters and Trappers Association and their Traditional
647 Ecological Knowledge relating to polar bears. The Nunavut field work was supported by

funds from the Nunavut Wildlife Management Board (NWMB), the Nunavut General Monitoring Plan (NGMP), the National Science and Engineering Research Council (NSERC), and Environment Canada (Gov. of Canada). We would also like to thank Marie Pages and Maxime Galan for assistance with cytochrome b barcode sequencing.

Competing interest.

We have no competing interests to declare.

References

1. Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI. Worlds within worlds: evolution of the vertebrate gut microbiota. *Nat Rev Microbiol* 2008; **6**: 776–788.

2. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, et al. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA* 2011; **108 Suppl 1**: 4516–4522.

3. Youngblut ND, de la Cuesta-Zuluaga J, Reischer GH, Dauser S, Schuster N, Walzer C, et al. Large scale metagenome assembly reveals novel animal-associated microbial genomes, biosynthetic gene clusters, and other genetic diversity. *mSystems* (in press).

4. Bäckhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI. Host-bacterial mutualism in the human intestine. *Science* 2005; **307**: 1915–1920.

5. Wolfgang MC, Kulasekara BR, Liang X, Boyd D, Wu K, Yang Q, *et al.* Conservation of genome content and virulence determinants among clinical and environmental isolates of Pseudomonas aeruginosa. *Proc Natl Acad Sci USA* 2003; **100**: 8484–8489.

6. Holden N, Pritchard L, Toth I. Colonization outwith the colon: plants as an alternative environmental reservoir for human pathogenic enterobacteria. *FEMS Microbiol Rev* 2009; **33**: 689–703.

7. French CT, Bulterys PL, Woodward CL, Tatters AO, Ng KR, Miller JF. Virulence from the rhizosphere: ecology and evolution of Burkholderia pseudomallei-complex species. *Curr Opin Microbiol* 2020; **54**: 18–32.

8. Waterfield NR, Wren BW, Ffrench-Constant RH. Invertebrates as a source of emerging human pathogens. *Nat Rev Microbiol* 2004; **2**: 833–841.

9. Sakib SN, Reddi G, Almagro-Moreno S. Environmental role of pathogenic traits in Vibrio cholerae. *J Bacteriol* 2018; **200**: e00795–17.

10. Dorr J, Hurek T. Type IV pili are involved in plant–microbe and fungus–

684 microbe interactions. *Mol Microbiol* 1998; **30**: 7–17.

685 11. van Hoek AHAM, Aarts HJM, Bouw E, van Overbeek WM, Franz E. The role
686 of rpoS in Escherichia coli O157 manure-amended soil survival and distribution of allelic
687 variations among bovine, food and clinical isolates. *FEMS Microbiol Lett* 2013; **338**: 18–
688 23.

689 12. Hingston P, Chen J, Dhillon BK, Laing C, Bertelli C, Gannon V, *et al.*
690 Genotypes associated with Listeria monocytogenes isolates displaying impaired or
691 enhanced tolerances to cold, salt, acid, or desiccation stress. *Front Microbiol* 2017; **8**:
692 369.

693 13. Wren BW. The yersiniae--a model genus to study the rapid evolution of
694 bacterial pathogens. *Nat Rev Microbiol* 2003; **1**: 55–64.

695 14. Gagneux S. Ecology and evolution of Mycobacterium tuberculosis. *Nat Rev*
696 *Microbiol* 2018; **16**: 202–213.

697 15. Orgeur M, Brosch R. Evolution of virulence in the Mycobacterium
698 tuberculosis complex. *Curr Opin Microbiol* 2018; **41**: 68–75.

699 16. Savin C, Criscuolo A, Guglielmini J, Le Guern A-S, Carniel E, Pizarro-Cerdá
700 J, *et al.* Genus-wide Yersinia core-genome multilocus sequence typing for species
701 identification and strain characterization. *Microb Genom* 2019; **5**.

702 17. Mikonranta L, Friman V-P, Laakso J. Life history trade-offs and relaxed
703 selection can decrease bacterial virulence in environmental reservoirs. *PLoS One* 2012;
704 **7**: e43801.

705 18. Granato ET, Ziegenhain C, Marvig RL, Kümmerli R. Low spatial structure
706 and selection against secreted virulence factors attenuates pathogenicity in
707 Pseudomonas aeruginosa. *ISME J* 2018; **12**: 2907–2918.

708 19. Apprill A, Robbins J, Eren AM, Pack AA, Reveillaud J, Mattila D, *et al.*
709 Humpback whale populations share a core skin bacterial community: towards a health
710 index for marine mammals? *PLoS One* 2014; **9**: e90785.

711 20. Apprill A, Miller CA, Moore MJ, Durban JW, Fearnbach H, Barrett-Lennard

712 LG. Extensive core microbiome in drone-captured whale blow supports a framework for
713 health monitoring. *mSystems* 2017; **2**.
714 21. Kudo T, Kidera A, Kida M, Kawauchi A, Shimizu R, Nakahara T, et al. Draft
715 genome sequences of Psychrobacter strains JCM 18900, JCM 18901, JCM 18902, and
716 JCM 18903, isolated preferentially from frozen aquatic organisms. *Genome Announc*
717 2014; **2**.
718 22. Banks JC, Craig Cary S, Hogg ID. Isolated faecal bacterial communities
719 found for Weddell seals, Leptonychotes weddellii, at White Island, McMurdo Sound,
720 Antarctica. *Polar Biol* 2014; **37**: 1857–1864.
721 23. Yassin AF, Busse H-J. Psychrobacter lutiphocae sp. nov., isolated from the
722 faeces of a seal. *Int J Syst Evol Microbiol* 2009; **59**: 2049–2053.
723 24. Kämpfer P, Jerzak L, Wilharm G, Golke J, Busse H-J, Glaeser SP.
724 Psychrobacter ciconiae sp. nov., isolated from white storks (Ciconia ciconia). *Int J Syst*
725 *Evol Microbiol* 2015; **65**: 772–777.
726 25. Kämpfer P, Glaeser SP, Irgang R, Fernández-Negrete G, Poblete-Morales
727 M, Fuentes-Messina D, *et al.* Psychrobacter pygoscelis sp. nov. isolated from the
728 penguin Pygoscelis papua. *Int J Syst Evol Microbiol* 2020; **70**: 211–219.
729 26. Svanevik CS, Lunestad BT. Characterisation of the microbiota of Atlantic
730 mackerel (Scomber scombrus). *Int J Food Microbiol* 2011; **151**: 164–170.
731 27. Yoon J-H, Lee C-H, Kang S-J, Oh T-K. Psychrobacter celer sp. nov., isolated
732 from sea water of the South Sea in Korea. *Int J Syst Evol Microbiol* 2005; **55**: 1885–
733 1890.
734 28. Bowman JP, Nichols DS, McMeekin TA. Psychrobacter glacincola sp. nov., a
735 halotolerant, psychrophilic bacterium isolated from Antarctic sea ice. *Syst Appl Microbiol*
736 1997; **20**: 209–215.
737 29. Matsuyama H, Minami H, Sakaki T, Kasahara H, Watanabe A, Onoda T, *et*
738 *al.* Psychrobacter oceani sp. nov., isolated from marine sediment. *Int J Syst Evol*
739 *Microbiol* 2015; **65**: 1450–1455.

30. Zeng Y-X, Yu Y, Liu Y, Li H-R. Psychrobacter glaciei sp. nov., isolated from the ice core of an Arctic glacier. *Int J Syst Evol Microbiol* 2016; **66**: 1792–1798.

31. Bakermans C, Ayala-del-Río HL, Ponder MA, Vishnivetskaya T, Gilichinsky D, Thomashow MF, et al. Psychrobacter cryohalolentis sp. nov. and Psychrobacter arcticus sp. nov., isolated from Siberian permafrost. *Int J Syst Evol Microbiol* 2006; **56**: 1285–1291.

32. Deschaght P, Janssens M, Vaneechoutte M, Wauters G. Psychrobacter isolates of human origin, other than Psychrobacter phenylpyruvicus, are predominantly Psychrobacter faecalis and Psychrobacter pulmonis, with emended description of P. faecalis. *Int J Syst Evol Microbiol* 2012; **62**: 671–674.

33. Le Guern R, Wallet F, Vega E, Courcol RJ, Loïez C. Psychrobacter sanguinis: an unusual bacterium for nosocomial meningitis. *J Clin Microbiol* 2014; **52**: 3475–3477.

34. Bakermans C. Adaptations to marine versus terrestrial low temperature environments as revealed by comparative genomic analyses of the genus Psychrobacter. *FEMS Microbiol Ecol* 2018.

35. Vela AI, Sánchez-Porro C, Aragón V, Olvera A, Domínguez L, Ventosa A, *et al.* Moraxella porci sp. nov., isolated from pigs. *Int J Syst Evol Microbiol* 2010; **60**: 2446–2450.

36. Vela AI, Arroyo E, Aragón V, Sánchez-Porro C, Latre MV, Cerdà-Cuéllar M, *et al*. Moraxella pluranimalium sp. nov., isolated from animal specimens. *Int J Syst Evol Microbiol* 2009; **59**: 671–674.

37. Lindqvist K. A Neisseria species associated with infectious keratoconjunctivitis of sheep: Neisseria ovis nov. spec. *The Journal of Infectious Diseases* March-April 1960; **106**: 162–165.

38. Xie C-H, Yokota A. Transfer of the misnamed [Alysiella] sp. IAM 14971 (=ATCC 29468) to the genus Moraxella as Moraxella oblonga sp. nov. *Int J Syst Evol Microbiol* 2005; **55**: 331–334.

768 39. Boevre K, Henriksen SD. A new Moraxella species, Moraxella osloensis, and
769 a revised description of Moraxella nonliquefaciens. *Int J Syst Evol Microbiol* 1967; **17**.
770 40. Catlin BW. Transfer of the organism named Neisseria Catarrhalis to
771 Branhamella gen. nov. *Int J Syst Bacteriol* 1970; **20**: 155–159.
772 41. Eyre JW. A clinical and bacteriological study of diplo-bacillary conjunctivitis.
773 *Journal of Pathology and Bacteriology* 1900; **6**: 1–13.
774 42. Vandamme P, Gillis M, Vancanneyt I. M, Hoste I. B, Kersters K, Falsen E.
775 Moraxella lincolnii sp. nov., isolated from the human respiratory tract, and reevaluation
776 of the taxonomic position of Moraxella osloensis. *Int J Syst Bacteriol* 1993; **43**: 474–481.
777 43. PJ Huntington, PJ Coloe, JD Bryden, and F MacDonald. Isolation of a
778 Moraxella sp from horses with conjunctivitis. *Australian Veterinury Journal* 1987; **64**:
779 118–119.
780 44. U. B. Über das Vorkommen von Neisserien bei einigen Tieren. *Zeitschrift f*
781 *Hygiene* 1962; **148**: 445–457.
782 45. Pelczar MJ Jr. Neisseria caviae nov spec. *J Bacteriol* 1953; **65**: 744.
783 46. Kodjo A, Tønjum T, Richard 2. Yves, Bq)vre K. Moraxella caprae sp. nov., a
784 new member of the classical Moraxellae with very close affinity to Moraxella bovis. *Int J*
785 *Syst Bacteriol* 1995; **45**: 467–471.
786 47. Jannes G, Vaneechou'tte M, La"o 2. Martine, Gillis 3. Monique, Vancanneyt
787 3. Marc, Vandamme 3. Peter, *et al.* Polyphasic taxonomy leading to the proposal of
788 Moraxella canis sp. nov. for Moraxella catarrhalis-like strains. *Int J Syst Bacteriol* 1993;
789 **43**: 438–449.
790 48. Angelos JA, Spinks PQ, Ball LM, George LW. Moraxella bovoculi sp. nov.,
791 isolated from calves with infectious bovine keratoconjunctivitis. *Int J Syst Evol Microbiol*
792 2007; **57**: 789–795.
793 49. Henriksen SD. Designation of a neotype strain for Moraxella bovis (Hauduroy
794 et al.) Murray. *Int J Syst Evol Microbiol* 1971; **21**: 28–28.
795 50. Kodjo A, Richard ". Yves, Tønjum T. Moraxella boevrei sp. nov., a new

796 Moraxella Species found in goats. *Int J Syst Bacteriol* 1997; **47**: 115–121.

797 51. K. Boevre, J.E. Fuglesang, N. Hagen, E. Jantzen, and L.O. Froeholm.

798 Moraxella atlantae sp. nov. and its distinction from Moraxella phenylpyrouvica. *Int J Syst*

799 *Bacteriol* 1976; **26**: 511–521.

800 52. Vaneechoutte M, Nemec A, Musílek M, van der Reijden TJK, van den

801 Barselaar M, Tjernberg I, et al. Description of Acinetobacter venetianus ex Di Cello *et al.*

802 1997 sp. nov. *Int J Syst Evol Microbiol* 2009; **59**: 1376–1381.

803 53. Nemec A, De Baere T, Tjernberg I, Vaneechoutte M, van der Reijden TJ,

804 Dijkshoorn L. Acinetobacter ursingii sp. nov. and Acinetobacter schindleri sp. nov.,

805 isolated from human clinical specimens. *Int J Syst Evol Microbiol* 2001; **51**: 1891–1899.

806 54. Nishimura Y, Ino T, Iizuka H. Acinetobacter radioresistens sp. nov. isolated

807 from cotton and soil. *Int J Syst Evol Microbiol* 1988; **38**: 209–211.

808 55. Li Y, Piao C-G, Ma Y-C, He W, Wang H-M, Chang J-P, et al. Acinetobacter

809 puyangensis sp. nov., isolated from the healthy and diseased part of Populus

810 xeuramericana canker bark. *Int J Syst Evol Microbiol* 2013; **63**: 2963–2969.

811 56. Yang Liu, Qiuhua Rao, Jiefeng Tu, Jiaonan Zhang, Minmin Huang, Bing Hu,

812 Qiu Lin and Tuyan Luo. Acinetobacter piscicola sp. nov., isolated from diseased farmed

813 Murray cod (Maccullochella peelii peelii). *Int J Syst Evol Microbiol* 2018; **68**: 905–910.

814 57. Choi JY, Ko G, Jheong W, Huys G, Seifert H, Dijkshoorn L, *et al.*

815 Acinetobacter kookii sp. nov., isolated from soil. *Int J Syst Evol Microbiol* 2013; **63**:

816 4402–4406.

817 58. Malhotra J, Anand S, Jindal S, Rajagopal R, Lal R. Acinetobacter indicus sp.

818 nov., isolated from a hexachlorocyclohexane dump site. *Int J Syst Evol Microbiol* 2012;

819 **62**: 2883–2890.

820 59. Nemec A, Musílek M, Maixnerová M, De Baere T, van der Reijden TJK,

821 Vaneechoutte M, et al. Acinetobacter beijerinckii sp. nov. and Acinetobacter gyllenbergii

822 sp. nov., haemolytic organisms isolated from humans. *Int J Syst Evol Microbiol* 2009;

823 **59**: 118–124.

824 60. Smet A, Cools P, Krizova L, Maixnerova M, Sedo O, Haesebrouck F, et al.
825 Acinetobacter gandensis sp. nov. isolated from horse and cattle. *Int J Syst Evol*
826 *Microbiol* 2014; **64**: 4007–4015.
827 61. Nemec A, Radolfova-Krizova L, Maixnerova M, Vrestiakova E, Jezek P,
828 Sedo O. Taxonomy of haemolytic and/or proteolytic strains of the genus Acinetobacter
829 with the proposal of Acinetobacter courvalinii sp. nov. (genomic species 14 *sensu*
830 Bouvet & Jeanjean), Acinetobacter dispersus sp. nov. (genomic species 17),
831 Acinetobacter modestus sp. nov., Acinetobacter proteolyticus sp. nov. and
832 Acinetobacter vivianii sp. nov. *Int J Syst Evol Microbiol* 2016; **66**: 1673–1685.
833 62. Nemec A, Radolfova-Krizova L, Maixnerova M, Sedo O. Acinetobacter
834 colistiniresistens sp. nov. (formerly genomic species 13 *sensu* Bouvet and Jeanjean and
835 genomic species 14 *sensu* Tjernberg and Ursing), isolated from human infections and
836 characterized by intrinsic resistance to polymyxins. *Int J Syst Evol Microbiol* 2017; **67**:
837 2134–2141.
838 63. Radolfova-Krizova L, Maixnerova M, Nemec A. Acinetobacter celticus sp.
839 nov., a psychrotolerant species widespread in natural soil and water ecosystems. *Int J*
840 *Syst Evol Microbiol* 2016; **66**: 5392–5398.
841 64. Álvarez-Pérez S, Lievens B, Jacquemyn H, Herrera CM. Acinetobacter
842 nectaris sp. nov. and Acinetobacter boissieri sp. nov., isolated from floral nectar of wild
843 Mediterranean insect-pollinated plants. *Int J Syst Evol Microbiol* 2013; **63**: 1532–1539.
844 65. Bouvet P and Grimont P. Taxonomy of the genus Acinetobacter with the
845 recognition of Acinetobacter baumannii sp. nov. Acinetobacter haemolyticus sp. nov.
846 Acinetobacter johnsonii sp. nov. and Acinetobacter junii sp. nov. and emended
847 descriptions of Acinetobacter calcoaceticus and Acinetobacter lwoffii. *Int J Syst*
848 *Bacteriol* 1986; **36**: 228–240.
849 66. Wirth SE, Ayala-del-Río HL, Cole JA, Kohlerschmidt DJ, Musser KA,
850 Sepúlveda-Torres L del C, et al. Psychrobacter sanguinis sp. nov., recovered from four
851 clinical specimens over a 4-year period. *Int J Syst Evol Microbiol* 2012; **62**: 49–54.

852 67. Baik KS, Park SC, Lim CH, Lee KH, Jeon DY, Kim CM, et al. Psychrobacter
853 aestuarii sp. nov., isolated from a tidal flat sediment. *Int J Syst Evol Microbiol* 2010; **60**:
854 1631–1636.
855 68. Bowman JP, Cavanagh J, Austin JJ, Sanderson K. Novel Psychrobacter
856 species from Antarctic ornithogenic soils. *Int J Syst Bacteriol* 1996; **46**: 841–848.
857 69. Kämpfer P, Albrecht A, Buczolits S, Busse H-J. Psychrobacter faecalis sp.
858 nov., a new species from a bioaerosol originating from pigeon faeces. *Syst Appl*
859 *Microbiol* 2002; **25**: 31–36.
860 70. Romanenko LA, Lysenko AM, Rohde M, Mikhailov VV, Stackebrandt E.
861 Psychrobacter maritimus sp. nov. and Psychrobacter arenosus sp. nov., isolated from
862 coastal sea ice and sediments of the Sea of Japan. *Int J Syst Evol Microbiol* 2004; **54**:
863 1741–1745.
864 71. Yumoto I, Hirota K, Sogabe Y, Nodasaka Y, Yokota Y, Hoshino T.
865 Psychrobacter okhotskensis sp. nov., a lipase-producing facultative psychrophile
866 isolated from the coast of the Okhotsk Sea. *Int J Syst Evol Microbiol* 2003; **53**: 1985–
867 1989.
868 72. Juni E, Heym GA. Psychrobacter immobilis gen. nov., sp. nov.: Genospecies
869 Composed of Gram-Negative, Aerobic, Oxidase-Positive Coccobacilli. *Int J Syst Evol*
870 *Microbiol* 1986; **36**: 388–391.
871 73. Yoon J-H, Lee C-H, Yeo S-H, Oh T-K. Psychrobacter aquimaris sp. nov. and
872 Psychrobacter namhaensis sp. nov., isolated from sea water of the South Sea in Korea.
873 *Int J Syst Evol Microbiol* 2005; **55**: 1007–1013.
874 74. Maruyama A, Honda D, Yamamoto H, Kitamura K, Higashihara T.
875 Phylogenetic analysis of psychrophilic bacteria isolated from the Japan Trench,
876 including a description of the deep-sea species Psychrobacter pacificensis sp. nov. *Int J*
877 *Syst Evol Microbiol* 2000; **50 Pt 2**: 835–846.
878 75. Yumoto I, Hirota K, Kimoto H, Nodasaka Y, Matsuyama H, Yoshimune K.
879 Psychrobacter piscatorii sp. nov., a psychrotolerant bacterium exhibiting high catalase

activity isolated from an oxidative environment. *Int J Syst Evol Microbiol* 2010; **60**: 205–208.

76. Poppel MT, Skiebe E, Laue M, Bergmann H, Ebersberger I, Garn T, *et al.* Acinetobacter equi sp. nov., isolated from horse faeces. *Int J Syst Evol Microbiol* 2016; **66**: 881–888.

77. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol* 2012; **3**: 217–223.

78. Karasov TL, Almario J, Friedemann C, Ding W, Giolai M, Heavens D, *et al.* Arabidopsis thaliana and Pseudomonas Pathogens Exhibit Stable Associations over Evolutionary Timescales. *Cell Host Microbe* 2018; **24**: 168–179.e4.

79. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 1982; **157**: 105–132.

80. Galan M, Pagès M, Cosson J-F. Next-generation sequencing for rodent barcoding: species identification from fresh, degraded and environmental samples. *PLoS One* 2012; **7**: e48374.

81. Goodrich JK, Waters JL, Poole AC, Sutter JL, Koren O, Blekhman R, *et al.* Human genetics shape the gut microbiome. *Cell* 2014; **159**: 789–799.

82. Dziewit L, Cegielski A, Romaniuk K, Uhrynowski W, Szych A, Niesiobedzki P, *et al.* Plasmid diversity in arctic strains of Psychrobacter spp. *Extremophiles* 2013; **17**: 433–444.

83. Collins C, Didelot X. A phylogenetic method to perform genome-wide association studies in microbes that accounts for population structure and recombination. *PLoS Comput Biol* 2018; **14**: e1005958.

84. Perez Vidakovics ML, Riesbeck K. Virulence mechanisms of Moraxella in the pathogenesis of infection. *Curr Opin Infect Dis* 2009; **22**: 279–285.

85. Zhou Q, Feng S, Zhang J, Jia A, Yang K, Xing K, *et al.* Two glycosyltransferase genes of Haemophilus parasuis SC096 implicated in lipooligosaccharide biosynthesis, serum resistance, adherence, and invasion. *Front Cell*

908 *Infect Microbiol* 2016; **6**: 100.

909 86. Weinberg ED. Iron availability and infection. *Biochim Biophys Acta* 2009;
910 **1790**: 600–605.

911 87. Collins T, Margesin R. Psychrophilic lifestyles: mechanisms of adaptation
912 and biotechnological tools. *Appl Microbiol Biotechnol* 2019; **103**: 2857–2871.

913 88. Feller G. Psychrophilic enzymes: from folding to function and biotechnology.
914 *Scientifica* 2013; **2013**: 512840.

915 89. Gerday C. Psychrophily and catalysis. *Biology* 2013; **2**: 719–741.

916 90. De Maayer P, Anderson D, Cary C, Cowan DA. Some like it cold:
917 understanding the survival strategies of psychrophiles. *EMBO Rep* 2014; **15**: 508–517.

918 91. Fields PA, Dong Y, Meng X, Somero GN. Adaptations of protein structure
919 and function to temperature: there is more than one way to 'skin a cat'. *J Exp Biol* 2015;
920 **218**: 1801–1811.

921 92. Bobay L-M, Ochman H. The evolution of bacterial genome architecture.
922 *Front Genet* 2017; **8**: 72.

923 93. Kuo C-H, Ochman H. Deletional bias across the three domains of life.
924 *Genome Biol Evol* 2009; **1**: 145–152.

925 94. Hughes DE, Pugh GW Jr. Isolation and description of a Moraxella from
926 horses with conjunctivitis. *Am J Vet Res* 1970; **31**: 457–462.

927 95. Whitman WB, Rainey F, Kämpfer P, Trujillo M, Chun J, DeVos P, *et al.* (eds).
928 Moraxella. *Bergey's Manual of Systematics of Archaea and Bacteria*. 2015. John Wiley
929 & Sons, Ltd, Chichester, UK, pp 1–17.

930 96. Feng S, Powell SM, Wilson R, Bowman JP. Extensive gene acquisition in the
931 extremely psychrophilic bacterial species Psychroflexus torquis and the link to sea-ice
932 ecosystem specialism. *Genome Biol Evol* 2014; **6**: 133–148.

933 97. Liu Y, Harrison PM, Kunin V, Gerstein M. Comprehensive analysis of
934 pseudogenes in prokaryotes: widespread gene decay and failure of putative horizontally
935 transferred genes. *Genome Biol* 2004; **5**: R64.

936 98. Bergholz PW, Bakermans C, Tiedje JM. Psychrobacter arcticus 273-4 uses
937 resource efficiency and molecular motion adaptations for subzero temperature growth. *J*
938 *Bacteriol* 2009; **191**: 2340–2352.
939 99. Hendry TA, Freed LL, Fader D, Fenolio D, Sutton TT, Lopez JV. Ongoing
940 transposon-mediated genome reduction in the luminous bacterial symbionts of deep-
941 sea ceratioid anglerfishes. *MBio* 2018; **9**.
942 100. Moran NA. Tracing the evolution of gene loss in obligate bacterial
943 symbionts. *Curr Opin Microbiol* 2003; **6**: 512–518.

944

945

Figure Legends

Figure 1. Genomic and phenotypic diversity in the family *Moraxellaceae*. A) The *Moraxellaceae* family phylogeny was constructed with 51 diverse *Moraxellaceae* genomes using the software PhyloPhlan, which constructs a phylogeny using fasttree with 1000 bootstraps, refined by RAxML under the PROTCATG model. Amino acid sequences from 400 marker genes were used in the alignment. Branches with bootstrap support greater than 70% are represented by filled circles. The scale bar represents the average amino acid substitutions per site. *Pseudomonas aeruginosa* was used as an outgroup. Clades are highlighted in colored blocks, and branches are colored by genus. Isolation source is depicted in a color strip, along with a heatmap of scaled notable genome characteristics that differ between the genera, with 0 representing the smallest value present and 1 the largest value (*P. aeruginosa* not included). GS = Genome size, ranging between 1.8 Mb and 4.5 Mb. PCoA = PC1 values from a PCoA based on gene presence/absence data. CD = genome coding density, ranging from 80% to 89%. Growth temperature range data was collected from type strain publications. B) PC1 and 2 of a PCoA analysis of a binary matrix of gene presence/absence for 51 species of *Moraxellaceae*, explaining 32% and 14% of the variation, respectively. Each genome is represented by one point, colored by genus. The P-clade *Moraxella* spp. are represented by yellow points with red outlines. The frequencies of genes associated with each COG category were associated with the PCoA axes as environmental vectors via the envfit function. All COG categories are significant (BH correction, p-value < 0.05), except L (replication, recombination and repair) and Z (cytoskeleton assembly and regulation) which are shown in pink vectors. COG categories with an $r^2 > 0.5$ are shown in dark grey, while categories with an $r^2 < 0.5$ are shown in light grey. C) PC2 and 3, explaining 14% and 5% of the variation.

Figure 2. *Psychrobacter's* restriction to cold temperatures is a newly emerged trait in the family *Moraxellaceae*. Continuous trait mapping growth temperature ranges of

973 51 species from the *Moraxellaceae* family, taken from type strain publications. Values at
974 nodes are imputed by maximum likelihood analysis. The phylogeny was constructed by
975 marker-gene analysis including 400 genes, as in Fig. 1. Genera are indicated with colored
976 boxes.
977
978 **Figure 3. *Psychrobacter* phenotypic and genomic diversity.** A) Using 85
979 *Psychrobacter* genomes, we constructed a genus-level phylogeny using fasttree with 1000
980 bootstraps, refined by RAxML under the PROTCATG model. Amino acid sequences from
981 400 marker genes were used in the alignment. Branches with bootstrap support greater than
982 70% are represented by filled circles. The scale bar represents the average amino acid
983 substitutions per site. *M. lincolnii* was used as an outgroup. Type strain isolate names are
984 indicated in bold and italicized type. Strains indicated with * next to their name exhibited
985 growth defects in liquid media, and were tested on solid agar media instead. Strains
986 indicated with ** exhibited growth defects on solid and liquid media, and were tested on solid
987 media supplemented with 0.1% Tween80. Strains indicated with a mouse silhouette were
988 later used in germ-free mouse colonization studies (Fig 5). Isolation source is depicted in
989 column 1 as a color strip. Columns 2 – 10 represent the growth probabilities of each strain
990 for each condition; media complexity is represented in yellow, salt concentration is
991 represented in blue, and temperature is represented in red. Type strain data supports our
992 temperature data except where indicated - colored triangles show conditions in which we
993 expected growth but did not observe it, while white triangles represent conditions in which
994 we observed growth we did not expect. The ecotype is shown in column 11 in a colorstrip,
995 and in the color of the branches. B) The first two PCs of a PCoA of a gene presence-
996 absence matrix of all 85 of the included accessions, colored by ecotype. The basal clade
997 strains are shown within the ellipse. The frequencies of genes associated with each COG
998 category were associated with the PCoA axes as environmental vectors. All COG categories
999 are significant (BH correction, p-value < 0.05), except G (carbohydrate transport and
1000 metabolism) and H (coenzyme transport and metabolism) which are shown in pink vectors.

COG categories with an $r^2 > 0.5$ are shown in dark grey, while categories with an $r^2 < 0.5$ are shown in light grey. C) The proportion of genes per genome falling in the highest quartile of "high cold adaptive amino acid traits" from each ecotype (n = 26 for FE, n = 59 for RE). D) the number of "core genes" (present in all 85 *Psychrobacter* accessions) that fall into the highest quartile of "high cold adaptive amino acid traits" (n = 26 for FE, n = 59 for RE). For all mean comparisons, the Wilcoxon rank sum test was used. ** indicates a p-value < 0.005, *** indicates a p-value < 0.0005.

Figure 4. Divergence of ecotypes could be driven by transposon-mediated genome reduction. A) proportion of Cluster of Orthologous Groups (COG) category "L" (replication, recombination, and repair related) genes per genome. B) copy numbers per genome of all the transposases in the *Psychrobacter* pan-genome. C) Proportion of predicted pseudogenes per genome. D) *Psychrobacter* accession genome size in megabases. For all mean comparisons, the Wilcoxon rank sum test was used. * indicates a p-value < 0.05, ** indicates a p-value < 0.005. For each comparison, n = 26 for FE strains and n = 59 for RE strains.

Figure 5. Psychrobacter strains occur and persist in the mammalian gut. A) Relative abundance at the genus level of 16S rRNA gene OTUs, clustered at 99% identity, of 86 polar bear fecal samples. *Psychrobacter* OTUs are colored red. B) *Psychrobacter* relative abundance in comparison to taxonomy of prey consumed by polar bears (n = 1 for *Actinoterygii*, n = 24 for *Aves*, n = 16 for mixed prey, and n = 32 for *Mammalia*). C) The average change in OD600 of 19 *Psychrobacter* accessions grown on 190 different substrates as sole carbon sources, compared across different classes of compounds. D) CFUs per gram of cecal contents of gnotobiotic mice is shown in comparison to accession phylogeny. 4 mice were tested per *Psychrobacter* strain, except *P. lutiphocae*, which was tested in 8 mice. FE strains are shown in yellow, while RE strains are shown in blue. The phylogeny was constructed as described in Fig. 3. Branches with bootstrap support higher

than 70% are indicated with a filled circle.

Means were compared using the Wilcoxon rank-sum test. *** indicates a p-value < 0.0005, ** indicates p-value < 0.005, * indicates p < 0.05.