Article

Current Biology

Evolutionary radiation of *Polaromonas* from mountain glaciers downstream

Highlights

- Last common ancestor of *Polaromonas* traced to glacier-fed streams
- Shared evolutionary history explains more gene content variation than habitat
- Habitat-transition nodes show elevated horizontal gene transfer and mobile elements
- Lake and wetland lineages gained phototrophy; glaciers gained anaerobic respiration

Authors

Grégoire Michoud, Aileen Geers, Hannes Peter, Amy C. Thorpe, Zhi-Ping Zhong, Virginia Rich, Tom J. Battin

Correspondence

gregoire.michoud@epfl.ch (G.M.), tom.battin@epfl.ch (T.J.B.)

In brief

Michoud et al. reconstruct the pangenome and evolutionary history of 282 *Polaromonas* genomes, predicting an ancestral origin in glacier-fed streams. Habitat transitions across aquatic and terrestrial ecosystems are accompanied by horizontal gene transfer and gene loss, which reveal how cryospheric environments may shape bacterial diversification and ecological success.

Michoud et al., 2026, Current Biology *36*, 1–13
August 3, 2026
https://doi.org/10.1016/j.cub.2026.06.070

CellPress

Please cite this article in press as: Michoud et al., Evolutionary radiation of *Polaromonas* from mountain glaciers downstream, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.06.070

Current Biology

CellPress
OPEN ACCESS

Article

Evolutionary radiation of *Polaromonas* from mountain glaciers downstream

Grégoire Michoud,[1,7,*] Aileen Geers,[1] Hannes Peter,[1] Amy C. Thorpe,[2] Zhi-Ping Zhong,[3,4,5,6] Virginia Rich,[4,5,6] and Tom J. Battin[1,*]

[1]River Ecosystems Laboratory, Ecole Polytechnique Fédérale de Lausanne, 1950 Sion, Switzerland
[2]UK Centre for Ecology & Hydrology, Benson Lane, Crowmarsh Gifford, Wallingford OX10 8BB, UK
[3]State Key Laboratory of Tibetan Plateau Earth System, Environment and Resources, Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100101, China
[4]Byrd Polar and Climate Research Center, Ohio State University, Columbus, OH 43210, USA
[5]Department of Microbiology, The Ohio State University, Columbus, OH 43210, USA
[6]Center of Microbiome Science, The Ohio State University, Columbus, OH 43210, USA
[7]Lead contact
*Correspondence: gregoire.michoud@epfl.ch (G.M.), tom.battin@epfl.ch (T.J.B.)
https://doi.org/10.1016/j.cub.2026.06.070

SUMMARY

Habitat transitions are central to microbial ecology and evolution and have been extensively studied across vastly different environments, such as between saline and non-saline environments. However, microbial habitat transitions along other large-scale environmental gradients remain poorly studied. This is particularly true for transitions involving the cryosphere, despite building evidence suggesting the Cryogenian as important for evolutionary radiation. Here, we investigated ecosystem transitions and the related genomic adaptations of the cosmopolitan cryospheric *Polaromonas* bacterium. We constructed a pangenome from 282 high-quality genomes, sourced from glaciers, glacier-fed streams (GFSs), lakes, wetlands, groundwater, rivers, and soils. Phylogenetic reconciliation suggested that the ancestral *Polaromonas* genome radiated from glacier ecosystems into various downstream environments through multiple independent transitions. These transitions were likely marked by extensive horizontal gene transfer and gene loss, with mobile genetic elements such as plasmids and prophages playing key roles in genomic diversification. Predicted ancestral genomes encoded versatile metabolic and stress-response capacities, which support adaptation to fluctuating and extreme conditions in the various cryospheric habitats. Compared to the ancestral *Polaromonas* genome, distinct genomic signatures were associated with specific habitats: GFS lineages possess expanded stress-tolerance repertoires, glacier lineages gained chemolithotrophic and anaerobic pathways, lake and wetland genomes acquired phototrophic functions, and soil lineages expanded substrate transport and stress tolerance. Together, our findings highlight the role of genomic plasticity in the ecological success of *Polaromonas* and also underscore the cryosphere as a potential evolutionary cradle from which lineages dispersed and adapted to downstream aquatic and terrestrial environments.

INTRODUCTION

Ecological or habitat transitions, such as the conquest of land by plants or transitions between marine and freshwater environments, are a major driver of evolution across the tree of life.[1–4] Plant terrestrialization, which necessitated genomic adaptations to high irradiance, desiccation, and oxygen stress,[5] represents a paragon among habitat transitions. Recent evidence points to the Cryogenian period (i.e., Snowball Earth) as potentially important for plant terrestrialization and evolution of multicellularity.[6,7] In contrast to eukaryotes, records of habitat transitions and their importance for the ecology and evolution of prokaryotes focused on a few events such as the marine-freshwater divide, soil and aquatic ecosystems, host-associated lifestyles, or endosymbiotic events.[2,8,9] Habitat transitions of bacteria are often associated with functional adaptations mediated by genome streamlining and horizontal gene transfer.[2] For instance, the transition of members of *Methylophilaceae* from freshwater sediments to ocean pelagic habitats,[10] or the transition from marine to freshwater environments of *Candidatus* Pelagibacterales (SAR11).[11] In general, the genes related to stress response, motility, energy, and carbon metabolism are often associated with bacterial habitat transitions.[2] Despite their important ramifications for bacterial ecology and evolution, how the genomic underpinnings of habitat transitions are related to the present-day success of bacterial clades remains generally understudied.

High dispersal capacity, metabolic versatility, stress tolerance, the ability to take up exogenous DNA, and a large genomic pool allowing for adaptation are key requirements for successful habitat transitions.[2] Moreover, cryospheric environments, characterized by high spatio-temporal heterogeneity and connectivity between frozen and liquid habitats, may represent hotspots

Please cite this article in press as: Michoud et al., Evolutionary radiation of *Polaromonas* from mountain glaciers downstream, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.06.070

for habitat transitions—as exemplified by plant terrestrialization.[7] Similar to global glaciation during Snowball Earth, the cryosphere continues to be a highly selective environment for microbial life.[12] Microbial clades specialized to the cryospheric conditions are adapted to cope with low temperatures, osmotic and radiative stress, and ultra-oligotrophy.[12]

Polaromonas is an iconic bacterial genus with a cosmopolitan distribution and is reported from numerous cryospheric (e.g., glaciers and permafrost soils)[13–16] and non-cryospheric (e.g., wetlands and groundwater) ecosystems.[17–19] *Polaromonas* is metabolically versatile,[20] possessing genes involved in both heterotrophic processes that are able to degrade organic carbon[21] and genes involved in the degradation of recalcitrant organic compounds such as hydrogen[22] or arsenic.[23] Some species are even able to fix nitrogen,[24] although this metabolic pathway is not widely distributed within the genus.[20] Further, in contrast to *Prochlorococcus*, a cosmopolitan marine bacterium,[25] the *Polaromonas* genome is not streamlined; it has an average size of 3–5 Mb and a GC content of 50%–60%,[21,26] typical for cryospheric taxa.[12] Comparative genomics has revealed the role of horizontal gene transfer in shaping the genetic diversity and adaptive potential of *Polaromonas* to cold environments.[27] Compared with mesophilic species (e.g., *Polaromonas naphthalenivorans* CJ2), cold-adapted *Polaromonas* possesses more (approximately 10%) horizontally transferred and transposase genes, which collectively suggests a greater potential for genome rearrangement and the acquisition of novel genetic material.[27]

The apparent ease with which *Polaromonas* disperses over long distances[20,28] may facilitate habitat transitions and explain its occurrence in air and snow at high altitudes.[29–31] Furthermore, local selection pressures and the interactions of *Polaromonas* with other bacteria and algae drive niche separation and microevolution within the genus.[32] Such microevolution may lead to the formation of *Polaromonas* phylotypes associated with distinct environments as reported from glacier habitats,[33] and further highlighting the adaptive capabilities of the genus. Indeed, a recent study has shown a clear functional selection between eight *Polaromonas* found in arctic environments compared with those found elsewhere (e.g., glaciers, groundwater, and wastewater), where the former tended to have extended genomes (on average, 500 kb) with an increase in genes related to osmoprotection and carbohydrate metabolism.[34] An additional study focusing on glacier-fed streams (GFSs) *Polaromonas* genomes against non-GFS strains showed an increase in genes related to defense mechanisms and energy production in the GFS *Polaromonas* genomes, which is probably linked to pronounced oligotrophy and low temperature in this peculiar environment.[35] These two studies, however, did not address the transition and evolutionary history of *Polaromonas* across cryospheric environments and from these to non-cryospheric ones. One hypothesis related to this radiation across environments is related to ecosystem relocation on Snowball Earth,[36] where the expansion of polar-alpine biomes across earth allowed the migration of adapted bacteria, thus allowing them to survive, evolve, and later repopulate post-glacial environments. The Cryogenian Snowball Earth, 650 million years ago, was important for the evolution and radiation of various lineages, such as Cyanobacteria and Chlorophytes,[36] and through habitat transitions into various environments that emerged upon deglaciation. The Cryogenian also left genomic and physiological imprints on bacteria, such as *Prochlorococcus*, which is a widespread primary producer in the modern ocean.[37] It seems, therefore, intuitive to hypothesize glaciers as cradles of microbial diversity. Ice-locked microorganisms adapt to the environment of glaciers from where they are released downstream with the meltwaters to colonize various aquatic and terrestrial ecosystems. Furthermore, glaciers can also be considered "refugia" for cryospheric microbes during periods of deglaciation, thus providing a longer-term cold habitat.[38] Therefore, we argue that *Polaromonas* is a good model system for studying bacterial habitat transitions, and even the potential role of the cryosphere for its evolution.

In this study, we explored the *Polaromonas* pangenome containing 282 high-quality genomes sourced from a wide range of cryospheric (e.g., glaciers and permafrost soils) and non-cryospheric (e.g., wetlands, lowland rivers, and groundwater) ecosystems. This enabled us to explore the breadth of genomic variation across *Polaromonas* populations and unravel the mechanisms that allow these bacteria to colonize diverse and often extreme environments. Using phylogenetic reconciliation,[39] we provide novel insights into the genome evolution of *Polaromonas*, assess its likely ancestral state, and the rearrangement of its genetic repertoire as it transitioned across habitats. Collectively, our findings shed new light on the ecology and evolution of a most successful bacterial clade, with its roots likely in the world's glaciers and their streams that are now changing at a rapid pace because of climate change. Our results support the notion of the cryosphere as potentially playing a role in microbial diversification with impacts on downstream ecosystems. Therefore, our findings provide a microbial perspective on how the cryosphere has shaped and continues to influence the global genetic and ecological landscape.

RESULTS AND DISCUSSION

Distribution and diversity of *Polaromonas*

First, we explored the spatial distribution of the *Polaromonas* genus to assess its capacity to colonize and dwell in various environments. To this end, we analyzed 16S rRNA gene data from the Microbe Atlas Project.[40] We found *Polaromonas* to be present across all continents and in a wide range of ecosystems and habitats, including rivers, lakes and reservoirs, groundwater, glacier ice and cryoconite, soils, sediments, and both animal and plant hosts (Figure S1). The majority of *Polaromonas* records were located in aquatic environments, including glaciers and ice (243, 63.4%), followed by soils and sediments (126, 32.9%), and animal and plant hosts (14, 3.65%). In oceans or saline lakes, *Polaromonas* was found only at low relative abundances (<0.1% relative abundance), with the notable exception of estuaries with a low temperature and low salinity (e.g., Finland and Antarctica),[28,41] thus suggesting that the genus has not globally passed the salinity barrier.[42]

In the next step, we compiled a list of genomes to investigate the pangenome of *Polaromonas*, including both metagenome-assembled genomes (MAGs), single amplified genomes (SAGs), and genomes sequenced from isolates. We collected genomes from the NCBI GenBank,[43] Tibetan aquatic

Please cite this article in press as: Michoud et al., Evolutionary radiation of *Polaromonas* from mountain glaciers downstream, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.06.070

CellPress
OPEN ACCESS

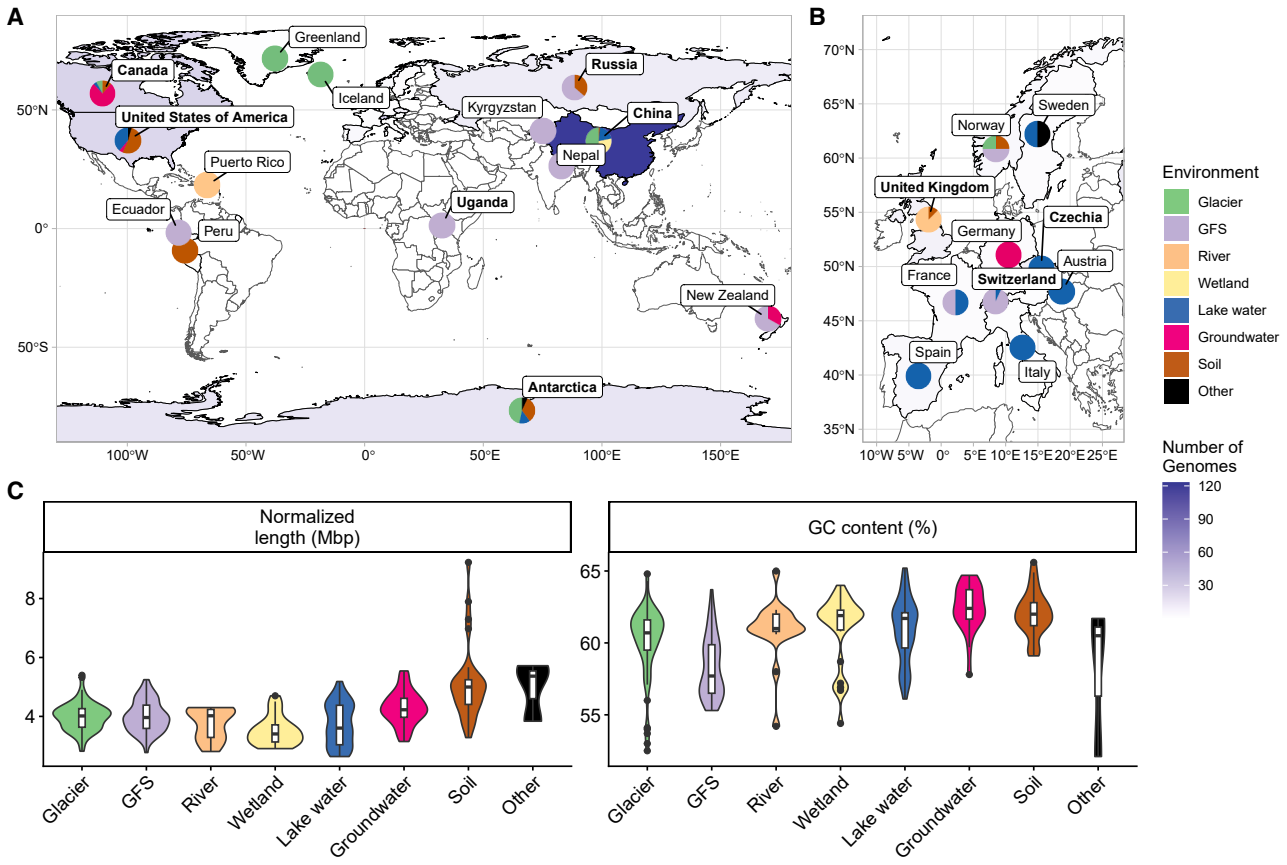

Figure 1. Locations of the origins of *Polaromonas* used in this study

(A and B) For each country or region, the shading corresponds to the number of genomes obtained, and the pie chart represents the environment from which these genomes were obtained. (B) represents European countries for visualization purposes. The "other" category corresponds to the cultured isolates obtained from other environments.

(C) Normalized length (by completeness) and GC content (%) of *Polaromonas* genomes across different environments. Box plots show median (center line), interquartile range (IQR, box), and whiskers extending to 1.5× IQR; violins show the full distribution density of values across the different environments.

See also Figure S1 and Tables S1 and S2.

microbiomes,[44] Tibetan glacier microbiomes,[45] GFS microbiomes,[16] the Guliya ice-cap,[46] and UK[19] and US river microbiomes.[18] These genomes were filtered for quality (STAR Methods), and taxonomies were confirmed using the Genome Taxonomy Database (GTDB).[47] Among the 282 non-redundant genomes retained, 21 grouped into 10 species-level taxa based on average nucleotide identity (average nucleotide identity $\geq$ 95%), while the remaining genomes appear to be monotypic, each representing a distinct species. All 282 genomes were obtained from a wide range of cryospheric and non-cryospheric environments, including glaciers (73), GFSs (54), lowland rivers (17), wetlands (26), lakes (55), groundwater (23), soils (31), and other habitats (e.g., where isolates were found) (3) (Tables S1 and S2; Figures 1A and 1B). We acknowledge that these definitions of environments lack precision because metadata often remains vague. For instance, for most glacially derived genomes, we were unable to differentiate between the various glacier habitats (e.g., englacial or supraglacial). Additionally, further increasing the number of habitat categories would reduce sample size and thus statistical power of the analyses.

Accounting for completeness (86.1% ± 9.9%) and contamination (1.4% ± 1.8%) *Polaromonas* genome length averaged 4.04 ± 0.8 Mbp, with an average GC content of 60.4% ± 2.5% (Table S2; Figure 1C). *Polaromonas* genomes differed in size between environments, with genomes sourced from soils being significantly larger (5.1 ± 1.26 Mbp [mean ± SD], Mann-Whitney U test, $p < 0.05$) than those from other environments, with the exception of groundwater genomes and type strains, which had similar sizes (Table S2; Figure 1C). These results were expected as "terrestrial" genomes tend to be larger than "aquatic" genomes, and genomes from isolates are larger than MAGs.[48] The distribution of GC content was similar to that previously reported and did not significantly differ between environments.[34]

Genomic differentiation across environments

To explore how *Polaromonas* genomes differed across the various environments, we analyzed the pangenome based on persistent, shell, and cloud genes (STAR Methods). We sourced a total of 958,785 genes from the 282 genomes, which we

Please cite this article in press as: Michoud et al., Evolutionary radiation of *Polaromonas* from mountain glaciers downstream, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.06.070

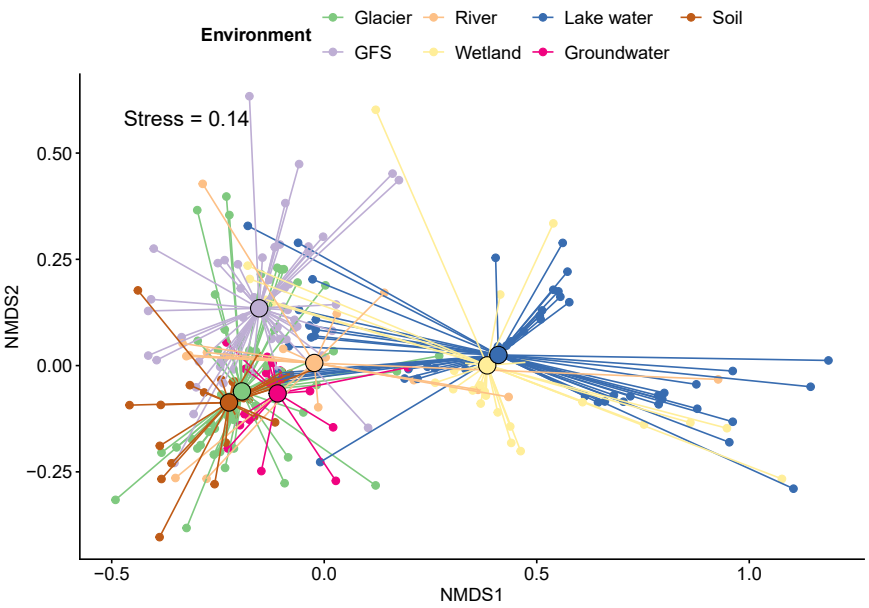

Figure 2. Non-metric multidimensional scaling ordination of persistent gene presence in different genomes shows that *Polaromonas* clusters by source environment (ANOSIM, R = 0.35, $p = 0.001$)

See also Figure S2 and Table S3.

arranged into 214,664 gene families, further divided into 1,905 persistent, 11,772 shell, and 200,987 cloud unique gene families (Figure S2A). On average, each persistent, shell, and cloud gene family was found in 61.3% ± 22.7%, 7.6% ± 4.9%, and 0.6% ± 0.6% of the genomes, respectively (Figure S2B). Based on the presence or absence of persistent genes, we observed a clear environmental segregation (Figure 2). This environmental segregation was consistently observed regardless of which gene categories were analyzed, but it was most evident for cloud (analysis of similarities [ANOSIM], R = 0.47, $p = 0.001$) and shell genes (ANOSIM, R = 0.51, $p = 0.001$), compared with persistent genes (ANOSIM, R = 0.35, $p = 0.001$) (Figures S2C and S2D). This pattern can be attributed to the fact that cloud and shell genes are less frequently shared across the pangenome[49] and are therefore likely less essential to the core functionality of the organism than persistent genes.[50] Hence, they are more likely to encode functions that facilitate adaptation to specific environmental conditions, thus driving the clearer segregation across environments. To evaluate potential confounding effects, we assessed the influence of phylogenetic relatedness on these patterns. Variation partitioning using distance-based redundancy analysis indicated that the phylogenetic structure explained a larger proportion of variation (R^2 = 26.7% for persistent genes) than the environmental category (R^2 = 3.3%), with a substantial shared fraction (R^2 = 20.2%) reflecting collinearity between the two (Table S3). Comparable patterns were observed for shell genes, whereas most variation in cloud genes remained unexplained.

To further disentangle the contributions of phylogeny and environment at the functional pathway level, we tested the association between functional pathways and habitat using phylogenetic generalized least squares (PGLS) using Pagel's λ (STAR Methods). Out of the 125 functional pathways, 36 showed a significant association with habitat; their jointly estimated λ values revealed three different categories (Table S4). First, three pathways (e.g., methane oxidation, aerobic ammonia oxidation, and cold-stress mRNA chaperone) were classified as habitat-driven ($\lambda \approx 0$, $p < 0.05$). This indicates that their distribution across genomes is independent of phylogenetic relatedness and therefore explained by ecological selection or habitat-specific horizontal gene transfer. Nine pathways showed habitat associations with an intermediate phylogenetic signal ($0.3 < \lambda < 0.6$, $p < 0.05$), which included thiosulfate oxidation, denitrification to N_2 and N_2O, the EMP glycolytic pathway, nitrogen transport, and formaldehyde oxidation. These represent habitat enrichments that are partially but not fully explained by clade structure. The remaining significant pathways ($\lambda \approx 1$, $p < 0.05$) (e.g., anoxygenic photosystem II, bacteriochlorophyll g, sulfur and sulfite oxidation, and several carbohydrate transport systems) showed strong phylogenetic conservation, meaning their observed habitat associations are substantially driven by clade-habitat co-occurrence rather than independent ecological adaptation.

Together, these results indicate that gene content variation in the *Polaromonas* genus is primarily structured by shared evolutionary history, with a smaller but detectable contribution of environmental differentiation. Accordingly, the observed environmental segregation should be interpreted cautiously and not solely attributed to adaptive processes. A similar environmental segregation was previously observed based on a pangenome that included 34 *Polaromonas* genomes with lower environmental resolution (GFS genomes versus non-GFS genomes)[35] and a study comparing polar to non-polar *Polaromonas*.[34] Interestingly, we found a clear clustering of genomes from lake water and wetlands on the one hand, and from GFSs and glaciers, soils, lowland rivers, and groundwater on the other (Figure 2). While this pattern is consistent with differentiation in gene repertoires across environments, it thus likely reflects a combination of ecological and phylogenetic signals rather than adaptation alone. Similar findings have been reported for transitions of Planctomycetes from soil to freshwater and *Methylophilaceae* from lakes to oceans.[10,51] In those cases, many changes occurred through gene loss via genome streamlining to adapt to mostly oligotrophic conditions. Here, however, we do not observe any signs of genome reduction, thus indicating different modes of evolution to adapt to the various environments.

Ancestral habitat and transitions

To further explore the importance of habitat transitions for the radiation of *Polaromonas*, we built a phylogenetic tree based on persistent and shell genes present in at least 30 genomes (i.e., ~10% of the pangenome) (STAR Methods). Our phylogenetic

Please cite this article in press as: Michoud et al., Evolutionary radiation of *Polaromonas* from mountain glaciers downstream, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.06.070

CellPress
OPEN ACCESS

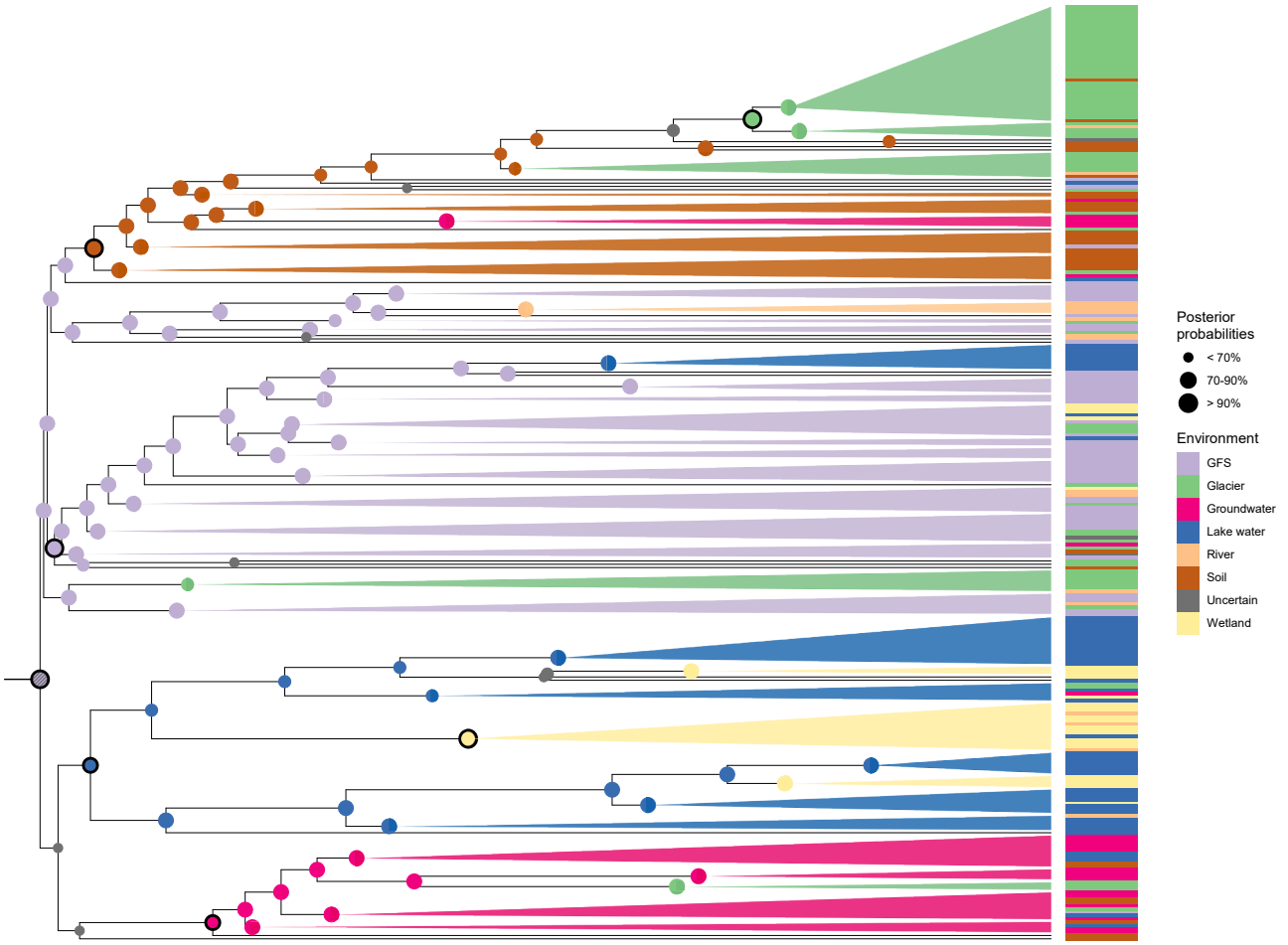

Figure 3. Phylogenetic tree of *Polaromonas* genomes

The color at each tip corresponds to the environment where the individual genomes were obtained, while the color at each node highlights the predicted environment where the *Polaromonas* ancestor was found via marginal posterior probabilities approximation using the PastML software.[53] Clade merging was also carried out using PastML, employing vertical merging, which groups adjacent sections of the tree where no state changes occur, and horizontal merging, which clusters independent events of the same type. Gray colors at the nodes indicate that there is uncertainty about the predicted environments. Node sizes correspond to the posterior probabilities as determined by PastML. Highlighted nodes (in black) indicate predicted major habitat transitions.
See also Figures S3 and S4 and Tables S4 and S5.

analysis revealed four main clusters, grouping most genomes from groundwater, lake water, and wetland, GFSs, and both soils and glaciers, respectively (Figure 3). This pattern suggests notable phylogenetic conservatism in habitat preference and a consistent grouping of closely related lineages in similar environments. Hence, environmental filtering, rather than dispersal, would play a major role in shaping microbial compositions. Indeed, if dispersal were more frequent and successful, we would expect "mixed" phylogenetic signatures where related lineages would be found across multiple environments without clear clustering. Instead, the environmental clustering indicates that selective pressures are sufficiently strong to maintain habitat-specific lineages, limiting successful establishment outside their environments. These results would support the niche conservatism hypothesis, which states that species tend to retain their ancestral ecological traits over time.[52] Furthermore, our results indicate that lineages often remain specialized to the environments occupied by their ancestors. As a result, closely related genomes are more likely to be found in similar habitats, while their transition across habitats seems rare.

Next, using PastML,[53] we predicted the potential ancestral habitat at each internal node of the phylogenetic tree and inferred the habitat transitions of *Polaromonas* (STAR Methods). Interestingly, PastML rooted the *Polaromonas* phylogenetic tree in GFSs with posterior probabilities of 92.9% (Figure 3). These results were also observed when estimating ancestral habitats using other statistical approaches (i.e., Bayesian stochastic mapping versus maximum-likelihood-based marginal reconstruction) (STAR Methods). Node-wise agreement was above 98.9% between the three methods, with probabilities of the root of the phylogenetic tree also being between 92% and 93% for the other methods tested. Despite this, all methods tend to overestimate the probabilities of predictions close to the root. We therefore urge caution in interpreting the predicted environment of the *Polaromonas* last common ancestor (LCA) and consider it uncertain.[54]

Please cite this article in press as: Michoud et al., Evolutionary radiation of *Polaromonas* from mountain glaciers downstream, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.06.070

Our analyses further predicted that *Polaromonas* transitioned from its putative ancestral environment in GFSs into groundwater, lakes, and wetlands, as well as soils and glaciers. The *Polaromonas* LCA has remained in GFSs since an early ancestral stage, while other lineages have radiated to soils and then to glaciers. Another lineage diverged into groundwater, while a last one transitioned from lakes into wetlands. To determine whether these patterns were sensitive to environmental misclassification, we performed a sensitivity analysis by randomly misclassifying up to 20% of the genomes' environment (Table S5). Node-state agreement with the original classification remained above 90% when 10% of the genomes were misclassified, and above 85% even when misclassification reached 20%. These results indicate that minor classification errors do not affect the observed phylogenetic and habitat-transition patterns.

While these broad patterns of habitat transitions appear relatively clear, we also detected signatures of multiple transitions occurring throughout the *Polaromonas* tree (Figure 3). Up to 10 MAGs from the same environments (e.g., rivers) were found throughout the phylogenetic tree, thus suggesting that some habitat transitions occurred several times. However, this pattern could also arise from missing intermediate lineages or undersampling of certain environments (e.g., lowland rivers), and therefore it cannot be considered definitive evidence of repeated and independent transitions. Nevertheless, numerous habitat transitions have been predicted to occur multiple times, which seems to be particularly true for host-associated bacteria, where some bacteria transition from their host to the environment very often, sometimes daily (e.g., *Vibrio fischeri* in a squid).[2,55]

Furthermore, we hypothesize that transitions are linked to cycles of deglaciations during which ice sheets created "climate refugia"[38] for cold-adapted microbes. This was already suggested for some plants and animals, especially in rock-covered glaciers,[56,57] where climate refugia are areas large enough to support populations whose habitat is at risk due to a changing climate. As glaciers retreated and expanded, dispersal strongly varied through hydrological mixing, which may have facilitated the genetic mixing and diversification of *Polaromonas* across some habitats. Notably, we found genomes from lowland rivers throughout the tree that were usually phylogenetically close to genomes from GFSs in several small MAG clusters (Figure 3). This indicates that lowland rivers act as collectors of *Polaromonas* MAGs from soil or groundwater, for instance.[58,59]

Gene re-arrangements

To gain mechanistic insights into the gene rearrangements that may accompany the habitat transitions of *Polaromonas*, we focused on the internal nodes where the major habitat transitions are predicted to occur, besides the root (LCA). These transitions occurred at several stages, including the adaptation of LCA lineages to GFSs, groundwater, lakes, and soils, as well as transitions of *Polaromonas* from lakes to wetlands and from soils to glaciers.

For each of these transitions, we determined the relative importance of duplications, transfers, losses, and origination (i.e., DTLO) for each predicted gene family (Figures 3 and S3). While gene inheritance is usually vertical from mother to daughter cells, prokaryotic genes are susceptible to relatively high frequencies of transfers and losses, less so for duplications.[39] Briefly, transfers either occurring through horizontal gene transfers or genetic recombination tend to homogenize gene pools.[60] Gene loss is also frequent in prokaryotes and tends to reduce the metabolic capability of the cell in the case of symbiosis, but can also act as an adaptive evolutionary force that is useful when organisms face abrupt environmental challenges, as encapsulated by the less-is-more hypothesis. This states that, for example, losing genes involved in the biosynthesis of compounds that are available in the environment will allow the cells to use their energy for other purposes, such as biofilm formation.[61]

We found that transfers (242.3 ± 149.4 per node) and losses (228.3 ± 160.1 per node) dominated DTLO events, compared with duplications (28.6 ± 42.7 per node) and originations (40.2 ± 107.3 per node). This pattern is consistent with general DTLO observations from prokaryotes (Figure S3).[39]

As the number of horizontal transfers was relatively high, we further looked for genomic evidence of such events. First, we predicted the presence of plasmids and viruses in each genome using Genomad.[62] Due to the fact that most of the genomes were assembled from metagenomes, we acknowledge possible biases associated with this analysis. Nevertheless, we found that on average 95.2% ± 5.0% (mean ± SD) of the genomes possessed at least one plasmid, which ranged from 80.6% to 100% of the genomes present in soils and rivers, respectively. Furthermore, many genomes possessed either a prophage or a virus (50.3% ± 14.5% of genomes); 80% of the wetland genomes are predicted to have been infected by viruses, while only 29% of glacier genomes were infected, potentially due to the higher cell densities in wetlands or soils than in glaciers (Figure S4A). As expected, most genes on these mobile genetic elements (MGEs) were predicted to have been transferred at the habitat transitions (Figure S4B). To test this observation, we performed statistical analyses to determine where transition nodes were enriched in horizontal transfers as well as MGEs. Compared with others, transition nodes exhibited significantly higher counts across all measures (p = 0.004–0.008), with large effect sizes (r = 0.57–0.62) and 95% confidence intervals ranging from 0.35–0.40 (lower bound) to 0.75–0.77 (upper bound) (Table S6). Together, these results could explain why the *Polaromonas* genus is microdiverse, which is where adaptations are observed at a finer phylogenetic level than species.[20] Indeed, the relevance of viruses for the diversity of gene pools across communities has been documented, and it has been suggested that the maintenance of high strain diversity prevents the accumulation of resources by specific strains.[63,64] The high abundance of MGEs and their presence at habitat transitions could facilitate the adaptation of *Polaromonas* species to changing environmental conditions, either by selecting against other populations or allowing selected species to diversify in the different environments.

The LCA of *Polaromonas*

Having unveiled the major habitat transitions of *Polaromonas*, we further assessed the genomic traits of its LCA (STAR Methods). First, we note that, unsurprisingly, the vast majority of DTLO events were originations (315) (Figure S3). Our results suggest a range of physiological and metabolic traits that enabled the LCA to adapt to diverse and fluctuating environmental

Please cite this article in press as: Michoud et al., Evolutionary radiation of *Polaromonas* from mountain glaciers downstream, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.06.070

CellPress
OPEN ACCESS

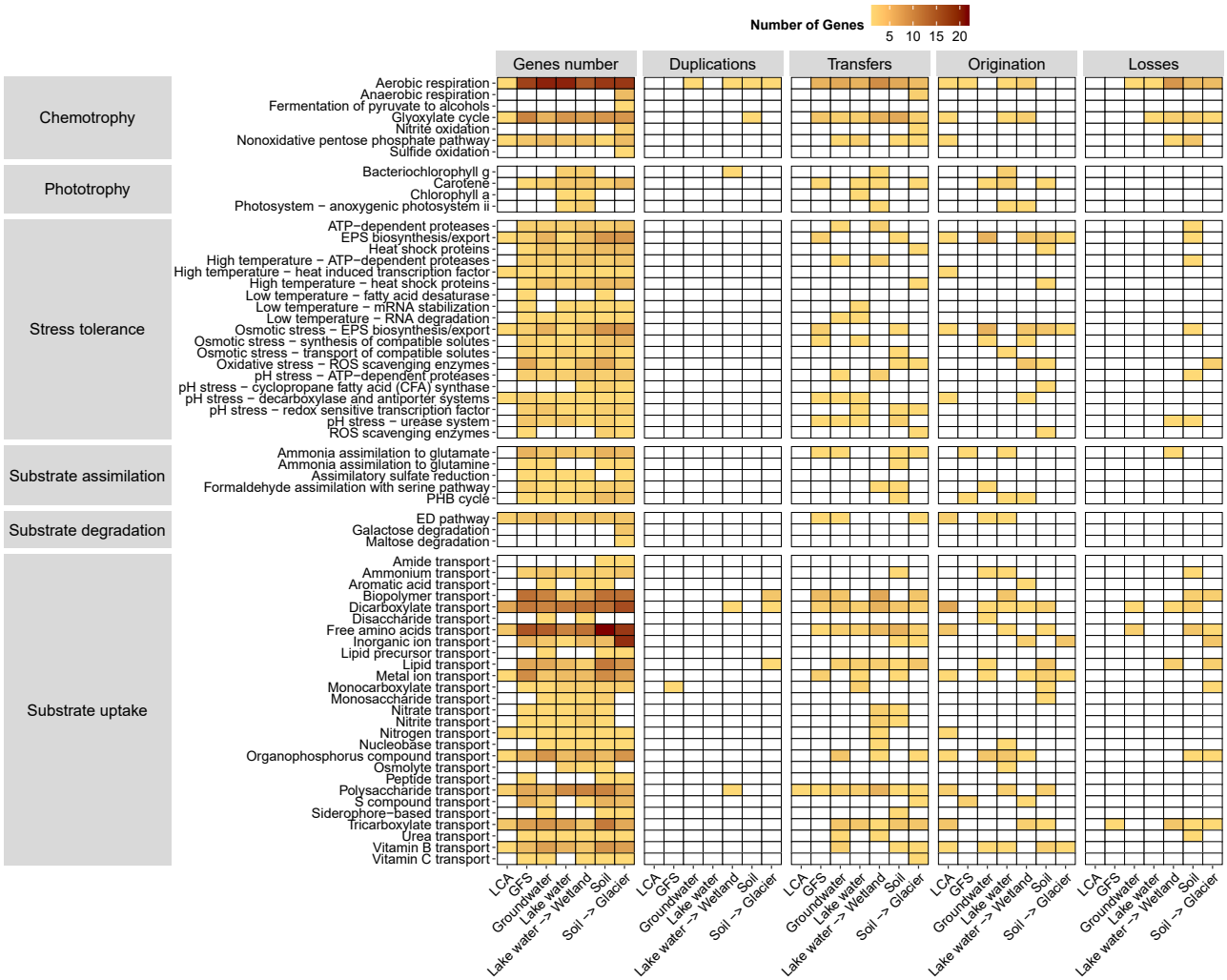

Figure 4. Heatmap representing duplications, transfers, originations, and losses of gene families associated with metabolic, stress response, and transport functions across six habitat adaptations and transitions of *Polaromonas*

Gene numbers correspond to numbers of representative genes per function (as determined by microtrait[65]) found in the different genomes, while color intensity reflects the number of genes per category.

See also Table S4.

conditions (Figures 4 and 5). Most likely, it displayed the capacity for aerobic respiration and heterotrophy (i.e., ETC complex IV), allowing for the exploitation of various organic energy sources. Furthermore, our predictions suggest that it is able to synthesize and export extracellular polymeric substances (EPSs), which are commonly associated with biofilm formation. It also possessed genes related to pH (decarboxylase and antiporter systems) and temperature stress (heat-induced transcriptional regulators). Finally, the *Polaromonas* LCA likely possessed diverse substrate uptake systems, including transporters for carboxylates (both dicarboxylates and tricarboxylates), simple and complex carbohydrates (including a complete Entner-Doudoroff pathway), free amino acids, metal ions, nitrogenous compounds, organophosphorus compounds, and vitamin B.

Together, these traits are likely consistent with a biofilm mode of life capable of exploiting a wide range of substrates. Indeed, these traits are advantageous for dwelling in GFSs, where biofilms reduce the turbulence-induced erosion of microbial cells and energy sources fluctuate over time.[16,35,66] Although environmental predictions suggest that the *Polaromonas* LCA dwelled in GFSs, the relatively small number of genes predicted for its genome (330) compared with those in other habitat transitions (735.7 ± 219.8) makes functional comparisons challenging. In fact, the absence of a gene in these predicted genomes cannot be taken as definitive evidence of its true absence. Hence, we did not consider that the LCA transitioned directly from GFSs to other ecosystems; rather, we considered adaptations of the LCA to specific environments (GFSs, soils, lakes, and groundwater), as well as transitions from lakes to wetlands and from soils to glaciers.

Genomic adaptations to habitat transitions

The traits of the *Polaromonas* LCA diversified with each transition downstream to aquatic and terrestrial ecosystems

Please cite this article in press as: Michoud et al., Evolutionary radiation of *Polaromonas* from mountain glaciers downstream, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.06.070

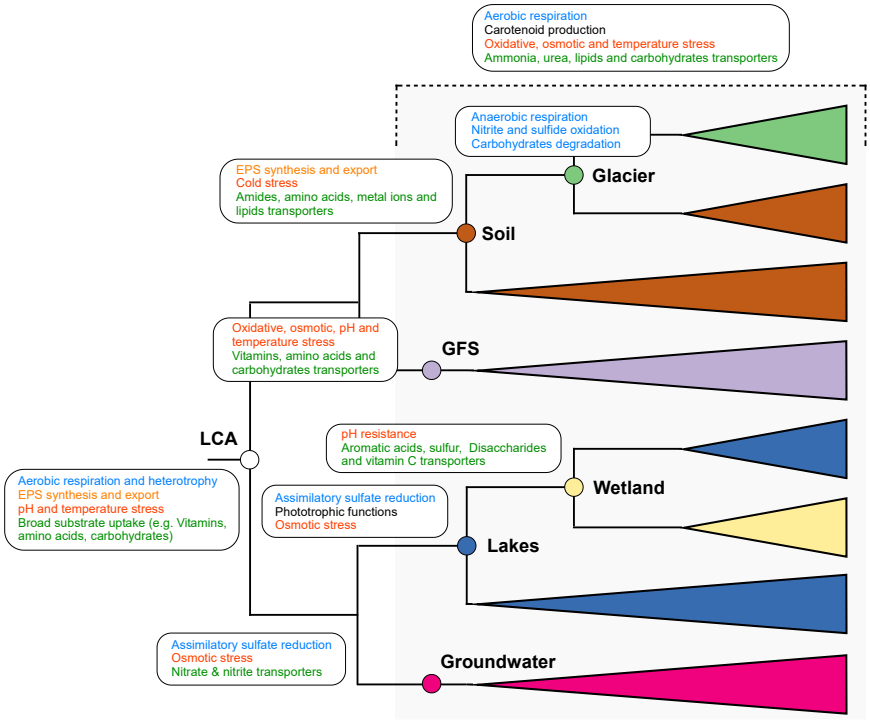

Figure 5. Schematic representation of functions gained among different habitat transitions

The gray square represents function gains that occurred in all habitat transitions.

(Figures 4 and 5). Overall, the predicted node-specific genomes at the various habitat adaptations and transitions (GFSs, groundwater, lakes, soils, transitions from lakes to wetlands, and transitions from soil to glaciers) tended to possess more metabolic genes compared with the LCA genome. Genomes downstream of the LCA were characterized by a more complete aerobic respiration mechanism with the four clusters of electron transport chain complexes I–IV. Basic chemotrophy functions were also present in all the genomes at the habitat transitions, such as the nonoxidative pentose phosphate pathway, the glyoxylate cycle and ammonia assimilation pathways (i.e., to glutamate). We also noted the presence of genes involved in the biosynthesis of a carotenoid, neurosporene, an intermediate used in the biosynthesis of lycopene and other carotenoids, which are involved in protection against UV radiation and oxidative stress.[67] In addition to the many transporters observed in the LCA genome, the genomes present at different habitat transitions possessed genes related to the substrate uptake of ammonium, urea, biopolymers, lipids, and polysaccharides. Furthermore, genes related to stress, such as heat shock proteins, ATP-dependent proteases, transporters of compatible solutes, and reactive oxygen species (ROS)-scavenging enzymes, may afford protection against high temperatures, elevated salt conditions, and oxidative damage. Collectively, this gene repertoire may allow *Polaromonas* species to exploit diverse carbon and nitrogen sources while withstanding thermal, osmotic, pH, and oxidative stresses. These traits facilitated the radiation of *Polaromonas* into the various aquatic and terrestrial ecosystems.

Despite the genetic repertoire that genomes at different habitat transitions shared, we note distinct genomic differentiations occurring that depend on the environment where those genomes were found (Figures 4 and 5). For example, genomes that transition to GFSs possess numerous genes related to the adaptation to cold temperatures—surprisingly also elevated temperatures—including cold shock proteins, fatty acid desaturases, and RNA-stabilizing enzymes, along with heat shock proteins. Furthermore, they also possess genes related to oxidative and osmotic stress, probably to counteract the effects of the high UV radiation at high altitudes, and redox shifts associated with the periodic drying of GFSs. An elevated number of genes related to the transport of amino acids, peptides, sugars, organophosphates, and vitamins collectively hint at interactions with algae and the recycling of nutrients within biofilms in GFSs.[16,35,68] Interestingly, a relatively high proportion of these genes were predicted to be horizontally transferred from other *Polaromonas* strains. We interpret this gene repertoire of *Polaromonas* as an adaptive response to the extreme and fluctuating environment (e.g., water flow, temperature, and light) in GFSs, driven by glacier melt dynamics.[15,69] We note that the environment of GFSs is more extreme and less stable compared with wetlands, soils, and groundwater, for instance.

Along the hydrological continuum, the transition into lakes from the *Polaromonas* ancestor was marked by several strikingly important genomic modifications that facilitate a planktonic mode of life (Figures 4 and 5). In fact, our results suggest that *Polaromonas* genomes are significantly enriched in phototrophic genes with the addition of functions related to the biosynthesis of both chlorophyll a and *g*, as well as the biosynthesis of lycopene and an anoxygenic photosystem. These modifications mostly occurred through originations and gene transfers, and were not found in the other habitat transitions, highlighting the plasticity of *Polaromonas*. We further note that anoxygenic photosystem systems are oxygen-independent, ecologically widespread, and critical for microbial energy harvesting in anaerobic, low-light environments. As such, they are well adapted to the pelagic environment in lakes and wetlands. We also observed the presence of osmolyte transporters (two genes with one origination), which were probably involved in salt stress regulation. Similar to the groundwater transition, genomes at the lake transition also possessed genes related to assimilatory sulfate reduction. Predicting that wetland *Polaromonas* genomes descended from lake genomes, we observed the gain of different transporters for aromatic acids, disaccharides, vitamin C, and sulfur compounds, thus extending its capacity for scavenging a wide array of organic molecules. The presence of sulfur transporters is not surprising, as wetlands are known to harbor very

Please cite this article in press as: Michoud et al., Evolutionary radiation of *Polaromonas* from mountain glaciers downstream, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.06.070

high concentrations of organic and sulfur compounds.[70] Wetland genomes also gained genes involved in pH resistance, such as cyclopropane fatty acid synthase for membrane stabilization. Overall, many of the genes that differed from those in lake water appear to have been horizontally transferred, suggesting adaptations to the fluctuating environment of wetlands (e.g., changing redox conditions linked to drying).

If, as predicted, the *Polaromonas* LCA dwelled in GFSs, it appears intuitive that it colonized groundwater through the continuous exchange between these water sources. We found that the transition of the *Polaromonas* LCA to groundwater occurred with the expansion of transporters related to nitrogen, particularly nitrate and nitrite transporters. Similarly, more stress-tolerance genes are present, along with low-temperature mRNA-degradation genes and explicit heat-induced transcription factor regulation. Further, similar to the lake transition, we found the potential for assimilatory sulfate reduction pathways, revealing a certain role for sulfur metabolism capabilities. Several of the genes related to heterotrophy (i.e., ETC complex I), pH stress, low temperature, and lipid and biopolymer transporters were acquired through horizontal transfer. Genes related to carotene, high temperature, and salt stress in particular are predicted to have originated at the specific transition to groundwater, rather than being inherited from the LCA.

The transition of the LCA to soils was marked by higher numbers of genes related to substrate uptake, such as amide, amino acid, lipid, or metal-ion transport systems in the genome. This likely reflects an adaptation to the nutrient heterogeneity and patchiness found in terrestrial environments, compared with the more homogeneous substrates in aquatic environments.[71] Furthermore, stress-tolerance mechanisms are more developed compared with other environments, including an increased number of genes involved in EPS biosynthesis and export capabilities. This suggests that *Polaromonas* is also able to form biofilms in soils, which would allow it to increase nutrient availability and resistance to drying, for instance.[72] We also noted the presence of fatty acid desaturases that have been shown to be involved in low-temperature stress.[73]

We noted pronounced genomic adaptation when *Polaromonas* transitioned between soils and glaciers. It is plausible that *Polaromonas* transitioned repeatedly between glaciers and proglacial soils during glaciation-deglaciation cycles. In fact, the phases of increased dispersal associated with these cycles may have fostered these habitat transitions. Glacier *Polaromonas* genomes obtained genes related to anaerobic respiration, nitrite oxidation, and sulfide oxidation, indicating an adaptation to oxygen-limited conditions and chemolithotrophic metabolism through horizontal transfer. Furthermore, we noted the presence of genes involved in galactose and maltose degradation, suggesting a specialized carbohydrate metabolism adapted to limited carbon sources. Carbohydrates may originate from snow and ice algae that can form massive blooms in these environments.[74] Interestingly, *Polaromonas* genomes in glaciers lost some genes related to monosaccharides and nitrate and nitrite transport, which likely reflects an adaptation to oligotrophic and oxidized conditions in glaciers where such substrates are probably scarcer than in other environments.[75] The presence of sugar degradation genes along with a loss of transporters for the same compounds may indicate that *Polaromonas* in glaciers may favor intracellular metabolisms over the maintenance of multiple transport systems.

In conclusion, our ancestral reconstruction suggests that *Polaromonas*, an iconic and cosmopolitan bacterial genus, radiated from the GFSs on the mountain tops downstream into various aquatic and terrestrial ecosystems. An unresolved question is whether this radiation reflects a single origin followed by dispersal, or multiple independent events occurring across geographically distinct regions. Nevertheless, the LCA of *Polaromonas* was likely a biofilm former—overall, a most successful mode of microbial life—with a versatile metabolism combining heterotrophic and respiratory capacities with broad substrate transport and stress-response systems. These traits reflect an adaptation to fluctuating and harsh environmental conditions typical of GFSs and glaciers. As *Polaromonas* lineages diversified downstream to colonize aquatic and terrestrial ecosystems, they tended to conserve genes involved in the stress response and EPS biosynthesis, which suggests that biofilm formation remains a main trait to adapt to diverse environmental conditions throughout the evolutionary history of this bacterial genus. Our analyses revealed that adaptations through gene gains and losses, as well as horizontal transfers, greatly facilitated habitat transitions. Examples include the acquisition of phototrophic functions and sulfur metabolism in lake and wetland lineages, as well as enhanced nitrogen and lipid uptake systems in groundwater and soils. The distinct adaptations accompanying the various transitions across aquatic and terrestrial ecosystems highlight the genomic plasticity of *Polaromonas* as a key to its success. Our findings also reveal the potential role of glaciers and streams as a potential cradle for bacterial evolution and diversity.

RESOURCE AVAILABILITY

Lead contact

Further information and requests should be directed to and will be fulfilled by the lead contact, Grégoire Michoud (gregoire.michoud@epfl.ch).

Materials availability

This study did not generate any materials or reagents.

Data and code availability

- The Guliya MAGs are deposited under accession number PRJNA568282 and on Figshare (https://figshare.com/s/525cd485afb5e87d0c90), while the other MAGs were obtained from public databases.
- The code to reproduce analyses and figures can be found at https://github.com/michoug/Polaromonas_Pangenome and at https://doi.org/10.5281/zenodo.18611200.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

The Vanishing Glaciers project is supported by The NOMIS Foundation (to T.J.B.). Partial support was provided by the Heising-Simons Foundation award #2022-4014 to V.R. and Z.-P.Z., and by Z.-P.Z.'s startup fund at the Chinese Academy of Sciences. We thank Susheel Bhanu Busi and Daniel Read at the UK Centre for Ecology and Hydrology for their supervision and methodological input into the generation of UK river metagenome-assembled genomes, and the UK Environment Agency for supporting sample collection and data generation. We also thank Lonnie G. Thompson, Ellen Mosley-Thompson, and Matthew B. Sullivan for their contributions in obtaining the Guliya glacier's *Polaromonas* MAGs.

AUTHOR CONTRIBUTIONS

G.M. was responsible for conceptualization, methodology, software, data curation, investigation, formal analysis, visualization, and writing of the original draft. A.G. was responsible for conceptualization, methodology, software, data curation, investigation, formal analysis, and visualization. H.P. was responsible for conceptualization, methodology, investigation, and writing of the original draft. A.C.T. was responsible for data curation. Z.-P.Z. was responsible for methodology, data curation, and investigation. V.R. was responsible for data curation and investigation. T.J.B. was responsible for conceptualization, methodology, investigation, writing of the original draft, supervision, project administration, and funding acquisition. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used ChatGPT in order to improve the orthography, grammar, and writing style. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- METHOD DETAILS
  - Microbe Atlas
  - Genome gathering
  - Pangenome analysis
  - Species tree estimation and reconciliation
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Sensitivity analysis of ancestral state reconstruction and enrichment of horizontal transfer events at transition nodes
  - Phylogeny-aware pathway-habitat association analysis

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2026.06.070.

Received: February 11, 2026
Revised: May 20, 2026
Accepted: June 25, 2026

REFERENCES

1. Jamy, M., Biwer, C., Vaulot, D., Obiol, A., Jing, H., Peura, S., Massana, R., and Burki, F. (2022). Global patterns and rates of habitat transitions across the eukaryotic tree of life. Nat. Ecol. Evol. *6*, 1458–1470. https://doi.org/10.1038/s41559-022-01838-4.
2. Jaffe, A.L., Castelle, C.J., and Banfield, J.F. (2023). Habitat transition in the evolution of bacteria and archaea. Annu. Rev. Microbiol. *77*, 193–212. https://doi.org/10.1146/annurev-micro-041320-032304.
3. Wolfe, J.M., Ballou, L., Luque, J., Watson-Zink, V.M., Ahyong, S.T., Barido-Sottani, J., Chan, T.-Y., Chu, K.H., Crandall, K.A., Daniels, S.R., et al. (2024). Convergent Adaptation of True Crabs (Decapoda: Brachyura) to a Gradient of Terrestrial Environments. Syst. Biol. *73*, 247–262. https://doi.org/10.1093/sysbio/syad066.
4. Panstruga, R., Fürst-Jansen, J., and De Vries, J. (2025). Plant terrestrialization: piecing together streptophyte trait evolution – an introduction to a Virtual Issue. New Phytol. *247*, 1963–1970. https://doi.org/10.1111/nph.70399.
5. Fürst-Jansen, J.M.R., De Vries, S., and De Vries, J. (2020). Evo-physio: on stress responses and the earliest land plants. J. Exp. Bot. *71*, 3254–3269. https://doi.org/10.1093/jxb/eraa007.
6. Simpson, C. (2021). Adaptation to a Viscous Snowball Earth Ocean as a Path to Complex Multicellularity. Am. Nat. *198*, 590–609. https://doi.org/10.1086/716634.
7. Žárský, J., Žárský, V., Hanáček, M., and Žárský, V. (2022). Cryogenian Glacial Habitats as a Plant Terrestrialisation Cradle – The Origin of the Anydrophytes and Zygnematophyceae Split. Front. Plant Sci. *12*, 735020. https://doi.org/10.3389/fpls.2021.735020.
8. Koskella, B., and Vos, M. (2015). Adaptation in natural microbial populations. Annu. Rev. Ecol. Evol. Syst. *46*, 503–522. https://doi.org/10.1146/annurev-ecolsys-112414-054458.
9. Moreno-Gámez, S. (2022). How bacteria navigate varying environments. Science *378*, 845. https://doi.org/10.1126/science.adf4444.
10. Salcher, M.M., Schaefle, D., Kaspar, M., Neuenschwander, S.M., and Ghai, R. (2019). Evolution in action: habitat transition from sediment to the pelagial leads to genome streamlining in Methylophilaceae. ISME J. *13*, 2764–2777. https://doi.org/10.1038/s41396-019-0471-3.
11. Eiler, A., Mondav, R., Sinclair, L., Fernandez-Vidal, L., Scofield, D.G., Schwientek, P., Martinez-Garcia, M., Torrents, D., McMahon, K.D., Andersson, S.G.E., et al. (2016). Tuning fresh: radiation through rewiring of central metabolism in streamlined bacteria. ISME J. *10*, 1902–1914. https://doi.org/10.1038/ismej.2015.260.
12. Bourquin, M., Busi, S.B., Fodelianakis, S., Peter, H., Washburne, A., Kohler, T.J., Ezzat, L., Michoud, G., Wilmes, P., and Battin, T.J. (2022). The microbiome of cryospheric ecosystems. Nat. Commun. *13*, 3087. https://doi.org/10.1038/s41467-022-30816-4.
13. Scheel, M., Zervas, A., Jacobsen, C.S., and Christensen, T.R. (2022). Microbial community changes in 26,500-year-old thawing permafrost. Front. Microbiol. *13*, 787146. https://doi.org/10.3389/FMICB.2022.787146.
14. Jia, P., Tian, M., Zhang, B., Wu, X., He, X., and Zhang, W. (2024). Habitat changes due to glacial freezing and melting reshape microbial networks. Environ. Int. *189*, 108788. https://doi.org/10.1016/J.ENVINT.2024.108788.
15. Ezzat, L., Peter, H., Bourquin, M., Busi, S.B., Michoud, G., Fodelianakis, S., Kohler, T.J., Lamy, T., Geers, A., Pramateftaki, P., et al. (2025). Diversity and biogeography of the bacterial microbiome in glacier-fed streams. Nature *637*, 622–630. https://doi.org/10.1038/s41586-024-08313-z.
16. Michoud, G., Peter, H., Busi, S.B., Bourquin, M., Kohler, T.J., Geers, A., Ezzat, L., Vanishing Glaciers Field Team, and Battin, T.J. (2025). Mapping the metagenomic diversity of the multi-kingdom glacier-fed stream microbiome. Nat. Microbiol. *10*, 217–230. https://doi.org/10.1038/s41564-024-01874-9.
17. Nemergut, D.R., Anderson, S.P., Cleveland, C.C., Martin, A.P., Miller, A.E., Seimon, A., and Schmidt, S.K. (2007). Microbial community succession in an unvegetated, recently deglaciated soil. Microb. Ecol. *53*, 110–122. https://doi.org/10.1007/s00248-006-9144-7.
18. Borton, M.A., McGivern, B.B., Willi, K.R., Woodcroft, B.J., Mosier, A.C., Singleton, D.M., Bambakidis, T., Pelly, A., Daly, R.A., Liu, F., et al. (2025). A functional microbiome catalogue crowdsourced from North American rivers. Nature *637*, 103–112. https://doi.org/10.1038/s41586-024-08240-z.
19. Thorpe, A.C., Busi, S.B., Warren, J., Hunt, L.H., Walsh, K., and Read, D.S. (2025). National-scale biogeography and function of river and stream bacterial biofilm communities. Nat. Commun. *16*, 10571. https://doi.org/10.1038/s41467-025-65620-3.
20. Darcy, J.L., Lynch, R.C., King, A.J., Robeson, M.S., and Schmidt, S.K. (2011). Global distribution of polaromonas phylotypes - evidence for a

highly successful dispersal capacity. PLOS One *6*, e23742. https://doi.org/10.1371/journal.pone.0023742.

21. Hwang, K., Choe, H., Nasir, A., and Kim, K.M. (2021). Complete genome of Polaromonas vacuolata KCTC 22033T isolated from beneath Antarctic Sea ice. Mar. Genomics *55*, 100790. https://doi.org/10.1016/j.margen.2020.100790.
22. Sizova, M., and Panikov, N. (2007). Polaromonas hydrogenivorans sp. nov., a psychrotolerant hydrogen-oxidizing bacterium from Alaskan soil. Int. J. Syst. Evol. Microbiol. *57*, 616–619. https://doi.org/10.1099/ijs.0.64350-0.
23. Osborne, T.H., Jamieson, H.E., Hudson-Edwards, K.A., Nordstrom, D.K., Walker, S.R., Ward, S.A., and Santini, J.M. (2010). Microbial oxidation of arsenite in a subarctic environment: diversity of arsenite oxidase genes and identification of a psychrotolerant arsenite oxidiser. BMC Microbiol. *10*, 205. https://doi.org/10.1186/1471-2180-10-205.
24. Nash, M.V., Anesio, A.M., Barker, G., Tranter, M., Varliero, G., Eloe-Fadrosh, E.A., Nielsen, T., Turpin-Jelfs, T., Benning, L.G., and Sánchez-Baracaldo, P. (2018). Metagenomic insights into diazotrophic communities across Arctic glacier forefields. FEMS Microbiol. Ecol. *94*, fiy114. https://doi.org/10.1093/femsec/fiy114.
25. Swan, B.K., Tupper, B., Sczyrba, A., Lauro, F.M., Martinez-Garcia, M., González, J.M., Luo, H., Wright, J.J., Landry, Z.C., Hanson, N.W., et al. (2013). Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. Proc. Natl. Acad. Sci. USA *110*, 11463–11468. https://doi.org/10.1073/pnas.1304246110.
26. Grigoryan, A.A., Bondici, V.F., Kryachko, Y., Khan, N.H., Lawrence, J.R., Wolfaardt, G.M., Withana Gamage, N., Shetty, D., and Korber, D.R. (2022). Draft Genome Sequence of Polaromonas eurypsychrophila AER18D-145, Isolated from a Uranium Tailings Management Facility in Northern Saskatchewan, Canada. Microbiol. Resour. Announc. *11*, e0001322. https://doi.org/10.1128/mra.00013-22.
27. Raymond-Bouchard, I., Goordial, J., Zolotarov, Y., Ronholm, J., Stromvik, M., Bakermans, C., and Whyte, L.G. (2018). Conserved genomic and amino acid traits of cold adaptation in subzero-growing Arctic permafrost bacteria. FEMS Microbiol. Ecol. *94*, 23. https://doi.org/10.1093/FEMSEC/FIY023.
28. Irgens, R.L., Gosink, J.J., and Staley, J.T. (1996). Polaromonas vacuolata gen. nov., sp. nov., a psychrophilic, marine, gas vacuolate bacterium from antarctica. Int. J. Syst. Bacteriol. *46*, 822–826. https://doi.org/10.1099/00207713-46-3-822.
29. Hervas, A., and Casamayor, E.O. (2009). High similarity between bacterioneuston and airborne bacterial community compositions in a high mountain lake area. FEMS Microbiol. Ecol. *67*, 219–228. https://doi.org/10.1111/j.1574-6941.2008.00617.x.
30. Liu, Y., Yao, T., Jiao, N., Kang, S., Xu, B., Zeng, Y., Huang, S., and Liu, X. (2009). Bacterial diversity in the snow over tibetan plateau glaciers. Extremophiles *13*, 411–423. https://doi.org/10.1007/s00792-009-0227-5.
31. Fahlgren, C., Bratbak, G., Sandaa, R.-A., Thyrhaug, R., and Zweifel, U.L. (2011). Diversity of airborne bacteria in samples collected using different devices for aerosol collection. Aerobiologia *27*, 107–120. https://doi.org/10.1007/s10453-010-9181-z.
32. Gawor, J., Grzesiak, J., Sasin-Kurowska, J., Borsuk, P., Gromadka, R., Górniak, D., Świątecki, A., Aleksandrzak-Piekarczyk, T., and Zdanowski, M.K. (2016). Evidence of adaptation, niche separation and microevolution within the genus Polaromonas on Arctic and Antarctic glacial surfaces. Extremophiles *20*, 403–413. https://doi.org/10.1007/S00792-016-0831-0.
33. Girard, C., Vincent, W.F., and Culley, A.I. (2023). Arctic bacterial diversity and connectivity in the coastal margin of the Last Ice Area. ISME Commun. *3*, 105. https://doi.org/10.1038/S43705-023-00313-W.
34. Du, Y., He, C., Lloyd, K.G., Vishnivetskaya, T.A., Cui, H., Li, B., Gong, D., Fan, X., Zhang, D., Jiang, H., et al. (2025). Comparative genomics reveals the high diversity and adaptation strategies of Polaromonas from polar environments. BMC Genomics *26*, 248. https://doi.org/10.1186/s12864-025-11410-6.
35. Busi, S.B., Bourquin, M., Fodelianakis, S., Michoud, G., Kohler, T.J., Peter, H., Pramateftaki, P., Styllas, M., Tolosano, M., De Staercke, V., et al. (2022). Genomic and metabolic adaptations of biofilms to ecological windows of opportunity in glacier-fed streams. Nat. Commun. *13*, 2168. https://doi.org/10.1038/s41467-022-29914-0.
36. Hoffman, P.F. (2025). Ecosystem relocation on Snowball Earth: Polar-alpine ancestry of the extant surface biosphere? Proc. Natl. Acad. Sci. USA *122*, e2414059122. https://doi.org/10.1073/pnas.2414059122.
37. Zhang, H., Sun, Y., Zeng, Q., Crowe, S.A., and Luo, H. (2021). Snowball Earth, population bottleneck and Prochlorococcus evolution. Proc. Biol. Sci. *288*, 20211956. https://doi.org/10.1098/rspb.2021.1956.
38. Brighenti, S., Hotaling, S., Finn, D.S., Fountain, A.G., Hayashi, M., Herbst, D., Saros, J.E., Tronstad, L.M., and Millar, C.I. (2021). Rock glaciers and related cold rocky landforms: Overlooked climate refugia for mountain biodiversity. Glob. Change Biol. *27*, 1504–1517. https://doi.org/10.1111/gcb.15510.
39. Williams, T.A., Davin, A.A., Szánthó, L.L., Stamatakis, A., Wahl, N.A., Woodcroft, B.J., Soo, R.M., Eme, L., Sheridan, P.O., Gubry-Rangin, C., et al. (2024). Phylogenetic reconciliation: making the most of genomes to understand microbial ecology and evolution. ISME J. *18*, wrae129. https://doi.org/10.1093/ismejo/wrae129.
40. Matias Rodrigues, J.F.M., Tackmann, J., Malfertheiner, L., Patsch, D., Perez-Molphe-Montoya, E., Näpflin, N., Gaio, D., Rot, G., Danaila, M., Peluso, M.E., et al. (2026). The MicrobeAtlas database: Global trends and insights into Earth's microbial ecosystems. Cell *189*, 2092–2107.e17. https://doi.org/10.1016/j.cell.2026.01.021.
41. Figueroa, D., Capo, E., Lindh, M.V., Rowe, O.F., Paczkowska, J., Pinhassi, J., and Andersson, A. (2021). Terrestrial dissolved organic matter inflow drives temporal dynamics of the bacterial community of a subarctic estuary (northern Baltic Sea). Environ. Microbiol. *23*, 4200–4213. https://doi.org/10.1111/1462-2920.15597.
42. Lozupone, C.A., and Knight, R. (2007). Global patterns in bacterial diversity. Proc. Natl. Acad. Sci. USA *104*, 11436–11440. https://doi.org/10.1073/pnas.0611525104.
43. O'Leary, N.A., Cox, E., Holmes, J.B., Anderson, W.R., Falk, R., Hem, V., Tsuchiya, M.T.N., Schuler, G.D., Zhang, X., Torcivia, J., et al. (2024). Exploring and retrieving sequence and metadata for species across the tree of life with NCBI Datasets. Sci. Data *11*, 732. https://doi.org/10.1038/s41597-024-03571-y.
44. Cheng, M., Luo, S., Zhang, P., Xiong, G., Chen, K., Jiang, C., Yang, F., Huang, H., Yang, P., Liu, G., et al. (2024). A genome and gene catalog of the aquatic microbiomes of the Tibetan Plateau. Nat. Commun. *15*, 1438. https://doi.org/10.1038/s41467-024-45895-8.
45. Liu, Y., Ji, M., Yu, T., Zaugg, J., Anesio, A.M., Zhang, Z., Hu, S., Hugenholtz, P., Liu, K., Liu, P., et al. (2022). A genome and gene catalog of glacier microbiomes. Nat. Biotechnol. *40*, 1341–1348. https://doi.org/10.1038/s41587-022-01367-2.
46. Zhong, Z.-P., Zablocki, O., Li, Y.-F., Van Etten, J.L., Mosley-Thompson, E., Rich, V.I., Thompson, L.G., and Sullivan, M.B. (2024). Glacier-preserved Tibetan Plateau viral community probably linked to warm–cold climate variations. Nat. Geosci. *17*, 912–919. https://doi.org/10.1038/s41561-024-01508-z.
47. Chaumeil, P.-A., Mussig, A.J., Hugenholtz, P., and Parks, D.H. (2019). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics *36*, 1925–1927. https://doi.org/10.1093/bioinformatics/btz848.
48. Rodríguez-Gijón, A., Nuy, J.K., Mehrshad, M., Buck, M., Schulz, F., Woyke, T., and Garcia, S.L. (2022). A genomic perspective across earth's microbiomes reveals that genome size in archaea and bacteria is linked to ecosystem type and trophic strategy. Front. Microbiol. *12*, 761869. https://doi.org/10.3389/fmicb.2021.761869.
49. Gautreau, G., Bazin, A., Gachet, M., Planel, R., Burlot, L., Dubois, M., Perrin, A., Médigue, C., Calteau, A., Cruveiller, S., et al. (2020).

PPanGGOLiN: Depicting microbial diversity via a partitioned pangenome graph. PLoS Comput. Biol. *16*, e1007732. https://doi.org/10.1371/JOURNAL.PCBI.1007732.

50. Acevedo-Rocha, C.G., Fang, G., Schmidt, M., Ussery, D.W., and Danchin, A. (2013). From essential to persistent genes: a functional approach to constructing synthetic life. Trends Genet. *29*, 273–279. https://doi.org/10.1016/j.tig.2012.11.001.
51. Andrei, A.-Ş., Salcher, M.M., Mehrshad, M., Rychtecký, P., Znachor, P., and Ghai, R. (2019). Niche-directed evolution modulates genome architecture in freshwater Planctomycetes. ISME J. *13*, 1056–1071. https://doi.org/10.1038/s41396-018-0332-5.
52. Wiens, J.J., and Graham, C.H. (2005). Niche Conservatism: Integrating Evolution, Ecology, and Conservation Biology. Annu. Rev. Ecol. Evol. Syst. *36*, 519–539. https://doi.org/10.1146/annurev.ecolsys.36.102803.095431.
53. Ishikawa, S.A., Zhukova, A., Iwasaki, W., and Gascuel, O. (2019). A fast likelihood method to reconstruct and visualize ancestral scenarios. Mol. Biol. Evol. *36*, 2069–2085. https://doi.org/10.1093/molbev/msz131.
54. Revell, L.J. (2025). Ancestral state reconstruction of phenotypic characters. Evol. Biol. *52*, 1–25. https://doi.org/10.1007/s11692-025-09645-y.
55. Nyholm, S.V., and McFall-Ngai, M.J. (1998). Sampling the light-organ microenvironment of euprymna scolopes: Description of a population of host cells in association with the bacterial symbiont vibrio fischeri. Biol. Bull. *195*, 89–97. https://doi.org/10.2307/1542815.
56. Gentili, R., Baroni, C., Caccianiga, M., Armiraglio, S., Ghiani, A., and Citterio, S. (2015). Potential warm-stage microrefugia for alpine plants: Feedback between geomorphological and biological processes. Ecol. Complexity *21*, 87–99. https://doi.org/10.1016/j.ecocom.2014.11.006.
57. Millar, C.I., Delany, D.L., Hersey, K.A., Jeffress, M.R., Smith, A.T., Van Gunst, K.J., and Westfall, R.D. (2018). Distribution, climatic relationships, and status of American pikas (*Ochotona princeps*) in the Great Basin, USA. Arct. Antarct. Alp. Res. *50*, e1436296. https://doi.org/10.1080/15230430.2018.1436296.
58. Crump, B.C., Amaral-Zettler, L.A., and Kling, G.W. (2012). Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. ISME J. *6*, 1629–1639. https://doi.org/10.1038/ismej.2012.9.
59. Caillon, F., and Schelker, J. (2020). Dynamic transfer of soil bacteria and dissolved organic carbon into small streams during hydrological events. Aquat. Sci. *82*, 41. https://doi.org/10.1007/s00027-020-0714-4.
60. Koonin, E.V., Makarova, K.S., and Aravind, L. (2001). Horizontal Gene Transfer in Prokaryotes: Quantification and Classification. Annu. Rev. Microbiol. *55*, 709–742. https://doi.org/10.1146/annurev.micro.55.1.709.
61. Albalat, R., and Cañestro, C. (2016). Evolution by gene loss. Nat. Rev. Genet. *17*, 379–391. https://doi.org/10.1038/nrg.2016.39.
62. Camargo, A.P., Roux, S., Schulz, F., Babinski, M., Xu, Y., Hu, B., Chain, P.S.G., Nayfach, S., and Kyrpides, N.C. (2024). Identification of mobile genetic elements with geNomad. Nat. Biotechnol. *42*, 1303–1312. https://doi.org/10.1038/s41587-023-01953-y.
63. Rodriguez-Valera, F., and Bellas, C. (2025). How Viruses Shape Microbial Plankton Microdiversity. Annu. Rev. Mar. Sci. *17*, 561–576. https://doi.org/10.1146/annurev-marine-040623-090847.
64. Gonzaga, A., Martin-Cuadrado, A.-B., López-Pérez, M., Megumi Mizuno, C., García-Heredia, I., Kimes, N.E., Lopez-García, P., Moreira, D., Ussery, D., Zaballos, M., et al. (2012). Polyclonality of Concurrent Natural Populations of Alteromonas macleodii. Genome Biol. Evol. *4*, 1360–1374. https://doi.org/10.1093/gbe/evs112.
65. Karaoz, U., and Brodie, E.L. (2022). microTrait: A Toolset for a Trait-Based Representation of Microbial Genomes. Front. Bioinform. *2*, 918853. https://doi.org/10.3389/fbinf.2022.918853.
66. Battin, T.J., Besemer, K., Bengtsson, M.M., Romani, A.M., and Packmann, A.I. (2016). The ecology and biogeochemistry of stream biofilms. Nat. Rev. Microbiol. *14*, 251–263. https://doi.org/10.1038/nrmicro.2016.15.
67. Reis-Mansur, M.C.P.P., Cardoso-Rurr, J.S., Silva, J.V.M.A., De Souza, G.R., Cardoso, V.D.S., Mansoldo, F.R.P., Pinheiro, Y., Schultz, J., Lopez Balottin, L.B., Da Silva, A.J.R., et al. (2019). Carotenoids from UV-resistant Antarctic Microbacterium sp. LEMMJ01. Sci. Rep. *9*, 9554. https://doi.org/10.1038/s41598-019-45840-6.
68. Michoud, G., Kohler, T.J., Ezzat, L., Peter, H., Nattabi, J.K., Nalwanga, R., Pramateftaki, P., Styllas, M., Tolosano, M., De Staercke, V., et al. (2023). The dark side of the moon: first insights into the microbiome structure and function of one of the last glacier-fed streams in Africa. R. Soc. Open Sci. *10*, 230329. https://doi.org/10.1098/rsos.230329.
69. Milner, A.M., and Petts, G.E. (1994). Glacial rivers: physical habitat and ecology. Freshw. Biol. *32*, 295–307. https://doi.org/10.1111/j.1365-2427.1994.tb01127.x.
70. Zeng, T., Arnold, W.A., and Toner, B.M. (2013). Microscale Characterization of Sulfur Speciation in Lake Sediments. Environ. Sci. Technol. *47*, 1287–1296. https://doi.org/10.1021/es303914q.
71. Nunan, N., Schmidt, H., and Raynaud, X. (2020). The ecology of heterogeneity: soil bacterial communities and C dynamics. Philos. Trans. R. Soc. Lond. B Biol. Sci. *375*, 20190249. https://doi.org/10.1098/rstb.2019.0249.
72. Musa, O.I., Akande, S.A., Ijah, U.J.J., Abioye, O.P., Maude, A.M., Samuel, J.O., Mustapha, A., Abdulrahim, A.-M., and Gusdanis, A.C.G. (2024). Biofilms communities in the soil: characteristic and interactions using mathematical model. Res. Microbiol. *175*, 104149. https://doi.org/10.1016/j.resmic.2023.104149.
73. Tao, Y., Sun, Y., Chai, Y., Duma, L., and Rossez, Y. (2025). Fatty acid desaturases in non-photosynthetic bacteria: classification, regulation, and roles in plasma membrane function and cellular homeostasis. Biochim. Biophys. Acta Mol. Cell Biol. Lipids *1870*, 159630. https://doi.org/10.1016/j.bbalip.2025.159630.
74. Krug, L., Erlacher, A., Markut, K., Berg, G., and Cernava, T. (2020). The microbiome of alpine snow algae shows a specific inter-kingdom connectivity and algae-bacteria interactions with supportive capacities. ISME J. *14*, 2197–2210. https://doi.org/10.1038/s41396-020-0677-4.
75. Wadham, J.L., Hawkings, J., Telling, J., Chandler, D., Alcock, J., O'Donnell, E., Kaur, P., Bagshaw, E., Tranter, M., Tedstone, A., et al. (2016). Sources, cycling and export of nitrogen on the Greenland Ice Sheet. Biogeosciences *13*, 6339–6352. https://doi.org/10.5194/bg-13-6339-2016.
76. Kang, D., Li, F., Kirton, E.S., Thomas, A., Egan, R.S., An, H., and Wang, Z. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ *7*, e7359. https://doi.org/10.7717/peerj.7359.
77. Wu, Y.W., Simmons, B.A., and Singer, S.W. (2015). MaxBin 2.0: An automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics *32*, 605–607. https://doi.org/10.1093/bioinformatics/btv638.
78. Imelfort, M., Parks, D.H., Woodcroft, B.J., Dennis, P., Hugenholtz, P., and Tyson, G.W. (2014). GroopM: An automated tool for the recovery of population genomes from related metagenomes. PeerJ. *2*, e603. https://doi.org/10.7717/peerj.603.
79. Sieber, C.M.K., Probst, A.J., Sharrar, A., Thomas, B.C., Hess, M., Tringe, S.G., and Banfield, J.F. (2018). Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. Nat. Microbiol. *3*, 836–843. https://doi.org/10.1038/s41564-018-0171-1.
80. Uritskiy, G.V., DiRuggiero, J., and Taylor, J. (2018). MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome *6*, 158. https://doi.org/10.1186/s40168-018-0541-1.
81. Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. *25*, 1043–1055. https://doi.org/10.1101/gr.186072.114.
82. Chklovski, A., Parks, D.H., Woodcroft, B.J., and Tyson, G.W. (2023). CheckM2: a rapid, scalable and accurate tool for assessing microbial

genome quality using machine learning. Nat. Methods *20*, 1203–1212. https://doi.org/10.1038/s41592-023-01940-w.

83. Aroney, S.T.N., Camargo, A.P., Tyson, G.W., and Woodcroft, B.J. (2024). Galah: More scalable dereplication for metagenome assembled genomes. Version 0.4.2. Zenodo. https://doi.org/10.5281/zenodo.10526086.
84. Schwengers, O., Jelonek, L., Dieckmann, M.A., Beyvers, S., Blom, J., and Goesmann, A. (2021). Bakta: Rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microb. Genomics *7*, 000685. https://doi.org/10.1099/MGEN.0.000685.
85. Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., Mende, D.R., Letunic, I., Rattei, T., Jensen, L.J., et al. (2019). EggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. *47*, D309–D314. https://doi.org/10.1093/NAR/GKY1085.
86. Shaw, J., and Yu, Y.W. (2023). Fast and robust metagenomic sequence comparison through sparse chaining with skani. Nat. Methods *20*, 1661–1665. https://doi.org/10.1038/s41592-023-02018-3.
87. Steinegger, M., and Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat. Biotechnol. *35*, 1026–1028. https://doi.org/10.1038/NBT.3988.
88. Nakamura, T., Yamada, K.D., Tomii, K., and Katoh, K. (2018). Parallelization of MAFFT for large-scale multiple sequence alignments. Bioinformatics *34*, 2490–2492. https://doi.org/10.1093/bioinformatics/bty121.
89. Huelsenbeck, J.P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics *17*, 754–755. https://doi.org/10.1093/bioinformatics/17.8.754.
90. Morel, B., Williams, T.A., Stamatakis, A., and Szöllősi, G.J. (2024). AleRax: a tool for gene and species tree co-estimation and reconciliation under a probabilistic model of gene duplication, transfer, and loss. Bioinformatics *40*, btae162. https://doi.org/10.1093/bioinformatics/btae162.
91. Revell, L.J. (2024). phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). PeerJ *12*, e16505. https://doi.org/10.7717/peerj.16505.
92. Paradis, E., and Schliep, K. (2019). ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics *35*, 526–528. https://doi.org/10.1093/bioinformatics/bty633.
93. Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N., and Pearse, W. (2011). caper: Comparative Analyses of Phylogenetics and Evolution in R. https://doi.org/10.32614/CRAN.package.caper.
94. Schwery, O., and O'Meara, B.C. (2016). *MonoPhy*: a simple R package to find and visualize monophyly issues. PeerJ Comput. Sci. *2*, e56. https://doi.org/10.7717/peerj-cs.56.
95. R Core Team (2021). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).
96. Landau, W. (2021). The targets R package: a dynamic Make-like function-oriented pipeline toolkit for reproducibility and high-performance computing. J. Open Source Softw. *6*, 2959. https://doi.org/10.21105/joss.02959.
97. Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., et al. (2019). Welcome to the Tidyverse. J. Open Source Softw. *4*, 1686. https://doi.org/10.21105/joss.01686.
98. Yu, G., Smith, D.K., Zhu, H., Guan, Y., and Lam, T.T.Y. (2017). ggtree : an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol. Evol. *8*, 28–36. https://doi.org/10.1111/2041-210X.12628.
99. Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., et al. (2020). Package "vegan" title community ecology package version 2.5-7. https://cran.r-project.org/web/packages/vegan/index.html.
100. Salter, S.J., Cox, M.J., Turek, E.M., Calus, S.T., Cookson, W.O., Moffatt, M.F., Turner, P., Parkhill, J., Loman, N.J., and Walker, A.W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. *12*, 87. https://doi.org/10.1186/s12915-014-0087-z.
101. Garner, R.E., Kraemer, S.A., Onana, V.E., Fradette, M., Varin, M.-P., Huot, Y., and Walsh, D.A. (2023). A genome catalogue of lake bacterial diversity and its drivers at continental scale. Nat. Microbiol. *8*, 1920–1934. https://doi.org/10.1038/s41564-023-01435-6.
102. Michoud, G., Kohler, T.J., Peter, H., Brandani, J., Busi, S.B., and Battin, T.J. (2023). Unexpected functional diversity of stream biofilms within and across proglacial floodplains despite close spatial proximity. Limnol. Oceanogr. *68*, 2183–2194. https://doi.org/10.1002/lno.12415.

Please cite this article in press as: Michoud et al., Evolutionary radiation of *Polaromonas* from mountain glaciers downstream, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.06.070

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Polaromonas MAGs – TPMC dataset	Cheng et al.[44]	https://download.cncb.ac.cn/bigd/TPMC/
Polaromonas MAGs – TG2G dataset	Liu et al.[45]	https://www.biosino.org/node/analysis/detail/OEZ008893
Polaromonas MAGs – UK rivers	Thorpe et al.[19]	EBI Bioproject: PRJEB85861
Polaromonas MAGs – GROW DB	Borton et al.[18]	NCBI Bioproject: PRJNA946291
Polaromonas MAGs – GFS dataset	Michoud et al.[16]	NCBI Bioproject: PRJNA781406
Polaromonas MAGs – Guliya glacier metagenomes	Zhong et al.[46]	NCBI BioProject: PRJNA568282, Figshare https://figshare.com/s/525cd485afb5e87d0c90
Polaromonas genomes – NCBI (downloaded 12 Dec 2024)	NCBI	https://www.ncbi.nlm.nih.gov/datasets/
MicrobeAtlas database	Rodrigues et al.[40]	https://microbeatlas.org
Software and algorithms		
MetaBAT2 (v2.16)	Kang et al.[76]	https://bitbucket.org/berkeleylab/metabat
MaxBin2 (v2.2.7)	Wu et al.[77]	https://sourceforge.net/projects/maxbin2/
GroopM (v0.3.5)	Imelfort et al.[78]	https://github.com/centre-for-microbiome-research/GroopM
UniteM (v0.0.18)	Parks	https://github.com/donovan-h-parks/UniteM
DAS Tool (v1.1.1)	Sieber et al.[79]	https://github.com/cmks/DAS_Tool
metaWRAP (v1.0.6)	Uritskiy et al.[80]	https://github.com/bxlab/metaWRAP
CheckM (v1.0.12)	Parks et al.[81]	https://github.com/Ecogenomics/CheckM
RefineM (v0.0.24)	Parks	https://github.com/donovan-h-parks/RefineM
GTDB-TK (v2.3.0)	Chaumeil et al.[47]	https://github.com/Ecogenomics/GTDBTk
CheckM2 (v1.0.1)	Chklovski et al.[82]	https://github.com/chklovski/CheckM2
galah (v0.4.2)	Aroney et al.[83]	https://github.com/wwood/galah
Bakta (v1.9.3)	Schwengers et al.[84]	https://github.com/oschwengers/bakta
microtrait (v1.0.0)	Karaoz and Brodie[65]	https://github.com/ukaraoz/microtrait
eggnog-mapper (v2.1.12)	Huerta-Cepas et al.[85]	https://github.com/eggnogdb/eggnog-mapper
Genomad (v1.11)	Camargo et al.[62]	https://github.com/apcamargo/genomad
skani (v0.3.1)	Shaw and Yu[86]	https://github.com/bluenote-1577/skani
PPanGGOLiN (v2.1.1)	Gautreau et al.[49]	https://github.com/labgem/PPanGGOLiN
mmseqs2 (v15)	Steinegger and Söding[87]	https://github.com/soedinglab/MMseqs2
MAFFT (v7.526)	Nakamura et al.[88]	https://mafft.cbrc.jp/alignment/software/
mrBayes (v3.2.7)	Huelsenbeck and Ronquist[89]	https://nbisweden.github.io/MrBayes/
AleRax (v1.0.0)	Morel et al.[90]	https://github.com/BenoitMorel/AleRax
PastML (v1.9.41)	Ishikawa et al.[53]	https://github.com/evolbioinfo/pastml
phytools (v2.5-2)	Revell[91]	https://github.com/liamrevell/phytools
ape (v5.8-1)	Paradis and Schliep[92]	https://cran.r-project.org/package=ape
caper (v1.0.4)	Orme et al.[93]	https://cran.r-project.org/package=caper
MonoPhy (v1.3.2)	Schwery and O'Meara[94]	https://cran.r-project.org/package=MonoPhy
R (v4.5.1)	R Core Team[95]	https://www.r-project.org/
targets (v1.11.3)	Landau[96]	https://docs.ropensci.org/targets/
tidyverse (v2.0.0)	Wickham et al.[97]	https://www.tidyverse.org/
ggtree (v3.16)	Yu et al.[98]	https://bioconductor.org/packages/ggtree
vegan (v2.6-10)	Oksanen et al.[99]	https://cran.r-project.org/package=vegan
Other		
Resource website for the publication	This paper	https://github.com/michoug/Polaromonas_Pangenome (10.5281/zenodo.18611200)

Please cite this article in press as: Michoud et al., Evolutionary radiation of *Polaromonas* from mountain glaciers downstream, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.06.070

METHOD DETAILS

Microbe Atlas

We used the Microbe Atlas database[40] to explore the spatial distribution of *Polaromonas* in diverse environments. As *Polaromonas* was found to be a potential contaminant in some kits, we chose a relatively high threshold of abundance to consider it present (> 0.1% relative abundance).[100]

Genome gathering

We obtained 632 *Polaromonas* genomes from different sources, including MAGs from the TPMC (160),[44] TG2G (52),[45] United Kingdom rivers (8),[19] GROW DB (4),[18] GFS (25),[16] and downloaded via the NCBI dataset software[43] on the 12th December 2024 (351). The 32 MAGs from Guliya glacier metagenomes[46] were reconstructed using an ensemble binning and refinement workflow. Contigs were binned with MetaBAT2, MaxBin2, and GroopM2 via UniteM (v0.0.18, https://github.com/donovan-h-parks/UniteM), incorporating sample-specific coverage information from mapped reads, and integrated using UniteM consensus, DAS Tool (v1.1.1),[79] and metaWRAP (v1.0.6)[80] bin_refinement. MAG completeness and contamination were assessed using CheckM (v1.0.12), and high-quality bins were further refined by removing GC- and coverage-based outlier contigs with RefineM (v0.0.24; https://github.com/donovan-h-parks/RefineM). Assembly methods are described previously.[46]

We assessed the taxonomy and quality of all MAGs with GTDB-TK (v2.3.0)[47] and CheckM2 (v1.0.1),[82] respectively. We then removed genomes that were not annotated as *Polaromonas* (52) and genomes whose completeness and contamination were less than 70% and more than 10%, respectively (248). In order to remove duplicate genomes, we dereplicated them at 99%, resulting in a total of 302 genomes using galah (v0.4.2).[83] As we were interested in only freshwater habitats, we removed strains which were isolated from mine drainage, activated sludge, landfill, permafrost, host-associated and bioreactor environments (20 genomes), with the exception of cultivated species (e.g., *Polaromonas aquatica* CCUG 39402^T (GenBank: GCA_042659145.1), *Polaromonas naphthalenivorans* CJ2^T (GenBank: GCA_000015505.1) and *Polaromonas vacuolata* KCTC 22033^T (GenBank: GCA_012584515.1)), resulting in a total of 282 genomes (Table S1). Interestingly, there were a number of genomes labelled as *Polaromonas* in the NCBI database, which were not in fact *Polaromonas*, but instead being classified as the genus JAAFIP01 or JAAFJR01 in the GTDB. Because of their distinct genomic content, we decided to exclude them from our *Polaromonas* pangenome analysis. We also removed twenty genomes from mine drainage, activated sludge, landfill, host-associated, and bioreactor environments to focus on the presence of this genus in natural environments. We used the MonoPhy R package (v1.3.2)[94] to confirm that the remaining genomes were part of a monophyletic group by comparing them with other close genera (e.g., *Rhodoferax*) (Figure S5). Due to the way that metadata is deposited in the different databases, we cannot always further detail the type of environments where the samples were taken. Indeed, some soil samples could be linked to permafrost as described elsewhere,[13] while it is usually impossible to find out if glacier samples are from supraglacial, englacial, or subglacial parts (Table S1). Further, many of the metadata linked to some of the samples were missing, such as temperature, salinity concentration, or carbon content. Due to recent global datasets being published,[16,18,44,45,101,102] most of the *Polaromonas* publicly available were obtained from samples coming from the Tibetan Plateau (China) (123), followed by Switzerland (29), Canada (25), and the United States (23) (Figure 1A).

Pangenome analysis

The selected 282 genomes were annotated with Baka (v1.9.3),[84] microtrait (v1.0.0)[65] and eggnog-mapper (v2.1.12).[85] Genomad (v1.11)[62] was used to identify mobile genetic elements on each genome, while skani (v0.3.1)[86] was used to determine the average nucleotide identity of each pair of MAGs. Then, PPanGGOLiN (v2.1.1)[49] was used to create the pangenome using default parameters with the exception of the number of partitions defined as 3 (-K). PPanGGOLiN via mmseqs2[87] (v15) clusters the proteins into gene families with an identity and coverage of 80%. These gene families are then divided into three categories: persistent, shell, and cloud. Persistent (1905) and shell genes families (11,722) were then aligned using the "ppanggolin msa" command via the MAFFT software (v7.526).[88] We analysed the pangenome based on persistent, shell, and cloud genes using the occurrence frequencies of genes and their genomic neighbourhood to divide them into persistent, shell, and cloud genes that are present in almost all genomes, at intermediate frequencies or low frequencies, respectively.[49]

Species tree estimation and reconciliation

We then used mrBayes (v3.2.7)[89] to create gene trees for each gene family using the GTR + Gamma evolutionary model. The Markov Chain Monte Carlo (MCMC) analysis was run for 300,000 generations, with trees sampled every 300 generations and convergence diagnostics assessed every 1,000 generations. We then inferred the species tree using AleRax (v1.0.0)[90] with the gene trees obtained by mrBayes. An initial species tree was constructed using the MiniNJ (Minimum Evolution Neighbor-Joining) method. To account for missing genes in MAGs, we included a fraction missing file with completeness values for each genome. To reduce runtime, we removed 0.5% of the largest families, and we only considered gene families present in at least 30 genomes (total 4,114). The final rooted species tree was inferred using maximum likelihood-based reconciliation approaches via the HYBRID method. Then, using this species tree, we performed a reconciliation analysis with the gene trees to determine DTLO (duplications, transfers, losses, and origination) parameters, which accounts for gene duplication, horizontal gene transfer, and gene loss events. The analysis was performed using a per-family model parameterization, allowing distinct evolutionary models for each gene family and using 100 sampled

Please cite this article in press as: Michoud et al., Evolutionary radiation of *Polaromonas* from mountain glaciers downstream, Current Biology (2026), https://doi.org/10.1016/j.cub.2026.06.070

gene trees to improve robustness. Ancestral state reconstruction to visualize habitat evolution was done using PastML (v1.9.41),[53] which calculates ancestral state at each internal node. Ancestral states were also inferred using stochastic character mapping as implemented in the R package phytools (v2.5-2),[91] generating 200 simulated mappings. Posterior probabilities of states at each node were summarized across simulations, and the most probable state was assigned to each node. We also reconstructed ancestral states via the *ace* function of the ape package (v5.8-1).[92] Node-wise concordance was high across methods, with 98.93% agreement between stochastic character mapping and PastML, and 99.28% agreement between PastML and ace. Agreement was calculated based on the most probable state at each node. All three approaches consistently inferred a glacier-fed state at the root, with posterior probabilities ranging from 92.2% to 92.89%, indicating strong support for this ancestral condition.

QUANTIFICATION AND STATISTICAL ANALYSIS

We used R (v4.5.1)[95] and the packages targets (v1.11.3),[96] tidyverse (v2.0.0),[97] ggtree (v3.16),[98] and vegan (v2.6-10)[99] to perform statistical analyses and plot figures. Statistical significance was defined as p-value below 0.05. Details of the statistical methods are provided in the main text, figure legends, and supplementary figures.

Sensitivity analysis of ancestral state reconstruction and enrichment of horizontal transfer events at transition nodes

To determine if potential misclassification of the environment of the genomes may influence ancestral reconstructions, we randomly misclassified them in the tree at rates of 1, 5, 10, 15, and 20% and repeated ancestral character estimation 100 times per rate using the *ace* function of the ape R package. We also tested whether transfer events and mobile genetic elements (MGEs) were enriched in transition nodes compared to background nodes. Counts of interest (including horizontal gene transfers and MGE abundances) were compared between groups using a one-sided Wilcoxon rank-sum test (alternative = "greater"). Effect sizes were estimated using the rank-biserial correlation, and uncertainty was quantified via bootstrap resampling (5,000 iterations), with 95% confidence intervals derived from the empirical distribution of resampled effects.

Phylogeny-aware pathway-habitat association analysis

To formally test associations between functional pathways (as determined by microtrait) and habitat while controlling for phylogenetic non-independence, genome size, and completeness, we used phylogenetic generalised least squares (PGLS) implemented in the caper package (v1.0.4).[93] The phylogenetic tree was used to construct a phylogenetic variance-covariance matrix for all subsequent models. For each of the 125 pathways, we fitted a PGLS model such as pathway $\sim$ habitat + log(genome size) + completeness. Habitat was encoded using sum-to-zero (effects) contrasts, making all coefficients and the likelihood ratio test independent of reference level choice. Pagel's λ was estimated by maximum likelihood for each pathway independently via the *pgls* function, allowing the degree of phylogenetic signal to be calibrated to the actual covariance structure of each trait rather than being assumed a priori. Statistical significance of the habitat term was assessed by a likelihood ratio test comparing the full model against a null model retaining only genome size and completeness as covariates. To aid interpretation, significant pathways were classified into three categories based on their estimated λ: habitat-driven ($\lambda \approx 0$), indicating pathway distributions essentially independent of phylogenetic relatedness; intermediate ($0.3 < \lambda < 0.6$), indicating habitat enrichments partially but not wholly explained by clade structure; and phylogenetically conserved ($\lambda \approx 1$), indicating that habitat associations are primarily driven by clade-habitat co-occurrence.