ORIGINAL ARTICLE

Complete ecological isolation and cryptic diversity in *Polynucleobacter* bacteria not resolved by 16S rRNA gene sequences

Martin W Hahn[1], Jitka Jezberová[1,2], Ulrike Koll[1], Tanja Saueressig-Beck[1] and Johanna Schmidt[1]

[1]*Research Institute for Limnology, University of Innsbruck, Mondsee, Austria and* [2]*Biology Centre of the ASCR, v.v.i., Institute of Hydrobiology, České Budějovice, Czech Republic*

Transplantation experiments and genome comparisons were used to determine if lineages of planktonic *Polynucleobacter* almost indistinguishable by their 16S ribosomal RNA (rRNA) sequences differ distinctively in their ecophysiological and genomic traits. The results of three transplantation experiments differing in complexity of biotic interactions revealed complete ecological isolation between some of the lineages. This pattern fits well to the previously detected environmental distribution of lineages along chemical gradients, as well as to differences in gene content putatively providing adaptation to chemically distinct habitats. Patterns of distribution of iron transporter genes across 209 *Polynucleobacter* strains obtained from freshwater systems and representing a broad pH spectrum further emphasize differences in habitat-specific adaptations. Genome comparisons of six strains sharing ⩾99% 16S rRNA similarities suggested that each strain represents a distinct species. Comparison of sequence diversity among genomes with sequence diversity among 240 cultivated *Polynucleobacter* strains indicated a large cryptic species complex not resolvable by 16S rRNA sequences. The revealed ecological isolation and cryptic diversity in *Polynucleobacter* bacteria is crucial in the interpretation of diversity studies on freshwater bacterioplankton based on ribosomal sequences.

Introduction

In comparison with marine systems, freshwater systems are characterized by a large spectrum of chemical conditions ranging from acidic softwaters heavily colored by humic substances to alkaline hardwaters poor in dissolved organic carbon (Stumm, 2007). In general, freshwater systems can largely differ in pH, in quality and quantity of dissolved inorganic ions and organic material. This enormous environmental and chemical variability should result in large numbers of ecological niches available for freshwater bacteria across the spectrum of habitats. Despite this expectation, the diversity of abundant taxa of freshwater bacteria across systems seems quite small (Newton *et al.*, 2011). One potential explanation for the unexpected low number of taxa is the lack of ecological discrimination among taxa in studies based on 16S ribosomal RNA (rRNA) sequences (Jaspers and Overmann, 2004).

We searched for cryptic and taxonomic diversity within subcluster PnecC of the genus *Polynucleobacter*, which represents a ubiquitous and frequently abundant fraction of freshwater bacterioplankton (Percent *et al.*, 2008; Jezberova *et al.*, 2010; Newton *et al.*, 2011). Planktonic strains affiliated with subcluster PnecC share 16S rRNA sequence similarities >99% (Hahn, 2003; Watanabe *et al.*, 2009) and are currently classified as members of the species *Polynucleobacter necessarius* (Hahn *et al.*, 2009). They inhabit a very broad spectrum of freshwaters differing strongly in chemical, climatic and other ecological conditions (Wu and Hahn, 2006; Percent *et al.*, 2008; Jezberova *et al.*, 2010; Ghai *et al.*, 2011; Hahn *et al.*, 2015), and are ecologically classified as oligotrophic ultramicrobacteria (Salcher, 2014). A previous study suggested ecological diversification among PnecC bacteria (Jezbera *et al.*, 2011), but the extent of this diversification remains unknown. In theory, this ecological diversification could result from slight adaptations that provide a particular strain with competitive advantages under certain ecological conditions or from a complete specialization of lineages to distinct environmental conditions. The latter evolutionary state may be represented by strains largely differing in gene content and even

representing well-separated species, whereas the former state may represent ecotypes with much smaller genetic differences.

To reveal the extent of diversification among PnecC bacteria, we performed transplantation experiments mainly with two groups of strains seemingly lacking any co-occurrence in nature (Jezbera et al., 2011). One group (F10) inhabits acidic softwater systems characterized by low concentrations of dissolved ions, whereas the second taxon (F17) dwells in alkaline hardwaters characterized by high concentrations of dissolved calcium carbonate. Three kinds of transplantation experiments differing in organismal complexity were performed, ranging from transplantation of pure culture strains into sterilized softwater or hardwater to transplantation of an entire community into a small pond. In order to test if the revealed growth limitations in foreign environments are restricted to F10 and F17 lineage strains, four strains representing other PnecC lineages, respectively, were included in one set of transplantation experiments. In addition to these ecophysiological experiments, we compared the whole-genome sequences of one F10 and one F17 lineage strain, and also included four additional publicly available PnecC genome sequences in these comparative analyses.

Materials and methods

Strains
Nine Polynucleobacter strains currently classified as P. necessarius subspecies asymbioticus (Supplementary Materials Table S1) and affiliated with subcluster PnecC were used for transplantation experiments. Four and one strains, respectively, were isolated from acidic Pond-1 (Hahn et al., 2005) and a similar acidic pond (pond Rechteckteich; Jezbera et al., 2011), one strain from a softwater lake characterized by circumneutral pH (Lake Unterer Klaffersee; Jezbera et al., 2011), and three strains from alkaline Lake Mondsee (Wu and Hahn, 2006). All strains had been previously characterized by sequencing of ribosomal markers and the glutamine synthetase gene (glnA), and all but one strain are affiliated

with previously characterized lineages (Jezbera et al., 2011).

Transplantation experiments
Three kinds of transplantation experiments differing in their organismal complexity were performed with pure culture strains or with an entire microbial community. The three kinds of experiments differed in the presence of competitors, algae and predators (Table 1).

(i) Home-and-away experiments. Reciprocal transplantation of strains were conducted to study growth performance of strains in sterile water of their home habitats and in sterile water of a habitat assumed not to be colonized by the strains. Sterile water of Lake Mondsee (Wu and Hahn, 2006) and Pond-1 (Hahn et al., 2005) were used for these experiments. As autoclaving of water from Lake Mondsee results in enormous carbonate precipitations, sterilization of water from both sources was performed after 0.2 µm filtration by gentle microwave treatment as described previously (Schauer and Hahn, 2005). The tested strains were pre-cultured in diluted ($10 \, mg \, l^{-1}$) NSY medium (Hahn et al., 2004), inoculated (5% of volume) in 50 ml of the sterilized lake or pond water, incubated at 15 °C in the dark without agitation, and transferred (5% of culture volume) into sterilized water at intervals of 3–5 days. All strains were tested in parallel in both waters (triplicates each), and batches without inoculation were run as negative controls. Viability of pre-cultures was tested by inoculation in $3 \, g \, l^{-1}$ NSY medium. Each experiment consisted of five transfers, each transfer resulting in a dilution of cultures by 1:20. Thus, the last sample quantified for bacterial numbers represented a 3.2×10^6 fold dilution of the inoculum. Treatments were sampled immediately before the transfers, and bacterial numbers were determined by direct counting using an epifluorescence microscope (Jezberova et al., 2010).

(ii) Transplantation of a Polynucleobacter F10 strain in Lake Mondsee water: microcosm experiment. Strain QLW-P1DMWA-1 representing lineage F10

Table 1 Differences between the performed transplantation experiments

Experiment	Organism(s) inoculated	Environment	Temperature (°C)	Presence of			
				Bacteria[a]	Algae	Predators[b]	Metazooplankton
Home-and-away	Nine strains[c]	Sterile lake/pond water	15	–	–	–	–
Microcosm experiment	QLW-P1DMWA-1	<100 µm Lake Mondsee water	15	+	+	+	–
Transplantation experiment I	Community Pond-1[d]	Garden pond	21.6–27.5	+	+	+	+
Transplantation experiment II	Community Pond-1[d]	Garden pond	15.6–22.4	+	+	+	+

[a]Prokaryotes affiliated with other taxa (that is, potential competitor and potential partners for beneficial interspecies interactions). [b]Protistan predators and other microzooplankton. [c]Pure culture strains, separately inoculated; Supplementary Materials Table S1. [d]Includes three Polynucleobacter lineages represented by strains investigated in the home-and-away experiments.

(Table 1, Supplementary Materials Table S1) was added to water from Lake Mondsee. Cells of the strain were harvested by centrifugation from cultures grown in NSY medium, washed with substrate-free inorganic medium (Hahn *et al.*, 2004) and added at a concentration of 10^5 cells ml^{-1} to 1000 ml of water from Lake Mondsee in 2 l bottles (triplicates). The water was freshly sampled from the lake and screened through a 100 μm sieve for removal of metazooplankton. As a negative control, 200 ml of <100 μm Lake Mondsee water not inoculated with the *Polynucleobacter* strain were incubated. Positive controls consisted of 200 ml of sterile (see above) Lake Mondsee and Pond-1 water, respectively, each inoculated with strain QLW-P1DMWA-1. All treatments were incubated at 15 °C in the light (16:8 h light–dark intervals), and transferred into the respective fresh media at 3-day intervals. Every third day, therefore, 50% of the treatment volume was replaced by fresh <100 μm Lake Mondsee water or by sterile water from Lake Mondsee or Pond-1, respectively. The experiment was sampled every third day. The abundance of strain QLW-P1DMWA-1 was determined by quantitative PCR as described previously (Hahn *et al.*, 2012).

(iii) Transplantation of a natural community in a garden pond. A natural plankton community dwelling in acidic Pond-1 (Hahn *et al.*, 2005) was transplanted in a small slightly alkaline garden pond (volume of about 1500 l; Supplementary Materials Table S2). The transplanted community contained *Polynucleobacter* lineages F10, F14 and F4 and other uncharacterized PnecC lineages, whereas the garden pond was colonized by lineage F17 and other PnecC lineages (Jezbera *et al.*, 2011). Two transplantation experiments were done during two consecutive summers. In each experiment, in total 200 l of untreated Pond-1 water were gently mixed into the pond water. As controls (triplicated), 700 ml of Pond-1 water were filled in 2 l Schott bottles and incubated floating at the water surface of the garden pond. The garden pond was sampled before and after the transplantation to determine various microbial and environmental parameters. This included determining total bacterial numbers by epifluorescence microscopy (Jezberova *et al.*, 2010), determining the percentage of *P. necessarius* bacteria (PnecC bacteria) by fluorescent *in situ* hybridization (Hahn *et al.*, 2005), and determining the percentage of F10 bacteria by quantitative PCR (Hahn *et al.*, 2012). The three control bottles were sampled once, 7 days after the start of the transplantation experiment. In both transplantation experiments performed in consecutive years, the introduced water from Pond-1 represented about 12% of the water contained in the garden pond. In both experiments, the F10 lineage (Figure 3) and two other *Polynucleobacter* lineages present in Pond-1 (data not shown)

could not be detected in the garden pond before inoculation with Pond-1 water.

Genome sequencing of strain MWH-MoK4
DNA used for genome sequencing was extracted from biomass of strain MWH-MoK4 grown in liquid NSY medium (Hahn *et al.*, 2004) as described previously for another *Polynucleobacter* strain (Meincke *et al.*, 2012). Construction of a mate-pair library and a fragment paired-end library, as well as sequencing of both libraries using a Roche GS FLX system and Titanium chemistry (Roche, Branford, CT, USA), was performed by Beckman Coulter Genomics (Danvers, MA, USA). *De novo* assembly of reads obtained from both libraries resulted in one scaffold and eight contigs. Gaps could be manually closed after elimination of low-quality sequences at the ends of contigs. The genome was sequenced at an average coverage of 38.6-fold. The closed genome sequence was annotated using the Integrated Microbial Genomes-Expert Review (IMG/ER) annotation pipeline (Markowitz *et al.*, 2012).

The genome sequence of strain MWH-MoK4 was compared with five publicly available genome sequences of PnecC bacteria (Supplementary Materials Table S3) using the software BLAST+ (Camacho *et al.*, 2009), IMG/ER (Markowitz *et al.*, 2012), JSpecies (Richter and Rosselló-Móra, 2009) and GGDC (Auch *et al.*, 2010). Genes with nucleotide sequences sharing >60% identity with sequence coverage >70% and e-values of E<10^{-5} in pairwise BLASTn comparisons were defined as homologous genes.
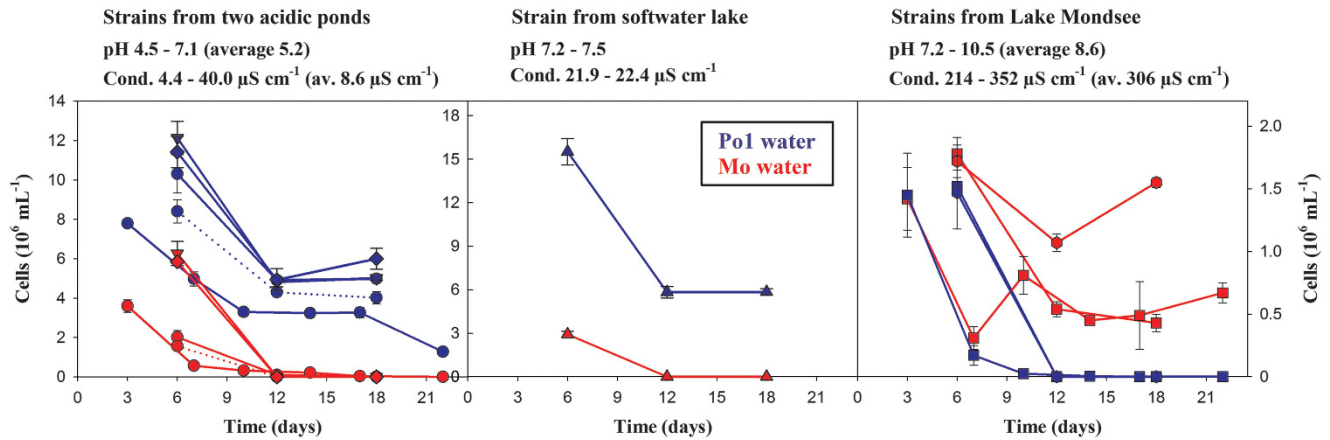
Nucleotide sequences
The genome sequence of strain MWH-MoK4 was deposited in GenBank under accession number CP007501. Accession numbers of partial iron transporter genes are LN898146–LN898169.

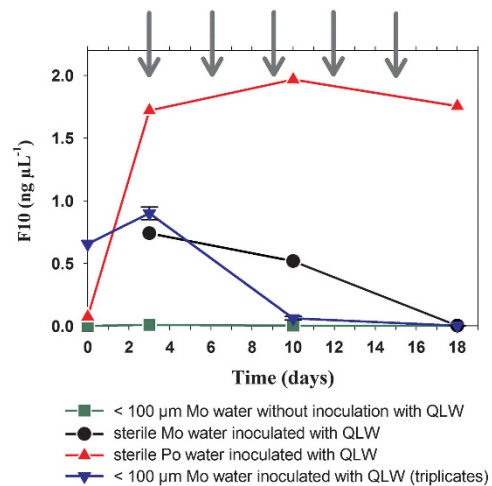## Results

Transplantation experiments
The growth potential of nine *Polynucleobacter* strains in a foreign environment was tested by reciprocal inoculation of strains in sterilized water from alkaline Lake Mondsee and acidic Pond-1, respectively (home-and-away experiments). Sterilization of Pond-1 water resulted in increase of pH from pH 5.0 to 6.3–6.6, whereas pH of Lake Mondsee water was only slightly changed (from 8.2 to 8.3–8.4) by the sterilization process. In each experiment, both triplicated treatments were initially inoculated with equal cell numbers of test strains. Bacterial numbers decreased in all experiments and treatments during the initial two or three transfer intervals, and more or less stabilized afterward (Figure 1). However, for all nine strains sustainable growth was observed in only one of the two tested types of waters. All strains grew

Strains from two acidic ponds
pH 4.5 - 7.1 (average 5.2)
Cond. 4.4 - 40.0 µS cm⁻¹ (av. 8.6 µS cm⁻¹)

Strain from softwater lake
pH 7.2 - 7.5
Cond. 21.9 - 22.4 µS cm⁻¹

Strains from Lake Mondsee
pH 7.2 - 10.5 (average 8.6)
Cond. 214 - 352 µS cm⁻¹ (av. 306 µS cm⁻¹)

Figure 1 Growth potential of nine PnecC strains (Supplementary Materials Table S1) in reciprocal transplantation (home-and-away) experiments. (Left panel) Three F10 lineage strains and two other strains (F4 and F5, respectively) all but one (dotted lines) originating from Pond-1 were tested for growth in sterile Pond-1 (blue) and sterile Lake Mondsee water (red). (Middle panel) Same experiment with a strain originating from a circumneutral softwater lake (Lake Unterer Klaffersee). (Right panel) Same experiment with three strains from Lake Mondsee affiliated with or closely related to lineage F17. The sterile media were inoculated at day 0, cultures were transferred in intervals of 3–5 days and sampled before transfers. In experiments with strains QLW-P1DMWA-1 and MWH-MoK4, cell numbers were determined for each transfer. In experiments with the other strains, cell numbers were determined for each second transfer. Data represent average cell numbers and standard deviations of three replicates.

sustainably only in the water from their home habitats, or in water sharing characteristics with their home habitat (strain MWH-RechtKolB). Interestingly, one strain (MWH-UK1W16) originating from a softwater lake (not tested for growth in water from that lake), with intermediate chemical characteristics compared with acidic Pond-1 and alkaline Lake Mondsee (Figure 1), grew only in sterile water from Pond-1 and not in water from Lake Mondsee. No growth was observed in all negative control treatments.

As in the home-and-away experiments, all strains were tested only for their growth potential in absence of other organisms, we tested exemplarily one strain for its growth potential in a foreign environment in the presence of other organisms. Prokaryotic or eukaryotic microorganisms could provide the strain with growth factors or services (Garcia et al., 2015) essential for growth in a foreign environment. In a microcosm experiment, strain QLW-P1DMWA-1 from acidic Pond-1 was grown in Lake Mondsee water retaining its entire microbial community. Only larger zooplankton and other macroorganisms were excluded by screening through 100 µm sieves. To enable photosynthetic activity of phytoplankton and activity of heterotrophic microbes, the treatments were incubated in the light, and half the volume of the treatments were replaced by freshwater from Lake Mondsee (<100 µm) every third day. The abundance of strain QLW-P1DMWA-1 increased slightly during the first 3 days of incubation (Figure 2), but in the further course of the experiment its abundance dropped below the detection limit of the quantitative PCR assay. A similar response by the strain was observed in a control treatment with sterile Lake Mondsee water, but the decline in the abundance of the strain was slower. By contrast, the strain was able to establish a stable and abundant

Figure 2 Results of the microcosm experiment transplanting strain QLW-P1DMWA-1 (QLW, representing lineage F10) in water from Lake Mondsee screened through a 100 µm sieve for removal of metazooplankton. This treatment was run in triplicates (blue triangles, average and s.d.). For comparison, the strain was also inoculated in sterile water from Lake Mondsee (black circles) and sterile water from Pond-1 (red triangles) but with different volumes and cell numbers. The treatments with 100 µm-screened and sterile Lake Mondsee water received the same inoculum volumes, but extraction of the 0 h sample failed in the latter treatment. In all, 100 µm-screened water of Lake Mondsee not inoculated with strain QLW-P1DMWA-1 served as negative control (green squares). In all treatments, 50% of medium was replaced each third day (gray arrows). Sampling on day 3 took place immediately before replacement of medium. The abundance of strain QLW-P1DMWA-1 was determined by quantitative PCR.

population in a control treatment with sterile Pond-1 water (Figure 2).

To test if the growth potential of strain QLW-P1DMWA-1 and related F10 strains is influenced by the presence of naturally co-occurring microorganisms,

large samples of the microbial community inhabiting Pond-1 were transplanted in a garden pond in two experiments conducted in consecutive years. This pond shares some ecological features with Pond-1, for example, shallowness and pronounced diurnal fluctuation of environmental conditions (Hahn *et al.*, 2012), but shares with Lake Mondsee a more alkaline chemistry (Pond-1, average pH 5.2; Lake Mondsee, average pH 8.6; garden pond, pH 7.4-7.8). The F10, F4 and F5 lineages typically present in Pond-1 (Hahn *et al.*, 2005; Jezbera *et al.*, 2011; Hahn *et al.*, 2012) were detected by quantitative PCR assays in the transplanted Pond-1 water but not in samples from the garden pond taken just before the start of both experiments (Figure 3). In both experiments, the transplantation resulted immediately in an increase of PnecC cell numbers and in appearance of a high abundance of F10 lineage bacteria. However, numbers of F10 bacteria (Figure 3) and the other two PnecC lineages from Pond-1 (data not shown) decreased steadily and dropped below the detection limits within 3–7 days after the transplantation. In both experiments, the F10 numbers decreased

exponentially but at different rates (Figure 4). Higher decay rates were observed in the transplantation experiment with higher water temperatures in the garden pond (Table 1).

Genomic differences between strain MWH-MoK4 and QLW-P1DMWA-1

The genomes of both strains representing the F17 and F10 lineages (Table 2) encode only a single ribosomal operon. The full-length 16S rRNA sequences (1527 bp) of the two strains differ in five positions, and thus share sequence similarities of 99.7% (Table 3). Four of the sequence differences represent two pairs of base-pairing positions of hairpin secondary structures. These base-pairing positions co-evolved in the hairpin structures in all 16S rRNA genes of *P. necessarius* (currently 231 sequences in Genbank) together, which decreases the phylogenetic information content of the five sequence differences. None of the five positions is lineage specific, for either the F10 or the F17 lineage. These base-pairing positions are located in the
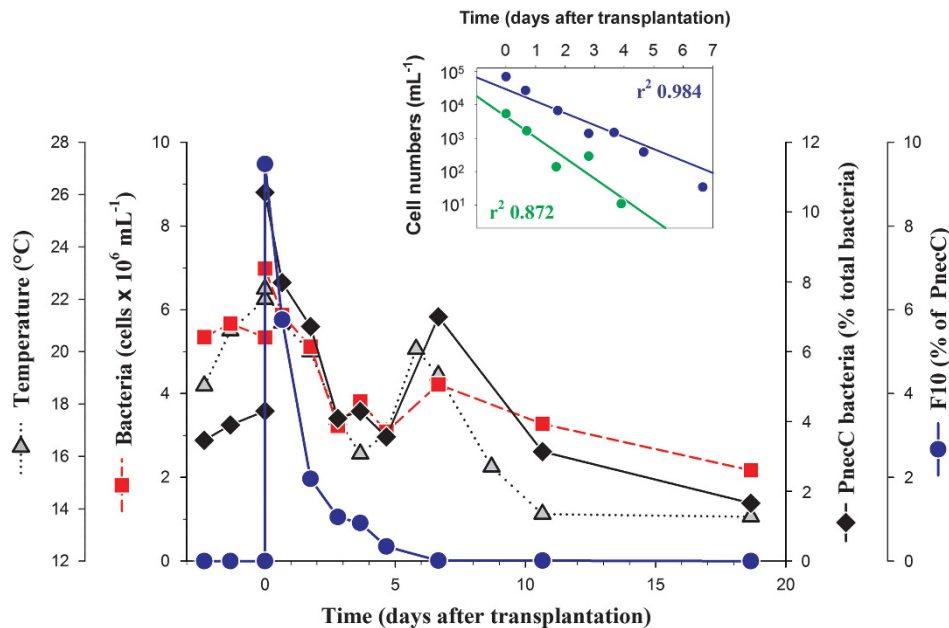

Figure 3 Results of transplantation experiment II. Water from Pond-1 was transplanted into a garden pond at day 0. F10 bacteria were not detected in the pond before the transplantation but represented about 10% of the PnecC bacteria immediately after the transplantation. The small graph presents cell numbers of F10 bacteria determined by qPCR assays in transplantation experiment I (green symbols) and II (blue symbols) in a semi-logarithmic plot. Lines represent linear regressions of data, respectively. Note the faster decline of cell numbers in experiment I, which was characterized by higher average water temperatures (Table 1).

Figure 4 Comparison of genomes of *Polynucleobacter* strains QLW-P1DMWA-1 (QLW) and MWH-MoK4 (MoK4) representing lineage F10 and F17, respectively. Graphs shown in the middle panel compare data of both taxa, graphs shown in the left and right panel show data on strain QLW and MoK4, respectively. (**a**) The F10 and F17 lineages seem not to co-occur in the same freshwater systems and seem to prefer habitats differing in pH values (data from Jezbera *et al*, 2011). (**b**) Acidic, dystrophic, shallow Pond-1 is the home habitat of strain QLW. (**c**) Alkaline, deep Lake Mondsee is the home habitat of strain MoK4. (**d**) Genome map of strain QLW, (**f**) genome map of strain MoK4. (**e**) Each strain shares about 70% of genes with the other strain, and about 30% of genes are unique to each strain. (**g**, **i**) GC content (mol%) of shared and unique genes. (**h**) Dot plot of positions of shared genes in the two genomes. Blue and red dots depict genes oriented in the same and the opposite orientation in the compared genomes. (**j**) Comparison of distribution of unique genes assigned to clusters of orthologous group (COG) categories shown as percentage of total numbers of unique genes in each genome.

variable regions V1 and V6 of the 16S rRNA gene, respectively, and the other sequence difference is located in position 546 (*Escherichia coli* numbering).

As such, typical cultivation-independent approaches investigating diversity of bacterial communities by pyrosequencing of partial 16S rRNA gene amplicons
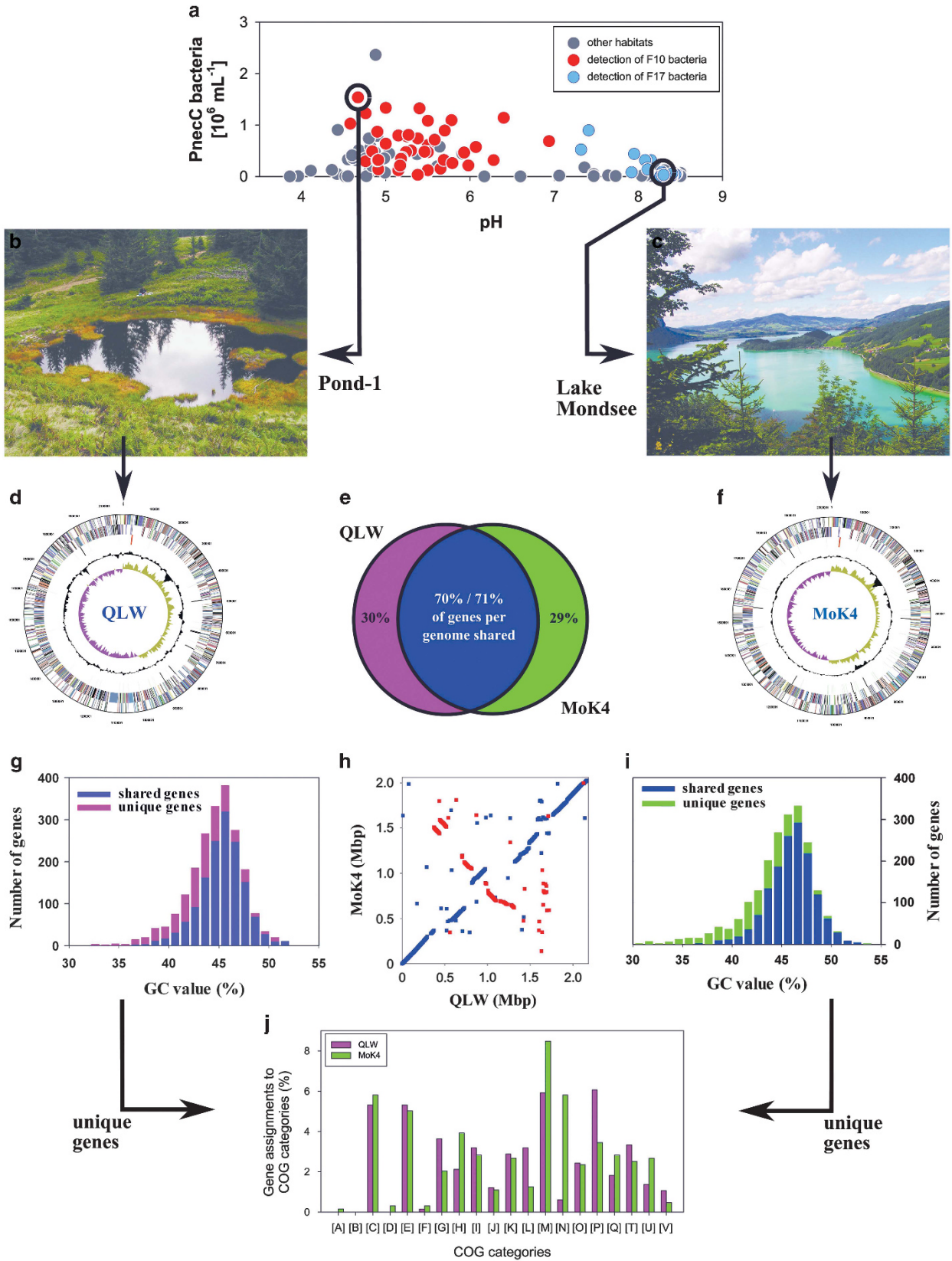
Table 2 General features of the genomes of two *Polynucleobacter* strains representing lineages F10 and F17, and examples for differences in gene content between the two genomes

Feature/genes	QLW-P1DMWA-1	MWH-MoK4
Origin	Pond-1	Lake Mondsee
RLBH group[a]	F10	F17
Genome size (Mbp)	2.16	2.03
GC content (mol%)	44.8	45.2
Coding bases (% of total)	93.1	93.5
Protein-coding genes	2088	2059
RNA-coding genes	48	48
Replicons	1	1
Ribosomal operons	1	1
Acc. number genome	CP000655	CP007501
Acc. number 16S rRNA	AJ879783	AJ550654
Inorganic nutrients		
Phosphate/phosphite/phosphonate ABC transporter	−	+
ABC-type Fe^{3+} transport system	−	+
ABC-type Fe^{3+}-hydroxamate transport system/ABC-type cobalamin/Fe^{3+}-siderophores transporter	−	+
feoAB genes (uptake of Fe^{2+})	+	−
ABC-type nitrate/nitrite/cyanate transporter	+	−
Nitrate reductase (assimilatory)	+	−
Nitrite reductase (assimilatory)	+	−
Cyanate lyase (releases NH_3 and CO_2 from cyanate)	+	−
Urease and ABC-type urease transporter	+	−
Amt family ammonia transporter	1 gene	2 genes
Regulation of K^+ transport (kdpD/kdpE)	+	−
Oxidative phosphorylation/energy metabolism		
Cytochrome bd-I terminal oxidase (CydAB)	+	−
Fumarate reductase	−	+
Carbon monoxide dehydrogenase	−	+
Acetate permease actP	+	−
Anoxygenic photosynthesis		
Photosynthesis gene cluster	−	+
Motility		
Flagella genes	−	+
Putative sigma-54	−	+
Oxidative stress		
Catalase	2 genes	−
Other		
Cellulose synthase operon protein C	+	−
Cellulose synthase catalytic subunit (UDP-forming)	+	−

Abbreviations: RLBH, reverse line blot hybridization; rRNA, ribosomal RNA.
[a]RLBH group (Jezbera *et al.*, 2011).

(Besemer *et al.*, 2012; Claesson *et al.*, 2012; Pinto *et al.*, 2012) would obtain sequences from the two *Polynucleobacter* strains containing only one or two of the five sequence differences, resulting in sequence similarities of 99.5% or 99.8% (assumed amplicon length of 400 bp).

As for the ribosomal genes, the two genomes also share high similarities in basic genomic features like genome size, gene number and GC content (Table 2) but largely differ in non-ribosomal sequence features (Table 3) and share only 71% of their genes (Figure 4e). Most genes not shared by the two genomes possess GC values below the average of the entire genomes (Figures 4g and i). The sets of unshared genes include many genes that may alter the phenotypes of the strains in ecologically significant ways (Table 2). The unshared (unique) genes

of the two genomes distribute differently (Figure 4j) across the clusters of orthologous groups categories G (carbohydrate metabolism and transport), H (coenzyme metabolism), L (replication and repair), M (cell wall/membrane/envelop biogenesis), N (cell motility) and P (inorganic ion transport and metabolism).

The two genomes differ in the presence of genes putatively encoding iron transporters. The strain isolated from acidic waters encodes a FeoAB transporter putatively enabling uptake of Fe^{2+}, whereas the strain obtained from an alkaline environment encodes an ABC transporter putatively enabling Fe^{3+} uptake from the periplasmic space (Table 2). We screened 209 PnecC strains for presence of these genes by PCR and found a strong trend of distribution of these genes across strains

Table 3 Sequence-based comparisons of strains MWH-MoK4 and QLW-P1DMWA-1

Compared feature	Value
16S rRNA similarity, entire gene[a] (%)	99.7
16S rRNA similarity, V3–V5 region[b] (%)	99.8
16S-23S ITS similarity (%)	96.1
glnA similarity[c] (%)	88.7
Shared genes (% of total)	71.0
AAI (shared genes/proteins) (%)	84.2
Two-way AAI (1557 proteins)[d] (%)	80.7
ANIb (JSpecies) (%)	75.6
Two-way ANI (Kostas lab) (%)	79.8
DNA–DNA hybridization estimates[e] (GGDC 2.0) (%)	19.3–29.3

Abbreviations: AAI, average amino-acid identity; ANI, average nucleotide identity; rRNA, ribosomal RNA.
[a]1527 bp.
[b]Partial gene encompassing the V3 to V5 region; typically sequenced in high-throughput studies; 584 bp.
[c]Partial glnA sequences, 603 bp.
[d]Average amino-acid identity calculator, Kostas lab.
[e]Three estimates by three different algorithms.

from different pH environments (Supplementary Materials Table S2). In 82% of the tested strains, only one of the two genes could be detected, and 8% of strains were negative for both genes. The feoB gene was mainly detected in strains obtained from freshwater systems with pH below 7, whereas the Fe^{3+} transporter gene was mainly detected in strains obtained from alkaline systems. Detection of both genes in single strains was most frequent in strains obtained from habitats with pH values in the range 6–7. Comparison of PCR results with complete genome sequences indicated that 8% of the detections were either false positive or negative. To confirm the observed trends, sequences obtained by PCR or genome sequencing from a total of 25 strains were analyzed (Figure 5a). This confirmed the trend observed by the analysis of detections by PCR and revealed in addition rather low sequence similarity values between the strains. The 18 partial feoB sequences showed an average sequence similarity of 78.3% (range 74.7–72.8%), and the 10 partial Fe^{3+} ABC transporter permease gene sequences were characterized by an average sequence similarity of 77.1% (range 72.5–80.8%).

Average nucleotide identities (ANIs) between six genomes of PnecC bacteria
Six genomes of putative PnecC bacteria (Supplementary Materials Table S3) were compared. The additional four genomes also encoded only one ribosomal operon each, respectively, but in one genome the operon was distributed over two contigs. A phylogenetic analysis with 16S rRNA gene sequences of the six strains confirmed their affiliation to subcluster PnecC (Supplementary Materials Table S1). All but one 16S rRNA sequence similarity values of the six strains were above 99% (Figure 5b). The one value slightly below 99% and most of the

other lower values resulted from comparison with the only obligate endosymbiotic strain, that is, STIR1 (Vannini *et al.*, 2007; Boscaro *et al.*, 2013), among the compared six strains. Pairwise comparison of the six genomes resulted in ANI values (Konstantinidis and Tiedje, 2005) ranging from 75% to 84% and an average value of 76.7% (Figure 5). The only value larger than 80% resulted from the comparisons of strains MWH-MoK4 and beta proteobacterium CB (Hao *et al.*, 2013). Obviously, ANI values and 16S rRNA sequence similarity values are not correlated with each other (Figure 5b).

To determine if the low ANI values of the six genome-sequenced strains could be exceptional among PnecC strains, partial sequences of glutamine synthetase (glnA) genes (600 bp) and complete 16S-23S ITS sequences (500 bp) of the six genome-sequenced strains were compared with sequences of 240 PnecC strains isolated from a broad spectrum of freshwater habitats (Jezbera *et al.*, 2011; Hahn *et al.*, 2015). The phylogenetic resolution of those two markers had been demonstrated previously to be much higher than the resolution of 16S rRNA sequences (Hahn *et al.*, 2012). This comparison did not reveal exceptional similarity values for the loci of the six genome-sequenced strains (Figures 5c and e). Actually, 84% and 62% of those 240 PnecC strains showed maximal glnA similarity values with any other strain equal or below the values representing the highest and second-highest glnA similarity found among the six genome-sequenced strains. This could indicate that many of the 240 PnecC strains also possess only low ANI values with all other strains of the culture collection.

Discussion

Ecophysiological and genomic traits of representatives of two groups (F10 and F17) of *Polynucleobacter* bacteria currently classified as *P. necessarius* (PnecC) were investigated. Furthermore, several strains representing other PnecC lineages were included in experiments and genomic analyses. The overall goal of these investigations was to reveal the extent of niche separation between PnecC strains. We found much larger separations than expected for strains sharing almost identical 16S rRNA genes and revealed at least for some of the investigated lineages complete niche separations, that is, complete ecological isolation of lineages.

Ecological isolation
The results indicate that ecological isolation of PnecC lineages took place between at least two groups of lineages. One group contains lineages F10, F4, F5 and F15, and the other group is represented by lineage F17. The former group could maintain populations in sterile Pond-1 water but not in sterile Lake Mondsee water, and the latter group showed

the opposite results. During the first days of the home-and-away experiments (Figure 1), all strains showed some growth in both sterile waters, but cell numbers dropped later either to zero or, in the home environments, to lower numbers. The initial growth may have resulted from the pre-cultivation in a rich
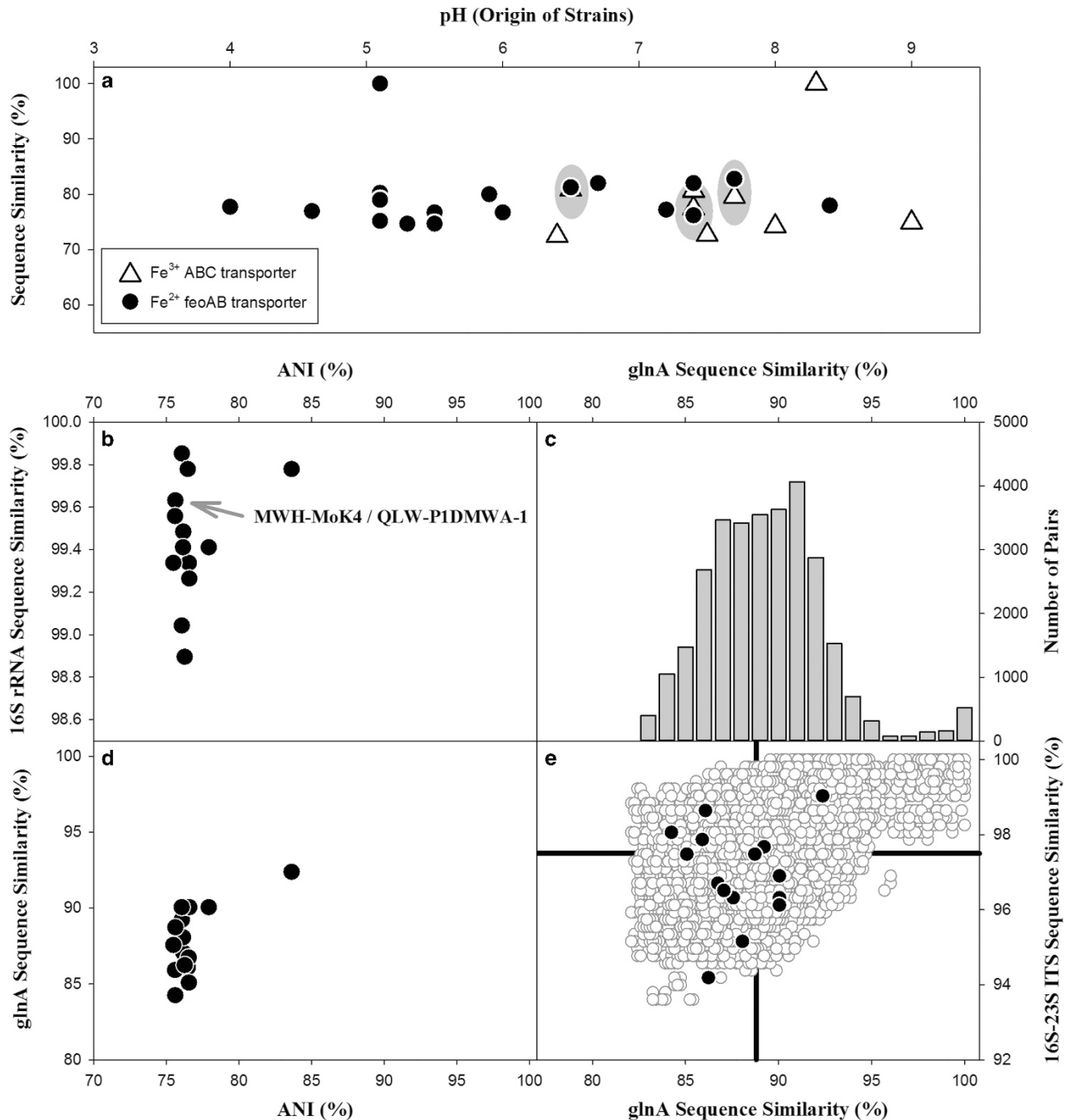
Figure 5 (**a**) pH of habitat of origin of 25 PnecC strains and presence of feoB Fe^{2+} and Fe^{3+} ABC transporter genes (compare Supplementary Materials Table S2). Partial sequences of genes were compared with the respective genes present in the genomes (Table 2) of strain QLW-P1DMWA-1 (feoB) and MWH-MoK4 (fpbB ABC transporter permease). The set of 25 strains also includes those two strains, whose self-comparisons resulted in the two plotted 100% similarity values, respectively. Both genes were detected in 3 of the 25 strains (indicated by gray ovals), respectively. (**b–e**) Distribution of ANI and 16S rRNA sequence similarity values (**b**) or partial glutamine synthetase (glnA) genes sequence (603 bp) similarities (**d**) of pairwise comparisons of six genomes of PnecC bacteria (Supplementary Materials Table S3). Note that two pairwise comparisons of ANI versus 16S rRNA similarity resulted in both parameters in almost identical values, indistinguishable in the shown plot. (**e**) Sequence similarities of glnA and 16S-23S ITS sequences (516 bp) of 240 PnecC strains (open circles) and the six genome-sequenced PnecC strains (closed circles). The plot depicts all pairwise comparisons of the 246 analyzed strains. The two black lines indicate the average sequence similarity values of the two markers. (**c**) Frequency distribution of glnA similarity values of the 246 PnecC strains.

complex medium and potential storage of limiting nutrients or transfer of NSY medium with the inoculum. Interestingly, strains able to grow in Pond-1 water maintained higher cell numbers in the second half of the experiments than strains able to grow in Lake Mondsee water. This result reflects differences in natural PnecC numbers observed between the two habitats or types of waters (Figure 4a; Hahn et al., 2005; Wu and Hahn, 2006). Disappearance of all strains in the respective away environments could be explained either by a much lower growth rate in these environments compared with the home environments or by complete growth inhibition after depletion of resources stored during pre-cultivation in NSY medium. In these home-and-away experiments, compensation for the performed 1:20 dilutions, which took place at intervals of 3–5 days, required average doubling times for cultures of below 17 h. Batch culture experiments with strain QLW-P1DMWA-1 in sterile water from Pond-1 (home environment) determined a doubling time of 4.3 h (data not shown). Growth rates of this strain in the foreign environment were thus either markedly lower or fell to zero after a few days of exposure to this environment.

Basically, the pattern of survival/no survival in the home-and-away experiments (Figure 1) could be explained by either abiotic limitation (for example, pH too high or too low) or by the lack of essential biotic interactions required by the Polynucleobacter strains under specific environmental conditions (Garcia et al., 2015). Such interactions could be, for instance, services such as synthesis of common goods (Cordero and Polz, 2014) including cofactors like cobalamine (Seth and Taga, 2014), exoenzymes (Morris et al., 2008) or metabolic precursors (Freilich et al., 2011). The microcosm (Figure 2) and the whole community transplantation experiments (Figure 3) suggest that the fitness of the tested strains in the tested foreign environments is not limited by growth supporting biotic interactions, which leaves only abiotic factors as explanation. However, identification of the limiting abiotic factor is difficult. The environmental distribution of the F10 and F17 lineages (Figure 4a), as well as of the other tested lineages (F4, F5 and F15; see Jezbera et al., 2011, for environmental distribution of these taxa) in combination with the pH values in the conducted experiments (Supplementary Materials Table S2) seems to suggest that pH is acting as limiting factor. In artificial media (NSY medium; Hahn et al., 2004), however, strains obtained from acidic environments tolerate pH increases of at least up to 7.8–8.0. Water chemistry of Pond-1 and Lake Mondsee differs not only in pH, but also in concentration of ions, and especially in concentration of dissolved $CaCO_3$ and humic substances. Several chemical differences between these habitats could thus be responsible for the observed ecophysiological differences between strains representing the two lineages. However, the presence/absence of genes encoding specific iron transporters in the genomes of the investigated F10 and F17 strains (Table 2) suggests that pH-related environmental parameters act at the least as fitness-limiting factors. Iron is an important micronutrient of bacteria, which exists in natural freshwaters in the oxidation states 2+ (ferrous) and 3+ (ferric). Solubility of ferrous and ferric iron mainly depends on pH, redox potential and concentrations of bicarbonate and complexing substances (Wetzel, 2001). The solubility of Fe^{2+} ions decreases in oxygenated natural waters strongly with increasing pH. This decrease in solubility together with increasing rates of oxidation of ferrous to ferric iron with increasing pH results in the absence of dissolved Fe^{2+} in alkaline freshwater systems (Wetzel, 2001; Stumm, 2007). In contrast to ferrous iron, solubility of ferric iron is extremely low in natural waters, but in the presence of complexing agents (for example, siderophores and humic substances) increased concentrations of complexed and bioavailable ferric iron occur. As a consequence, bioavailability of dissolved iron differs strongly between acidic and alkaline freshwater systems (Wetzel, 2001). In acidic and circumneutral softwater or humic freshwater systems, dissolved Fe^{2+} ions are usually available, whereas in alkaline hardwater systems the bioavailability of iron is frequently very low and acquisition of iron requires siderophores or other high-affinity Fe^{3+} chelating compounds (Zane and Butler, 2013). In principle, prokaryotes can acquire iron in the ferrous and ferric oxidation state, but different transport systems are used for different iron ions (Zane and Butler, 2013). Ferrous iron can be transported by the FeoAB transporter as Fe^{2+} ions, whereas ferric iron is usually acquired as chelated Fe^{3+}, that is, Fe^{3+} ions bound to siderophores or other chelating molecules. Several transport systems involved in uptake of Fe^{3+} ions released from chelating agents or together with the chelating agents have been characterized (Zane and Butler, 2013). Interestingly, five of the six investigated PnecC genomes either encode transporters putatively involved in Fe^{2+} uptake (feoAB) or in ferric iron uptake (ABC-type Fe^{3+} transporter and ABC-type cobalamin/Fe^{3+}-siderophores transport system). The only exception is beta proteobacterium CB (Hao et al., 2013), which encodes both types of putative iron acquisition systems. The contrasting presence of Fe^{2+} and Fe^{3+} transporters in the genomes of strains QLW-P1DMWA-1 and MWH-MoK4 (Table 2) matches well to the expected occurrence and availability of different iron species under the acidic and alkaline conditions in their home habitats (Figure 4a). In general, the observed distribution of transporter genes in the six genomes suggests that the majority of the genome-sequenced strains are adapted to a limited pH range, that is, either to acidic/circumneutral or to alkaline environments, which fits well to tendencies previously revealed by a cultivation-independent method for several PnecC lineages (Jezbera et al., 2011), as well as to the results

1652

of the home-and-away experiments. Currently, not many other examples illustrating niche separation among closely related strains on a genomic level are known for free-living bacteria. One example is the more frequent presence of an operon enabling glucose oxidation in costal SAR11 group Ia strains compared with open-ocean strains of the same group (Schwalbach et al., 2010). The frequency distribution of this trait correlates with ocean productivity, which putatively reflects the availability of glucose for pelagic marine bacteria. By contrast, the literature is rich in examples on genomic differences, among closely related pathogenic strains, which enable niche partitioning regarding access to different host species or tissues, or enabling resistance to different antibiotics.

Detailed comparison of the genomes representing lineages F10 and F17 (Table 2) indicates that the ecological diversification of these two lineages is not restricted to iron acquisition. The observed differences in gene content suggest differences in use of several nutrient sources and differences in survival strategies. The potential to produce flagella may release strain MWH-MoK4 from the passive lifestyle previously suggested for strain QLW-P1DMWA-1 (Hahn et al., 2012), and could enable the colonization of surfaces, aggregates or other nutrient-rich microhabitats (Grossart, 2010). However, motility of strain MWH-MoK4 has so far not been observed (Hahn et al., 2009) and the use of motility may be limited because of the putative lack of genes encoding a chemotaxis system. Both strains encode genes that may increase fitness under low oxygen concentrations (cytochrome bd-I oxidase) or under anoxic conditions (fumarate reductase). Key genes of anoxygenic photosynthesis were previously discovered in Polynucleobacter bacteria (Martinez-Garcia et al., 2012). Strain MWH-MoK4 encodes a large photosynthesis gene cluster containing all key genes of anoxygenic photosynthesis. This strain, however, is the only strain among the six strains represented by genome sequences that contains such a gene cluster (Table 2). This gene cluster potentially enables utilization of an energy source unavailable to the other PnecC strains. Cultures of strain MWH-MoK4, however, showed no pigmentation indicating expression of those genes, and the strain also lacks the typical genes for inorganic carbon assimilation. Interpreting the ecological role and significance of several of the observed genomic differences among the strains requires additional investigations. An extensive ecological diversification of strains beyond simple adaptation to a few chemical differences between the inhabited ecosystems is, nonetheless, obvious.

Cryptic ecological and taxonomic diversity among PnecC bacteria

The revealed complete ecological isolation of some PnecC lineages that separates those lineages in different habitat types, which may be only a few kilometers apart, is expected to completely prevent direct gene flow between affected taxa. Theoretically, loss and acquisition of a few genes restricting distribution of a population to a certain habitat type (pH range) could drive speciation in PnecC bacteria. Another barrier to gene flow that is likely to be formidable is the low sequence similarity in homologous genes among lineages (Table 3 and Figure 5). This sequence divergence is expected to strongly limit homologous recombination among lineages (Fraser et al., 2007).

In fact, the ANI between the six compared PnecC genomes (Figure 5b) is far below the threshold of about 95–96% that was suggested for demarcating prokaryotic species (Konstantinidis and Tiedje, 2005; Konstantinidis et al., 2006; Kim et al., 2014). Representation of different species by strains sharing >99% similarity of 16S rRNA genes is not unusual among bacteria (Jaspers and Overmann 2004; Konstantinidis and Tiedje, 2005; Stackebrandt and Ebers, 2006; Kim et al., 2014). Representing distinct undescribed species by all six randomly selected Polynucleobacter strains, as well as the low sequence similarities of marker genes in the majority of the 240 PnecC strains (Figures 5c and e), suggests, however, that the PnecC cluster represents a large cryptic species complex. If such large species richness in a ⩾99% 16S rRNA similarity taxon is usual for pelagic freshwater bacteria is currently unknown. For comparison, marine strains currently classified as Alteromonas macleodii represent a taxon ecologically diversified in at least two ecotypes preferring water layers in different depths (López-López et al., 2005). This taxon is characterized by a 16S rRNA similarity of >98% (López-López et al., 2005), and includes genomically well-separated strains (Ivars-Martinez et al., 2008). In a recent study, Varghese et al. (2015) sorted all genomes contained in the IMG database (Markowitz et al., 2012) in species-like taxa called cliques. These assignments basically used similar but more sophisticated criteria compared with those suggested by Konstantinidis and Tiedje (2005) previously. The genomes of 14 strains contained in IMG and currently classified as A. macleodii were sorted into three cliques, whereas all Polynucleobacter genomes considered in the analyses were sorted into individual cliques (Supplementary Materials Table S3; Varghese et al., 2015). By contrast, the 1243 genomes of strains classified as E. coli were sorted in only three cliques, of which one clique accommodates all but five of those E. coli strains. This major clique (ID 1014) also accommodates several strains currently classified as Shigella sonnei, S. dysenteriae, S. flexneri or S. boydii (Varghese et al., 2015).

The new findings on Polynucleobacter diversity request a taxonomic revision and further investigations aimed at a sound estimation of the taxonomic and ecological diversity within cluster PnecC. Furthermore, these new insights into the diversity

of PnecC bacteria raise the question of whether each of the other four previously characterized *Polynucleobacter* subclusters (Hahn, 2003; Wu and Hahn, 2006) also represents a cryptic species complex.

Implications of the revealed cryptic diversity

The transplantation experiments and previous cultivation-independent investigations of environmental distribution of PnecC lineages (Jezbera *et al.*, 2011) show that the revealed cryptic diversity has not just an arbitrary taxonomic relevance but a strong ecological significance. Independent of the extent of the cryptic diversity among PnecC bacteria, the lack of discrimination of ecologically distinct taxa by 16S rRNA sequences may, for instance, result in misinterpretation of patterns (for example, biogeography) and processes (for example, community assembly, metapopulation dynamics), or, if investigations are exclusively based on operational taxonomic units defined by thresholds of ribosomal sequence similarities, may result in ecological misclassification of taxa (for example, generalists versus specialists). Owing to the frequently high abundance of *Polynucleobacter* bacteria in freshwater systems (Percent *et al.*, 2008; Jezberova *et al.*, 2010; Newton *et al.*, 2011; Besemer *et al.*, 2012) ribosomal data of this taxon are included in many ecological studies (Van der Gucht *et al.*, 2007; Newton *et al.*, 2011; Besemer *et al.*, 2012; Livermore and Jones, 2015) and may influence the results of such studies. The ecological importance of bacterial diversity not resolvable by ribosomal markers has been documented by several studies (Jaspers and Overmann, 2004; Hahn and Pockl, 2005; Sikorski and Nevo, 2007) but is largely ignored in many studies using next-generation sequencing of partial 16S rRNA sequences. Here we revealed that this resolution bias is in abundant freshwater bacteria of high significance.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

References

Auch AF, von Jan M, Klenk HP, Goker M. (2010). Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci* **2**: 117–134.

Besemer K, Peter H, Logue JB, Langenheder S, Lindstrom ES, Tranvik LJ *et al.* (2012). Unraveling assembly of stream biofilm communities. *ISME J* **6**: 1459–1468.

Boscaro V, Felletti M, Vannini C, Ackerman MS, Chain PSG, Malfatti S *et al.* (2013). *Polynucleobacter necessarius*, a model for genome reduction in both free-living and symbiotic bacteria. *Proc Natl Acad Sci USA* **110**: 18590–18595.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K *et al.* (2009). BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421.

Claesson MJ, Jeffery IB, Conde S, Power SE, O'Connor EM, Cusack S *et al.* (2012). Gut microbiota composition correlates with diet and health in the elderly. *Nature* **488**: 178–184.

Cordero OX, Polz MF. (2014). Explaining microbial genomic diversity in light of evolutionary ecology. *Nat Rev Micro* **12**: 263–273.

Fraser C, Hanage WP, Spratt BG. (2007). Recombination and the nature of bacterial speciation. *Science* **315**: 476–480.

Freilich S, Zarecki R, Eilam O, Segal ES, Henry CS, Kupiec M *et al.* (2011). Competitive and cooperative metabolic interactions in bacterial communities. *Nat Commun* **2**: 589.

Garcia SL, Buck M, McMahon KD, Grossart HP, Eiler A, Warnecke F. (2015). Auxotrophy and intra-population complementary in the 'interactome' of a cultivated freshwater model community. *Mol Ecol* **24**: 4449–4459.

Ghai R, Rodriguez-Valera F, McMahon KD, Toyama D, Rinke R, Souza de Oliveira TC *et al.* (2011). Metagenomics of the water column in the pristine upper course of the Amazon River. *Plos One* **6**: e23785.

Grossart HP. (2010). Ecological consequences of bacterioplankton lifestyles: changes in concepts are needed. *Environ Microbiol Rep* **2**: 706–714.

Hahn MW. (2003). Isolation of strains belonging to the cosmopolitan *Polynucleobacter necessarius* cluster from freshwater habitats located in three climatic zones. *Appl Environ Microbiol* **69**: 5248–5254.

Hahn MW, Stadler P, Wu QL, Pockl M. (2004). The filtration-acclimatization method for isolation of an important fraction of the not readily cultivable bacteria. *J Microbiol Methods* **57**: 379–390.

Hahn MW, Pockl M. (2005). Ecotypes of planktonic *Actinobacteria* with identical 16S rRNA genes adapted to thermal niches in temperate, subtropical, and tropical freshwater habitats. *Appl Environ Microbiol* **71**: 766–773.

Hahn MW, Pockl M, Wu QL. (2005). Low intra-specific diversity in a *Polynucleobacter* subcluster population numerically dominating bacterioplankton of a freshwater pond. *Appl Environ Microbiol* **71**: 4539–4547.

Hahn MW, Lang E, Brandt U, Wu QL, Scheuerl T. (2009). Emended description of the genus *Polynucleobacter* and the species *Polynucleobacter necessarius* and proposal of two subspecies, *P. necessarius* subsp. *necessarius* subsp nov and *P. necessarius* subsp *asymbioticus* subsp nov. *Int J Syst Evol Microbiol* **59**: 2002–2009.

Hahn MW, Scheuerl T, Jezberova J, Koll U, Jezbera J, Simek K *et al.* (2012). The passive yet successful way

of planktonic life: Genomic and experimental analysis of the ecology of a free-living *Polynucleobacter* population. *Plos One* **7**: e32772.

Hahn MW, Koll U, Jezberova J, Camacho A. (2015). Global phylogeography of pelagic *Polynucleobacter* bacteria: Restricted geographic distribution of subgroups, isolation by distance, and influence of climate. *Environ Microbiol* **17**: 829–840.

Hao Z, Li L, Liu J, Ren Y, Wang L, Bartlam M *et al.* (2013). Genome sequence of a freshwater low-nucleic-acid-content bacterium, Betaproteobacterium strain CB. *Genome Announcements* **1**: e0013513.

Ivars-Martinez E, Martin-Cuadrado A-B, D'Auria G, Mira A, Ferriera S, Johnson J *et al.* (2008). Comparative genomics of two ecotypes of the marine planktonic copiotroph Alteromonas macleodii suggests alternative lifestyles associated with different kinds of particulate organic matter. *ISME J* **2**: 1194–1212.

Jaspers E, Overmann J. (2004). Ecological significance of microdiversity: Identical 16S rRNA gene sequences can be found in bacteria with highly divergent genomes and ecophysiologies. *Appl Environ Microbiol* **70**: 4831–4839.

Jezbera J, Jezberova J, Brandt U, Hahn MW. (2011). Ubiquity of *Polynucleobacter necessarius* subspecies *asymbioticus* results from ecological diversification. *Environ Microbiol* **13**: 922–931.

Jezberova J, Jezbera J, Brandt U, Lindstrom ES, Langenheder S, Hahn MW. (2010). Ubiquity of *Polynucleobacter necessarius* ssp. *asymbioticus* in lentic freshwater habitats of a heterogenous 2000 km² area. *Environ Microbiol* **12**: 658–669.

Kim M, Oh H-S, Park S-C, Chun J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* **64**: 346–351.

Konstantinidis K, Tiedje J. (2005). Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci USA* **102**: 2567–2572.

Konstantinidis KT, Ramette A, Tiedje JM. (2006). The bacterial species definition in the genomic era. *Philos Trans R Soc B Biol Sci* **361**: 1929–1940.

Livermore JA, Jones SE. (2015). Local-global overlap in diversity informs mechanisms of bacterial biogeography. *ISME J.* **9**: 2413–2422.

López-López A, Bartual SG, Stal L, Onyshchenko O, Rodríguez-Valera F. (2005). Genetic analysis of housekeeping genes reveals a deep-sea ecotype of Alteromonas macleodii in the Mediterranean Sea. *Environ Microbiol* **7**: 649–659.

Markowitz VM, Chen IMA, Palaniappan K, Chu K, Szeto E, Grechkin Y *et al.* (2012). IMG: the integrated microbial genomes database and comparative analysis system. *Nucleic Acids Res* **40**: D115–D122.

Martinez-Garcia M, Swan BK, Poulton NJ, Gomez ML, Masland D, Sieracki ME *et al.* (2012). High-throughput single-cell sequencing identifies photoheterotrophs and chemoautotrophs in freshwater bacterioplankton. *ISME J* **6**: 113–123.

Meincke L, Copeland A, Lapidus A, Lucas S, Berry KW, Del Rio TG *et al.* (2012). Complete genome sequence of *Polynucleobacter necessarius* subsp. *asymbioticus* type strain (QLW-P1DMWA-1^T). *Stand Genomic Sci* **6**: 74–83.

Morris JJ, Kirkegaard R, Szul MJ, Johnson ZI, Zinser ER. (2008). Facilitation of robust growth of *Prochlorococcus* colonies and dilute liquid cultures by "helper" heterotrophic bacteria. *Appl Environ Microbiol* **74**: 4530–4534.

Newton RJ, Jones SE, Eiler A, McMahon KD, Bertilsson S. (2011). A guide to the natural history of freshwater lake bacteria. *Microbiol Mol Biol Rev* **75**: 14–49.

Percent SF, Frischer ME, Vescio PA, Duffy EB, Milano V, McLellan M *et al.* (2008). Bacterial community structure of acid-impacted lakes: What controls diversity? *Appl Environ Microbiol* **74**: 1856–1868.

Pinto AJ, Xi C, Raskin L. (2012). Bacterial community structure in the drinking water microbiome is governed by filtration processes. *Environ Sci Technol* **46**: 8851–8859.

Richter M, Rosselló-Móra R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* **106**: 19126–19131.

Salcher MM. (2014). Same same but different: ecological niche partitioning of planktonic freshwater prokaryotes. *J Limnol* **73**: 74–87.

Schauer M, Hahn MW. (2005). Diversity and phylogenetic affiliations of morphologically conspicuous large filamentous bacteria occurring in the pelagic zones of a broad spectrum of freshwater habitats. *Appl Environ Microbiol* **71**: 1931–1940.

Schwalbach MS, Tripp HJ, Steindler L, Smith DP, Giovannoni SJ. (2010). The presence of the glycolysis operon in SAR11 genomes is positively correlated with ocean productivity. *Environ Microbiol* **12**: 490–500.

Seth EC, Taga ME. (2014). Nutrient cross-feeding in the microbial world. *Front Microbiol* **5**: 350.

Sikorski J, Nevo E. (2007). Patterns of thermal adaptation of *Bacillus* simplex to the microclimatically contrasting slopes of 'Evolution Canyons' I and II, Israel. *Environ Microbiol* **9**: 716–726.

Stackebrandt E, Ebers J. (2006). Taxonomic parameters revisited: tarnished gold standards. *Microbiol Today* **33**: 152–155.

Stumm W. (2007). Chemical processes regulating the composition of lake waters. *The Lakes Handbook*, Volume 1. Blackwell Science Ltd, Malden, CA, pp 79–106.

Van der Gucht K, Cottenie K, Muylaert K, Vloemans N, Cousin S, Declerck S *et al.* (2007). The power of species sorting: Local factors drive bacterial community composition over a wide range of spatial scales. *Proc Natl Acad Sci USA* **104**: 20404–20409.

Vannini C, Poeckl M, Petroni G, Wu QL, Lang E, Stackebrandt E *et al.* (2007). Endosymbiosis in statu nascendi: close phylogenetic relationship between obligately endosymbiotic and obligately free-living *Polynucleobacter* strains (*Betaproteobacteria*). *Environ Microbiol* **9**: 347–359.

Varghese NJ, Mukherjee S, Ivanova N, Konstantinidis KT, Mavrommatis K, Kyrpides NC *et al.* (2015). Microbial species delineation using whole genome sequences. *Nucleic Acids Res* **43**: 6761–6771.

Watanabe K, Komatsu N, Ishii Y, Negishi M. (2009). Effective isolation of bacterioplankton genus *Polynucleobacter* from freshwater environments grown on photochemically degraded dissolved organic matter. *FEMS Microbiol Ecol* **67**: 57–68.

Wetzel RG. (2001). 14 - Iron, sulfur, and silica cycles. In: Wetzel RG (ed). *Limnology*, 3rd edn. Academic Press: San Diego, pp 289–330.

Wu QL, Hahn MW. (2006). Differences in structure and dynamics of *Polynucleobacter* communities in a temperate and a subtropical lake, revealed at three phylogenetic levels. *FEMS Microbiol Ecol* **57**: 67–79.

Zane HK, Butler A. (2013). 3.01 - Fe acquisition. In: Poeppelmeier JR (ed). *Comprehensive Inorganic Chemistry II*, 2nd edn. Elsevier: Amsterdam, pp 1–20.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)