

# The Passive Yet Successful Way of Planktonic Life: Genomic and Experimental Analysis of the Ecology of a Free-Living *Polynucleobacter* Population

Martin W. Hahn<sup>1\*</sup>, Thomas Scheuerl<sup>1</sup>, Jitka Jezberová<sup>1,2</sup>, Ulrike Koll<sup>1</sup>, Jan Jezbera<sup>1,2</sup>, Karel Šimek<sup>2</sup>, Claudia Vannini<sup>3</sup>, Giulio Petroni<sup>3</sup>, Qinglong L. Wu<sup>1,4</sup>

**1** Institute for Limnology, Austrian Academy of Sciences, Mondsee, Austria, **2** Institute of Hydrobiology, Biology Centre of the AS CR, v.v.i., České Budějovice, Czech Republic, **3** Biology Department, Protistology-Zoology Unit, University of Pisa, Pisa, Italy, **4** State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography & Limnology, Chinese Academy of Sciences, Nanjing, People's Republic of China

## Abstract

**Background:** The bacterial taxon *Polynucleobacter necessarius* subspecies *asymbioticus* represents a group of planktonic freshwater bacteria with cosmopolitan and ubiquitous distribution in standing freshwater habitats. These bacteria comprise <1% to 70% (on average about 20%) of total bacterioplankton cells in various freshwater habitats. The ubiquity of this taxon was recently explained by intra-taxon ecological diversification, i.e. specialization of lineages to specific environmental conditions; however, details on specific adaptations are not known. Here we investigated by means of genomic and experimental analyses the ecological adaptation of a persistent population dwelling in a small acidic pond.

**Findings:** The investigated population (F10 lineage) contributed on average 11% to total bacterioplankton in the pond during the vegetation periods (ice-free period, usually May to November). Only a low degree of genetic diversification of the population could be revealed. These bacteria are characterized by a small genome size (2.1 Mb), a relatively small number of genes involved in transduction of environmental signals, and the lack of motility and quorum sensing. Experiments indicated that these bacteria live as chemoorganotrophs by mainly utilizing low-molecular-weight substrates derived from photooxidation of humic substances.

**Conclusions:** Evolutionary genome streamlining resulted in a highly passive lifestyle so far only known among free-living bacteria from pelagic marine taxa dwelling in environmentally stable nutrient-poor off-shore systems. Surprisingly, such a lifestyle is also successful in a highly dynamic and nutrient-rich environment such as the water column of the investigated pond, which was undergoing complete mixis and pronounced stratification in diurnal cycles. Obviously, metabolic and ecological versatility is not a prerequisite for long-lasting establishment of abundant bacterial populations under highly dynamic environmental conditions. Caution should be exercised when generalizing the obtained insights into the ecology and adaptation of the investigated lineage to other *Polynucleobacter* lineages.

**Citation:** Hahn MW, Scheuerl T, Jezberová J, Koll U, Jezbera J, 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(3): e32772. doi:10.1371/journal.pone.0032772

**Editor:** Jonathan H. Badger, J. Craig Venter Institute, United States of America

**Received:** October 20, 2011; **Accepted:** January 30, 2012; **Published:** March 20, 2012

**Copyright:** © 2012 Hahn et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The study was supported by the JGI Community Sequencing Program (CSP) and the Austrian Science Fund (FWF, grants P19853 and I482-B09). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: martin.hahn@oeaw.ac.at

## Introduction

The current picture on diversity and function of freshwater bacteria is highly fragmentary. The most abundant taxa of freshwater bacterioplankton could be identified and, at least for some of those taxa, the respective geographic distributions and habitat preferences could tentatively be revealed (e.g. [1,2]). However, insights into the respective ecological functions have been/were obtained only for a few taxa (e.g. [3]), and little is known about the population structure and microdiversity of abundant taxa (e.g. [4,5]). Recent investigations have revealed quite pronounced ecological differences between seemingly closely related strains. Such strains are assigned to the same species-like freshwater taxon because they completely lack differences or differ only marginally in phylogenetic markers (i.e., ribosomal sequenc-

es), yet exhibit significant differences in habitat preferences [6], lifestyle [7] or phenotypic traits such as substrate utilization [8] and thermal adaptation [9]. These observations challenge the ecological uniformity of the lowest rank taxa (i.e. tribes) currently considered by freshwater microbiologists [2]. Conclusions from inter-habitat comparisons in particular could be biased by ecological differences between populations assigned to the same taxon. Detailed inter-habitat comparisons on diversity [6] and function of populations assigned to the same taxon are needed in order to test the comparability and generalizability of findings obtained at the currently lowest level of taxonomic rank (i.e., tribe [2]). Here we prepared the ground for extending such detailed comparisons by performing an in-depth investigation into the ecology and diversity of a single well-characterized population of *Polynucleobacter necessarius* subspecies (ssp.) *asymbioticus* [10]. The

extensive knowledge of ecological function, population structure and genomic features established in this study for a particular population will enable future inter-habitat comparisons with other *Polynucleobacter* populations.

The subspecies *P.n. sp. asymbioticus* (*Burkholderiaceae*, *Betaproteobacteria*) [10], aka PnecC bacteria [11] or PnecC tribe [2], represents a group of abundant bacteria [12] with cosmopolitan [2,10,11,13,14,15,16] and ubiquitous distribution in standing freshwater systems [12]. This subspecies inhabits an ecologically broad range of habitat types including small puddles, ponds and lakes with acidic, circum-neutral or alkaline conditions [1,11,12,13,17]. By contrast, to the best of our knowledge, detection of *Polynucleobacter* bacteria in marine off-shore or in soil systems has never been reported. Extrapolation of *P.n. sp. asymbioticus* abundance data obtained in a systematic but geographically limited survey of more than 100 standing freshwater systems estimated an average global contribution of 20% to bacterioplankton cells in freshwater systems [12], but pronounced inter-habitat differences in relative abundance of these bacteria were observed. Contributions of up to 70% to bacterioplankton in dystrophic ponds were observed [9,12], while in other systems cell numbers of *P.n. sp. asymbioticus* were close to the detection limit [12,18]. Both the ubiquitous distribution across different types of freshwater systems and the high average relative abundance are remarkable for a phylogenetically narrow group characterized by >99% 16S rRNA sequence similarity. However, a recent investigation employing methods further resolving the subspecies suggested that the ubiquity of the taxon resulted from ecological diversification within the taxon into lineages differing in ecological preferences and distribution across freshwater habitats (ubiquity by diversification hypothesis [6]).

It was believed that strains affiliated with subspecies *P.n. sp. asymbioticus* represent chemoorganotrophs with a free-living, planktonic lifestyle [2,10,11,15]. The latter was confirmed by fluorescence *in situ* hybridization (FISH) with *Polynucleobacter*-specific probes [4,12,17]. However, recently published findings suggest that at least some strains of the subspecies possess in addition to their chemorganotrophic metabolism a potential for phototrophy [19]. The second subspecies of *P. necessarius*, i.e. *P.n. sp. necessarius*, harbors exclusively obligate symbionts of ciliates [7,10], which implies fundamental differences in ecology and function between the two subspecies. The majority of *Polynucleobacter* strains possess small cell sizes and are classified as ultramicrobacteria (<0.1  $\mu\text{m}^3$  cell volume) [11,20]. Cells of such strains can pass through even 0.2  $\mu\text{m}$  pore size filters [21]. Currently, in practice we lack detailed knowledge of ecological adaptation of bacteria belonging to the subspecies *P.n. sp. asymbioticus* to their life in the water columns of different freshwater habitats and of lineage-specific differences in adaptation resulting in niche separation [6,10].

The major aims of this study were (i) to reveal the ecological function of a well-characterized, persistent *P.n. sp. asymbioticus* population, (ii) to obtain deeper insights in the lifestyle and population structure of the investigated bacteria, and (iii) to establish a set of reference data for future comprehensive comparisons with other populations of the subspecies. Our investigation combined a broad array of approaches: monitoring a natural population, genome analysis of one strain representing this population, phenotypic characterizations of the genome-sequenced strain, and experimental investigations with cultured and uncultured representatives of the population.

We focused on the so-called G1 (genotype 1) population [4], which represents a stable and abundant *P.n. sp. asymbioticus* population in a small shallow pond (Pond-1). The G1 population

forms, together with related strains known from other habitats, the so-called F10 lineage [6]. For the sake of clarity, we will use throughout this paper the more general term “F10 lineage population” for the investigated population. An overview of the current taxonomy of *Polynucleobacter* bacteria and a list of used abbreviations are provided as Table S1 and S2, respectively. The investigated F10 lineage population has been monitored from 2003 and is currently represented by a total of twelve cultivated strains obtained from Pond-1 over a period of four years. One of these strains was genome-sequenced by the DOE Joint Genome Institute, and an ecological interpretation of its genome is presented here; the description of sequencing and annotation, as well as a basic characterization of the genome, has been presented elsewhere [22].

## Results

The results section is organized into three parts. The first focuses on the environmental conditions in the pond persistently inhabited by the investigated population (F10 lineage population), as well as on the microdiversity and population structure of this population. The second part presents an ecological interpretation of the genome annotation of one strain representing the F10 lineage population as well as empirical characterizations and validations of metabolic traits of the strain. Finally, the third part presents experiments intended to further reveal and confirm the ecological function of the investigated population.

### Habitat characteristics, persistence and population structure

#### Annual and diurnal dynamics of ecological conditions in

**Pond-1.** Pond-1 (Fig. 1) is a shallow (maximum depth of 1.2 m) small pond (surface area 200 m<sup>2</sup>) located (47°44'23.42"N, 13°18'6.07"E) at an altitude of 1300 m in the Austrian Alps near the city of Salzburg [4]. The habitat is seasonally ice-covered for periods of up to six months, resulting in oxygen depletion in the small water body remaining beneath the thick ice layer. The pond lacks any surface water inlet and is mainly fed by water percolating from the slopes surrounding the pond on three sides as well as by precipitation. The pond water is characterized by low ion concentrations (average conductivity 10.5  $\mu\text{S cm}^{-1}$ ), acidic pH (range 4.3–6.0), and dissolved organic carbon (DOC) concentrations in the range of 3–9 mg L<sup>-1</sup>. The DOC is mainly composed of humic substances (HS). The environmental conditions in the pond show pronounced annual (Fig. 2) and even diurnal (Fig. 3) variations. Over an eight-year sampling period, water temperatures ranging from 0.1 (below ice cover) to 21.3°C and oxygen concentrations ranging from 0 (below ice cover) to almost 160% saturation were observed. During the vegetation period (ice-free period, usually May to November) the entire water column of the pond is oxic. Monitoring of environmental parameters over a 23-hour period in August 2010 revealed pronounced stratification of the water column with temperature differences of almost 5°C across a vertical distance of only 45 cm. During the night the upper water layers cooled, resulting in complete mixing of the water column by the end of the night (Fig. 3). The measured oxygen concentrations and conductivity indicated that water layers close to the bottom of the pond were influenced by the inflow of percolating water, while conditions in the upper water layers were mainly influenced by solar energy input.

In general, this habitat can be characterized as a small dystrophic polymictic pond undergoing rapid diurnal and pronounced annual changes in physical and chemical parameters.



**Figure 1. Shallow dystrophic Pond-1.** This pond is persistently inhabited by the investigated *Polynucleobacter* lineage (F10 lineage). The images show the pond in autumn (left) and winter (right), when the pond is hidden below a thick ice and snow cover. The latter situation lasts for up to six months.

doi:10.1371/journal.pone.0032772.g001

This dynamic system shares many characteristics with shallow ponds and lakes in the taiga and tundra (e.g., [13]).

**Habitat persistence and genetic diversity of the F10 lineage population in Pond-1.** The genome-sequenced strain QLW-P1DMWA-1 and two other strains with identical 16S rRNA genes, 16S–23S ITS sequences and identical genetic fingerprints were isolated from Pond-1 in 2003 [4]. These strains were initially characterized as genotype G1 [4], which belong to the so-called F10 lineage [6]. Sequencing of two phylogenetic markers (16S–23S ITS, glutamine synthetase gene (*glnA*)) of a total of 176 *P. necessarius* ssp. *asymbioticus* strains indicated that the F10 lineage is genetically distinct from all other lineages within the subspecies (Fig. 4; [4,6]; Hahn et al., unpublished data). Nine additional strains affiliated with the F10 lineage were isolated from Pond-1 in 2004, 2005, and 2007. All twelve F10 lineage strains isolated from the pond shared identical 16S–23S ITS and *glnA* sequences. Multilocus sequence analysis of 12 loci, as well as genetic fingerprinting by three independent methods, indicated only low genetic diversification among the twelve strains (Fig. 5). Only 15 variable positions (0.17% of all positions) were detected in the set of eleven partially sequenced genes and the completely sequenced 16S–23S ITS sequence, which represented for each strain a total sequence length of 8935 bp (0.4% of genome size). These few polymorphic sites represented single nucleotide polymorphisms located in only two genes (*icd2* and *Pnuc\_1095*) but widely spread over the sequenced parts (878 bp and 1676 bp) of these genes. Only four of the 15 polymorphic sites represented non-synonymous substitutions (Fig. 5). In the case of both genes, only two alleles were found among the twelve strains. Additionally, sequencing of gene *Pnuc\_1240* in all twelve strains revealed the presence of an insertion sequence element (IS) in strain P1-KOL5. Finally, genetic fingerprinting of the strains with three different methods revealed a difference between two strains (P1-KOL8 and P1-05-86) not distinguishable by the sequence data (Fig. 5). Thus, six different genotypes could be identified among the twelve strains, but the observed genetic differences are very small. No genetic differences could be revealed for seven strains (including the genome-sequenced strain) isolated in 2003, 2005, and 2007.

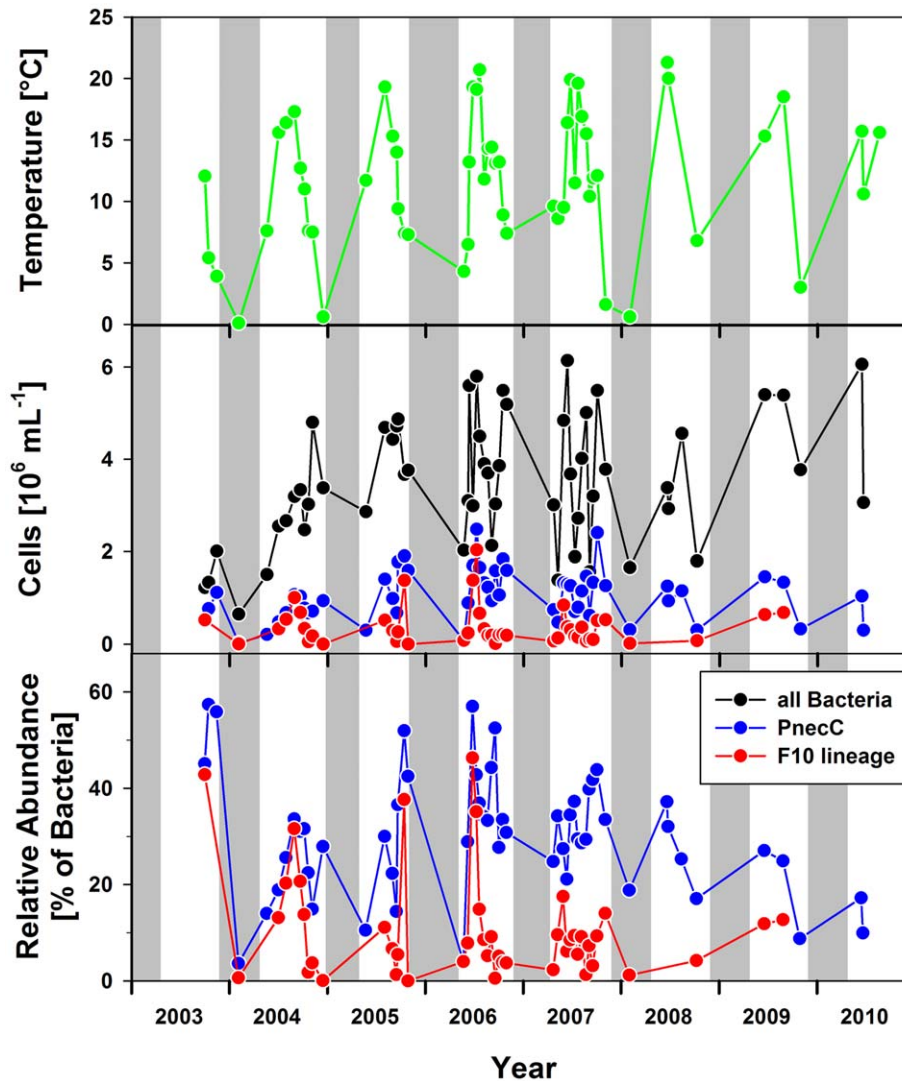
Besides repeated isolation of strains, application of cultivation-independent methods demonstrated the persistence of the F10

lineage in Pond-1 over the entire investigation period, i.e. 2003 to 2009 (Fig. 2). Reverse line blot hybridization (RLBH) with lineage-specific probes [6] and lineage-specific qPCR assays (Fig. 2) detected the lineage in the majority of investigated samples from Pond-1. The F10-specific RLBH assay confirmed the presence of this lineage in 41 out of 42 seasonal samples taken over the period from October 2003 to February 2008. The F10-specific qPCR assay was applied in total to 45 samples collected over a period of seven years and detected the lineage in all but one sample. *P.n. ssp. asymbioticus* and the F10 lineage were even detected in the ice-covered water body of the pond at mid-winter (two samplings during winter season). During the vegetation period, the F10 lineage contributed on average to 11.3% of total bacterial numbers and 35.5% of *P. necessarius* cells in the water column of the pond. However, both the relative contribution of the F10 lineage to the bacterial community and the absolute numbers of F10 lineage cells were highly variable over time but did not follow a clear seasonal pattern. In some years the F10 lineage population peaked during autumn but spring and summer peaks were also observed (Fig. 2). Thus, the investigated population did not show a clear recurrent seasonal pattern as was observed in Lake Mondsee for another *Polynucleobacter* population (or assemblage) representing another species or species assemblage (*P. acidiphobus* and/or *P. difficilis*) [23].

The pond does not receive water from tributaries but rather water that percolates from nearby slopes. Analysis of DNA extracted from water samples from a natural spring located on this slope at a distance of 350 meters by qPCR detected no *P.n. ssp. asymbioticus* or the F10 lineage. Consequently, the F10 lineage population can be viewed as a resident population persisting with seasonally varying population size in Pond-1 over periods of at least several years.

**Influence of protistan predation on the *P.n. ssp. asymbioticus* population of Pond-1.** Grazing experiments with fluorescently labeled bacteria (FLB) were conducted. Lab-grown cultures of strain QLW-P1DMWA-1 (*Polynucleobacter*-FLB) and samples of natural bacterioplankton (bacterioplankton-FLB) were fluorescently stained, and uptake of both FLB types by protistan predators present in Pond-1 was microscopically determined. At the date of the experiment, the potentially





**Figure 2. Annual cycles of water temperature and bacterial parameters of Pond-1.** Water temperatures (measured in depths of about 30 cm; upper graph), abundance of the entire *P.n. ssp. asymbioticus* assemblage (PnecC) determined by FISH and the F10 lineage based on qPCR measurements (middle graph), and relative abundances of these two taxa (bottom graph). The gray areas indicate periods of ice-cover. Note that some samples were not investigated by qPCR.  
doi:10.1371/journal.pone.0032772.g002

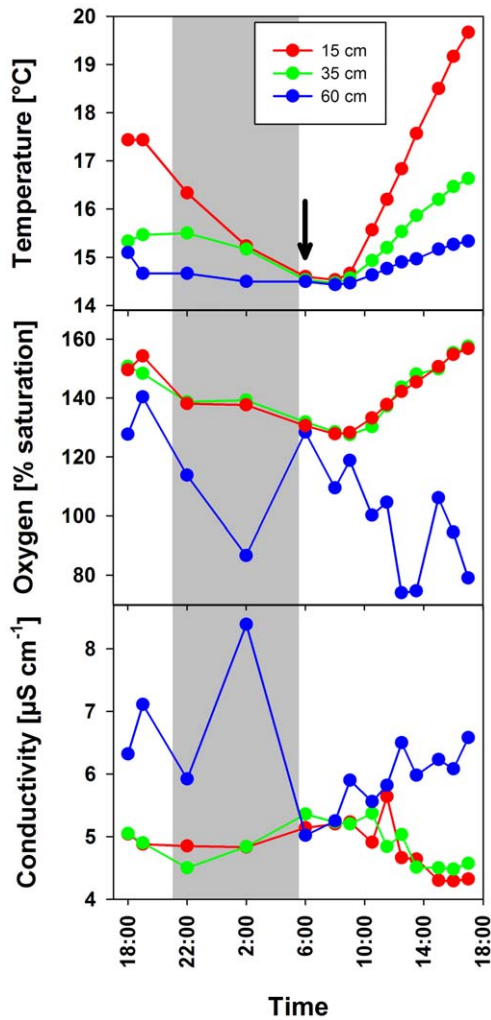
bacterivorous protistan community of Pond-1 consisted mainly of heterotrophic nanoflagellates (HNF,  $3.7 \times 10^2$  cells  $\text{mL}^{-1}$ ), small *Cyclidium*-like scuticociliates ( $3.1 \times 10^2$  cells  $\text{mL}^{-1}$ ), and *Dinobryon* spp. ( $2.5 \times 10^2$  cells  $\text{mL}^{-1}$ ). All three groups of grazers predated on *Polynucleobacter*-FLB and bacterioplankton-FLB at quite similar grazing rates. The strongest grazing impact was observed for the highly abundant *Cyclidium* sp., which consumed on average 205 bacteria  $\text{cell}^{-1} \text{h}^{-1}$ , while predation rates by HNF and *Dinobryon* sp. were 8.0 and 5.1 bacteria  $\text{cell}^{-1} \text{h}^{-1}$ , respectively. In total, contributions of ciliates, HNF and *Dinobryon* sp. were 93.7%, 4.4%, and 1.9% to the total bacterivory in the water column of Pond-1, respectively. Predation by these protists removed 33% of the bacterial standing stock per day, which indicated a very fast population turnover time of the target bacteria considering a rather low *in situ* temperature of 9°C. Furthermore, FISH analysis conducted directly inside flagellate food vacuoles clearly confirmed the presence of the targeted *P.n. ssp. asymbioticus* cells among food items ingested in food vacuoles (data not shown).

#### Growth of the *P.n. ssp. asymbioticus* community in Pond-1.

Three processes – predation, viral lysis, and export via outflow – are expected to represent the major loss factors of the *Polynucleobacter* community in Pond-1. These losses have to be compensated by growth since an import of *Polynucleobacter* cells from the terrestrial environment is highly unlikely. Only loss by protistan predation has been quantified so far. Based on the estimated grazing mortality of *Polynucleobacter* bacteria of 33% of cells per day, a minimum growth rate of these bacteria in the same order of magnitude could be expected.

#### Phenotypic and genomic traits of strain QLW-P1DMWA-1

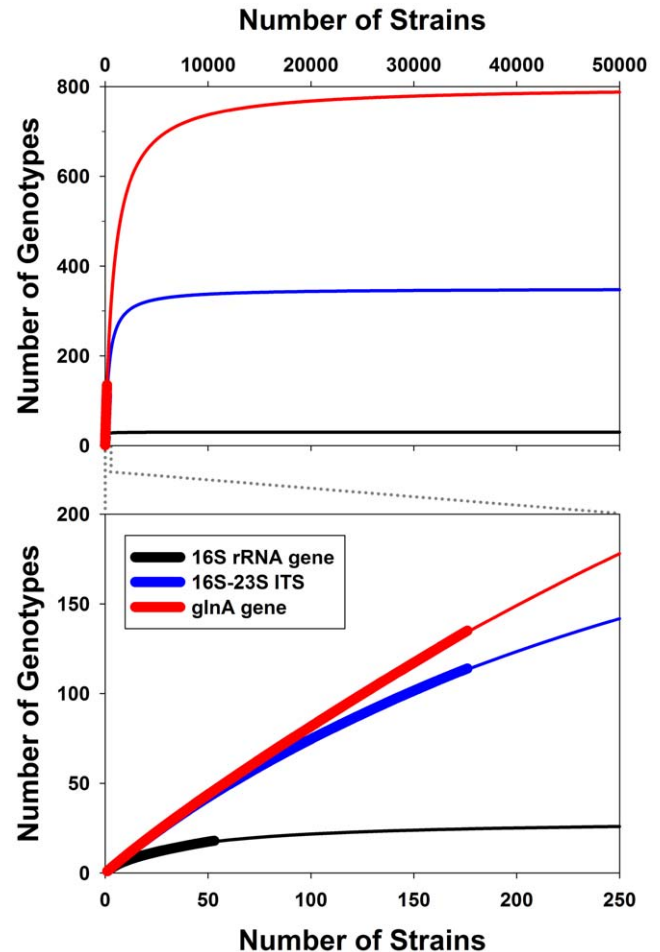
**Substrate assimilation tests.** In total, 93 substances were empirically tested for assimilation by pure cultures of strain QLW-P1DMWA-1 (Table S3). Sixty percent of the substances did not support growth of the strain, and among the substances yielding growth only assimilation of ten (11% of total) resulted in efficiencies >5% (Fig. 6). Nine out of these ten substances were



**Figure 3. Diurnal cycles of environmental parameters in three different depths of Pond-1.** Water temperature, oxygen concentration, and conductivity were measured in August 2010. The days before the recording of the diurnal cycle were characterized by cooler dull weather, while the weather conditions on the first day of measurement improved and the second day was characterized by sunny and calm weather. The arrow indicates the point in time of vertical mixing. Note that maximum temperature and oxygen values, as well as minimal conductivity values measured during the diurnal cycle were recorded in late afternoon. Since the routine sampling of the pond (Fig. 2) always took place during late-morning hours, the widths of annual variation of environmental parameters were most likely underestimated. doi:10.1371/journal.pone.0032772.g003

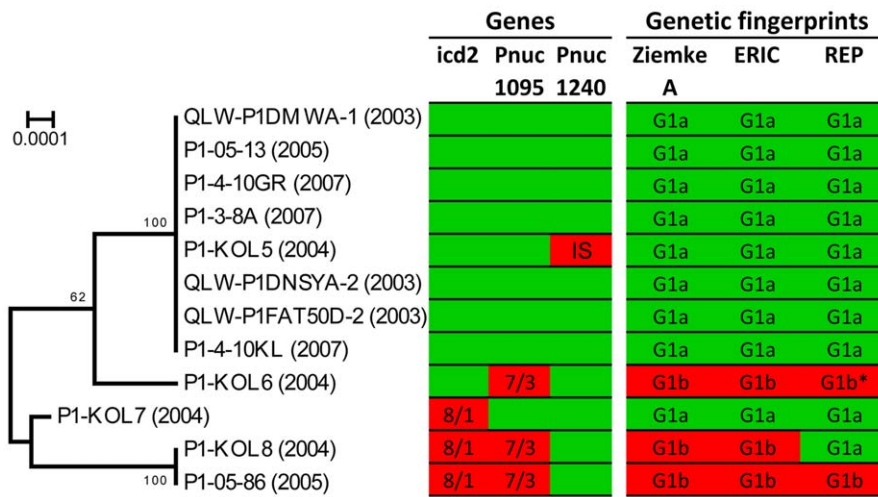
carboxylic acids with two to five carbon atoms. Five of these compounds represent intermediates of the central metabolism, especially of the tricarboxylic acid cycle (TCA). Furthermore, two amino acids and one typical fermentation product were included. In addition, the efficiency for growth on photooxidation products of humic substances (HS) was estimated to be 9%. However, no significant growth was observed for untreated HS (see below). Interestingly, no growth or only weak growth was observed for all tested carbohydrates, as well as for most of the amino acids.

**Nitrogen and sulfur metabolism.** The genome of QLW-P1DMWA-1 encoded a conserved cluster of 13 genes (Pnuc\_1190 - Pnuc\_1202) also found in *Ralstonia*, *Cupriavidus* and *Burkholderia* strains. This cluster encoded a putative ABC transporter for urea, a urease consisting of three subunits as well as some accessory



**Figure 4. Diversity of *P. necessarius* ssp. *asymbioticus* as represented by a culture collection of 176 strains.** The graphs show collector's curves depicting genotype numbers for the global non-redundant culture collection of *P.n. ssp. asymbioticus* strains (broad lines) after randomization of sampling order, as well as further extrapolations for continued collection of strains (thin lines). This non-redundant culture collection excludes all but one strain of the same genotype (ITS and *glnA* marker considered) isolated from the same habitat. Therefore, the collection includes only one (i.e. QLW-P1DMWA-1) of the twelve F10 lineage strains isolated from Pond-1 (compare Fig. 5). The collection consists currently of 176 strains isolated from various freshwater systems located all over the world. The sequencing of 16S rRNA genes of newly isolated strains was stopped after 53 strains. The length of the sequenced stretches of the markers 16S rRNA gene, 16S–23S ITS, and *glnA* were 1500 bp, 500 bp, and 600 bp, respectively. Note that genotype or species-richness predictions based exclusively on observed data usually underestimated the true richness. Randomization of sampling order was performed using the software EstimateS [75]. doi:10.1371/journal.pone.0032772.g004

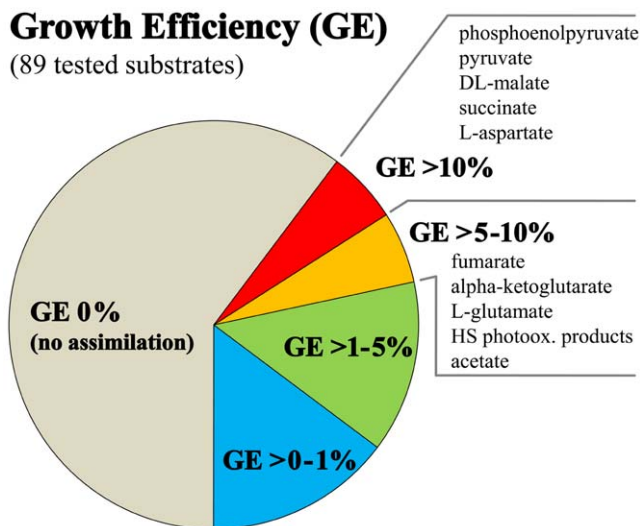
genes related to urease activity. Growth with urea as sole nitrogen source could be observed in experiments with acetate as sole energy and carbon source. A region of 16 genes (Pnuc\_0987 - Pnuc\_1003) encoded genes predominantly involved in nitrogen assimilation. This includes putative transporters for amino acids and nitrate, a cyanate transporter and a cyanate hydratase, as well as reductases for assimilation of nitrate and nitrite. This region consists of conserved modules of three or four genes, which are also found in other *Betaproteobacteria* or other *Proteobacteria*. Growth experiments revealed that strain QLW-P1DMWA-1 is able to grow with combinations of nitrate and sulfate, nitrate and



**Figure 5. Microdiversity of the F10 lineage population of Pond-1.** The genetic diversity of the twelve F10 lineage strains isolated from Pond-1 over a period of four years (2003–2007) is illustrated. **Left panel.** NJ tree based on concatenated nucleotide sequences of twelve loci (Table S4) with a total length of 9031 alignment positions. **Middle panel.** Only three out of 13 sequenced loci showed sequence polymorphism. The *icd2* and *Pnuc\_1095* genes were present with two alleles differing in the sequenced parts in total in eight and seven positions (polymorphic positions), respectively. Numbers depict the total numbers of polymorphic sites (first number) and the number of non-synonymous sites (second number) among the polymorphic sites. The locus *Pnuc\_1240* of strain P1-Kol5 differed from all other sequenced strains in the presence of an insertion element (IS, insertion sequence). Sequences of *Pnuc\_1240* were not considered in the calculation of the presented phylogenetic tree. **Right panel.** Results from genetic fingerprinting with the independent methods RAPD (Ziemke A), ERIC, and REP-PCR. Each method resulted in different fingerprints, but each method revealed basically only two types of fingerprints (G1a and G1b). The REP-PCR fingerprint G1b\* differed only weakly from G1b. Colors green and red indicate that the respective strains share or do not share, respectively, a particular trait with the genome-sequenced strain QLW-P1DMWA-1. doi:10.1371/journal.pone.0032772.g005

thiosulfate, urea and sulfate, urea and thiosulfate as sole nitrogen and sulfur sources, respectively. *Pnuc\_1350* encoding a putative rhodanase could be responsible for the capability to use thiosulfate as sole sulfur source. Furthermore, a region of the genome containing genes *Pnuc\_1476* to *Pnuc\_1494* encodes several

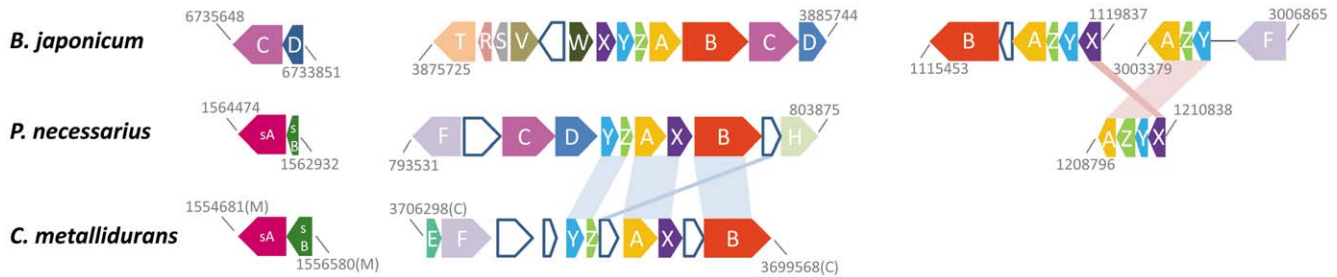
functions involved in assimilation of inorganic sulfur. The presence of SOX genes (Fig. 7) suggested that the strain is able to gain energy from oxidization of inorganic sulfur species such as H<sub>2</sub>S. However, this trait has not yet been experimentally confirmed.



**Figure 6. Growth efficiencies of strain QLW-P1DMWA-1 on various tested substances.** Growth efficiencies (GE) of the genome-sequenced strain were determined for 89 substances by batch culture experiments. GE describes the percentage of carbon contained in the respective tested substance transferred by the strain to carbon in its biomass. Only ten substrates resulted in GE>5%. These substances are listed with decreasing efficiencies (top to bottom). doi:10.1371/journal.pone.0032772.g006

**Genome characteristics and genome evolution.** A phylogenetic analysis based on multilocus sequences of protein-encoding housekeeping genes (Fig. 8) confirmed previous phylogenetic analyses solely based on 16S rRNA gene sequences [10,11]. Accordingly, the genus *Polynucleobacter* is affiliated with the family *Burkholderiaceae*, and the closest related described genera are *Ralstonia* and *Cupriavidus*. The genus *Polynucleobacter* is currently the sole taxon within the family *Burkholderiaceae*, which contains bacteria with a planktonic lifestyle. All related taxa either represent free-living soil bacteria or bacteria associated with plants or animals. The 2.15 Mb genome of strain QLW-P1DMWA-1 consists of a single replicon (Table 1) and is significantly smaller than the genomes of other currently sequenced strains belonging to the *Burkholderiaceae* (Fig. 9). Genome sequences of 78 strains affiliated to this family are currently available in the Integrated Microbial Genomes (IMG) system [24]. The size range (excluding two *Polynucleobacter* genomes) of these genomes is 3.3 to 9.7 Mb, and the average size (without *Polynucleobacter*) is 6.9 Mb. Furthermore, the genome of the strain differs remarkably in several parameters including, for example, G+C content of DNA (Fig. 9). Interestingly, the genome encoded unusually high percentages of genes that could be classified in Enzyme Commission (EC) number categories or assigned to one of the Kyoto Encyclopedia of Genes and Genomes (KEGG) categories (Fig. 9). The genome was characterized by the lack of many genes present in related organisms (see Text S1). The ecologically most interesting traits lacking were motility and photosynthesis (compare reference [19]). Importantly, the genome of QLW-P1DMWA-1, for a free-living strain, encoded an unusually low absolute and relative number of signal transduction genes (Figs. 8, 9, 10). This is a feature





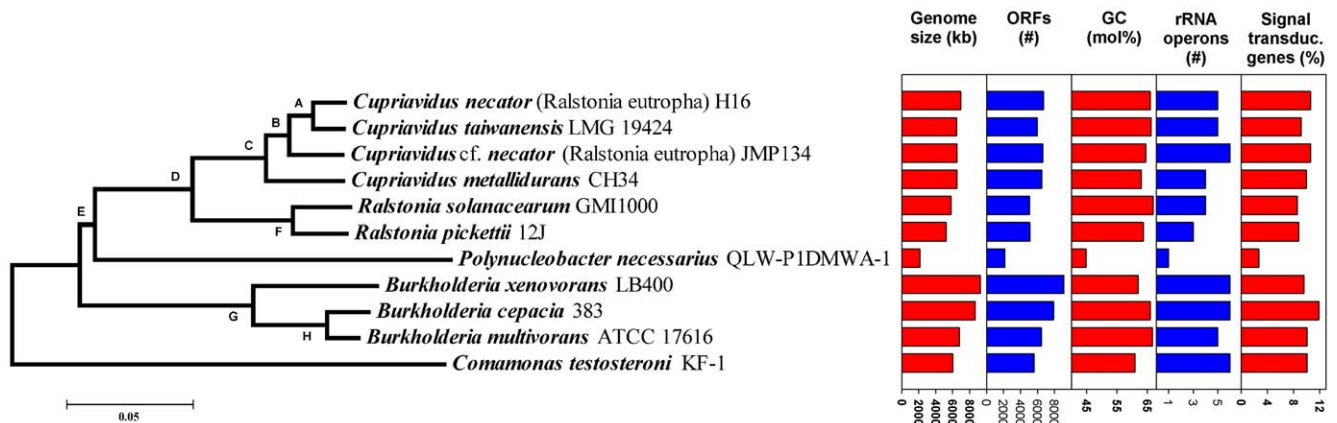
**Figure 7. Comparison of SOX gene clusters.** The SOX gene clusters of strain QLW-P1DMWA-1 (*Burkholderiaceae*, *Betaproteobacteria*), *Bradyrhizobium japonicum* (*Alphaproteobacteria*) and *Cupriavidus metallidurans* (*Burkholderiaceae*, *Betaproteobacteria*) are compared. The three clusters of the *Polynucleobacter* strain are represented by the genes Pnuc\_1486, Pnuc\_1487, and Pnuc\_0799 - Pnuc\_0809, and Pnuc\_1154 - Pnuc\_1157. Start- and end-positions of the clusters in the respective genome sequences are depicted. Shaded areas indicate if *Polynucleobacter* genes are closer related to *Cupriavidus* or *Bradyrhizobium* genes. doi:10.1371/journal.pone.0032772.g007

shared with probably all marine planktonic bacteria with streamlined genomes, e.g. SAR11 bacteria (including *Candidatus Pelagibacter ubique*) and *Prochlorococcus* spp. (Fig. 10B). The genome annotation revealed no reliable indications of an adaptation to a host-associated lifestyle.

For 278 (13.3%) of the protein-encoding ORFs of the genome, analysis employing the IMG system indicated a putative origin by horizontal gene transfer. The majority of these genes (74.2%) were putatively obtained from other *Proteobacteria*. The three classes *Alphaproteobacteria* (22.9%), *Betaproteobacteria* (20.8%), and *Gammaproteobacteria* (21.1%) seem to have contributed about equally to the total numbers of horizontally acquired genes. For example, the genome contained genes involved in oxidization of inorganic sulfur (SOX genes, Fig. 7). Sulfur oxidization pathways are also present in *Cupriavidus* and *Ralstonia* strains (e.g. [25]). However, the genome of QLW-P1DMWA-1 contained particular SOX genes not present in those taxa but in *Alphaproteobacteria* such as *Bradyrhizobium japonicum* [26].

Despite the pronounced differences in ecological adaptation and habitat preferences between QLW-P1DMWA-1 and other genome-sequenced members of the family *Burkholderiaceae*, genes characterized by both a beneficial role in adaptation to a planktonic lifestyle and absence in other *Burkholderiaceae* remain hard to identify. A more detailed analysis of this topic is presented in the Text S2.

**Metabolic reconstruction.** The genome of strain QLW-P1DMWA-1 (F10 lineage) encoded genes for a complete tricarboxylic acid cycle, Embden-Meyerhof-Parnas pathway (glycolysis), a complete glyoxylate cycle, and encoded an acetyl-CoA synthetase enabling utilization of acetate. The glycolysis and gluconeogenesis pathways seemed to be completely encoded by the genome, however, genes encoding kinases for phosphorylation of  $\alpha$ - or  $\beta$ -D-glucose seemed to be lacking. Since kinases for other hexoses as well as a phosphotransferase system (PTS) were also missing, it seems to be unlikely that the bacterium is utilizing this pathway for degradation of hexoses. The strain lacked the oxidative part of the pentose phosphate pathway and an Entner-Doudoroff pathway. Genes of pathways for *de novo* synthesis of all but one common cofactor and vitamins were found in the genome. In several cases the majority of genes required for a complete biosynthesis pathway were annotated, but one or two genes seemed to be lacking (e.g. thiamine and biotin pathways). The only fully absent pathway was the one for biosynthesis of vitamin B12. Since the genome of the strain encoded at least two B12-binding enzymes (i.e. Methylmalonyl-CoA mutase (Pnuc\_0910), methionine synthase (Pnuc\_1979)), this vitamin is most likely acquired by using transporters. Furthermore, genes for *de novo* synthesis of all twenty common proteinogenic amino acids apart from selenocystein could be identified in the genome. A tRNA gene for selenocystein is also lacking, while at least one tRNA gene for



**Figure 8. Phylogenetic position of the genome-sequenced *Polynucleobacter* strain.** Reconstruction of the phylogenetic position of the genome-sequenced strain *P.n. ssp. asymbioticus* QLW-P1DMWA-1 and genome characteristics of the strains used for the phylogenetic analysis. The presented tree is based on concatenated alignments of amino acid sequences of proteins encoded by eight housekeeping genes and was constructed using the NJ method. Letters A to H refer to bootstrap values obtained for the respective nodes by three different treeing methods in analyses of nucleotide and amino acid sequence data sets. The bootstrap values are presented in Table S7. doi:10.1371/journal.pone.0032772.g008

**Table 1.** Basic features characterizing the genome of *P. necessarius* ssp. *asymbioticus* strain QLW-P1DMWA-1.

Feature	QLW-P1DMWA-1
Genome size (Mb)	2.16
G+C content (mol %)	44.8
Number of replicons	1
Extrachromosomal elements	0
Coding region (% of bp)	93.1
Total genes	2136
Protein-coding genes	2088
rRNA operons	1

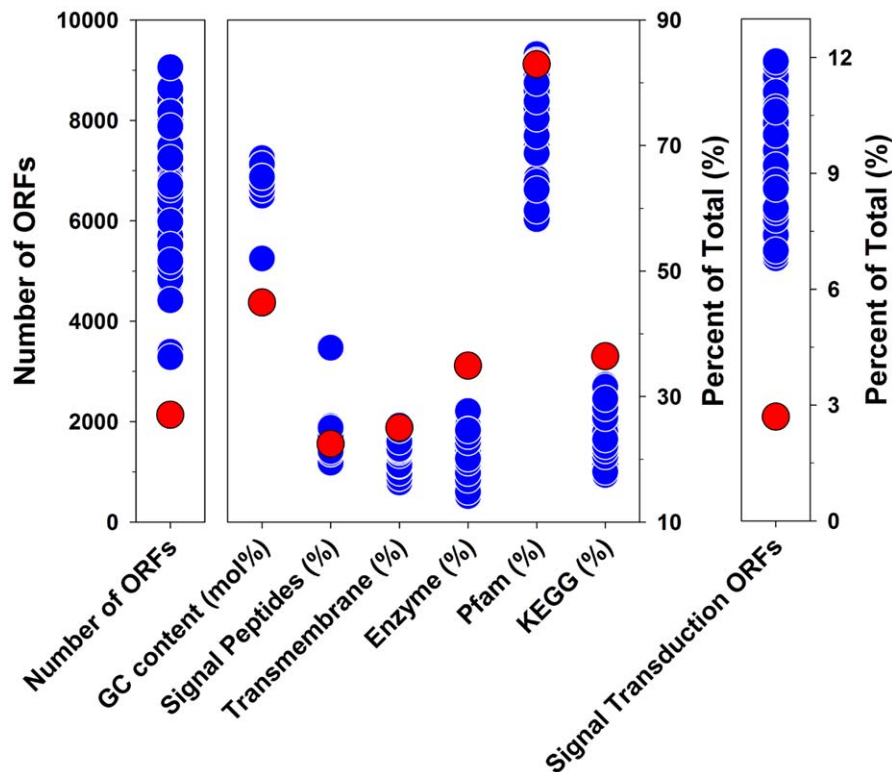
doi:10.1371/journal.pone.0032772.t001

each common amino acid is encoded. Previously demonstrated aerobic growth of strain QLW-P1DMWA-1 on acetate as sole carbon and energy source in an inorganic medium exclusively supplemented with vitamin B12 [10] suggested that all other vitamins and cofactors, as well as amino acids and other substances required for growth, can be synthesized from precursors provided by the TCA.

In previous investigations, anaerobic growth of strain QLW-P1DMWA-1 could be observed [10]. The genome annotation

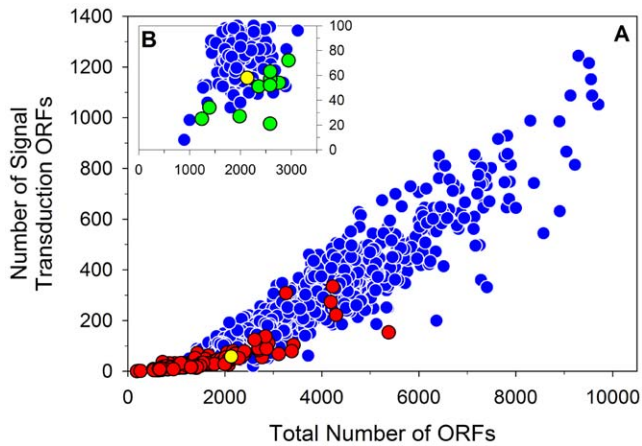
revealed no reliable indications of the kind of anaerobic metabolism performed by the strain. Nitrate respiration seemed unsupported by the annotation. Further investigations must reveal if anaerobic growth results from anaerobic respiration or from fermentation. However, survival or growth under anoxic conditions is expected to be of ecological relevance due to the lack of oxygen in the water body of the pond under longer-lasting ice-covered conditions (Figs. 1 and 2).

Analysis by TransportDB resulted in no indications of specific putative substrates for the majority of putative transporters. However, based on the analysis, many putative transporters could be assigned to functional groups. These assignments indicated that only a small number (three genes, 1.4% of genes putatively involved in transport) could be related to 'import of carbohydrates', whereas in total 17.7% of transporter proteins were assigned to the two categories 'import of carboxylic acids' (4.5%) and 'import of amino acids' (13.2%). Twenty-five and 36.4% of the transporter genes were assigned to the 'import of inorganic nutrients' and 'export and efflux' categories, respectively. The assignment of the three genes to the 'import of carbohydrates' category could be misleading since one assignment is based only on a low E-value (Pnuc\_1331) and one other assignments represented a protein (major facilitator family transporter) putatively involved in transport of a broader spectrum of substrates including sugar phosphates, tartrate and phthalate (Pnuc\_1635). Actually, the genome seemed to be lacking any ABC transporters for carbohydrates, as well as a phosphotransferase system (PTS).



**Figure 9. Variability of genomic traits among genome-sequenced members of the family *Burkholderiaceae*.** The depicted traits are the total number of ORFs, G+C content of the genomes, percentage of CDS encoding signal peptides, transmembrane domains, genes assigned to Enzyme Commission (EC) number classes, proteins assigned to protein families (Pfam), proteins assigned to KEGG categories, and proteins involved in signal transduction. For most traits, 78 genomes of strains affiliated with the family were analyzed using the IMG system [24]. The sole exception is the analysis of signal transduction genes, which was performed using the MIST2 database [69], which provided data on 37 strains affiliated with the *Burkholderiaceae*. For all depicted parameters, all available genomes of family members were considered but not the genome of the endosymbiotic *Polynucleobacter* strain STIR1. Data of strain QLW-P1DMWA-1 are indicated as red dots. doi:10.1371/journal.pone.0032772.g009





**Figure 10. Gene number of bacterial genomes versus number of encoded signal transduction genes.** The plot depicts the total number of annotated signal transduction genes in particular bacterial genomes versus the total number of genes of the respective genomes. Data on *Archaea* are not shown, and multiple genomes per species or species-like taxa were excluded. Strain *P. necessarius* QLW-P1DMWA-1 is highlighted as a yellow dot. (A) Entire data set consisting of 904 analyzed bacterial genomes. Red dots indicate bacteria obligately associated as symbionts or pathogens with a host; blue dots indicate free-living or facultative host-associated bacteria. (B) Plot of genomes of bacteria classified as free-living or facultative host-associated (shown in (A) as blue dots) with  $\leq 100$  signal transduction genes. All blue dots represent bacteria either facultatively associated with hosts (e.g., dwelling in the oral cavity or in the intestinal tract) or dwelling as free-living organisms in extreme environments (permanently anoxic, low or high temperature, low pH). The yellow (strain QLW-P1DMWA-1) and the green dots represent the only exclusively free-living taxa inhabiting non-extreme systems. Almost all of them are phototrophic (*Synechococcus* and *Prochlorococcus*) or predominantly heterotrophic taxa (*Pelagibacter* and *Puniceispirillum*) dwelling as planktonic bacteria in marine systems. The signal transduction gene data were obtained from the database MIST2 [69]. doi:10.1371/journal.pone.0032772.g010

By contrast, most genomes of *Cupriavidus* and *Ralstonia* strains encode putative transporters for carbohydrates.

**Oxidative stress.** Photochemical reactions of UV light with dissolved HS results in the formation of reactive oxygen species (ROS; [27,28,29]), which have an inhibitory effect on growth of bacterioplankton [30,31]. A recent study revealed that at least some *P. necessarius* strains were well adapted to oxidative stress by singlet oxygen [31]. An analysis of the genome of QLW-P1DMWA-1 for traits involved in resistance to oxidative stress revealed no unusual adaptations of the strain. The genome encoded two catalase genes (Pnuc\_2044, Pnuc\_2053) and two superoxide dismutase genes (Pnuc\_1385, Pnuc\_1626). In both cases, one of the isoenzymes is putatively exported while the other seems to be a cytoplasmic enzyme. The two encoded superoxide dismutases represent one enzyme of the Fe/Mn and one of the Cu/Zn type. The presence of both two catalase and two superoxide dismutase genes in one genome is not unusual among *Betaproteobacteria*. The analysis of 153 betaproteobacterial genomes revealed the presence of 1 to 6 superoxide dismutase genes per genome. Fifty-five percent of these genomes encoded two such genes, while only 29% encoded only one isoenzyme. Furthermore, the genome encoded a putative glutathione peroxidase (Pnuc\_1970), two methionine-S-sulfoxide reductases (Pnuc\_0662, Pnuc\_0972), two putative rubredoxins (Pnuc\_0238, Pnuc\_1377), and one putative rubrerythrin (Pnuc\_1376).

### Ecological function: major substrate sources

The high average relative and absolute cell numbers of the F10 lineage population in Pond-1 (Fig. 2) indicated that it is rather unlikely that the population is mainly utilizing rare substrate sources. Therefore, the search for their substrate sources should be focused on pools with high substrate concentrations or on pools filled with high fluxes. In case of Pond-1, two basically different substrate pools could potentially provide sufficient amounts of substrates for supporting the large F10 lineage population. On one hand, a low-molecular-weight pool of substances derived from algal or macrophyte (floating mats) primary production and especially algal exudation products could serve as major substrate source of the population [32]. On the other hand, HS forming usually 70–80% of DOC in freshwater systems [33] are present in Pond-1 with estimated average concentrations of about 3–9 mg L<sup>-1</sup> (average 5 mg L<sup>-1</sup>). We tested experimentally if F10 lineage strains are able to benefit from dissolved HS or algal exudates.

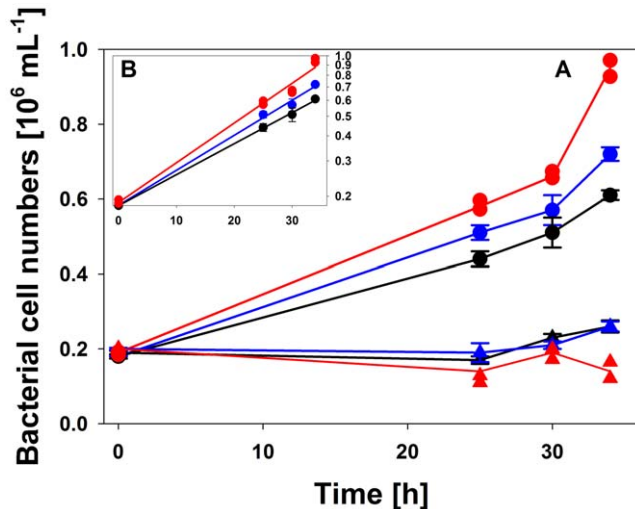
**Humic substances.** Experiments with pure cultures did not indicate a potential for substantial enzymatic degradation and assimilation of HS by strain QLW-P1DMWA-1 (Table S3). Furthermore, screening of the genome for genes potentially involved in HS degradation yielded no indications of such a capability. The genome lacks genes for mono- and dioxygenases that could be involved in cleavage of aromatic rings of HS. This finding is in agreement with the lack of assimilation of phenol and benzoic acid (Table S3). Assimilation of phenylacetate could be explained by cleavage and assimilation of the acetyl residue by an esterase. A gene putatively encoding an exported arylesterase (Pnuc\_0938) could be involved in such a reaction. Four other genes were found, which encode putatively exported esterases. However, further assimilation experiments with four other esters (ethylpyruvate,  $\alpha$ -terpinylacetate, isoamylacetate, and methylacetate) resulted in significant (t-test,  $P = 0.01$ ) assimilation only for methylacetate. In order to avoid toxic effects of the tested esters, these experiments were performed with low substrate and NSY medium concentrations (15 mg L<sup>-1</sup> and 5 mg L<sup>-1</sup>, respectively). Despite the low concentrations, the other three tested substances resulted in significant inhibition effects on the growth of QLW-P1DMWA-1. However, the observed assimilation of phenylacetate and methylacetate may indicate a potential for strain QLW-P1DMWA-1 to utilize at least some aliphatic residues of HS.

Instead of direct enzymatic degradation of HS, the strain could indirectly benefit under *in situ* conditions from HS by utilization of photooxidation products generated by the impact of UV and visible light (VIS) on HS [34], which was indeed suggested for *Polynucleobacter* bacteria previously [15,17]. We performed two experiments to test this hypothesis. In the first, strain QLW-P1DMWA-1 was grown on light- or dark-treated commercially available HS. In the second, humic water from Pond-1 containing the natural bacterial community (including the F10 lineage) of the pond was incubated under different light treatments (microcosm experiment) and growth of *Polynucleobacter* bacteria and the remaining bacterial community were determined.

In the first experiment increases in cell numbers of QLW-P1DMWA-1 were observed in the light and dark treatments. However, the increase in the dark treatments was only weak and did not significantly differ from the control treatments (data not shown), while the increase in cell numbers in the light treatments significantly differed from the dark and the control treatments (t-test,  $P < 0.005$ ). By considering typical photooxidation-caused production rates of low-molecular-weight (LMW) substances from HS [34], a growth efficiency of 9% – i.e., transformation of 9% of

the carbon supplied as HS to cellular carbon – could be calculated for the light-treated HS (Table S3).

In the second experiment (microcosm experiment) the influence of two light qualities (UV+VIS, and VIS only) and the absence of light (dark treatment) on growth of a bacterial community obtained from Pond-1 were investigated. Due to filtration capacity limitations, the absolute and relative *in situ* abundances of F10 lineage bacteria present in the pond at the time the water used for this experiment was sampled could not be determined. However, six days after sampling for the experiment, the relative abundance of *P.n. ssp. asymbioticus* cells in Pond-1 was 43.8% of total bacterial numbers and the F10 lineage comprised 21.2% of *P.n. ssp. asymbioticus* cells. In the experiment, *P.n. ssp. asymbioticus* cell numbers increased exponentially in all three treatments (Fig. 11B), while the other bacteria showed no significant growth (UV treatment) or only very weak and delayed growth (dark and VIS treatments) (Fig. 11A). Interestingly, growth of *Polynucleobacter* bacteria was strongest in the UV treatment (doubling time 15.2 h), but surprisingly, even in the dark treatment *Polynucleobacter* bacteria showed relatively strong growth (doubling time 19.5 h).



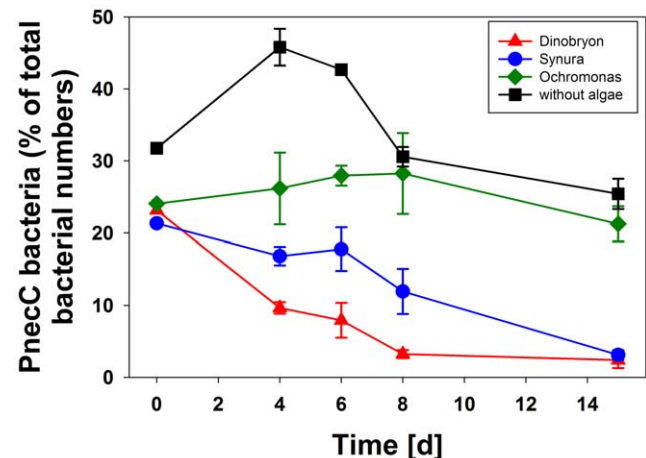
**Figure 11. Microcosm experiment on the influence of light on growth of bacteria from Pond-1.** Growth of a natural bacterial community originating from Pond-1 was investigated in three different light treatments (UV+VIS, VIS only, and no light). Water from Pond-1 was filtered through membrane filters with a nominal pore size of 0.8  $\mu\text{m}$  (exclusion of potential predators, algae, and other larger organisms) mixed with two volumes of 0.2  $\mu\text{m}$ -filtered pond water (almost free of bacteria) and incubated under *in situ* conditions in floating bottles covered with foil blocking all light (i.e., dark treatment), with foil blocking (>90%) UV light resulting in VIS light only (i.e., VIS only treatment), or in untreated bottles (penetration of >70% UV light), i.e. UV treatment. The bacterial community consisted initially of 48% *P.n. ssp. asymbioticus* and 52% other bacteria. The treatments were incubated under *in situ* conditions in the pond for a period of 34 hours. (A) Linear plots of results, (B) Linear regressions of semi-logarithmic plots of the results obtained for *Polynucleobacter*.  $R^2$  values of the regression lines were 0.999, 0.996, and 0.989 for the dark, VIS, and UV treatments, respectively. Filled circles, *P.n. ssp. asymbioticus*; filled triangles, other bacteria; red, with UV light; blue, VIS light only, black, no light. All shown data, apart of the data of the UV treatment (red symbols and lines), represent average data and SD (bars) from three replicates. One of three parallels of the UV treatment showed no growth at all. Presentation of average data was therefore omitted and only particular data for each of the two replicates with growth are shown.

doi:10.1371/journal.pone.0032772.g011

**Algal exudation products.** *Polynucleobacter* bacteria sampled from Pond-1 were tested for their growth response to the presence of algae representing abundant phytoplankton species in Pond-1. The results of this experiment are described in detail in Text S3, and Fig. 12 presents a graphical overview. This experiment revealed no indications of measurable growth of F10 lineage strains or other *Polynucleobacter* strains from Pond-1 on algal exudates. In order to further test the utilization of exudation products, experiments on assimilation of glycolate, which is a typical photorespiration exudate of algae [35,36], were performed with pure cultures of strain QLW-P1DMWA-1. In four experiments, no assimilation of this substance was observed, though slight inhibition of growth was noted. The substrate concentration used (6 mM) was in the range employed by other studies of the utilization of this substance by bacteria, i.e. 1–100 mM [37]. Our own experiments with *Cupriavidus necator* DSM 13513, *C. metallidurans* DSM 2839, and *C. basilensis* DSM 11853 under the same experimental conditions resulted in significant assimilation of glycolate. Furthermore, experiments with 19 other *Polynucleobacter* strains, representing all five currently known species, resulted only in one case in pronounced and, in another case, weak assimilation of this substrate (data not shown). The genome of strain QLW-P1DMWA-1 encoded a putative glycolate oxidase consisting of three subunits (Pnuc\_1781, 1782, 1783). The encoded enzyme, however, could also represent a lactate dehydrogenase. In addition, a glycolate permease gene as found in *Ralstonia/Cupriavidus* strains is missing in QLW-P1DMWA-1.

## Discussion

Free-living bacteria not associated with any host are traditionally viewed by microbiologists as organisms characterized by high



**Figure 12. Experiment on potential influence of algal exudates on growth of *P.n. ssp. asymbioticus* bacteria.** The influence of the presence of growing algae populations on the development of *P.n. ssp. asymbioticus* bacteria was investigated in batch culture experiments. Grazer-free (0.8  $\mu\text{m}$  filtered) water from Pond-1 was mixed with different algal cultures representing species typically abundant in Pond-1. The treatments were incubated with continuous illumination for 15 days at a constant temperature of 16°C. The graph shows mean values of relative abundance of *P.n. ssp. asymbioticus* bacteria and standard deviations from three replicates. The initial total bacterial numbers in all treatments that had received algae was  $4.7 \times 10^6$  cells  $\text{mL}^{-1}$ , and *P.n. ssp. asymbioticus* bacteria comprised 23% of bacterial cells. Cell numbers of the three different algal species developed differently over the course of the experiment, but all three showed growth over periods of six to 15 days.

doi:10.1371/journal.pone.0032772.g012

dispersal rates [38,39] and high potential to adapt to newly invaded environments [40,41]. Consequently, it could be expected that bacterial populations thriving in particular habitats are characterized by relatively high diversity and frequent changes in population structure, both resulting from frequent invasion by new genotypes and rapid ecological diversification due to adaptation to local ecological conditions. Here we investigated structure and dynamics of a *P.n. ssp. asymbioticus* population inhabiting a small acidic pond. Our investigation revealed a population persisting in the habitat over a period of at least seven years and having a relatively stable population structure characterized by a low genetic diversity. On the other hand, this low-diversity population is contributing during the ice-free seasons on average to 11% of the bacterial cells in the water column of the pond. In order to reveal the traits responsible for this ecological success, we characterized the ecology of the population by genome analysis of a representative strain and experimental analysis of the physiology of strains representing the population. These investigations revealed a highly passive lifestyle, which results in exclusion of any opportunity for targeted access and exploitation of trophic hot spots in the water column. The revealed passive lifestyle stands in strong contrast to the ecological success of the population under the volatile environmental conditions of its home habitat, and the low diversity and stable structure of the population contrasts with the expected population dynamics resulting from frequent invasion by new genotypes.

### Genome evolution

Reductive evolution of genome size, also known as genome streamlining, is well documented for 'true picoplankton' [42] bacteria inhabiting the upper water layers of the sea [43,44,45]. This includes *Candidatus Pelagibacter ubique* [43] affiliated with the ubiquitous SAR11 clade (*Alphaproteobacteria*), as well as some *Cyanobacteria* affiliated with the genus *Prochlorococcus* [44]. These as well as other taxa of the 'true picoplankton' fraction of the marine plankton represent abundant groups with cosmopolitan distributions [42] and an oligotrophic physiology [46]. Here we have demonstrated that abundant (Fig. 2) freshwater bacteria passed through a reductive evolutionary process highly similar to that of true marine picoplankton. The ancestors of the investigated *Polynucleobacter* bacteria most likely shared a copiotrophic lifestyle, broad metabolic capabilities, adaptations for dwelling in soils or in association with hosts, as well as a larger genome size with the recent *Burkholderiaceae* taxa (Figs. 5 and 6). Similarly as with *Candidatus Pelagibacter ubique* [43], reductive evolution resulted in a sparser metabolic network with a smaller number of entries for substrates. The reduced catabolic versatility of QLW-P1DMWA-1 is not only indicated by the relatively small proportion of utilized substrates (Fig. 6) and the specialization for carboxylic acids but also in the lower (by about a factor of ten) amount of biomass produced on complex media as compared to other *Burkholderiaceae* species [10]. The high proportions of QLW-P1DMWA-1 genes assigned to EC, Pfam and KEGG categories (Fig. 9) could be interpreted as a trend towards elimination of 'exotic' (not well characterized) accessory genes that increase the fitness of the strain only in rare situations. As in true picoplankton bacteria, the decreased catabolic versatility is contrasted with maintenance of basic anabolic pathways, resulting, in the case of QLW-P1DMWA-1, in an almost complete prototrophy. As reported for true picoplankton bacteria [42,43] motility, chemotaxis, and quorum sensing are absent, and the absolute and relative number of signal transduction genes are low. The genome shares several characteristics with genomes of marine oligotrophs [46]. This includes, for instance, the presence of only a single ribosomal

operon, and underrepresentation of genes assigned to the COG categories N (cell motility and secretion), T (signal transduction), V (defense), and K (transcription). An overrepresentation of categories Q (secondary metabolites biosynthesis, transport and catabolism) and I (lipid transport and metabolism), however, was not observed (data not shown). Altogether, streamlining of the QLW-P1DMWA-1 genome resulted in a highly passive oligotrophic lifestyle with little metabolic plasticity.

### Contrast between passive lifestyle and environmental dynamics in the home habitat

A low number of encoded signal transduction genes (Fig. 10) resulting in only minor capabilities to sense changes in environmental conditions and other features of passive lifestyles fit well to true planktonic life in stable environments. *Candidatus Pelagibacter ubique* and *Prochlorococcus* strains inhabit the upper water layers of oceans, where they experience such stable physicochemical conditions [42,43]. Apart from potential diurnal changes in conditions for photosynthesis and influences of micro-patchiness of nutrient distribution [47,48], only rather small fluctuations in growth conditions may be experienced by these marine bacteria over time scales of days and weeks. By contrast, the investigated *Polynucleobacter* population persists in a physicochemically and biologically highly dynamic system. Pronounced seasonal differences in oxygen concentration, light availability and temperature occur in annual cycles. Additionally, short-term temporal changes in weather conditions immediately impact the conditions in this small and shallow pond, and sunny weather conditions result in diurnal stratification and mixing cycles (Fig. 3). Mixing of the water column distributes the F10 lineage cells randomly over the depth profile of the pond. The vertical position of particular cells in the water column at the beginning of the stratification process result in strong differences in growth conditions between particular strains for periods of up to 24 hours. Cells placed in the upper water layers experience pronounced increases in water temperature and intensive exposure to sunlight, while cells placed in deeper water layers are much less impacted by solar radiation. As discussed below, such different positions in the water column should result in differences in availability of nutrients (photooxidation products), as well as in different levels of oxidative stress [31]. It is surprising that bacteria characterized by such a passive lifestyle are able to establish under such dynamic environmental conditions such large populations (Fig. 2). Previously published data [6] as well as qPCR data (Jezberová et al., unpublished results) suggest that the F10 lineage is abundant in many habitats with similarly unstable conditions as characterized for Pond-1; the success of this lineage in Pond-1, thus, does not seem to be exceptional. The less extreme genome streamlining of QLW-P1DMWA-1 compared to the three genome-sequenced *Pelagibacter* strains – indicated, for instance by genome sizes of 2.1 Mb versus 1.3–1.5 Mb – may be a consequence of adaptation to a passive lifestyle in a rather unstable environment.

### Ecological function of F10 lineage strains

Results from genome analysis and experimental characterization of metabolic traits indicated that strain QLW-P1DMWA-1 mainly acts as a chemoorganoheterotroph using organic substrates as energy and carbon sources as well as electron donors. In addition, the strain seems able to utilize inorganic sulfur species as electron donors. However, it is not known if reduced sulfur species serve as important energy sources for the F10 lineage population in Pond-1. It is likely that reduced sulfur played no significant role in the microcosm experiment shown in Fig. 11, since the microcosms received only water, and no sediment samples. The



experimental conditions thus primarily mimicked the conditions in the water column of the pond but largely excluded the activity of sulfur reducers and fluxes of reduced sulfur from the sediment to the water.

In dystrophic freshwater systems such as Pond-1, the DOC pool of the water body is usually mainly composed of high-molecular-weight (HMW) compounds, i.e. by HS [33]. In contrast to the minor fraction of LMW compounds, the HMW substances represent recalcitrant compounds characterized by long turn-over times [49]. It was found that the absolute and relative abundance of *P.n. ssp. asymbioticus* cells in a large set of investigated habitats was positively correlated with concentrations of HS [12]. Explanations for this correlation could be direct or indirect utilization of HS by *P.n. ssp. asymbioticus* bacteria. Screening of the genome of QLW-P1DMWA-1 for genes typically involved in microbial degradation of aromatic substances and HS, as well as assimilation experiments with model substances, revealed no indication of any pronounced capabilities in HS degradation (Table S3). Some indications of capabilities of strain QLW-P1DMWA-1 to cleave and utilize at least some aliphatic residues of HS were observed. However, assimilation experiments with HS did not indicate that such capabilities result in significant growth of the strain.

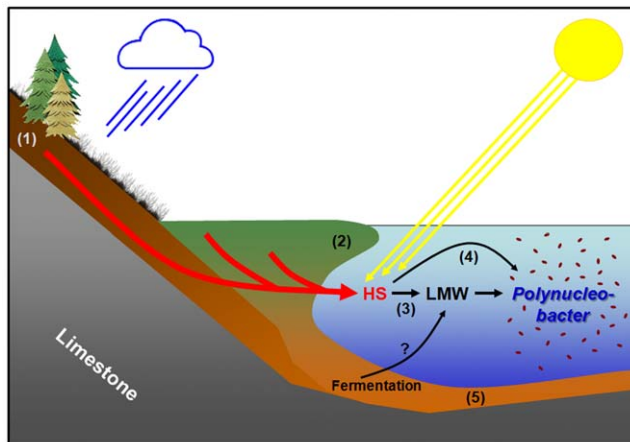
Indirect utilization of HS as photooxidation products was suggested for *Polynucleobacter* bacteria previously [15,17]. Photooxidation of HS is caused by excitation of HS by UV light and to a lesser extent by visible light [50]. The main products of the photooxidation are LMW organic substances usually mainly composed of carboxylic acids, carbon monoxide, carbon dioxide and other inorganic compounds [34,51,52]. Several of the substances assimilated by strain QLW-P1DMWA-1 with high efficiencies in our laboratory experiments, e.g. pyruvate, acetate, and succinate, were reported as photooxidation products [34], and strong uptake of acetate by *P.n. ssp. asymbioticus* cells in a humic lake was observed previously [17]. The revealed preference of the strain for carboxylic acids (Fig. 6) fits well in general to the reported spectrum of photooxidation products [34]. However, some substances known to be produced by photooxidation of HS, e.g. citrate, oxalate, or glyoxal, were not assimilated. Experiments with light-treated HS solutions and the microcosm experiment (Fig. 11) further support the notion that F10 lineage bacteria are able to utilize photooxidation products. In the microcosm experiment, the *P.n. ssp. asymbioticus* population responded best to the UV treatment while the other bacteria showed the opposite response. Surprisingly, *P.n. ssp. asymbioticus* bacteria also grew well but more slowly in the dark treatment. This could indicate that despite the storage of the water samples used for the experiment in the dark for 24 h, the water had still contained enough substrates available for *P.n. ssp. asymbioticus* and also other bacteria. The delayed growth of the other bacteria could have been caused either by the presence of reactive oxygen species (ROS) with long half-lives, e.g. hydrogen peroxide ( $H_2O_2$ ), for which half-lives in lake water of 10–22 h have been determined [53], or longer-lasting inhibitory effects caused by singlet oxygen ( $^1O_2$ ) through direct cell damage by oxidation of lipids, nucleic acids and proteins [54]. Both variants of ROS are formed during photooxidation of HS [27,28,29]. Thus, the delayed growth of the other bacteria could have been a result of the environmental conditions at the time the water samples used in the experiment were taken.

Even algae in fairly good physiological state release a part of their photosynthetic products to the surrounding water [35] under natural conditions. These released organic substances represent an important carbon and energy source for freshwater and marine bacterioplankton [55]. The quality and quantity of the released

organic material is species-specific and strongly influenced by growth conditions and growth phase of the algae. However, carbohydrates (mainly polysaccharides) and glycolate usually constitute the major fraction of the released substances [35]. Sugar alcohols, other organic acids, peptides and amino acids were found to be released by many algal species but constituted frequently only small fractions of algal exudates. The most frequently released amino acids were found to be Ser, Gly, Lys, Ala, Glu, Asp, Orn, and His. Our assimilation experiments indicated that strain QLW-P1DMWA-1 did not – or only to a low extent – assimilate most of the mentioned algal exudation products. Several monosaccharides and glycolate were not utilized, whereas some sugars were utilized with low efficiency. The almost complete lack of genes putatively encoding transporters for carbohydrates, as well as the lack of genes for specific phosphorylation of monosaccharides suggests that strain QLW-P1DMWA-1 is not adapted for utilization of glucose, fructose, mannose and other carbohydrates. The observed weak utilization of such substances could represent laboratory artifacts caused by high substrate concentrations in the assimilation experiments. Interestingly, Paver & Kent [36] suggested that *Polynucleobacter* bacteria represent a large part of the freshwater bacterioplankton with the genetic potential to use glycolate. They used bacteria-specific primers for specific amplification of glycolate oxidase genes (*glcD*) encoding one subunit of the glycolate oxidase, which is a key enzyme in the assimilation of glycolate [37]. Forty-eight percent of the *glcD* sequences obtained from five lakes clustered with the putative *glcD* gene (*Pnuc\_1781*) of strain QLW-P1DMWA-1 [36].

Altogether, our results do not indicate that *P.n. ssp. asymbioticus* strains affiliated with the F10 lineage are specially adapted to utilization of algal exudation products. Neither the experimentally revealed assimilation patterns nor the growth experiment performed in the presence of algae (Fig. 12) indicate such an adaptation. In addition, besides the presence of putative glycolate oxidase genes, there is a lack of genomic data further supporting this kind of adaptation. However, we cannot rule out that other lineages affiliated with *P.n. ssp. asymbioticus* or other *Polynucleobacter* taxa mainly grow on substances released by phytoplankton. Previous investigations on *P. acidiphobus*/*P. difficilis* assemblages (*PnecB*) in Lake Mondsee indicated a relationship to phytoplankton dynamics [23], and recent experimental investigations support this hypothesis [3]. In contrast to our investigations on the F10 lineage, other studies on different freshwater bacteria, i.e. strains affiliated with the important genus *Limnohabitans*, resulted in a more obvious match between substrate assimilation pattern [56], successful growth on algal exudates [3], and correlation between phytoplankton and *Limnohabitans* dynamics [57,58]. Even weak glycolate assimilation was demonstrated for *Limnohabitans planktonicus* [56].

Based on the presented results, we propose a model of the ecological function of the F10 lineage population in Pond-1 that is centered on the utilization of photooxidation products of HS (Fig. 13). The hydrological conditions of Pond-1 indicated that the majority of HS dissolved in the water body is of terrestrial origin. HS are transported to the pond by water percolating from the surrounding slopes, as well as from the *Sphagnum*- and *Carex*-dominated floating mats. In the pond, HS are temporarily exposed to solar radiation and a slow photodegradation takes place. The resulting LMW substances are utilized by the F10 and potentially also by the few other *P.n. ssp. asymbioticus* lineages present in the pond [4]. Due to the limited versatility caused by the evolutionary genome streamlining, the F10 lineage bacteria can most likely consume only a fraction of the released photooxidation products.



**Figure 13. Conceptual model on carbon fluxes utilized by the F10 lineage population in Pond-1.** The model suggests origin and transformation of major carbon sources utilized by the F10 lineage population inhabiting Pond-1. Humic substances (HS) leached from soils (1) and floating mats (2) (*Schwinggrasen*, formed mainly of *Sphagnum* and *Carex*) are transported by percolating water to the pond. *P.n. ssp. asymbioticus* bacteria affiliated with the F10 lineage mainly utilize low-molecular-weight (LMW) photooxidation products of HS (3). Direct utilization of aliphatic residues of HS by enzymatic cleavage is also expected to take place (4). Another potential substrate source could be fermentation products released by bioturbation and other processes from the sediment (5) to the water column of the pond. doi:10.1371/journal.pone.0032772.g013

This limited versatility may enable the coexistence of different lineages by substrate partitioning [59]. For instance, the F10 lineage bacteria seem unable to consume citrate, which was reported to be a product of photodegradation [34]. In addition to utilization of photooxidation products, F10 lineage bacteria may cleave and utilize aliphatic residues of HS. The LMW pool exploited by the F10 lineage bacteria may also be filled by other fluxes, e.g. the release of fermentation products from anoxic sediment or the cleavage of HS by microbial exoenzymes released by other organisms. Additionally, even algal exudation may deliver some substances utilized by this lineage. All data obtained so far indicate that the F10 lineage bacteria float passively in the water column of the pond and mainly wait for the delivery of substrates by the photooxidation processes. The performed predation experiments suggest that the *P.n. ssp. asymbioticus* population including the F10 lineage serves as a food source for heterotrophic and mixotrophic protists, which should result in a transfer of HS-derived organic material in the pelagic food web of the pond.

### Why is the passive lifestyle of the investigated population so successful?

The above-presented hypothesis on the ecological function of the investigated F10 lineage population gives no indication at all that these bacteria represent specialists able to monopolize a trophic niche. In contrast, it has to be assumed that a broad diversity of other bacteria should also be able to exploit the substrate sources utilized by the F10 lineage population. Most likely, non-trophic adaptations - or at least combinations of trophic and non-trophic adaptations - are responsible for the ecological success of the population. What kind of adaptations could make the F10 lineage population so successful in Pond-1 and other *P.n. ssp. asymbioticus* populations so successful in other habitats? Results from several studies indicate that a rather passive and oligotrophic lifestyle is a common trait of most or all *P.n. ssp. asymbioticus*

bacteria (e.g. [10,11,15]). Furthermore, the commonly relatively low G+C content of *Polynucleobacter* genomes (<50 mol% vs. >60 mol% for the majority of other *Burkholderiaceae*) as well as previous estimations of genome sizes [7] suggest that reduced genome sizes represent a common feature of *Polynucleobacter* bacteria. Interestingly, the sibling group of *Polynucleobacter* formed by the two genera *Cupriavidus* and *Ralstonia* (Fig. 8) is absent or plays only a minor role in the water columns of freshwater bodies [1,2]. Members of these two genera are generally characterized by genome sizes about three times the size of *Polynucleobacter* genomes. If microevolutionary adaptations, e.g. adaptations to the spectrum of major substrates available in the water columns of freshwater systems, were enough to become successful planktonic players, one would expect *Cupriavidus/Ralstonia*-like bacteria to contribute significantly to the composition of freshwater bacterioplankton. Instead, the genome-streamlined sibling group *Polynucleobacter* is the one that represents a prominent fraction of freshwater bacterioplankton [12]. Macroevolutionary genome economization seems to be a prerequisite for a successful colonization of the water column of freshwater systems by *Burkholderiaceae* bacteria. It is believed that, in prokaryotes, small genome size is a prerequisite for maintenance of small cell sizes by metabolically active and growing cells [42]. Small cell sizes are known to be advantageous in substrate acquisition due to the increased surface-to-volume ratio [60], as well as in the interaction with many protistan predators [61,62] and viruses [63]. Future investigations will have to reveal if optimization of substrate uptake or reduction of predation mortality is the more important driver in the evolution of *P.n. ssp. asymbioticus* bacteria. However, it is likely that additional but so far undiscovered or only superficially investigated adaptations also contribute to the ecological success of these bacteria. Resistance to oxidative stress [31] could be one such trait but does not have to be the only relevant adaptation. The passive lifestyle of the F10 lineage bacteria could, rather, be interpreted as a byproduct of the genome size reduction than as a strategy responsible for the success of the population.

### Could the obtained genomic and ecological insights be generalized to other *P. necessarius* lineages?

Currently, we can only speculate if generalization of the obtained insights in the ecology and adaptation of the investigated lineage to other lineages of *P.n. ssp. asymbioticus* or other *Polynucleobacter* species is suitable. To what extent differences in adaptations are involved in the ecological diversification of *P.n. ssp. asymbioticus* bacteria is still unknown [6]. This diversification may be restricted to non-trophic adaptations such as pH or temperature adaptation. In this case, it could be that all *P.n. ssp. asymbioticus* strains exploit more or less the same trophic niche, i.e. utilization of LMW products of HS photodegradation. On the other hand, adaptation to different trophic niches can currently not be excluded. Reports on photoheterotrophy in *Polynucleobacter* bacteria [19] and indications of potential utilization of major algal exudation products [36] may suggest a pronounced trophic diversification.

### Conclusions

Evolutionary genome streamlining of F10 lineage bacteria resulted in a highly passive lifestyle so far known only among free-living bacteria of pelagic marine taxa (true picoplankton bacteria) dwelling in environmentally stable, nutrient-poor, off-shore systems. Surprisingly, such a lifestyle is also successful in a highly dynamic and nutrient-rich environment such as the water column of the investigated pond. Obviously, neither metabolic and ecological versatility nor a pronounced microdiversity of popula-

tions are prerequisites for long-lasting establishment of abundant bacterial populations under highly dynamic environmental conditions. We are only just beginning to understand the diversity and ecology of *Polynucleobacter* bacteria, therefore, caution should be exercised when generalizing the obtained insights in ecology and adaptation of the investigated lineage to other *Polynucleobacter* lineages.

## Materials and Methods

### Sampling of Pond-1

The pond was sampled over a period of six years (autumn 2003 to autumn 2009), as described elsewhere [4]. Most of the samplings were performed during the vegetation periods (May to November), but two samplings were performed in deep winter (2004 and 2008). No specific permits were required for the described field studies.

### Quantification of bacteria

Total bacterial numbers were determined by epifluorescence microscopic analysis of DAPI-stained samples as described previously [4]. Total numbers of *P. necessarius* bacteria were determined by fluorescence *in situ* hybridization (FISH) using the specific probe PnecC-16S-445 as described previously [4]. The relative contribution of the *Polynucleobacter* F10 lineage to the total number of *P. necessarius* bacteria was determined by quantitative PCR (qPCR). Quantitative PCR was performed on an Eppendorf Mastercycler (ep realplex Thermal Cyclers) using TaqMan probes labeled with 6-FAM (5'-end) and 5-TAMRA (3'-end) and template DNA extracted from water samples as described previously [6]. We have designed two quantitative PCR assays, one specific for the F10 lineage, and one for the entire taxon *P. necessarius*. The F10 lineage assay detected a 97 bp fragment and was performed with the forward primer G1\_TF (5'-GCTGTTTAGGTGAAGAAAG-TGTGACA-3'), reverse primer G1\_TR (5'-CTTCTGTGAAT-GGCGCTTTCTG-3') and probe G1\_TM (5'-6FAM-TTTGCCGAACACCCGCTTATCAATACACTG-TMR-3'). The *P. necessarius* assay detected a fragment of a total length of 142 bp and consisted of the forward primer C\_TF (5'-GRGTAAAGATT-GAATCATCAATCAG-3'), reverse primer C\_TR (5'-CTA-TAACGAGCACCATTGCTAGY-3') and the probe C\_TM (5'-6FAM-CTTACAGTTTGGATTACGGCAAACATGTC-TMR-3'). The 12  $\mu$ L PCR reactions contained PCR water, 1 $\times$  buffer (TaqMan Universal PCR Master Mix - Applied Biosystems), 900 nM forward primer, 900 nM reverse primer and 150 nM probe. The PCR cycle consisted of 10 min at 95°C followed by 50 cycles: 95°C for 15 s, 60°C for 60 s. Fluorescence signal results were analyzed using the Realplex software. CT value threshold was set for 100, and CT values were recalculated to gene copies using standards. The ratio  $P_{F10}/P_{all}$  of copy numbers of markers specific for the F10 lineage ( $P_{F10}$ ) and markers specific for all *P. necessarius* ( $P_{all}$ ) was determined. Absolute numbers of F10 lineage cells were calculated by multiplication of the absolute number of *P. necessarius* cells, as determined by FISH, by the ratio  $P_{F10}/P_{all}$ .

### Grazing experiments

Two grazing experiments with two different kinds of fluorescently labeled bacteria (FLB) were performed for estimation of grazing rates of protists on the entire bacterial community and the *Polynucleobacter* assemblage in Pond-1. FLB were produced using a sample of natural bacterioplankton (Řimov Reservoir, Czech Republic), as well as a pure culture of strain QLW-P1DMWA-1 by following previously published protocols ([64]; for specific details see [65]). In each experiment FLB (15–20% of total bacterial

numbers in a sample) were added to a water sample from Pond-1 (20 September 2005) and incubated under *in situ* conditions (temperature 9°C) in the laboratory. Subsamples were taken at 10–15 minute intervals (ciliates) and 60 minute intervals (flagellates) after addition of the FLB and fixed by adding 0.5% of alkaline Lugol's solution, immediately followed by 2% (vol/vol) (final concentration) formaldehyde and several drops of 3% (wt/vol) sodium thiosulfate to clear the color of the Lugol's solution [64]. The numbers of ingested FLB per protistan cell were enumerated using an epifluorescence microscope [65]. Food vacuole content of the protozoa was also checked for presence of FISH-positive cells using the FISH probe PnecC-16S-445 [4] as described elsewhere [66].

### Isolation of *Polynucleobacter* strains

Bacterial strains were either isolated by the filtration acclimatization method [67] or by the dilution acclimatization method [4]. Pure cultures of strains were stored at  $-70^{\circ}\text{C}$ . *Polynucleobacter* strains and F10 lineage strains were identified by sequencing of the 16S–23S ITS [4,6].

### Multilocus sequence analyses of F10 strains

The complete genome sequence (accession number CP000655, [22]) of *P.n. ssp. asymbioticus* strain QLW-P1DMWA-1 [4,10] was used to develop primers for multilocus sequence analyses (Table S4) of F10 strains isolated from Pond-1. Furthermore, all twelve F10 strains (including QLW-P1DMWA-1) isolated from the pond were characterized by genetic fingerprinting with three independent methods, as described previously [4].

### Phylogenetic analysis

The phylogenetic position of the genome-sequenced strain QLW-P1DMWA-1 was reconstructed using a multilocus sequence analysis approach. The sequences of eight protein-encoding housekeeping genes (Table S5) of the strain and ten genome-sequenced reference strains (Table S6) were used for this analysis. Phylogenetic trees were constructed by the neighbor-joining (NJ), the maximum likelihood (ML), and the maximum parsimony (MP) methods, using the software MEGA4 and MEGA5 [68]. Tree constructions were based on either nucleotide sequences or amino acid sequences. For comparison, the phylogeny of the 16S rRNA gene of the eleven strains was reconstructed using the same treeing methods.

### Comparative analysis of genomic traits

Comparative analyses of the genome of strain QLW-P1DMWA-1 were mainly performed using the integrated microbial genomes (IMG) system [24]. Data on signal transduction genes of the strain and reference bacteria were obtained from the Microbial Signal Transduction Database MIST2 [69]. A functional interpretation of putative transporter proteins was performed by comparing the amino acid sequences of putative transporter genes to the database TransportDB [70]. Because this comparative analysis, for the majority of putative transporters, gave no indications of their specific putative function (i.e., of their specific putative substrates), putative transporter genes were assigned to seven more general categories of putative functions (i.e., import of amino acids, import of carbohydrates, import of carboxylic acids, import of other organic substances, import of inorganic nutrients, export and efflux, and unknown transport function).

### Substrate assimilation tests with strain QLW-P1DMWA-1

Assimilation tests with a total of 93 substrates were performed in batch culture experiments. Previous experiments have revealed



that growth of QLW-P1DMWA-1 on single carbon and energy sources is unreliable and weak [10]. Therefore, assimilation experiments were performed by comparing optical density ( $OD_{575\text{ nm}}$ ) values established in one-tenth-strength ( $0.3\text{ g L}^{-1}$ ) liquid NSY medium [67] with and without  $0.5\text{ g test substance L}^{-1}$  (pH 7.2). The NSY medium was supplemented with thiamine ( $80\text{ }\mu\text{g L}^{-1}$ ), biotine ( $0.4\text{ }\mu\text{g L}^{-1}$ ) and cyanocobalamin ( $0.4\text{ }\mu\text{g L}^{-1}$ ). Photometrically determined  $\Delta OD_{575\text{ nm}}$  values (i.e.,  $OD_{\text{NSY+substrate}} - OD_{\text{NSY}}$ ) represented biomass generated by the strain by utilizing the tested substrate. Growth efficiencies – i.e. the fraction (%) of test substrate carbon transformed to bacterial biomass carbon – for the various tested substrates were calculated by transferring  $\Delta OD_{575\text{ nm}}$  values to cell numbers based on an empirical conversion factor, multiplication of cell numbers with epifluorescent microscopically determined average cell volumes, and conversion of cell volumes to carbon units using the conversion factor  $308\text{ fg C }\mu\text{m}^{-3}$  suggested by Fry [71]. According to Loferer-Kröbächer et al. [72], this factor is the most probable volume-to-carbon conversion factor for cells with sizes in the range observed for strain QLW-P1DMWA-1. Assimilation experiments were performed in triplicates at room temperature, and for each experiment OD values or cell numbers (see below) were determined twice after 7 and 11 days of incubation. Cultures were carefully checked for contaminations by streaking samples taken at the start and end of the experiments on agar plates. Experiments lacking positive  $\Delta OD_{575\text{ nm}}$  values were repeated at least once. Experiments resulting in negative  $\Delta OD_{575\text{ nm}}$  values, which potentially indicates toxic effects of the test substance on the bacterial strain, were repeated with decreased concentrations of test substances ( $15\text{ mg L}^{-1}$ ) and NSY medium ( $5\text{ mg L}^{-1}$ ). In these experiments, growth was recorded by epifluorescence microscopical cell counts. Assimilation of four esters was exclusively tested at these low substrate concentrations. The tested substances are listed in Table S3.

### Experiment on influence of light conditions on growth of strain QLW-P1DMWA-1 on dissolved humic acids

For testing the utilization of photooxidation products of humic substances (HS) by strain QLW-P1DMWA-1, commercially available HS (Sigma-Aldrich) were used. Light treatment and growth experiment were performed separately. HS were dissolved ( $500\text{ mg L}^{-1}$ ) in Milli-Q water, filled in 1 L Pyrex bottles (for light transmission properties of the bottle see below), and the solutions were sterilized by microwaving and exposed for three days to natural sunlight (February 2008, sunny weather; total light exposure was 21 hours). Triplicated dark (covered with aluminum foil) and light treatments (no cover) were exposed in parallel. Samples for measurement of fluorescence excitation emission matrix (EEM) were taken before and after exposure. For testing growth of QLW-P1DMWA-1, nine bottles were set up and inoculated with the strain. Three bottles received aliquots from the light-treated HS solutions, three received aliquots from the dark-treated bottles, and three received no HS (negative control). All bottles contained IBM medium (not supporting growth of the bacteria) supplemented with a vitamin mix (see above), and the calculated HS concentrations were  $30\text{ mg L}^{-1}$ . Bottles were incubated in the dark at room temperature for 18 days. Bacterial cell numbers and cell sizes were determined microscopically.

### Experiment on influence of light conditions on growth of the natural *Polynucleobacter* population of Pond-1

Water sampled from Pond-1 before sunrise was filtered either through  $0.8\text{ }\mu\text{m}$  (removal of algae, protistan predators and other larger

particles) or  $0.2\text{ }\mu\text{m}$  (removal of almost all bacteria and larger particles) polycarbonate membrane filters and stored in the dark for 24 h to minimize inhibitory effects of potentially present reactive oxygen species (ROS), e.g.  $\text{H}_2\text{O}_2$  [30]. The next day, the differently filtered water samples were mixed in a ratio of 1:2 ( $0.8\text{ }\mu\text{m}:0.2\text{ }\mu\text{m}$  filtrates), and 150 mL volumes of the mixture were distributed in 1 L borosilicate bottles. Three different light treatments were created by covering bottles with aluminum foil (no light), by covering bottles with UV filter foil (e-color #226, Rosco) (visible light only), and by using uncovered bottles (UV and visible light). Transmission measurements with a Xenon (75 W) light source revealed that the UV filter foil absorbed  $>90\%$  of light in the 350–380 nm range (UV) and 10% in the 390–780 nm range (visible light, VIS). The uncovered borosilicate bottles absorbed 65–70% of UV light and 65% (380 nm) to 10% (700 nm) of the visible light. The transmissibility of the bottles increased steadily from 380 to 700 nm. Three bottles of each treatment were incubated in the pond for 34 h. During the incubation, the bottles floated at the water surface and were kept in position by a rope.

### Experiment on potential utilization of algal exudates as major substrate source

Batch experiments with a natural bacterial community (including the F10 lineage) obtained from Pond-1 and three species of algae were performed under controlled laboratory conditions. Non-axenic cultures of three algal species, i.e. non-phagotrophic *Synura* sp. LO2 34 K-F, putatively mixotrophic *Dinobryon* sp. LO2 26 K-S, and putatively mixotrophic *Ochromonas* sp. LO2 44 K-D, isolated by Steffen Jost from a dystrophic pond located at a distance of 120 m from Pond-1 were used as potential substrate producers for bacteria. The two ponds shared highly similar physicochemical characteristics (Hahn et al, unpublished data), and all three taxa represented typical species of both ponds (Alexandra Pitt, personal communication). Bacteria from Pond-1 (sample of 3<sup>rd</sup> November 2007) were separated from protistan predators and other eukaryotes by filtration through a  $0.8\text{ }\mu\text{m}$  membrane filter [73]. Four treatments, of three different algal species and one control without algae, were each set up in triplicates. In the treatments with algae, 200 ml of  $0.8\text{ }\mu\text{m}$  filtered pond water was mixed with 30 ml of algal cultures and distributed equally in three 100 mL Erlenmeyer flasks. The controls received no algae. The treatments were incubated in a cultivation room at  $16^\circ\text{C}$  with continuous illumination. Samples for determination of total bacterial numbers, *Polynucleobacter necessarius* numbers, and algal numbers were taken after 0, 4, 6, 8 and 15 days of incubation. Algal cells were counted after fixation with Lugol's solution using an inverted light microscope [74].

### DNA sequences accession numbers

The GenBank/EMBL/DDBJ accession numbers for the sequences used in the microdiversity study on the F10 lineage population in Pond-1 are HE646440 - HE646569, AJ879799, and AJ879800. All relevant sequences of strain QLW-P1DMWA-1 were taken from the deposited genome sequence (for gene tags see Table S4).

### Supporting Information

**Table S1** Current taxonomy of *Polynucleobacter* bacteria. The genus *Polynucleobacter* currently harbors five species (Hahn et al., 2009, 2010, 2011a, 2011b, 2012). These five species represent the four *Polynucleobacter* tribes (PncA to PncD) suggested by the taxonomy in Newton et al. (2011). Some of the strains affiliated with the subspecies *P. necessarius asymbioticus* were recently assigned

to lineages putatively representing differently adapted ecotypes (Jezbera et al., 2011). Lineages F10, F5 and F4 inhabit Pond-1 while other lineages and genotypes are absent from this habitat (Hahn et al., 2005). The populations of these three lineages dwelling in Pond-1 were previously described as G1, G2 and G3 populations (Hahn et al., 2005). The current study is focusing exclusively on the F10 lineage (G1) population of Pond-1. The subspecies *P. necessarius* ssp. *necessarius* contains exclusively endosymbionts of ciliates (Hahn et al., 2009), thus contains only strains fundamentally differing in lifestyle from all other currently established *Polynucleobacter* taxa (Vannini et al., 2007).

(DOCX)

**Table S2** Abbreviations

(DOCX)

**Table S3** Results of the assimilation experiments performed with 93 substrates. Substrates that did not support growth of the strain are highlighted in red. Substrates yielding growth efficiencies  $\geq 5\%$  are highlighted in green. Growth efficiencies were not determined for five of the listed substrates.

(DOCX)

**Table S4** Loci used for the multilocus sequence analysis of the eleven F10 lineage strains isolated from Pond-1 (compare Fig. 5). The sequences of Pnuc\_1240 were not considered for the calculation of the tree shown in Fig. 5.

(DOCX)

**Table S5** Loci used for the multilocus sequence analysis shown in Fig. 8.

(DOCX)

**Table S6** Genomes and strains considered for the multilocus sequence analysis presented in Fig. 8.

(DOCX)

**References**

- Zwart G, Crump BC, Kamst-van Agterveld MP, Hagen F, Han SK (2002) Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. *Aquat Microb Ecol* 28: 141–155.
- 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.
- Šimek K, Kasalický V, Zapomělová E, Horňák K (2011) Algal-derived substrates select for distinct betaproteobacterial lineages and contribute to niche separation in *Limmohabitans* strains. *Appl Environ Microbiol* 77: 7307–7315.
- Hahn MW, Pöckl M, Wu QL (2005) Low intraspecific diversity in a *Polynucleobacter* subcluster population numerically dominating bacterioplankton of a freshwater pond. *Appl Environ Microbiol* 71: 4539–4547.
- Jogler M, Siemens H, Chen H, Bunk B, Sikorski J, et al. (2011) Identification and targeted cultivation of abundant novel freshwater sphingomonads and analysis of their population substructure. *Appl Environ Microbiol* 77: 7355–7364.
- Jezbera J, Jezberová J, Brandt U, Hahn MW (2011) Ubiquity of *Polynucleobacter necessarius* subspecies *asymbioticus* results from ecological diversification. *Environ Microbiol* 13: 922–931.
- Vannini C, Pöckl M, Petroni G, Wu QL, Lang 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.
- Jaspers E, Overmann J (2004) Ecological significance of microdiversity: identical 16S rRNA gene sequences can be found in bacteria with highly divergent genomes and ecophysologies. *Appl Environ Microbiol* 70: 4831–4839.
- Hahn MW, Pöckl 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, Lang E, Brandt U, Wu QL, Scheuerl T (2009) Emended description of the genus *Polynucleobacter* and the species *P. necessarius* and proposal of two subspecies, *P. necessarius* subspecies *necessarius* subsp. nov. and *P. necessarius* subsp. *asymbioticus* subsp. nov. *Int J Syst Evol Microbiol* 59: 2002–2009.
- 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.
- Jezberová J, Jezbera J, Brandt U, Lindström ES, Langenheder S, et al. (2010) Ubiquity of *Polynucleobacter necessarius* ssp. *asymbioticus* in lentic freshwater habitats of a heterogeneous 2000 km<sup>2</sup> area. *Environ Microbiol* 12: 658–669.
- Crump BC, Kling GW, Bahr M, Hobbie JE (2003) Bacterioplankton community shifts in an arctic lake correlate with seasonal changes in organic matter source. *Appl Environ Microbiol* 69: 2253–2268.
- Pearce DA, van der Gast CJ, Lawley B, Ellis-Evans JC (2003) Bacterioplankton community diversity in a maritime antarctic lake, determined by culture-dependent and culture-independent techniques. *FEMS Microbiol Ecol* 45: 59–70.
- 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.
- Ghai R, Rodriguez-Valera F, McMahon KD, Toyama D, Rinke R, et al. (2011) Metagenomics of the water column in the pristine upper course of the Amazon River. *PLoS ONE* 6(8): e23785.
- Buck U, Grossart HP, Amann R, Pernthaler J (2009) Substrate incorporation patterns of bacterioplankton populations in stratified and mixed waters of a humic lake. *Environ Microbiol* 11: 1854–1865.
- Salcher MM, Pernthaler J, Zeder M, Penner R, Posch T (2008) Spatio-temporal niche separation of planktonic *Betaproteobacteria* in an oligo-mesotrophic lake. *Environ Microbiol* 10: 2074–2086.
- Martinez-Garcia M, Swan BK, Poulton NJ, Gomez ML, Masland D, et al. (2012) High-throughput single-cell sequencing identifies photoheterotrophs and chemoautotrophs in freshwater bacterioplankton. *ISME J* 6: 113–123.
- Boenigk J, Stadler P, Wiedroither A, Hahn MW (2004) Strain-specific differences in the grazing sensitivities of closely related ultramicrobacteria affiliated with the *Polynucleobacter* cluster. *Appl Environ Microbiol* 70: 5787–5793.
- Hahn MW (2004) Broad diversity of viable bacteria in 'sterile' (0.2 μm) filtered water. *Res Microbiol* 155: 688–691.
- Meincke L, Copeland A, Lapidus A, Lucas S, Berry KW, et al. (2012) Complete genome sequence of *Polynucleobacter necessarius* subsp. *asymbioticus* type strain (QLW-P1DMWA-1<sup>T</sup>). *Stand Genomic Sci* 6(1). In press.

**Table S7** Bootstrap values obtained in phylogenetic reconstructions using the concatenated amino acid (AA) or nucleotide sequences (nt) of eight housekeeping genes (Fig. 8). For comparison, phylogenetic reconstructions solely based on the 16S rRNA gene sequences of the same set of organisms were performed (tree not shown). Note the strong support for node E placing the *Polynucleobacter* strain next to the *Ralstonia/Cupriavidus* clade.

(DOCX)

**Text S1** Lack of genes well known from other *Burkholderiaceae* bacteria.

(DOCX)

**Text S2** Search for genes involved in adaptations to a planktonic lifestyle.

(DOCX)

**Text S3** Experiment on growth support of *Polynucleobacter* bacteria by algal exudates.

(DOCX)

**Acknowledgments**

We thank S.O. Jost for providing us with cultures of algae, M. Musso (University of Salzburg) for determination of the light transparency of various types of glass bottles, A. Camacho (University of Valencia) for chemical characterization of samples, A. Pitt (University of Salzburg) for providing phytoplankton data, J. Glaeser (University of Giessen) for constructive comments on an earlier version of the manuscript, and A. Baer (Banska Bystrica, Slovakia) for correcting the English in the manuscript.

**Author Contributions**

Conceived and designed the experiments: MWH TS J. Jezberova J. Jezbera KŠ. Performed the experiments: TS UK J. Jezberova J. Jezbera KS QLW MWH. Analyzed the data: MWH TS J. Jezberova J. Jezbera KŠ CV GP. Wrote the paper: MWH KŠ J. Jezbera J. Jezberova.

23. Wu QL, Hahn MW (2006) High predictability of the seasonal dynamics of a species-like *Polynucleobacter* population in a freshwater lake. *Environ Microbiol* 8: 1660–1666.
24. Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, et al. (2010) The integrated microbial genomes system: an expanding comparative analysis resource. *Nucleic Acids Res* 38: D382–D390.
25. Janssen PJ, van Houdt R, Moors H, Monsieurs P, Morin N, et al. (2010) The complete genome sequence of *Cupriavidus metallidurans* strain CH34, a master survivalist in harsh and anthropogenic environments. *PLoS ONE* 5: e10433.
26. Masuda S, Eda S, Ikeda S, Mitsui H, Minamisawa K (2010) Thiosulfate-dependent chemolithoautotrophic growth of *Bradyrhizobium japonicum*. *Appl Environ Microbiol* 76: 2402–2409.
27. Zepp RG, Wolfe NL, Baughman GL, Hollis RC (1977) Singlet oxygen in natural waters. *Nature* 267: 421–423.
28. Baxter RM, Carey JH (1983) Evidence for photochemical generation of superoxide ion in humic waters. *Nature* 306: 575–576.
29. Cooper WJ (1989) Sunlight-induced photochemistry of HS in natural waters: major reactive species. *Adv Chem Ser* 219: 332–362.
30. Anesio AM, Granéli W, Aiken GR, Kieber DJ, Mopper K (2005) Effect of humic substance photodegradation on bacterial growth and respiration in lake water. *Appl Environ Microbiol* 71: 6267–6275.
31. Glaeser SP, Grossart HP, Glaeser J (2010) Singlet oxygen, a neglected but important environmental factor: short-term and long-term effects on bacterioplankton composition in a humic lake. *Environ Microbiol* 12: 3124–36.
32. Coveney MF, Wetzel RG (1989) Bacterial metabolism of algal extracellular carbon. *Hydrobiol* 173: 141–149.
33. Wetzel RG (2001) *Limnology, lake and river ecosystems*. Third Edition. San Diego, San Francisco and London: Academic press. 1006 p.
34. Moran MA, Zepp RG (1997) Role of photoreactions in the formation of biologically labile compounds from dissolved organic matter. *Limnol Oceanogr* 42: 1307–1316.
35. Mykeltad SM (2000) Dissolved organic carbon from phytoplankton. In *Marine Chemistry* Wangersky PJ, ed 5: 111–148.
36. Paver SF, Kent AD (2010) Temporal patterns in glycolate-utilizing bacterial community composition correlate with phytoplankton population dynamics in humic lakes. *Microb Ecol* 60: 406–418.
37. Lau WWY, Armbrust EV (2006) Detection of glycolate oxidase gene *glcD* diversity among cultured and environmental marine bacteria. *Environ Microbiol* 8: 1688–1702.
38. Gorbushina AA, Kort R, Schulte A, Lazarus D, Schnetger B, et al. (2007) Life in Darwin's dust: intercontinental transport and survival of microbes in the nineteenth century. *Environ Microbiol* 9: 2911–2922.
39. Hervás A, Camarero L, Reche I, Casamayor EO (2009) Viability and potential for immigration of airborne bacteria from Africa that reach high mountain lakes in Europe. *Environ Microbiol* 11: 1462–2920.
40. Rainey PB, Travisano M (1998) Adaptive radiation in a heterogeneous environment. *Nature* 394: 69–72.
41. Barrick JE, Yu DS, Yoon SH, Jeong H, Oh TK, et al. (2009) Genome evolution and adaptation in a long-term experiment with *Escherichia coli*. *Nature* 461: 1243–1247.
42. Yooshep S, Nealon KH, Rusch DB, McCrow JP, Dupont CL, et al. (2010) Genomic and functional adaptation in surface ocean planktonic prokaryotes. *Nature* 468: 60–66.
43. Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, et al. (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* 309: 1242–1245.
44. Scanlan D, Ostrowski M, Mazard S, Dufresne A, Garczarek L, et al. (2009) Ecological genomics of marine picocyanobacteria. *Microbiol Mol Biol Rev* 73: 249–299.
45. Dufresne A, Garczarek L, Partensky F (2005) Accelerated evolution associated with genome reduction in a free-living prokaryote. *Genome Biol* 6: R14.
46. Lauro FM, McDougald D, Thomas T, Williams TJ, Egan S, et al. (2009) The genomic basis of trophic strategy in marine bacteria. *Proc Natl Acad Sci U S A* 106: 15527–15533.
47. Azam F (1998) Microbial control of oceanic carbon flux: the plot thickens. *Science* 280: 694–696.
48. Fenchel T (2002) Microbial behaviour in a heterogeneous world. *Science* 296: 1068–1071.
49. Tranvik LJ (1998) Degradation of dissolved organic matter in humic waters by bacteria. In: Hessen DO, Tranvik LJ, eds. *Aquatic humic substances: ecology and biogeochemistry*, Ecological Studies 133, Springer-Verlag, Berlin Heidelberg. pp 259–283.
50. Miller WL (1998) Effects of UV radiation on aquatic humus: Photochemical principles and experimental considerations. In: Hessen DO, Tranvik LJ, eds. *Aquatic humic substances: ecology and biogeochemistry*, Ecological Studies 133, Springer-Verlag, Berlin Heidelberg. pp 125–143.
51. Mopper K, Zhou X, Kieber RJ, Kieber DJ, Sikorski RJ, et al. (1991) Photochemical degradation of dissolved organic carbon and its impact on the oceanic carbon cycle. *Nature* 353: 60–62.
52. Kieber DJ, McDaniel J, Mopper K (1989) Photochemical source of biological substrates in sea water: implications for carbon cycling. *Nature* 341: 637–639.
53. Cooper WJ, Lean DRS, Carey JH (1989) Spatial and temporal patterns of hydrogen peroxide in lake waters. *Canadian Journal of Fisheries and Aquatic Sciences* 46: 1227–1231.
54. Briviba K, Klotz LO, Sies H (1997) Toxic and signaling effects of photochemically or chemically generated singlet oxygen in biological systems. *Biol Chem* 378: 1259–1265.
55. Cole JJ, Findlay S, Pace ML (1988) Bacterial production in fresh and saltwater ecosystems: A cross-system overview. *Mar Ecol Prog Ser* 43: 1–10.
56. Kasalický V, Jezbera J, Šimek K, Hahn MW (2010) *Limnolobus planktonicus* sp. nov., and *Limnolobus parvus* sp. nov., two novel planktonic *Betaproteobacteria* isolated from a freshwater reservoir. *Int J Syst Evol Microbiol* 60: 2710–2714.
57. Pérez MT, Sommaruga R (2006) Differential effect of algal- and soil-derived dissolved organic matter on alpine lake bacterial community composition and activity. *Limnol Oceanogr* 51: 2527–2537.
58. Šimek K, Horňák K, Jezbera J, Nedoma J, Znachor P, et al. (2008) Spatio-temporal patterns of bacterioplankton production and community composition related to phytoplankton composition and protistan bacterivory in a dam reservoir. *Aquat Microb Ecol* 51: 249–262.
59. Hunt DE, David LA, Gevers D, Preheim SP, Alm EJ, et al. (2008) Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* 320: 1081–1085.
60. Young KD (2006) The selective value of bacterial shape. *Microbiol Mol Biol Rev* 70: 660–703.
61. Hahn MW, Höfle MG (2001) Grazing of protozoa and its effect on populations of aquatic bacteria. *FEMS Microbiol Ecol* 35: 113–121.
62. Pernthaler J, Amann R (2005) Fate of heterotrophic microbes in pelagic habitats: focus on populations. *Microbiol Mol Biol Rev* 69: 440–461.
63. Weinbauer MG (2004) Ecology of prokaryotic viruses. *FEMS Microbiol Rev* 28: 127–181.
64. Sherr EB, Sherr BF (1993) Protistan grazing rates via uptake of fluorescently labelled prey. p. 695–701. In P. Kemp, BF. Sherr, EB. Sherr, J. Cole, eds. *Handbook of methods in aquatic microbial ecology*, Lewis, Boca Raton, Fla.
65. Šimek K, Pernthaler J, Weinbauer MG, Horňák K, Dolan JR, et al. (2001) Changes in bacterial community composition, dynamics, and viral mortality rates associated with enhanced flagellate grazing in a mesoeutrophic reservoir. *Appl Environ Microbiol* 67: 2723–2733.
66. Jezbera J, Horňák K, Šimek K (2005) Food selection by bacterivorous protists: insight from the analysis of the food vacuole content by means of fluorescent *in situ* hybridization. *FEMS Microb Ecol* 52: 351–363.
67. Hahn MW, Stadler P, Wu QL, Pöckl M (2004) The filtration–acclimatization method for isolation of an important fraction of the not readily cultivable bacteria. *J Microbiol Meth* 57: 379–390.
68. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731–2739.
69. Ulrich LE, Zhulin IB (2010) The MiST2 database: a comprehensive genomics resource on microbial signal transduction. *Nucleic Acids Res* 38: D401–407.
70. Ren Q, Chen K, Paulsen IT (2007) TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. *Nucleic Acids Res* 35: D274–D279.
71. Fry JC (1988) Determination of biomass. In: Austin B, ed. *Methods in aquatic bacteriology*, J Wiley & Sons Ltd, New York. pp 27–72.
72. Loferer-Krößbacher M, Klima J, Psenner R (1998) Determination of bacterial cell dry mass by transmission electron microscopy and densitometric image analyses. *Appl Environ Microbiol* 64: 688–694.
73. Šimek K, Horňák K, Jezbera J, Nedoma J, Vrba J, et al. (2006) Maximum growth rates and possible life strategies of different bacterioplankton groups in relation to phosphorus availability in a freshwater reservoir. *Environ Microbiol* 8: 1613–1624.
74. Lund JWG, Kipling C, Le Cren ED (1958) The inverted microscope method of estimating algal numbers and the statistical basis of estimation by counting. *Hydrobiol* 11: 14–70.
75. Colwell RK (2009) EstimateS: Statistical estimation of species richness and shared species from samples. Version 8.2. User's Guide and application published at: <http://purl.oclc.org/estimates>.



*Supplementary Materials for*

**The passive yet successful way of planktonic life: genomic and experimental analysis of the ecology of a free-living *Polynucleobacter (Betaproteobacteria)* population.**

Martin W. Hahn\*, Thomas Scheuerl, Jitka Jezberová, Ulrike Koll, Jan Jezbera, Karel Šimek,  
Claudia Vannini, Giulio Petroni, and Qinglong L. Wu

\* Corresponding author; email, [martin.hahn@oeaw.ac.at](mailto:martin.hahn@oeaw.ac.at)

**This supplementary information file includes (ordered by appearance in the paper):**

Supplementary Tables (S1-S7)

Supplementary Texts (S1-S3)

References for Supplementary Information

**Table S1.** Current taxonomy of *Polynucleobacter* bacteria. The genus *Polynucleobacter* currently harbors five species (Hahn et al., 2009, 2010, 2011a, 2011b, 2012). These five species represent the four *Polynucleobacter* tribes (PnecA to PnecD) suggested by the taxonomy in Newton et al., (2011). Some of the strains affiliated with the subspecies *P. necessarius asymbioticus* were recently assigned to lineages putatively representing differently adapted ecotypes (Jezbera et al., 2011). Lineages F10, F5 and F4 inhabit Pond-1 while other lineages and genotypes are absent from this habitat (Hahn et al., 2005). The populations of these three lineages dwelling in Pond-1 were previously described as G1, G2 and G3 populations (Hahn et al., 2005). The current study is focusing exclusively on the F10 lineage (G1) population of Pond-1. The subspecies *P. necessarius* ssp. *necessarius* contains exclusively endosymbionts of ciliates (Hahn et al., 2009), thus contains only strains fundamentally differing in lifestyle from all other currently established *Polynucleobacter* taxa (Vannini et al., 2007).

<i>Polynucleobacter acidiphobus</i>	(PnecB1)	} (PnecB)
<i>Polynucleobacter difficilis</i>	(PnecB2)	
<i>Polynucleobacter cosmopolitanus</i>		(PnecD)
<i>Polynucleobacter necessarius</i>		(PnecC)
<b><i>Polynucleobacter necessarius</i> subsp. <i>asymbioticus</i></b>		
<b>Lineage F10 (including G1)</b>		
Lineage F5 (including G2)		
Lineage F4 (including G3)		
Other lineages		
Strains not yet assigned to any lineage		
<i>Polynucleobacter necessarius</i> subsp. <i>necessarius</i>		
<i>Polynucleobacter rarus</i>		(PnecA)

## Table S2. Abbreviations

**ABC transporters**, ATP-binding cassette transporter, family of transporter proteins

**aka**, also known as

**Ala**, alanine

**Asp**, aspartate

**bp**, base pairs

**COG**, clusters of orthologous groups of proteins

**CRISPR**, clustered regularly interspaced short palindromic repeats

**CT value**, cycle threshold value

**DAPI**, 4',6-diamidino-2-phenylindole, a fluorescent dye

**DNA**, deoxyribonucleic acid

**DOC**, dissolved organic carbon

**DSM**, Deutsche Sammlung von Mikroorganismen (German Collection of Microorganisms and Cell Cultures, DSMZ)

**EC**, Enzyme Commission

**EEM**, fluorescence excitation emission matrix

**F10 lineage**, phylogenetic group within PnecC (see Supplementary Materials Table S1)

**FAM, 6FAM**, 6-carboxyfluorescein, a fluorescent dye

**FISH**, fluorescence *in situ* hybridization

**FLB**, fluorescently labeled bacteria

**G+C content** (= GC value), guanine-cytosine content of DNA (usually given as mol%)

**G1**, genotype 1, subgroup of F10 lineage (see Supplementary Materials Table S1)

**GE**, growth efficiency

**glcD**, glycolate oxidase gene

**glnA**, glutamine synthetase gene

**Glu**, glutamate

**Gly**, glycine

**His**, histidine

**HMW**, high-molecular-weight

**HNF**, heterotrophic nanoflagellates

**HS**, humic substances

**IBM**, inorganic basal medium

**IMG system**, Integrated Microbial Genomes system

**ITS**, intergenic transcribed spacer (aka, internal transcribed spacer)

**kb**, kilo base pairs ( $10^3$  bp)

**KEGG**, Kyoto Encyclopedia of Genes and Genomes

**LMW**, low-molecular-weight

**Lys**, lysine

**Mb**, mega base pairs ( $10^6$  bp)

**ML**, maximum likelihood

**MP**, maximum parsimony

**NJ**, neighbor-joining

**NSY medium**, nutrient broth-soytone-yeast extract medium

**OD**, optical density

**ORF**, open reading frame

**Orn**, ornithine

**PCR**, polymerase chain reaction

**Pfam**, database of protein families

**PnecA, PnecB, PnecC, PnecD**, phylogenetic clusters representing subgroups within the genus *Polynucleobacter* (see Supplementary Materials Table S1)

**PnecC**, *Polynucleobacter* cluster C, currently more or less identical with the species *P. necessarius*

**PTS**, phosphotransferase system

**QLW-P1DMWA-1**, strain designation of the genome sequenced bacterium

**qPCR**, quantitative PCR

**RLBH**, reverse line blot hybridization

**ROS**, reactive oxygen species

**rRNA**, ribosomal ribonucleic acid

**SAR11**, clade of marine *Alphaproteobacteria* including *Candidatus Pelagibacter ubique*

**Ser**, serine

**SOX genes**, genes of the SOX locus encoding parts of the sulfur-oxidizing (Sox) enzyme system first characterized in *Paracoccus*

**ssp.**, subspecies

**TAMRA, 5-TAMRA**, 5-carboxytetramethylrhodamine, a quencher

**TCA**, tricarboxylic acid cycle

**tRNA**, transfer ribonucleic acid

**UV**, ultra violet light

**VIS**, visible light



**Table S3.** Results of the assimilation experiments performed with 93 substrates. Substrates that did not support growth of the strain are highlighted in red. Substrates yielding growth efficiencies  $\geq 5\%$  are highlighted in green. Growth efficiencies were not determined for five of the listed substrates.

<b>Substrate</b>	<b>Growth efficiency (%)<sup>#</sup></b>
<b>Carbohydrates and derivatives</b>	
D-Glucose	0.4
D-Arabinose	0.0
D-Erythrose	0.0
D-Lyxose	1.0
D-Galactose	0.6
D-Mannose	1.2
D-Xylose	0.0
D-Fructose	0.6
D-Ribose	0.0
L-Fucose	0.5
L-Sorbose	0.0
L-Rhamnose	1.7
N-Acetyl-D-glucosamine	0.4
Pectin	+ <sup>&amp;</sup>
<b>Alcohols, sugar alcohols, sugar acids</b>	
D-Galacturonic acid	1.4
D-Gluconate	0.0
D-Sorbitol	0.4
Ethanol	0.0
Glycerate	0.0
Glycerol	0.0
Methanol	0.0
<b>Amino acids</b>	
Glycine	0.0
L-Alanin	0.0
L-Arginie	0.0
L-Asparagine	0.5
L-Aspartate	10.0
L-Cysteine	2.4
L-Glutamine	1.1
L-Glutamate	8.8
L-Histidine	0.0
L-Isoleucine	0.0
L-Leucine	0.0
L-Lysine	0.4
L-Methionine	0.0

L-Phenylalanine	0.0
L-Proline	0.0
L-Serine	0.0
L-Threonine	0.0
L-Tryptophan	0.0
L-Tyrosin	0.9
L-Valine	0.0

#### Fatty acids, carboxylic acids, keto acids

Acetate	7.9
3-Methylbutyrate	0.0
Butyrate	1.8
Capric acid	0.0
Caprylic acid	0.0
DL-Lactate	0.0
n-Caproic acid	0.0
Propionate	3.9
Formate	0.0
6-Phosphogluconate	0.0
Citrate	0.0
DL-Malate	12.0
Fumarate	9.4
Glycolate	0.0
Glyoxylate	2.1
Isobutyrate	3.8
Malonate	2.3
Oxalate	0.0
Oxaloacetate	0.0
Succinate	11.7
alpha-Ketoglutarate	9.2
Levulinate	1.0
Phosphoenolpyruvate	31.5
Pyruvate	21.0

#### Aldehydes

Acetaldehyde	3.9
Butyraldehyde	0.0
Formaldehyde	0.0
Glyoxal	0.0
Propionaldehyde	0.0

#### Ketones

Aceton	0.1
1,3-Dihydroxyacetone	0.0
2-Butanone	0.0
Methylglyoxal	0.0

#### Amines

Choline	0.0
Glycine betaine	0.8
L-Carnitine	0.0
Methylamine	0.0
Putrescine	0.0

Spermidine	0.0
<b>Nucleotides</b>	
Adenine	0.0
Cytosine	0.0
Guanin	0.0
Thymine	0.0
<b>Aromatic substances</b>	
Phenylacetate	3.2
Phenol	0.0
Benzoic acid	0.0
<b>Esters (experiments with 15 mg L<sup>-1</sup> substrate concentration)</b>	
Phenylacetate	+ <sup>§§</sup>
Ethylpyruvate	- <sup>§§</sup>
$\alpha$ -terpinylacetate	- <sup>§§</sup>
Isoamylacetate	- <sup>§§</sup>
Methylacetate	+ <sup>§§</sup>
<b>Humic substances and photooxidation products</b>	
Humic Substances	0.0
Photooxidation products of humic substances <sup>§</sup>	8.7
<b>Bacteriological complex medium</b>	
10% NSY	4.4

<sup>%</sup> carbon content unknown, assimilation of substance was observed

<sup>§</sup> a carbon content of 22% of dry weight was determined

<sup>§§</sup> growth efficiency was not determined

### **Supplementary Text S1. Lack of genes well known from other *Burkholderiaceae* bacteria.**

The genome of *Polynucleobacter necessarius* ssp. *asymbioticus* strain QLW-P1DMWA-1 lacks any genes for flagella-driven motility and chemotaxis, confirming the previously reported phenotypic lack of motility (Hahn et al., 2009). The genome also encodes no genes for harvesting light energy, such as genes for anoxygenic photosynthesis (Yurkov & Beatty, 1998) or proteorhodopsin genes (Beja et al., 2000). Furthermore, genes known from related species – such as genes for oxidizing molecular hydrogen (*Cupriavidus necator* H16, synonym *Ralstonia eutropha*, (Pohlmann et al., 2006); *C. metallidurans* CH34, (Janssen et al., 2010)), genes potentially involved in oxidization of carbon monoxide (e.g., *C. metallidurans*), genes for cleavage of aromatic compounds (e.g., *C. necator* JMP134, (Lykidis et al., 2010)) genes for nitrate respiration (e.g., *C. necator*), genes for nitrogen fixation (*C. taiwanense* LMG19424, (Amadou et al., 2008)), as well as genes for CO<sub>2</sub> fixation (*C. necator*) – are lacking. Furthermore, the genome seems to lack any ABC transporters for carbohydrates, as well as a phosphotransferase system (PTS). By contrast, most genomes of *Cupriavidus* and *Ralstonia* strains encode putative transporters for carbohydrates.

According to CRISPRFinder (Grissa et al., 2007) the genome of QLW-P1DMWA-1 lacks any CRISPR structures (Clustered Regularly Interspaced Short Palindromic Repeats). However, with the exception of three strains, such sequences are also largely absent in the genomes of other *Burkholderiaceae* species.

### **Supplementary Text S2. Search for genes involved in adaptations to a planktonic lifestyle.**

The comparison of the QLW-P1DMWA-1 genome with all genomes of *Cupriavidus* (n = 4) and *Ralstonia* strains (n = 5) available at the time of analysis in the IMG system identified 38 clusters of orthologous groups (COGs), which are overrepresented by gene numbers (7 COGs representing 23 genes) or exclusively present in the *Polynucleobacter* genome (31 COGs with 36 genes). The majority of COGs belong to the categories S (function unknown, 17 COGs, 17 genes), M (cell envelope biogenesis, outer membrane, 4 COGs, 10 genes), R (general function prediction only, 4 COGs, 8 genes), and C (energy production and conversion, 4 COGs, 4 genes). Among the 59 genes representing the 38 COGs an overrepresentation of transmembrane proteins (42.4% versus 25.1% in entire genome),



proteins with signal peptides (33.9% versus 22.5% in entire genome), and genes with either low ( $\leq 40\%$ ) or high ( $\geq 50\%$ ) GC content (in total 27.1% versus 9.3% of genes in genome) was observed. The latter may indicate a relatively recent horizontal gene transfer of at least some of the genes with rather atypical GC content. A few of the genes found in this analysis may represent prophage genes. None of the genes seem to indicate an advantageous adaptation to a planktonic lifestyle.

One interesting gene regarding adaptation to planktonic life was, however, overlooked by the COG as well as by a similar protein families (Pfam) analysis. QLW-P1DMWA-1 encodes an open reading frame (Pnuc\_1095) of a length (31290 bp) only rarely found among other bacteria (Reva & Tümmler, 2008). Among the 580 sequenced prokaryotic genomes only 14% (representing 47 taxa) encoded giant genes ( $> 20$  kb). Like other giant genes investigated by Reva & Tümmler, Pnuc\_1095 encodes a putative protein, which is acidic, threonine-rich, lacks cysteine and is characterized by amino acid repeats, however Pnuc\_1095 shares only low sequence identity with other giant genes. Analysis with the service PRED-TMMB (<http://biophysics.biol.uoa.gr/PRED-TMMB/>) indicates that Pnuc\_1095 encodes an outer membrane protein. This is also indicated by the putative presence of a C-terminal outer membrane autotransporter barrel. Comparative analysis of the gene sequence yielded no reliable hints on the function of the putatively encoded 10429 amino acids protein. Most of the giant genes identified so far were present in environmental bacteria, including several planktonic bacteria. For instance, *Candidatus Pelagibacter ubique*, some *Synechococcus* strains and a few *Prochlorococcus* strains encoded such giant genes (Reva & Tümmler, 2008; Scanlan et al., 2009). In one *Synechococcus* strain the encoded protein enables a flagella-independent motility. In most of the other cases, it is assumed that the giant genes encode surface proteins, which shield the cell against negative environmental influences, which may result in resistance to viral infection or protection against grazing (Scanlan et al., 2009).

### ***Supplementary Text S3. Experiment on growth support of Polynucleobacter bacteria by algal exudates.***

An experiment on potential utilization of algal exudates as major substrate source for *Polynucleobacter* bacteria originating from Pond-1 was conducted under laboratory conditions as described in the Materials and Methods section. A 0.8  $\mu\text{m}$ -filtered (predator-

free) water sample from Pond-1 was used for this experiment. On the day of sampling, PnecC bacteria comprised 33.5% of total bacterial numbers, and the F10 lineage population comprised 41.9 % of PnecC numbers in Pond-1 (Fig. 2). Relative abundance of PnecC bacteria changed slightly during the preparation of the experiment (Fig. 12). Subsamples of the filtered water sample were mixed with algal cultures and incubated in the light. All three algal species (representing species present in Pond-1) increased in numbers during the course of the experiment but peaked after different incubation periods. *Dinobryon* sp. showed the strongest growth, with increases in cell numbers from initially  $5.4 \times 10^3$  cells mL<sup>-1</sup> (t = 0 d) to  $104.0 \times 10^3$  cells mL<sup>-1</sup> (t = 8 d). After the peak, numbers fell slightly to  $84.7 \times 10^3$  cells mL<sup>-1</sup> at the end of the experiment (t = 15 d). *Synura* sp. numbers increased from initially  $45.6 \times 10^3$  to  $70.6 \times 10^3$  cells mL<sup>-1</sup> (t = 6 d) and fell slightly afterwards to  $67.4 \times 10^3$  cells mL<sup>-1</sup> (t = 15 d). *Ochromonas* sp. numbers increased over the entire duration of the experiment from initially  $13.5 \times 10^3$  to  $27.3 \times 10^3$  cells mL<sup>-1</sup>(t = 15 d). Total bacterial numbers changed little in the *Synura* and *Dinobryon* treatments during the experiment but increased in the *Ochromonas* treatment from initially  $4.6 \times 10^6$  cells mL<sup>-1</sup> to  $9.0 \times 10^6$  cells mL<sup>-1</sup> (t = 4 d) and decreased thereafter to  $3.0 \times 10^6$  cells mL<sup>-1</sup> until the end of the experiment. In all three algal treatments the final (t = 15 d) total bacterial numbers were lower than the initial numbers (57 %, 81 %, and 66 % of initial cell numbers in *Dinobryon*, *Synura*, and *Ochromonas* treatments, respectively) while the numbers increased in the algae-free control treatment (136 % of initial numbers). The decrease of bacterial numbers in the *Dinobryon* and the *Ochromonas* treatment but not in the *Synura* treatment could have partially been caused by potential mixotrophic capabilities of the two former algae. By contrast, *Synura* is completely lacking any phagotrophic capabilities. Apart from the control treatment, the relative abundance of *Polynucleobacter* bacteria did not increase significantly during the experiment (Fig. 12). In contrast to the control treatment, relative PnecC numbers steadily decreased in the *Dinobryon* and *Synura* treatments. Solely the *Ochromonas* treatment showed a more or less unchanged proportion of PnecC bacteria.

**Table S4.** Loci used for the multilocus sequence analysis of the eleven F10 lineage strains isolated from Pond-1 (compare Fig. 5). The sequences of Pnuc\_1240 were not considered for the calculation of the tree shown in Fig. 5.

Gene/Spacer	Gene tag	Fragment length <sup>§</sup> (nt)	# Polymorphic sites	Position in genome
gyrB	Pnuc_0003	873	-	2851-5349
16S-23S ITS <sup>§</sup>	-	506	-	38362-38868
icd2	Pnuc_0366	878	8	360802-363036
hemF	Pnuc_0616	742	-	604744-605676
ilvD	Pnuc_0884	428	-	876398-878146
-	Pnuc_1095	1676	7	1102872-1134161
-	Pnuc_1239	298	-	1302926-1304254
-	Pnuc_1240	-	insertion of IS	1304359-1305312
-	Pnuc_1241	264	-	1305340-1306224
glnA	Pnuc_1255	615	-	1320425-1321840
iscS	Pnuc_1493	730	-	1568270-1569529
aspS	Pnuc_1892	995	-	1967417-1969216
-	Pnuc_2073	930	-	2140109-2142313

<sup>§</sup> intergenic transcribed spacer (ITS) including tRNA-Ile and tRNA-Ala genes

<sup>§</sup> fragments used for multi locus sequence analysis

**Table S5.** Loci used for the multilocus sequence analysis shown in Fig. 8.

Gene	Encoded product	Locus tag <sup>§</sup>	Gene length <sup>§</sup> (nt)	Position in genome <sup>§</sup>
dnaA	Chromosomal replication initiator protein	Pnuc_0001	1428	33
gyrB	DNA gyrase subunit B	Pnuc_0003	2499	2851
atpD	F1 sector of membrane-bound ATP synthase	Pnuc_0024	1542	21873
rpoB	DNA-directed RNA polymerase subunitB	Pnuc_0046	4101	47471
gltA	Citrate synthase	Pnuc_0763	1314	763471
glnA	Glutamine synthetase	Pnuc_1255	1416	1320425
lpxB	Lipid A disaccharide synthetase	Pnuc_1438	1206	1508834
recA	Recombination protein RecA	Pnuc_1833	1092	1913680

<sup>§</sup> genome annotation of QLW-P1DMWA-1

<sup>§</sup> genome sequence of QLW-P1DMWA-1 (total length 2159490 nt)



**Table S6.** Genomes and strains considered for the multilocus sequence analysis presented in Fig. 8.

Species (synonym)	Strain	Family	Accession numbers
<i>Polynucleobacter necessarius</i> ssp. <i>asymbioticus</i>	QLW-P1DMWA-1 <sup>T</sup>	<i>Burkholderiaceae</i>	NC_009379
<i>Burkholderia cepacia</i>	383 (= R18194)	<i>Burkholderiaceae</i>	NC_007509 - NC_007511
<i>Burkholderia multivorans</i>	ATCC 17616	<i>Burkholderiaceae</i>	NC_010801, NC_010802, NC_010804, NC_010805
<i>Burkholderia xenovorans</i> ( <i>B. fungorum</i> )	LB400	<i>Burkholderiaceae</i>	NC_007951 - NC_007953
<i>Ralstonia pickettii</i>	12J	<i>Burkholderiaceae</i>	NC_010678, NC_010682, NC_010683
<i>Ralstonia solanacearum</i>	GMI1000	<i>Burkholderiaceae</i>	NC_003295, NC_003296
<i>Cupriavidus necator</i> ( <i>Ralstonia eutropha</i> )	H16	<i>Burkholderiaceae</i>	NC_008313, NC_008314
<i>Cupriavidus</i> cf. <i>necator</i> ( <i>Ralstonia eutropha</i> )	JMP134	<i>Burkholderiaceae</i>	NC_007336, NC_007347, NC_007348
<i>Cupriavidus taiwanensis</i>	LMG19424	<i>Burkholderiaceae</i>	NC_010528 - NC_010530
<i>Cupriavidus metallidurans</i>	CH34	<i>Burkholderiaceae</i>	NC_007971 - NC_007974
<i>Comamonas testosteroni</i>	KF-1	<i>Comamonadaceae</i>	NZ_AAUIJ00000000

**Table S7.** Bootstrap values obtained in phylogenetic reconstructions using the concatenated amino acid (AA) or nucleotide sequences (nt) of eight housekeeping genes (Fig. 8). For comparison, phylogenetic reconstructions solely based on the 16S rRNA gene sequences of the same set of organisms were performed (tree not shown). Note the strong support for node E placing the *Polynucleobacter* strain next to the *Ralstonia/Cupriavidus* clade.

Node	Eight concatenated genes (AA)			Eight concatenated genes (nt)			16S rRNA gene		
	(NJ <sup>&amp;</sup> )	(MP)	(ML <sup>§</sup> )	(NJ <sup>§</sup> )	(MP)	(ML <sup>§</sup> )	(NJ <sup>§</sup> )	(MP)	(ML <sup>§</sup> )
A	100	99	100	100	100	100	40	-	-
B	100	97	99	100	99	99	44	-	-
C	100	100	100	100	100	100	79	38	33
D	100	100	100	100	100	100	95	-	40
E	100	100	100	100	100	100	100	100	98
F	66	-	67	-	-	60	100	-	99
G	100	100	100	100	100	100	100	99	98
H	100	100	100	100	100	100	100	95	46

<sup>§</sup> using the best model (GRT+I+G) as identified by MEGA5 for ML, gamma distribution with 5 categories

<sup>&</sup> gamma distribution, 5 categories, pair-wise gap deletion

<sup>§</sup> model rtREV+G+I+F

## References for Supplementary Materials

**Amadou C, Pascal G, Mangenot S, Glew M, Bontemps C, et al. (2008)** Genome sequence of the beta-rhizobium *Cupriavidus taiwanensis* and comparative genomics of rhizobia. *Genome Res* 18: 1472-1483.

**Beja O, Aravind L, Koonin EV, Suzuki MT, Hadd A, et al. (2000)** Bacterial rhodopsin: Evidence for a new type of phototrophy in the sea. *Science* 289: 1902–1904.

**Grissa I, Vergnaud G, Pourcel C (2007)** CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* 35: W52-W57.

- Hahn MW, Lang E, Brandt U, Wu QL, Scheuerl T (2009)** Emended description of the genus *Polynucleobacter* and the species *P. necessarius* and proposal of two subspecies, *P. necessarius* subspecies *necessarius* subsp. nov. and *P. necessarius* subsp. *asymbiomaticus* subsp. nov. Int J Syst Evol Microbiol 59: 2002-2009.
- Hahn MW, Lang E, Brandt U, Lünsdorf H., Wu QL, et al. (2010)** *Polynucleobacter cosmopolitanus* sp. nov., free-living planktonic bacteria inhabiting freshwater lakes and rivers. Int J Syst Evol Microbiol 60: 166-173.
- Hahn MW, Lang E, Tarao M, Brandt, U (2011a)** *Polynucleobacter rarus* sp. nov., a free-living planktonic bacterium isolated from an acidic lake. Int J Syst Evol Microbiol 61: 781-787.
- Hahn MW, Lang E, Brandt U, Spröer C (2011b)** *Polynucleobacter acidiphobus* sp. nov., a representative of an abundant group of planktonic freshwater bacteria. Int J Syst Evol Microbiol 61: 788-794.
- Hahn MW, Minasyan A, Lang E, Koll U, Spröer C (2012)** *Polynucleobacter difficilis* sp. nov., a planktonic freshwater bacterium affiliated with subcluster B1 of the genus *Polynucleobacter*. Int J Syst Evol Microbiol. *in press*.
- Lykidis A, Pérez-Pantoja D, Ledger T, Mavromatis K, Anderson IJ, et al. (2010)** The complete multipartite genome sequence of *Cupriavidus necator* JMP134, a versatile pollutant degrader. PLoS ONE 5: e9729.
- Pohlmann A, Fricke W, Reinecke F, Kusian B, Liesegang H, et al. (2006)** Genome sequence of the bioplastic-producing Knallgas bacterium *Ralstonia eutropha* H16. Nature Biotechnol 24: 1257–1262.
- Reva O, Tümmler B (2008)** Think big – giant genes in bacteria. Environ Microbiol. 10: 768-77.
- Yurkov VV, Beatty JT (1998)** Aerobic anoxygenic phototrophic bacteria. Microbiol Mol Biol Rev 62: 695–724.