Published in final edited form as:
Int J Syst Evol Microbiol. 2017 December ; 67(12): 5087–5093. doi:10.1099/ijsem.0.002421.

Reclassification of a *Polynucleobacter cosmopolitanus* strain isolated from tropical Lake Victoria as *Polynucleobacter victoriensis* sp. nov.

Martin W. Hahn[1], **Johanna Schmidt**[1], **Grace Ssanyu Asiyo**[2], **Nikos C. Kyrpides**[3], **Tanja Woyke**[3], and **William B. Whitman**[4]

[1]Research Department for Limnology, University of Innsbruck, Mondseestrasse 9, A-5310 Mondsee, Austria [2]Department of Biological Sciences, Kyambogo University, P.O Box 1, Kyambogo, Uganda [3]DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598-1698, USA [4]Department of Microbiology, 527 Biological Sciences Building, University of Georgia, Athens, GA

Abstract

The genus *Polynucleobacter* (*Burkholderiaceae*) is phylogenetically subdivided into at least four subclusters. One of those, i.e. subcluster PnecC, was recognized as a cryptic species complex. Here we test by comparative genome analyses if subcluster PnecD currently solely represented by the species *P. cosmopolitanus* also represents such a cryptic species complex. The genome sequences of the two *P. cosmopolitanus* strains MWH-MoIso2^T and MWH-VicM1, were determined. The latter strain was also characterized in the previous description of *P. cosmopolitanus*. These two strains originate from a temperate lake located in Austria and from the large tropical Lake Victoria located in East Africa, respectively. Strains MWH-MoIso2^T and MWH-VicM1 possess quite small genomes of 1.78 and 1.63 Mbp, respectively, and share similar G+C values of 44.1 and 43.1 mol%, respectively. Both strains encode only a single copy of the ribosomal operon, and their 16S rRNA genes differ only in four positions equalling a sequence similarity of 99.74 %. Both genomes possess characteristics indicating evolutionary genome streamlining such as high coding densities of 93.9 and 94.6 % of bases, respectively. Genome comparisons based on average nucleotide identity (ANI) values of the two strains resulted in ANI values of 78.4 %, suggesting that each of the strains represents a separate species. Our investigation suggests that PnecD represents an additional cryptic species complex within the genus *Polynucleobacter* that was not resolved by 16S rRNA gene sequence analyses. We propose reclassification of strain MWH-VicM1^T (=DSM 21486^T=JCM 32005^T) as *Polynucleobacter victoriensis* sp. nov.

DDBJ/EMBL/GenBank accession numbers
Polynucleobacter victoriensis sp. nov. strain MWH-VicM1^T: AJ550651 (16S rRNA gene) and FYEX00000000.1 (whole genome)

Conflicts of Interest
The authors declare the absence of any conflict of interest.

Ethical Statement
The presented study does not include any experimental work with humans or vertebrates.

Europe PMC Funders Author Manuscripts

The genus *Polynucleobacter* and the species *Polynucleobacter necessarius* were described by Klaus Heckmann and Helmut J. Schmidt as obligate endosymbionts of ciliates [1]. Diversity studies by cultivation-independent methods revealed that bacteria closely related to *P. necessarius* are typically present in the water column of lakes [2], ponds [3] and running waters [4], and that these bacteria frequently occur in such systems with high cell numbers [5]. Isolation of strains [6], investigations of environmental samples by fluorescent *in situ* hybridization (FISH) [3, 7], and cultivation experiments with endosymbiotic *P. necessarius* [8] indicated that *Polynucleobacter* bacteria dwelling in the water column of freshwater systems represent free-living bacteria and not obligate symbionts of ciliates [9]. Recent genome comparisons between an endosymbiotic and a free-living *Polynucleobacter* strain revealed lifestyle-specific genomic signatures, including, for instance, the presence of a large number of pseudogenes and about 30% reduction in genome size of the obligate endosymbiont [10].

A large number of *Polynucleobacter* strains could be isolated from freshwater systems and were characterized by sequencing of phylogenetic markers [6, 11, 12]. Based on phylogenetic analyses of 16S rRNA genes, the strains were grouped into the subclusters PnecA, PnecB (including PnecB1 and PnecB2), PnecC and PnecD [3, 6]. These subclusters are characterized by inter- and intra-subcluster 16S rRNA sequence similarity values of < 98 % and ≥ 99 %, respectively [6, 9]. Recent investigations demonstrated that subcluster PnecC represents a cryptic species complex of currently seven described species and an unknown but presumably large number of undescribed species [13–16], which cannot be resolved by analysis of 16S rRNA gene sequences. Currently, it is not known if the other three *Polynucleobacter* subclusters also represent such cryptic species complexes.

Subcluster PnecD is currently represented exclusively by the species *P. cosmopolitanus* [17]. The description of this species included the characterization of five strains forming a monophyletic lineage and sharing 16S rRNA sequence similarities ≥ 99.3 %. Due to generally weak growth of all *Polynucleobacter* strains in artificial medium (see Suppl. Mat. Fig. S2 in [9]) and the difficulty in extracting sufficient genomic DNA for pairwise DNA-DNA reassociation experiments, investigations on genomic coherence of the five strains were omitted. Instead the five strains were preliminarily described as members of the species *P. cosmopolitanus* [17].

In the study presented here, we investigated the genomic similarity of two strains previously included in the species *P. cosmopolitanus* [17]. These strains are the type strain MWH-MoIso2^{T} isolated from Lake Mondsee, Austria, and the tropical strain MWH-VicM1 isolated from Lake Victoria near Kampala, Uganda [6]. Both strains were cultivated by using the filtration-acclimatization method and NSY medium [18] as described previously [6]. A comprehensive phenotypic and chemotaxonomic characterization of both strains was included in the previous species description [17]. Based on large genomic differences revealed by comparative genome analyses, we conclude that the two strains represent different species and propose to establish for strain MWH-VicM1^{T} the new species *P. victoriensis* sp. nov.

Genomic characterization of strains MWH-MoIso2^T and MWH-VicM1^T

Genomic DNA of both strains used for genome sequencing, respectively, was extracted from biomass grown in liquid NSY medium as described previously [19]. DNA of strain MWH-MoIso2^T was sequenced by Illumina MiSeq. Paired-end sequencing (2 x 150 bp) of a fragment library resulted in about 1.2 x 10^6 quality filtered reads with a mean length of 148 bp. Sequence assembly and subsequent closure of some gaps resulted in six contigs of a total length of 1.78 Mbp (Table 1). Sequencing coverage was about 100-fold.

Strain MWH-VicM1^T was sequenced at the DOE-Joint Genome Institute as part of the Genomic Encyclopedia of Type Strains, Phase III (KMG-III) study [20] using the Illumina HiSeq–2000 1TB platform. Paired-end sequencing (2 x 150 bp) of a fragment library resulted in about 7.4 x 10^6 quality filtered reads. Assembly of reads resulted in three contigs with a total sequence length of 1.63 Mbp and a sequencing coverage of about 680-fold.

The obtained genome sequences of the two strains were annotated using the IMG/ER annotation pipeline [21]. The IMG Genome IDs of the two genomes are 2642422582 and 2710264786. Both genome sequences were also deposited in DDBJ/EMBL/GenBank (Accession Numbers NJGG00000000 and FYEX00000000). Gene finding using the JGI annotation pipeline [22, 23] resulted in 1822 and 1677 open reading frames in strains MWH-MoIso2^T and MWH-VicM1^T, respectively (Table 1).

Despite both genomes are not closed, the obtained genome sizes are expected to be quite close to the real genome size of the two strains, respectively. Due to the high sequencing coverage, the whole or almost the whole genome of the strains should be represented by reads. Previous independent sequencing of two very closely related *P. asymbioticus* strains isolated from the same habitat with a time interval of four years and differing only in 22 single nucleotide polymorphisms clearly resulted in the same genome size despite different sequencing technologies and strategies were used [24]. On the other hand, many *Polynucleobacter* strains contain a few repetitive sequences, which includes insertion elements, giant genes with repetitive sequence elements and the translation elongation factor Tu gene found to be present in all so far investigated *Polynucleobacter* genomes with two identical copies. Such repetitive sequences frequently do not assemble properly and form contigs in the size range of 0.1 – 2.0 kbp. Such contigs originating from repetitive sequences are usually characterized by coverage values of two, three or four-fold of the average coverage of the whole genome. Contigs smaller than 1 kbp are discarded by the IMG annotation pipeline, thus do not contribute to the final genome size values of draft genomes. In general, such repetitive sequences not assembling properly, result in underestimations of the real genome size, however, analyses of several *Polynucleobacter* genome assemblies suggest that the underestimation is usually in the range of about 1 kbp to about 5 kbp, which equals in typical *Polynucleobacter* genomes in underestimations of genome size by less than 0.4%. Thus, the genome size data given in Table 1 are assumed to represent quite well the genome sizes of the investigated strains.

Among the free-living *Polynucleobacter* strains with sequenced genomes, strain MWH-VicM1^T has the smallest genome size. Both genomes show signatures of evolutionary

genome streamlining previously described for another free- living *Polynucleobacter* strain [25] but lack large numbers of pseudogenes, a known signature for reductive genome evolution in obligate endosymbionts [10, 26]. Genome streamlining in these two investigated PnecD strains is indicated, for instance, by the high coding densities of the genomes of strains MWH-MoIso2^T and MWH-VicM1^T of 93.9 and 94.6 %, respectively. Such high coding densities were not reported, so far, for other members of the family *Burkholderiaceae* not affiliated with the genus *Polynucleobacter*. The IMG system [21] contains currently (March 2017) 811 genomes of *Burkoholderiaceae* (single cell genomes were excluded) bacteria not affiliated with the genus *Polynucleobacter*. Among these genomes, only seven genomes (0.9 %) had coding densities > 90 %, and the average coding density was 85.1 %. By contrast, genomes of free-living *Polynucleobacter* strains usually possess coding densities > 92 % [13, 14, 16, 19, 24, 25, 27].

The two genomes presented here are the first *Polynucleobacter* genomes of strains affiliated with subcluster PnecD [6]. All previously published *Polynucleobacter* genomes represent strains belonging to subcluster PnecC [13, 14, 16, 24, 25]. Sizes of the two PnecD genomes presented in this study and the eighteen previously published genomes of free-living *Polynucleobacter* strains affiliated with subcluster PnecC (Fig. 1 and [24]) differ sizes. All published genomes of free-living PnecC strains are characterized by genome sizes > 2 Mbp, while the two PnecD genomes possess sizes of ≤ 1.78 Mbp (Table 1). Unpublished genome size data of PnecC strains disproof a systematic difference in genome size between PnecD and PnecC strains (Hahn et al. unpublished data). Interestingly, the genome of the free-living (planktonic) strain MWH-VicM1^T is only 66 kbp larger than the genome of the obligate endosymbiont *P. necessarius* STIR1 [10] and is even smaller than genomes of a few other endosymbiotic *Polynucleobacter* strains recently presented [26]. This suggests that genome size in obligate endosymbiotic and free-living *Polynucleobacter* strains are not necessarily very different. Thus genome size is neither a systematic trait discriminating all PnecC and PnecD strains, nor discriminates all free-living and endosymbiotic *Polynucleobacter* strains.

Both of the PnecD strains lack a gene cluster encoding an anoxygenic photosynthesis system previously found in the type strains of *P. duraquae* [14] and *P. wuianus* [16]. Interestingly, *P. cosmopolitanus* strain MWH-MoIso2^T encodes a special variant of rhodopsins (xanthorhodopsins) which is also found in the obligate methylotroph "*Candidatus* Methylopumilus turicensis" [28] with an amino acid identity of 82%. These xanthorhodopsins are light driven proton pumps tuned to green light, which could represent an adaptation to the light conditions in more productive systems [29]. Rhodopsin genes were only found in one [27] of the eighteen genomes of planktonic *Polynucleobacter* strains and in none of the endosymbiotic strains [10, 26] investigated previously. Strain MWH-MoIso2^T encodes flagella, which strain MWH-VicM1^T lacks. Both strains encode an ABC-type Fe^{3+} transport system but no feoAB transporters for uptake of Fe^{2+} ions. This reflects very well the alkaline pH conditions of their habitats [6, 13]. Neither strain encodes a cytochrome bd-I terminal oxidase (CydAB) [14–16] or a fumarate reductase (FrdABCD) found in some PnecC genomes [14–16]. Thus, both PnecD strains lack these features typically suggesting adaptation to low oxygen concentrations and facultative anoxybiosis, respectively.

Analysis of genome similarity between *P. cosmopolitanus* MWH-MoIso2^T and strain MWH-VicM1^T by the average nucleotide identity (ANI) [30, 31] using the IMG system [21] resulted in a value of 78 % ANI (Fig. 1). By contrast, pairwise ANI comparisons with genomes of PnecC strains resulted in values of about 72 % ANI, and ANI values shared by strain MWH-VicM1^T with the type strains of two *Cupriavidus* species were about 70%. The alignment fractions from which these ANI values resulted were 78%, 41-46% and 12-15% of the genome size of strain MWH-VicM1^T for the comparisons with strain MWH-MoIso2^T (intra PnecD comparison), for comparisons with PnecC strains, and for comparisons with *Cupriavidus* spp. strains, respectively. Both, the ANI values [24] and the alignment fraction data indicate that strain MWH-VicM1^T shares more homologous genes with the other PnecD strain than with PnecC strains or the closest relatives outside of the genus *Polynucleobacter*.

Phylogeny

As shown previously, phylogenetic reconstructions based on 16S rRNA gene sequences placed strains MWH-MoIso2^T and MWH-VicM1^T in subcluster PnecD of the genus *Polynucleobacter* [6], however the sequences of this ribosomal gene are too similar (99.74%) to provide a suitable phylogenetic resolution (Fig. 1B). In order to establish a phylogenetic reconstruction with a higher resolution, a multilocus sequence analysis based on eight housekeeping genes defined previously [14, 27] was performed. In all but one case (endosymbiotic *P. necessarius* strain Ammermann, [14]), these genes were extracted from the genomes of the type strains of all previously described PnecC species, from two PnecC strains representing two separate undescribed PnecC species [32, 33], and the two PnecD strains investigated here. The genomes of two *Cupriavidus* strains [34, 35] served as an outgroup. Concatenation of partial gene sequences resulted in an alignment length of 6249 alignment positions. Neighbour joining (NJ), maximum likelihood (ML), and maximum parsimony (MP) trees were calculated by using the software MEGA7 [36] (Fig. 1B). The obtained multilocus trees confirm the separation of the investigated strains in the two subclusters, PnecC and PnecD, and demonstrate that the phylogenetic distance between the two PnecD strains is similar to distances between distinct PnecC species.

Ecology and biogeography

Strain MWH-VicM1^T was isolated from a tropical lake. Like PnecC strains also isolated from tropical freshwater systems [37] it lacked the ability to grow at temperatures of 4-6°C. On the other hand, the strain showed a higher maximum growth temperature (38°C) than all other PnecD or PnecC strains isolated from habitats located in temperate climatic zones [9, 15–17, 27, 38–40]. Both traits hint on a thermal adaptation of strain MWH-VicM1^T to tropical or at least to warmer climate conditions [41]. For tropical PnecC strains a restricted geographic range was indicated by cultivation-independent investigations [37]. The clade formed by these tropical strains could not be detected in Central European habitats but was detected in Ugandan habitats. Similarities between thermal adaptation and biogeography of PnecC strains and strain MWH-VicM1^T are obvious. However, characterization of additional strains of the new species are needed to determine if the geographic range of the new species is really restricted in a similar way as in the PnecC strains.

Cryptic species complex

The ANI results of the comparisons of the genomes of strains MWH-MoIso2^T and MWH-VicM1^T stand in strong contrast to the sequence similarity of their 16S rRNA genes (Fig. 1). A very similar discrepancy between ANI values and 16S rRNA gene similarities is well documented for several strains affiliated with subcluster PnecC [13, 27]. Only two of the five strains included in the description of *P. cosmopolitanus* [17] where genome sequenced so far, thus, one can only speculate if the remaining three strains also represent separate species. A recent investigation on four free-living *Polynucleobacter* strains included in the emended description of *P. necessarius* [9] revealed that these strains represent four separate species [14]. Sequence comparisons on partial glutamine synthetase (*glnA*) gene sequences may provide hints on whether or not the three other *P. cosmopolitanus* strains [17] actually represent novel species. A previous investigation on PnecC strains suggested that strains sharing ANI values < 85% also share *glnA* sequence similarities < 94% [13]. Analyses of the five *glnA* sequences (accession numbers FR732027, FR732028, FR732031, FR732033, and FR821089) of PnecD strains included in the previous description of the species *P. cosmopolitanus* resulted in an average sequence similarity of 92.3 %, a minimum value of 88.9 % and a maximum value of 97.7 %. All of the ten pairwise comparisons but the maximum value were below 94% sequence similarity. The partial *glnA* sequences of strains MWH-MoIso2^T and MWH-VicM1^T shared a similarity of 92.7%, thus slightly above the average value of all comparisons. These results indicate that it is quite likely that not only strains MWH-MoIso2^T and MWH-VicM1^T represent separate species and support the assumption that subcluster PnecD represents a second cryptic species complex within the genus *Polynucleobacter*.

Proposal of the new species *Polynucleobacter victoriensis* sp. nov.

The performed ANI analyses and the phylogenetic reconstructions show that the previously described species *P. cosmopolitanus* contains at least two distinct species. The determined value of 78.4 % ANI (Fig. 1A) lies far below the threshold range of 93-96 % ANI suggested to demarcate strains representing separate species [31, 42, 43]. Furthermore, comparison of the revealed phylogenetic distances between the two investigated PnecD strains with distances between distinct species within subcluster PneC confirm the separation of the two strains in two species (Fig. 1A). We propose to establish the new species *P. victoriensis* sp. nov. to harbour strain MWH-VicM1^T. Future genome sequence-based investigations have to reveal if the other three strains included in the description of *P. cosmopolitanus* really belong to this species, or if they also belong to the new species proposed here, or if they represent separate, so far undescribed species

Strain MWH-VicM1^T can be discriminated from *Polynucleobacter* strains not affiliated with subcluster PnecD by the presence of the fatty acid $C_{12:0}$ 3-OH, which was in *Polynucleobacter* strains so far detected exclusively in all investigated strains affiliated with subcluster PnecD [17]. A genotypic trait of all so far investigated strains affiliated with subcluster PnecD is the presence of the signature sequence 5’-AA(T/G)CCCT(A/T)AGGGGGAAA-3’ within the 16S rRNA gene (*E. coli* positions 181–197) [44]. Strain MWH-VicM1^T can be discriminated from the type strain of *P. cosmopolitanus*, which also

belongs to subcluster PnecD, by its ability to assimilate malonic acid, as well as by growth at 38 °C and lack of growth at 5 °C [17].

Description of *Polynucleobacter victoriensis* sp. nov.

Polynucleobacter victoriensis (vic.to.ri.en'sis. N.L. masc. adj. *victoriensis* of or belonging to Lake Victoria, the lake from which the type strain was isolated).

The description is based on phenotypical data of Hahn (2003) [6] and Hahn et al. (2010) [17], on chemotaxonomical data of Hahn et al. (2010) [17] and on genomic data presented in this study (Table 1). Contains free-living *Polynucleobacter* strains dwelling in the water column of freshwater systems. Cells are short curved rods, 0.4 - 1.1 μm in length and 0.3 – 0.5 μm in width, depending on cultivation conditions. Chemo-organotrophic, aerobic, weak anaerobic growth was observed. Colonies grown on NSY agar are non-pigmented, circular and convex with smooth surface. Growth occurs up to 38 °C but not at 5 °C. Growth occurs in 0 – 0.5% (w/v) NaCl. Assimilates acetate, propionate, pyruvate, malate, malonate, fumarate, succinate, oxaloacetate, L-alanine, and L-cysteine. Weak assimilation of D-galactose, and D-galacturonic acid. Does not assimilate glycolate, glyoxylate, oxalate, citrate, levulinate, D-fructose, D-fucose, D-glucose, D-lyxose, D-mannose, D-sorbitol, L-aspartate, L-asparagine, L-glutamate, L-serine, or betaine. Major fatty acids of the strain are summed feature 3 including $C_{16:1}$ ω7c and iso-$C_{15:1}$-2OH, $C_{18:1}$ ω7c, $C_{16:0}$ and $C_{12:0}$ 3-OH. The type strain is MWH-VicM1^T (=DSM 21486^T = JCM 32005^T), which was isolated from Lake Victoria near Kampala, Uganda. The species epithet indicates the origin of the type strain but does not indicate that the distribution of the taxon is restricted to a certain geographic area or a certain freshwater system. The draft genome of the type strain is characterized by a size of 1.63 Mbp and a G+C content of 43.1 mol%.

Acknowledgements

Funding Information

This study was supported by the Austrian Science Fund (FWF) project I482-B09 and the European Science Foundation (ESF) project FREDI. Genome sequencing of strain MWH-VicM1^T was conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Abbreviations

ANI	average nucleotide identity
FISH	fluorescent in situ hybridization
NSY medium	nutrient broth soytone yeast extract medium
PnecD	subcluster PnecD of the genus Polynucleobacter
glnA	glutamine synthetase gene
n.a.	not available

Europe PMC Funders Author Manuscripts

References

1. Heckmann K, Schmidt HJ. *Polynucleobacter necessarius* gen. nov., sp. nov., an obligately endosymbiotic bacterium living in the cytoplasm of *Euplotes aediculatus*. Int J Syst Bacteriol. 1987; 37:456–457.
2. Bahr M, Hobbie JE, Sogin ML. Bacterial diversity in an arctic lake: A freshwater SAR11 cluster. Aquat Microb Ecol. 1996; 11:271–277.
3. Wu QL, Hahn MW. Differences in structure and dynamics of *Polynucleobacter* communities in a temperate and a subtropical lake, revealed at three phylogenetic levels. FEMS Microbiol Ecol. 2006; 57:67–79. [PubMed: 16819951]
4. Crump BC, Peterson BJ, Raymond PA, Amon RMW, Rinehart A, et al. Circumpolar synchrony in big river bacterioplankton. Proc Natl Acad Sci USA. 2009; 106:21208–21212. [PubMed: 19940248]
5. Jezberova J, Jezbera J, Brandt U, Lindstrom ES, Langenheder S, et al. Ubiquity of *Polynucleobacter necessarius* ssp. *asymbioticus* in lentic freshwater habitats of a heterogenous 2000 km^2 area. Environ Microbiol. 2010; 12:658–669. [PubMed: 20041938]
6. Hahn MW. Isolation of strains belonging to the cosmopolitan *Polynucleobacter necessarius* cluster from freshwater habitats located in three climatic zones. Appl Environ Microbiol. 2003; 69:5248–5254. [PubMed: 12957910]
7. Burkert U, Warnecke F, Babenzien D, Zwirnmann E, Pernthaler J. Members of a readily enriched beta-proteobacterial clade are common in surface waters of a humic lake. Appl Environ Microbiol. 2003; 69:6550–6559. [PubMed: 14602613]
8. Vannini C, Poeckl M, Petroni G, Wu QL, Lang E, et al. Endosymbiosis in statu nascendi: close phylogenetic relationship between obligately endosymbiotic and obligately free-living *Polynucleobacter* strains (*Betaproteobacteria*). Environ Microbiol. 2007; 9:347–359. [PubMed: 17222133]
9. Hahn MW, Lang E, Brandt U, Wu QL, Scheuerl T. Emended description of the genus *Polynucleobacter* and the species *Polynucleobacter necessarius* and proposal of two subspecies, *P. necessarius* subsp. *necessarius* subsp. nov. and *P. necessarius* subsp. *asymbioticus* subsp. nov. Int J Syst Evol Microbiol. 2009; 59:2002–2009. [PubMed: 19567561]
10. Boscaro V, Felletti M, Vannini C, Ackerman MS, Chain PSG, et al. *Polynucleobacter necessarius*, a model for genome reduction in both free-living and symbiotic bacteria. Proc Natl Acad Sci USA. 2013; 110:18590–18595. [PubMed: 24167248]
11. Watanabe K, Komatsu N, Ishii Y, Negishi M. Effective isolation of bacterioplankton genus *Polynucleobacter* from freshwater environments grown on photochemically degraded dissolved organic matter. FEMS Microbiol Ecol. 2009; 67:57–68. [PubMed: 19049496]
12. Watanabe K, Komatsu N, Kitamura T, Ishii Y, Park H-D, et al. Ecological niche separation in the *Polynucleobacter* subclusters linked to quality of dissolved organic matter: a demonstration using a high sensitivity cultivation-based approach. Environ Microbiol. 2012; 14:2511–2525. [PubMed: 22759205]
13. Hahn MW, Jezberová J, Koll U, Saueressig-Beck T, Schmidt J. Complete ecological isolation and cryptic diversity in *Polynucleobacter* bacteria not resolved by 16S rRNA gene sequences. ISME J. 2016; 10:1642–1655. [PubMed: 26943621]
14. Hahn MW, Schmidt J, Pitt A, Taipale SJ, Lang E. Reclassification of four *Polynucleobacter necessarius* strains as *Polynucleobacter asymbioticus* comb. nov., *Polynucleobacter duraquae* sp. nov., *Polynucleobacter yangtzensis* sp. nov., and *Polynucleobacter sinensis* sp. nov., and emended description of the species *Polynucleobacter necessarius*. Int J Syst Evol Microbiol. 2016; 66:2883–2892. [PubMed: 27064460]
15. Hahn M, Karbon G, Koll U, Schmidt J, Lang E. *Polynucleobacter sphagniphilus* sp. nov. a planktonic freshwater bacterium isolated from an acidic and humic freshwater habitat. Int J Syst Evol Microbiol. (in press).
16. Hahn MW, Huymann LR, Koll U, Schmidt J, Lang E, et al. *Polynucleobacter wuianus* sp. nov., a free-living freshwater bacterium affiliated with the cryptic species complex PnecC. Int J Syst Evol Microbiol. 2017; 67:379–385. [PubMed: 27902302]

17. Hahn MW, Lang E, Brandt U, Luensdorf H, Wu QL, et al. *Polynucleobacter cosmopolitanus* sp. nov., free-living planktonic bacteria inhabiting freshwater lakes and rivers. Int J Syst Evol Microbiol. 2010; 60:166–173. [PubMed: 19648339]
18. Hahn MW, Stadler P, Wu QL, Pockl M. The filtration-acclimatization method for isolation of an important fraction of the not readily cultivable bacteria. J Microbiol Methods. 2004; 57:379–390. [PubMed: 15134885]
19. Meincke L, Copeland A, Lapidus A, Lucas S, Berry KW, et al. Complete genome sequence of *Polynucleobacter necessarius* subsp. *asymbioticus* type strain (QLW-P1DMWA-1^T). Stand Genomic Sci. 2012; 6:74–83. [PubMed: 22675600]
20. Whitman WB, Woyke T, Klenk H-P, Zhou Y, Lilburn TG, et al. Genomic Encyclopedia of Bacterial and Archaeal Type Strains, Phase III: the genomes of soil and plant-associated and newly described type strains. Stand Genomic Sci. 2015; 10:26–26. [PubMed: 26203337]
21. Markowitz VM, Chen IMA, Palaniappan K, Chu K, Szeto E, et al. IMG: the integrated microbial genomes database and comparative analysis system. Nucleic Acids Res. 2012; 40:D115–D122. [PubMed: 22194640]
22. Huntemann M, Ivanova NN, Mavromatis K, Tripp HJ, Paez-Espino D, et al. The standard operating procedure of the DOE-JGI Microbial Genome Annotation Pipeline (MGAP v.4). Stand Genomic Sci. 2015; 10:86. [PubMed: 26512311]
23. Chen IMA, Markowitz VM, Palaniappan K, Szeto E, Chu K, et al. Supporting community annotation and user collaboration in the integrated microbial genomes (IMG) system. BMC Genomics. 2016; 17:307. [PubMed: 27118214]
24. Hoetzinger M, Schmidt J, Jezberová J, Koll U, Hahn MW. Microdiversification of a pelagic *Polynucleobacter* species is mainly driven by acquisition of genomic islands from a partially interspecific gene pool. Appl Environ Microbiol. 2017; 83:e02266–16. [PubMed: 27836842]
25. Hahn MW, Scheuerl T, Jezberova J, Koll U, Jezbera J, et al. The passive yet successful way of planktonic life: Genomic and experimental analysis of the ecology of a free-living *Polynucleobacter* population. Plos One. 2012; 7
26. Boscaro V, Kolisko M, Felletti M, Vannini C, Lynn DH, et al. Parallel genome reduction in symbionts descended from closely related free-living bacteria. Nature Ecology & Evolution. 2017; 1:1160–1167. [PubMed: 29046583]
27. Hahn MW, Koll U, Karbon G, Schmidt J, Lang E. *Polynucleobacter aenigmaticus* sp. nov. isolated from the permanently anoxic monimolimnion of a temperate meromictic lake. Int J Syst Evol Microbiol. (in press).
28. Salcher MM, Neuenschwander SM, Posch T, Pernthaler J. The ecology of pelagic freshwater methylotrophs assessed by a high-resolution monitoring and isolation campaign. ISME J. 2015; 9:2442–53. [PubMed: 25942006]
29. Giovannoni SJ, Hayakawa DH, Tripp HJ, Stingl U, Givan SA, et al. The small genome of an abundant coastal ocean methylotroph. Environ Microbiol. 2008; 10:1771–82. [PubMed: 18393994]
30. Konstantinidis K, Tiedje J. Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci USA. 2005; 102:2567–2572. [PubMed: 15701695]
31. Kim M, Oh H-S, Park S-C, Chun J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int J Syst Evol Microbiol. 2014; 64:346–351. [PubMed: 24505072]
32. Hao Z, Li L, Liu J, Ren Y, Wang L, et al. Genome sequence of a freshwater low-nucleic-acid-content bacterium, Betaproteobacterium strain CB. Genome announcements. 2013; 1:e0013513–e0013513. [PubMed: 23599288]
33. Zaremba-Niedzwiedzka K, Caceres EF, Saw JH, Backstrom D, Juzokaite L, et al. Asgard archaea illuminate the origin of eukaryotic cellular complexity. Nature. 2017; 541:353–358. [PubMed: 28077874]
34. Janssen PJ, Van Houdt R, Moors H, Monsieurs P, Morin N, et al. The complete genome sequence of *Cupriavidus metallidurans* strain CH34, a master survivalist in harsh and anthropogenic environments. Plos One. 2010; 5

35. Amadou C, Pascal G, Mangenot S, Glew M, Bontemps C, et al. Genome sequence of the beta-rhizobium Cupriavidus taiwanensis and comparative genomics of rhizobia. Genome Res. 2008; 18:1472–83. [PubMed: 18490699]
36. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016; 33:1870–1874. [PubMed: 27004904]
37. Hahn MW, Koll U, Jezberova J, Camacho A. Global phylogeography of pelagic *Polynucleobacter* bacteria: Restricted geographic distribution of subgroups, isolation by distance, and influence of climate. Environ Microbiol. 2015; 17:829–840. [PubMed: 24920455]
38. Hahn MW, Lang E, Tarao M, Brandt U. *Polynucleobacter rarus* sp. nov., a free-living planktonic bacterium isolated from an acidic lake. Int J Syst Evol Microbiol. 2011; 61:781–787. [PubMed: 20435748]
39. Hahn MW, Minasyan A, Lang E, Koll U, Sproeer C. *Polynucleobacter difficilis* sp. nov., a planktonic freshwater bacterium affiliated with subcluster B1 of the genus *Polynucleobacter*. Int J Syst Evol Microbiol. 2012; 62:376–383. [PubMed: 21441373]
40. Hahn MW, Lang E, Brandt U, Sproeer C. *Polynucleobacter acidiphobus* sp. nov., a representative of an abundant group of planktonic freshwater bacteria. Int J Syst Evol Microbiol. 2011; 61:788–794. [PubMed: 20435747]
41. Hahn MW, Pockl M. Ecotypes of planktonic *Actinobacteria* with identical 16S rRNA genes adapted to thermal niches in temperate, subtropical, and tropical freshwater habitats. Appl Environ Microbiol. 2005; 71:766–773. [PubMed: 15691929]
42. Konstantinidis KT, Tiedje JM. Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci U S A. 2005; 102:2567–72. [PubMed: 15701695]
43. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, et al. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol. 2007; 57:81–91. [PubMed: 17220447]
44. Hahn MW, Pockl M, Wu QL. Low intraspecific diversity in a *Polynucleobacter* subcluster population numerically dominating bacterioplankton of a freshwater pond. Appl Environ Microbiol. 2005; 71:4539–4547. [PubMed: 16085847]

Europe PMC Funders Author Manuscripts

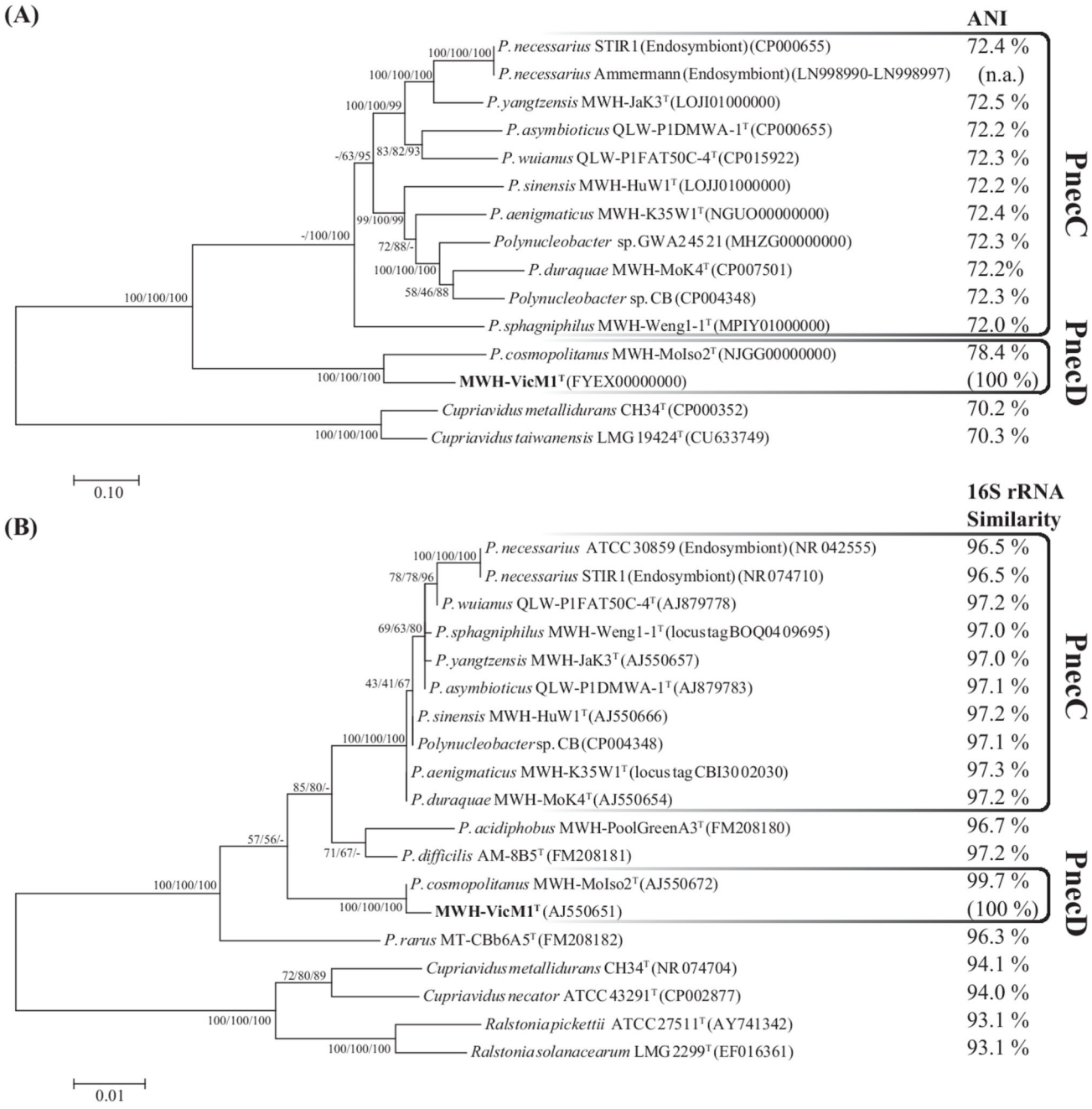

Fig. 1.

Phylogenetic position of strain MWH-VicM1^T. (A) Maximum Likelihood (ML) tree calculated with eight concatenated housekeeping gene sequences. Sequences of the type material of *P. necessarius* [1] could not be included due to the unavailability of the culture containing this endosymbiont [14]. Instead sequences of two other *P. necessarius* strains were included of which one is most likely identical with the type material of the species [14]. Bootstrap values are shown from left to right for neighbour joining (NJ), ML, and maximum parsimony (MP) trees calculated with the same sequence set. Pairwise average

nucleotide identity (ANI) values for whole genome comparisons of strain MWH-VicM1^T with the other shown taxa are given. Bar, 0.1 substitutions per nucleotide position. (B) ML tree calculated with 16S rRNA gene sequences and pairwise 16S rRNA sequence similarity values of strain MWH-VicM1^T with the other shown taxa. Bootstrap values are shown from left to right for NJ, ML, and MP trees calculated with the same sequence set. Bar, 0.01 substitutions per nucleotide position. N.a., not available.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Table 1

Genome characteristics of the investigated *Polynucleobacter* strains.

Species	Strain	*Polynucleobacter* subcluster	Life style	Genome size (Mbp)	Scaffolds	G+C content(mol%)	DDBJ/EMBL/GenBank accession number	Reference
P. victoriensis sp. nov.	MWH-VicM1^T (=DSM 21486^T)	PnecD	FL	1.63	3	43.1	FYEX00000000	This study
P. cosmopolitanus	MWH-MoIso2^T (=DSM 21490^T)	PnecD	FL	1.78	6	44.1	NJGG00000000	This study
P. necessarius	STIR1 [host, *Euplotes aediculatus*]	PnecC	E	1.56	1	45.6	CP001010	Boscaro *et al.*, 2013
P. aenigmaticus	MWH-K35W1^T (=DSM 24006^T)	PnecC	FL	2.14	37	46.0	NGUO00000000	Hahn et al., in press
P. asymbioticus	QLW-P1DMWA-1^T (=DSM 18221^T)	PnecC	FL	2.16	1	44.8	CP000655	Meincke *et al.*, 2012
P. duraquae	MWH-MoK4^T (=DSM 21495^T)	PnecC	FL	2.03	1	45.2	CP007501	Hahn *et al.*, 2016b
P. sinensis	MWH-HuW1^T (=DSM 21492^T)	PnecC	FL	2.32	19	45.5	LOJJ01000000	Hahn *et al.*, 2016a
P. yangtzensis	MWH-JaK3^T (=DSM 21493^T)	PnecC	FL	2.05	42	45.4	LOJI01000000	Hahn *et al.*, 2016a
P. wuianus	QLW-P1FAT50C-4^T (=DSM 24008^T)	PnecC	FL	2.23	1	44.9	CP015922	Hahn *et al.*, 2017
P. sphagniphilus	MWH-Weng1-1^T (=DSM 24018^T)	PnecC	FL	2.04	17	45.6	MPIY01000000	Hahn et al., in press

FL, free-living; E, endosymbiotic