# Quantification of archaea-driven freshwater nitrification from single cell to ecosystem levels

Franziska Klotz<sup>1</sup>, Katharina Kitzinger<sup>2</sup>, David Kamanda Ngugi<sup>3</sup>, Petra Büsing<sup>3</sup>, Sten Littmann<sup>2</sup>, Marcel M. M. Kuypers<sup>2</sup>, Bernhard Schink<sup>1</sup>, and Michael Pester<sup>1,3,4</sup>

- 1. Department of Biology, University of Konstanz, Universitätsstrasse 10, Konstanz, D-78457, Germany
- 2. Max Planck Institute for Marine Microbiology, Celsiusstrasse 1, D-28359 Bremen, Germany.
- 3. Leibniz Institute DSMZ German Collection of Microorganisms and Cell Cultures, Inhoffenstr. 7B, D-38124 Braunschweig, Germany
- 4. Technical University of Braunschweig, Institute for Microbiology, Spielmannstrasse 7, D-38106 Braunschweig, Germany

## Supplementary Methods

Measurements of environmental parameters

Vertical profiles of temperature and oxygen were measured down to the lake sediment with a multisampling probe (RBR Ltd.; Ottawa, Canada, Sea & Sun Technology GmbH, Trappenkamp, Germany or bbe Moldaenke GmbH, Schwentinental, Germany) at a resolution of 0.5 m and on the following dates (yyyy-mm-dd): 2017-11-08, 2017-11-21, 2017-12-05, 2018-01-09, 2018-01-23, 2018-02-06, 2018-02-20, 2018-03-13, 2018-03-27, 2018-04-10, 2018-04-24, 2018-05-22, 2018-06-05, 2018-06-19, 2018-07-03, 2018-07-31, 2018-08-14, 2018-09-21, 2018-10-09, 2018-10-23, 2018-11-22, 2018-12-04, 2018-12-18, 2019-01-11, 2019-02-12, 2019-02-26, 2019-03-12, 2019-03-27, 2019-04-09, 2019-04-25, 2019-06-18, 2019-07-02, 2019-07-31, 2019-08-13, 2019-08-28, 2019-09-24, 2019-10-08, 2019-10-23, 2019-11-04, 2019-11-19.

Vertical profiles of nitrate and total ammonium (NH<sub>4</sub><sup>+</sup> + NH<sub>3</sub>) were measured for the same dates as used for qPCR analyses. Nitrate and total ammonium concentrations were determined using the auto-analyzer and Seal analytics methods G-172-96 Rev. 12 and G-171-96 Rev. 14, respectively (SEAL Analytical GmbH, Norderstedt, Germany). For these measurements, 20 ml water each from 13 depths at 1, 5, 10, 15, 20, 25, 30, 40, 50, 60, 85, 110 and 135 m were taken, filtered through a Chromafil<sup>®</sup> GF/PET-20/25 filter (pore size 1.0 and 0.2  $\mu$ m, VWR, Vienna, Austria) and stored at –20°C until analysis. Samples obtained between July and November 2019 were measured by an alternative method: Nitrate was measured by ion chromatography (S150 Chromatography System, SYKAM) and

total ammonium was measured fluorometrically by the *ortho*-phthaldialdehyde method [1]. Chlorophyll *a* was sampled from 22 depths over a gradient of 0–60 m and analyzed spectrophotometrically after extraction in hot ethanol as described previously [2], but without correcting for pheopigments.

#### DNA and RNA extraction

For DNA extraction, 0.22 µm-filters (47 mm diameter) were placed in 2 ml-screw cap tubes and vortexed for 15 min in a solution containing 500 µl TE-buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA), 12.5 µl 20% sodium lauryl sulfate-solution (SLS; Sigma-Aldrich, Taufkirchen, Germany), 500 µl phenol-chloroform-isoamylalcohol 25:24:1 (Carl Roth GmbH), and 200 µl pre-combusted zirconium beads (0.1 mm in diameter, Carl Roth GmbH, Karlsruhe, Germany). After centrifugation (4°C, 10 min, 18,620 ×*g*), the supernatant was washed once with 500 µl chloroform-isoamylalcohol 24:1 (Carl Roth GmbH). This mixture was centrifuged again and DNA precipitated from the separated supernatant over night at  $-20^{\circ}$ C using a mixture of 0.1 volume 3 M sodium acetate (pH 4.8), 2.5 volume molecular grade ethanol (Carl Roth GmbH) and 1 µl glycogen (20 mg ml<sup>-1</sup>, Thermo Fisher Scientific, Waltham, MA, USA). Precipitated DNA was centrifuged, washed twice with 80% molecular grade ethanol, and dissolved in nuclease-free water (MP biomedicals, Eschwege, Germany). RNA was removed by an RNase ONE<sup>TM</sup> ribonuclease treatment following kit instructions (Promega, Fitchburg, WI, USA) and DNA samples were stored at  $-20^{\circ}$ C until processing.

For RNA extraction, 0.22  $\mu$ m-filters (142 mm diameter) filters were cut with a sterilized scissor into thirds and extracted as described above except for the following modifications: filters were extracted in extraction buffer (50 mM sodium acetate and 10 mM EDTA, pH 4.2) with 0.025% SLS (Sigma-Aldrich) and phenol-chloroform-isoamylalcohol 25:24:1 (Roti-Aqua-P/C/I 4.5-5.0, Carl Roth GmbH). Washing of the aqueous phase with chloroform-isoamylalcohol 24:1 was done in the presence of 0.1 volume 3 M sodium acetate. RNA was finally precipitated with 1 volume ice-cold isopropanol in the presence of 1  $\mu$ l glycogen (35 mg ml<sup>-1</sup>, RNase-free, VWR), washed as stated above, and eluted in nuclease-free water (MP biomedicals). DNA was digested with the TURBO DNA-free<sup>TM</sup> kit (Thermo Fisher Scientific) and RNA samples were stored afterwards at -80°C until sequencing.

#### Determination of AOA abundance by CARD-FISH

Before CARD-FISH, cells on the filter sections were immobilized by embedding in 0.1% low-gelling agarose (Metaphor). CARD-FISH was performed using a specific HRP-labeled oligonucleotide probe for *Nitrososphaeria* (HRP-labeled Thaum726 [GCTTTCATCCCTCACCGTC] and unlabeled competitors [Thaum726\_compA: GCTTTCGTCCCTCACCGTC, Thaum726\_compB: GCTTTCATCCCTCACCGTC]) [3, 4]

as described previously [5]. Negative controls using NonEUB [6] to exclude unspecific signals were performed according to a defined protocol [7]. Briefly, endogenous peroxidases were inactivated by incubation in 0.01 M HCl for 10 min. Cells were permeabilized by HCl (0.1 M HCl for 1 min) and subsequently washed with MilliQ water. Filter pieces were hybridized with HRP probes and the respective competitor probes at 25% formamide concentration at 46°C for up to 3 h. After a 5 min washing step at 48°C and HRP probe equilibration in 1× PBS for 5 to 15 min, signal amplification was performed with OregonGreen488-labeled tyramides at 48°C for 30 min. Cells were counterstained with 4',6-diamidino-2-phenylindole (DAPI, 10 µg ml<sup>-1</sup>, 5 min at room temperature). Filter sections were mounted onto glass slides, and embedded in a 4:1 mixture of Citifluor AF1 and Vectashield (Citifluor Ltd, London, UK; Vector Laboratories, Burlingame, CA, USA). *Nitrososphaeria* and DAPI signals were counted on an Axiophot or Axioplan 2 microscope (Zeiss, Germany).

#### Next generation sequencing and bioinformatics processing

Metagenome sequencing libraries were prepared with the NEBNext<sup>®</sup> Ultra<sup>™</sup> DNA Library Prep Kit for Illumina<sup>®</sup> (New England Biolabs GmbH, Frankfurt am Main, Germany) and sequenced on an Illumina NextSeq500 sequencer using 2  $\times$  150 bp. This resulted in 0.6–2.7  $\times$  10<sup>8</sup> reads per metagenome with an average of 1.3 × 10<sup>8</sup> reads (8.8–41 Gbp, average 20 Gbp). Raw Illumina reads were quality checked with FastQC v.0.11.8 [8] and subsequently quality filtered and trimmed using Sickle v1.33 [9]. Thereafter, reads were assembled with Megahit v1.1.2 [10] and binned with maxbin2 v2.2.4 [11]. A co-assembly of the November 2017, December 2017, and February 2018 metagenomes resulted in the best AOA bin. To refine this bin, DNA from November 2017 was sequenced in addition by PacBio sequencing on a Sequel instrument (Pacific Biosciences, Menlo Park, CA, USA) using circular consensus sequencing with a target length of 2 kbp. Raw PacBio reads were quality controlled with smrt analysis using an accuracy of 0.999; the number of resulting CCS bases was 0.55 Gbp. The original Illumina AOA bin was used as trusted contigs in spades v3.11.1 [12] and re-assembled with Sequel-reads for gap closure. A subsequent binning in metabat2 v2.12.1 [13] resulted in the final MAG AOA-LC4. MAGs related to AOB, NOB or comammox could not be further refined by long PacBio reads. All MAGs were tested for completeness, strain heterogeneity, and contamination using CheckM v1.0.7 [14] and for their index of replication using iRep v1.10 [15]. MAGs and single contigs were screened for the presence of the functional marker genes amoA and nxrB by both blastp v2.10.1 [16] (e-value threshold  $1^{-10}$ ) and hmm-search v3.3 [17] (e-value threshold  $1^{-5}$ ). The latter was based on hmm-models retrieved from the fungene database with manual curation [amoA\_AOA.hmm (Feifei Liu), amoA\_AOB.hmm (RDP) amoA\_comammox.hmm (Yang Ouyang), nxrB.hmm (RDP), fungene.cme.msu.edu] [18]. MAG AOA-LC4 was annotated using the Microscope platform [19]. The automated annotation was manually refined using annotation rules laid out before [20]. Additional

MAGs were annotated with PROKKA v1.12 [21] and curated manually for their functional marker genes *amoABC*, *hao*, and *nxrAB*, where appropriate.

For metatranscriptome sequencing, messenger RNA (mRNA) was enriched from total RNA extracts by depleting ribosomal RNA with the Ribo-off rRNA Depletion Kit for bacteria (Vazyme, Nanjing, China). Thereafter, the sequencing library was prepared with the TruSeq® Stranded mRNA Library Prep (Illumina) and sequenced on a NextSeq500 sequencer using 2 × 150 bp. The sequencing depth ranged between  $0.7-1.9 \times 10^8$  reads per metatranscriptome with an average of  $1.2 \times 10^8$  reads (10–27 Gbp, average 17 Gbp). Raw reads were quality filtered and trimmed using trimmomatic v0.38 [22] and the fastx toolkit v0.0.14 (hannonlab.cshl.edu/fastx\_toolkit). Residual ribosomal reads were removed using SortMeRNA v2.1b [23]. Curated metatranscriptome reads were mapped against MAGs and contigs of interest using bowtie2 v2.30 [24] to determine the transcription levels of individual genes. Subsequent network analysis of co-transcribed genes of MAG AOA-LC4 was based on genes with transcription values higher than the median transcription of all AOA-LC4 genes (35.35 FPKM). Transcription values were correlated pairwise using Spearman correlation; only significant (FDR-corrected p-value < 0.05) correlations with a correlation coefficient of  $r_s$  > 0.8 were further processed. For the final network construction, only genes, which correlated in their transcriptional response to at least two of the either *amoA*, *amoB* or *amoC* were taken into account. The network was created with the R package igraph v1.2.5 [25] and refined with cytoscape v3.8.1 [26].

A gene-centric analysis was performed to gain an overview of all ammonia transporter (*amt*) and urea transporter (*dur3*) or urea transport system substrate-binding protein (*urtA*) gene sequences in Lake Constance. Therefore, single assemblies of all 9 metagenomes were annotated with DRAM [27]. The retrieved *amt*, *dur3* and *urtA* sequences were de-replicated with cd-hit-est [28, 29]. Curated metatranscriptomes were mapped onto unique *amt*, *dur3* and *urtA* sequences using bowtie2 v2.30 [24] to compare their transcription levels with AOA-LC4.

#### Phylogenetic analyses

Phylogenomic analyses of the nitrifying MAGs were performed on the basis of concatenated amino acid alignments of 122 translated archaeal or 120 bacterial single copy genes [30, 31]. Alignments were generated using GTDB-Tk v0.3.2 [32] and maximum likelihood trees were constructed using IQ-tree v1.6.12 [33]. Branch support was tested with the Shimodaira–Hasegawa approximate likelihood-ratio test [34] and ultrafast bootstrap [33] options in IQ-tree. Genome-wide average nucleic and amino acid identities (ANI and AAI, respectively) were calculated with the online tool (enve-omics.ce.gatech.edu) developed previously [35] using default settings. Maximum likelihood trees for AOB-*amoA*, comammox-*amoA* and NOB/comammox-*nxrB* genes were constructed in IQ-tree based on manually curated alignments established in ARB v6.0.3 [36].

## Supplementary Results

#### Phylogenetic analysis of bacterial ammonia oxidizers

Phylogenomic maximum likelihood tree construction revealed that MAG AOB-LC263 formed a stable cluster with other freshwater MAGs, which represented a sister clade to bona fide Nitrosospira species (Fig. S3). This was corroborated by phylogenetic analysis of its *amoA* gene (Fig. S5). Closest relatives of AOB-LC263 were MAGs retrieved from Lake Baikal, the Great Lakes, and Lake Biwa. Based on the currently proposed species and genus delineation thresholds of ca. <95% ANI and <65% AAI, respectively [37], AOB-LC263 would represent a new species and genus within the Nitrosomonadaceae (Fig. S4). The phylogenetic affiliation of contigs AOB-LC199628 and AOB-LC368213 could only be assessed based on their amoA genes. Contig AOB-LC199628 clustered in a stable clade consisting of environmental sequences that was distinct from the AOB-LC263 and Nitrosospira clusters. Its closest cultured relative was Nitrosospira sp. Np39-19 as based on 78.3% amoA nucleotide identity. Contig AOB-LC368213 clustered within sequences affiliated with Nitrosomonas species with its closest cultured relative being Nitrosomonas ureae Nm10 with 89.9% amoA nucleotide identity (Fig. S5). Comparison of the retrieved amoA sequences to a previous bacterial amoA clone library obtained from Lake Constance waters [38] revealed that the amoA gene of contig AOB-LC199628 was 100% identical (nucleic acid identity) to clone BmcYyy23.2 (MH780622.1) from OTU1 [38]. Furthermore, the amoA of MAG AOB-LC263 was 99.8% identical to clone BmcYyy33 (MH780602.1) from OTU2. Since only two OTUs were detected previously [38], contig AOB-LC368213 had no representatives in the earlier *amoA* clone library.

#### Phylogenetic analysis of nitrite-oxidizing bacteria and comammox bacteria

Phylogenomic maximum likelihood tree construction placed the two MAGs NOB-LC29 and NOB-LC32 into *Nitrospira* lineage II but outside the intra-lineage comammox clades A and B (Fig. S6). This was corroborated by phylogenetic analysis of their *nxrB* genes (Fig. S8). The two MAGs shared an ANI and AAI of 85% and 85%, respectively, indicating that they represent two separate species within the genus *Nitrospira* (Fig. S7). Phylogenomic tree construction of MAG COM-LC224 placed it into comammox clade B within *Nitrospira* lineage II (Fig. S6), which was corroborated by phylogenetic placement of its single genes *nxrB* and *amoA* (Fig. S8 and S9). Interestingly, COM-LC224 showed <65% AAI to both type and *Candidatus* species within the genus *Nitrospira* (Fig. S7), but at the same time exhibited AAI values of >65% with freshwater *Nitrospira* like NOB-LC29 and NOB–LC32. Without further data, its affiliation at the taxonomic rank of a genus is currently inconclusive.

### References

1. Holmes RM, Aminot A, Kérouel R, Hooker BA, Peterson BJ. A simple and precise method for measuring ammonium in marine and freshwater ecosystems. Can J Fish Aquat Sci

1999;56:1801–1808.

- 2. Tilzer MM. The importance of fractional light absorption by photosynthetic pigments for phytoplankton productivity in Lake Constance. Limnol Oceanogr 1983;28:833–846.
- 3. Beam JP. Geobiological interactions of archaeal populations in acidic and alkaline geothermal springs of Yellowstone National Park, WY, USA. 2015. Montana State University.
- 4. Sauder LA, Albertsen M, Engel K, Schwarz J, Nielsen PH, Wagner M, et al. Cultivation and characterization of *Candidatus* Nitrosocosmicus exaquare, an ammonia-oxidizing archaeon from a municipal wastewater treatment system. ISME J 2017;11:1142–1157.
- 5. Kitzinger K, Padilla CC, Marchant HK, Hach PF, Herbold CW, Kidane AT, et al. Cyanate and urea are substrates for nitrification by Thaumarchaeota in the marine environment. Nat Microbiol 2019;4:234–243.
- 6. Wallner G, Amann R, Beisker W. Optimizing fluorescent in situ hybridization with rRNAtargeted oligonucleotide probes for flow cytometric identification of microorganisms. Cytometry 1993;14:136–143.
- Pernthaler A, Pernthaler J, Amann R. Sensitive multi-color fluorescence in situ hybridization for the identification of environmental microorganisms. In: Kowalchuk G, de Bruijn FJ, Head IM, Akkermans ADL, van Elsas JD (eds). Molecular Microbial Ecology Manual, 2nd ed. 2004. Kluwer Academic Publishers, Dordrecht, Boston, London, pp 711–726.
- 8. Andrews S. FASTQC: A quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. [Software]. 2010.
- 9. Joshi N, Fass J. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]. 2011.
- 10. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 2015;31:1674–1676.
- 11. Wu YW, Simmons BA, Singer SW. MaxBin 2.0: An automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 2016;32:605–607.
- 12. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. Genome Res 2017;27:824–834.
- 13. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 2019;7:e7359.
- 14. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 2015;25:1043–1055.
- 15. Brown CT, Olm MR, Thomas BC, Banfield JF. Measurement of bacterial replication rates in microbial communities. Nat Biotechnol 2016;34:1256–1263.
- 16. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215:403–410.
- 17. Eddy SR. A new generation of homology search tools based on probabilistic inference. Genome Inform 2009;23:205–211.
- 18. Fish JA, Chai B, Wang Q, Sun Y, Brown CT, Tiedje JM, et al. FunGene: the functional gene pipeline and repository. Front Microbiol 2013;4:291.

- 19. Vallenet D, Calteau A, Cruveiller S, Gachet M, Lajus A, Josso A, et al. MicroScope in 2017: an expanding and evolving integrated resource for community expertise of microbial genomes. Nucleic Acids Res 2016;45:D517–D528.
- 20. Hausmann B, Pelikan C, Rattei T, Loy A, Pester M. Long-term transcriptional activity at zero growth of a cosmopolitan rare biosphere member. mBio 2019;10:e02189-18.
- 21. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics 2014;30:2068–2069.
- 22. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30:2114–2120.
- 23. Kopylova E, Noé L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 2012;28:3211–3217.
- 24. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357–359.
- 25. Csardi G, Nepusz T. The igraph software package for complex network research. InterJournal, Complex Syst 2006;1695:1–9.
- 26. Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics 2011;27:431–432.
- 27. Shaffer M, Borton MA, McGivern BB, Zayed AA, La Rosa SL 0003 3527 8101, Solden LM, et al. DRAM for distilling microbial metabolism to automate the curation of microbiome function. Nucleic Acids Res 2020;48:8883–8900.
- 28. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 2006;22:1658–1659.
- 29. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 2012;28:3150–3152.
- Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol 2018;36:996.
- 31. Parks DH, Chuvochina M, Chaumeil P-A, Rinke C, Mussig AJ, Hugenholtz P. A complete domain-to-species taxonomy for Bacteria and Archaea. Nat Biotechnol 2020;38:1079–1086.
- 32. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics 2019;36:1925–1927.
- 33. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 2014;32:268–274.
- Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 2010;59:307–321.
- 35. Rodriguez-R LM, Konstantinidis K. Bypassing cultivation to identify bacterial species: Cultureindependent genomic approaches identify credibly distinct clusters, avoid cultivation bias, and provide true insights into microbial species. Microbe Mag 2014;9:111–118.
- 36. Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, et al. ARB: a software environment for sequence data. Nucleic Acids Res 2004;32:1363–1371.
- 37. Konstantinidis KT, Rosselló-Móra R, Amann R. Uncultivated microbes in need of their own

taxonomy. ISME J 2017;11:2399–2406.

- 38. Herber J, Klotz F, Frommeyer B, Weis S, Straile D, Kolar A, et al. A single Thaumarchaeon drives nitrification in deep oligotrophic Lake Constance. Environ Microbiol 2020;22:212–228.
- 39. Spasov E, Tsuji JM, Hug LA, Doxey AC, Sauder LA, Parker WJ, et al. High functional diversity among *Nitrospira* populations that dominate rotating biological contactor microbial communities in a municipal wastewater treatment plant. ISME J 2020;14:1857–1872.
- 40. Sakoula D, Koch H, Frank J, Jetten MSM, van Kessel MAHJ, Lücker S. Enrichment and physiological characterization of a novel comammox *Nitrospira* indicates ammonium inhibition of complete nitrification. ISME J 2021;15:1010–1024.
- 41. Daims H, Lebedeva E V., Pjevac P, Han P, Herbold C, Albertsen M, et al. Complete nitrification by *Nitrospira* bacteria. Nature 2015;528:504–509.
- 42. Pjevac P, Schauberger C, Poghosyan L, Herbold CW, van Kessel MAHJ, Daebeler A, et al. *AmoA*-targeted polymerase chain reaction primers for the specific detection and quantification of comammox *Nitrospira* in the environment. Front Microbiol 2017;8:1–11.
- 43. Van Kessel MAHJ, Speth DR, Albertsen M, Nielsen PH, Op Den Camp HJM, Kartal B, et al. Complete nitrification by a single microorganism. Nature 2015;528:555–559.

# Supplementary Tables

**Table S1.** Overview of metagenome-assembled genomes (MAGs) and contigs related to the nitrifying community in the hypolimnion of Lake Constance.

**Table S2.** Annotation and seasonally resolved transcription of MAG AOA-LC4 genes involved in nitrogen metabolism, vitamin synthesis, carbon fixation, cell division, replication, transport systems, respiration and the TCA-cycle.

**Table S3.** NCBI accession numbers or Taxon IDs of all species, MAGs and clones, which are part of the phylogenetic trees of *Nitrososphaeria*, *Nitrosomonadaceae* and *Nitrospira* (Figures 2, S3, S5, S6, S8, and S9).

# Supplementary Figures



**Figure S1.** AOA abundance in hypolimnetic water from 85 m depth as measured by archaeal *amoA*targeted qPCR or CARD-FISH using a *Nitrososphaeria*-specific probe (probe Thaum726), which currently encompasses all AOA. Samples were taken on June 18<sup>th</sup> and November 5<sup>th</sup> 2019. CARD-FISH was performed on water samples used for nitrification rate measurements after 67 h (June) or 48 h (November) of incubation at 4°C in the dark. Hybridized filters were counted either 10 times (June) or 25 times (November) independently.

#### Supplementary Information

ANI

100	100	100	86	86	86	84	85	85	79	79	79	80	80	80	81	81	81	Nitrosopumilus sp. BACL13 MAG-120910-bin56	100
100	100	100	85	85	85	83	84	84	78	78	78	79	79	79	80	80	79	Nitrosopumilus sp. BACL13 MAG-121220-bin23	95
100	100	100	85	85	85	83	84	84	78	78	78	79	79	79	80	80	79	Nitrosopumilus sp. UBA241	00
86	85	85	100	99	99	84	84	84	77	77	78	78	78	79	79	79	79	Nitrosopumilus sp. Baikal-deep-G182	90
86	85	85	99	100	99	84	84	84	78	77	77	79	79	79	79	79	79	Nitrosopumilus sp. Casp-thauma1	85
86	85	85	99	99	100	84	84	84	78	78	78	78	79	79	79	80	80	AOA-LC4	80
84	83	83	84	84	84	100	89	89	77	78	78	79	79	79	79	80	79	Nitrosopumilus sp. Nsub	
85	84	84	84	84	84	89	100	93	78	78	78	79	79	79	79	80	79	Candidatus Nitrosomarinus catalina SPOT01	
85	84	84	84	84	84	89	93	100	78	78	78	80	79	79	80	80	80	Nitrosopumilus sp. MED-G94	
79	78	78	77	78	78	77	78	78	100	82	82	78	78	78	78	78	78	Candidatus Nitrosarchaeum koreense MY1	
79	78	78	77	77	78	78	78	78	82	100	96	78	78	78	78	78	78	Candidatus Nitrosarchaeum limniae BG20	
79	78	78	78	77	78	78	78	78	82	96	100	78	78	78	78	78	78	Candidatus Nitrosarchaeum limniae SFB1	
80	79	79	78	79	78	79	79	80	78	78	78	100	85	85	80	80	80	Nitrosopumilus maritimus SCM1	
80	79	79	78	79	79	79	79	79	78	78	78	85	100	86	80	80	80	Candidatus Nitrosonumilus koreense AR1	
80	79	79	79	79	79	79	79	79	78	78	78	85	86	100	80	80	80	Nitrosopumilus piranensis D3C	
81	80	80	79	79	79	79	79	80	78	78	78	80	80	80	100	81	81	Nitrosopumilus adriaticus NE5	
81	80	80	79	79	80	80	80	80	78	78	78	80	80	80	81	100	82	Candidatus Nitrosonumilus sediminis AR2	
81	79	79	79	79	80	79	79	80	78	78	78	80	80	80	81	82	100	Candidatus Nitrosopumilus salarius BD31	
								Δ	Δι										
_										_	_	_				_	_		- 100
100	98	98	84	84	85	85	86	86	75	75	75	78	79	79	80	80	78	Nitrosopumilus sp. BACL13 MAG-120910-bin56	100
98	100	100	84	84	84	83	84	85	73	73	73	77	77	77	77	78	77	Nitrosopumilus sp. BACL13 MAG-121220-bin23	95
98	100	100	84	84	84	83	84	84	73	72	73	76	77	76	77	78	77	Nitrosopumilus sp. UBA241	00
84	84	84	100	99	99	85	85	85	74	74	74	77	77	77	77	78	77	Nitrosopumilus sp. Baikal-deep-G182	90
84	84	84	99	100	99	84	85	85	74	73	73	77	77	77	77	78	77	Nitrosopumilus sp. Casp-thauma1	85
85	84	84	99	99	100	84	85	85	74	73	73	77	77	77	77	78	77	AOA-LC4	80
85	83	83	85	84	84	100	92	92	73	73	72	77	77	77	77	78	76	<i>Nitrosopumilus</i> sp. Nsub	00
86	84	84	85	85	85	92	100	95	73	73	73	77	77	77	77	78	77	Candidatus Nitrosomarinus catalina SPOT01	75
86	85	84	85	85	85	92	95	100	74	73	74	78	78	78	78	79	77	Nitrosopumilus sp. MED-G94	
75	73	73	74	74	74	73	73	74	100	83	83	74	75	75	75	75	75	Candidatus Nitrosarchaeum koreense MY1	
75	73	72	74	73	73	73	73	73	83	100	96	74	74	73	74	74	75	Candidatus Nitrosarchaeum limniae BG20	
75	73	73	74	73	73	72	73	74	83	96	100	74	74	74	74	74	74	Candidatus Nitrosarchaeum limniae SFB1	
78	77	76	77	77	77	77	77	78	74	74	74	100	87	87	80	80	79	Nitrosopumilus maritimus SCM1	
79	77	77	77	77	77	77	77	78	75	74	74	87	100	88	80	80	79	Candidatus Nitrosopumilus koreense AR1	
79	77	76	77	77	77	77	77	78	75	73	74	87	88	100	80	80	79	Nitrosopumilus piranensis D3C	
80	77	77	77	77	77	77	77	78	75	74	74	80	80	80	100	81	80	Nitrosopumilus adriaticus NF5	
80	78	78	78	78	78	78	78	79	75	74	74	80	80	80	81	100	81	Candidatus Nitrosopumilus sediminis AR2	
78	77	77	77	77	77	76	77	77	75	75	74	79	79	79	80	81	100	Candidatus Nitrosopumilus salarius BD31	
Ş	Ş	Z	Ş	Z	≥	Z	Ω	Ş	S	Q	Ω	Z	Ω	Z	Z	Ω	Ω		
itro.	itro.	itro.	itro.	itro.	Ş	itro.	anc	itro.	anc	anc	anc	itro.	anc	itro.	itro,	anc	anc		
Soc	sop	los	sop	los	Ľ	los	lida	sop	fida	lida	lida	sop	lida	sop	sop	lide	lida		
ŭ	nn	nn	ŭ'n	nn	4	nn	atus	un	atus	atus	atus	ŭ	atus	nn	nn	atus	atus		
ıilu	ıilu	ıilu	ıilu	nilu		ıilu	z	ıilu	z	z	z	ıilu	z	ıilu	ıilu	z	z		
ss	ss	ss	ss	ss		ss	itro	ss	itro	itro	itro	sn	itro	sр	s a	itro	itro		
p.	р. Е	р <u>.</u>	p.	p.		р. 7	sor	<u>p</u> . 1	sar	sar	sar	ıan	los	iraı	dria	los	los		
Ă	Ă	Ē,	Bail	Cag		Jsr	nai	≦	ch	ch	ch	tim	nuc	her	atic	un	un		
5	5	ş	ka -	Ϋ́		Ъ	Ϊ	P	aer	aer	aer	sni	<u>ni</u>	isis	Sn	nii	nilu		
3	13	4	de	tha			S S	<u>7</u> 65	Ш	m	Ш	SC	N N	ő	Ę	s s	s si		
MΑ	Ă٨		ep-	um.			áta	4	kor	lim	lim	ž	Ore	õ	σi	edi	ala		
Ģ	Ģ		Ġ	а1			alin		ee.	nia	nia	-	en			ni.	niu		
120	121		82				a S		nse	е В	e S		se.			nis.	в		
91	22						PC		Ξ	ŝ	ËB		AR			AR	D3		
9-0	0-Ь						OL(		Σ	ŏ	2		-			Ň	-		
in5	in2						٦												
-	č.S.																		

**Figure S2.** Pairwise genome-wide average nucleotide identities (ANI) and average amino acid identities (AAI) of MAG AOA-LC4 (shown in bold) in comparison to representatives of the family *Nitrosopumilaceae*. MAG AOA-LC4 represents a novel species together with Casp-thauma1 and

Baikal-Deep-G182 in the genus *Nitrosopumilus* based on the species-level threshold of 95% for ANI [37] and genus-threshold of 65% for AAI [37].



**Figure S3.** Phylogeny of MAG AOB-LC263 in relation to closely related freshwater MAGs and pure cultures of the ammonia-oxidizing bacteria within the genera *Nitrosomonas* and *Nitrosospira*. The phylogenomic maximum likelihood tree was constructed using the IQ-tree algorithm [33] on the basis of a concatenated amino acid alignment of 120 translated single copy genes that were established by the GTDB-based taxonomy for phylogenetic inference of bacteria [30, 31]. Branch

support was tested with the Shimodaira–Hasegawa approximate likelihood-ratio test (SH-aLRT; 1000 replicates) and ultrafast bootstraps (1000 replicates) within IQ-tree. Branch support was set as significant at  $\geq$ 80% for SH-aLRT and  $\geq$ 95% for ultrafast bootstrap values (black semi-circles for significant and white for non-significant). MAGs or species with freshwater-origin are colored blue. *Methylotenera mobilis* (NCBI accession number GCA\_000023705.1), *Methylovorus glucosetrophus* (NC\_012969.1) and *Methylobacillus flagellates* (GCA\_000013705.1) were used as outgroup. The scale bar indicates 10% estimated amino acid sequence divergence. All accession numbers can be found in Table S3.



100	94	100	99	100	100	10	84	10	10	73	73	74	74	
94	100	94	94	94	94	10	78	71	10	73	73	73	74	1
100	94	100	99	100	100	10	89	10	10	74	74	74	75	
99	94	99	100	99	99	10	85	10	10	10	73	73	74	1
100	94	100	99	100	100	10	81	10	10	72	73	73	74	
100	94	100	99	100	100	10	86	82	10	74	74	74	74	1
10	10	10	10	10	10	100	99	99	74	10	10	10	10	
84	78	89	85	81	86	99	100	99	74	10	10	78	79	
10	71	10	10	10	82	99	99	100	74	10	10	10	10	1
10	10	10	10	10	10	74	74	74	100	10	10	10	10	1
73	73	74	10	72	74	10	10	10	10	100	78	77	77	
73	73	74	73	73	74	10	10	10	10	78	100	79	78	1
74	73	74	73	73	74	10	78	10	10	77	79	100	79	
74	74	75	74	74	74	10	79	10	10	77	78	79	100	

100 AOB-LC263 Great Lakes MAG MC4 S2 CONCOCT 122 Lake Baikal MAG Baikal-deep-G101 80 Great Lakes MAG SU08M 6M CONCOCT 0731 Lake Biwa LBMAG hypo.bin2 60 Great Lakes MAG MC14 S12 M CONCOCT 099 Lake Baikal MAG Baikal-deepG99 Lake Biwa LBMAG hypo.bin3 40 Great Lakes MAG MC17 S15 bin 110 Saline water MAG UBA6933 20 Nitrosospira multiformis C-71 Nitrosospira tenuis Nv1 Nitrosospira briensis C-128 Nitrosospira lacus APG3

AAI

100	88	03	0/	05	0/	68	67	67	64	62	63	63	64	AOB   C263	100
88	100	03	94	95	04 Q/	71	71	70	67	70	73	73	74	Great Lakes MAG MC4 S2 CONCOCT 122	
93	93	100	97	qq	qq	70	70	70	66	70	72	72	74	Lake Baikal MAG Baikal-deen-G101	
04	95	97	100	00	90	70	70	70	67	71	72	72	74	Great Lakes MAG SU08M 6M CONCOCT 0731	90
05	95	91	00	100	00	70	70	70	68	72	73	73	75	Lake Biwa I BMAG hypo bin2	
90	90 Q/	00	99	00	100	71	71	71	67	72	72	72	74	Great Lakes MAG MC14 S12 M CONCOCT 099	~~
68	71	70	70	71	71	100	08	08	68	68	70	72	74	Lake Baikal, MAG Baikal-deenG99	80
67	71	70	70	71	71	08	100	90	68	68	60	70	71	Lake Biwa I BMAG hypo bin3	
67	70	70	70	71	71	08	00	100	68	68	60	60	70	Great Lakes MAG MC17 S15 bin 110	70
64	67	66	67	68	67	68	68	68	100	65	66	66	67	Saline water MAG LIBA6033	70
62	70	71	71	72	71	68	68	68	65	100	75	76	75	Nitrosospira multiformis C-71	~~
63	73	72	73	72	72	70	60	60	66	75	100	70	77	Nitrosospira Inducionnis C-11	60
63	73	72	73	73	72	70	70	60	66	76	77	100	78	Nitrosospira terruis INT	
64	74	74	73	75	74	70	70	70	67	75	77	78	100	Nitrosospira briensis C-120	
<b>&gt;</b>	<u> </u>		0	- -	( <sup>1</sup>			0	(0	-	-	70	>	Nillosospila lacus Al Oo	
OB-LC263	ireat Lakes MAG MC4 S2 CONCOCT 122	ake Baikal MAG Baikal-deep-G101.	ireat Lakes MAG SU08M 6M CONCOCT 0731	.ake Biwa LBMAG hypo.bin2	ireat Lakes MAG MC14 S12 M CONCOCT 099	.ake Baikal MAG Baikal-deepG99	ake Biwa LBMAG hypo.bin3	Breat Lakes MAG MC17 S15 bin 110	aline water MAG UBA6933	Vitrosospira multiformis C-71	Vitrosospira tenuis Nv1	litrosospira briensis C-128	Vitrosospira lacus APG3		

**Figure S4.** Pairwise genome-wide average nucleotide identities (ANI) and average amino acid (AAI) identities for MAG AOB-LC263 (shown in bold) in comparison to closely related MAGs and representatives of the genus *Nitrosospira*. AOB-LC263 represents a novel genus compared to described *Nitrosospira* species based on the threshold of 65% for AAI [37].



**Figure S5.** Phylogeny of MAG AOB-LC263 and contigs AOB-LC199628 and AOB-LC368213 in relation to ammonia-oxidizing bacteria and environmental sequences affiliated with the family *Nitrosomonadaceae* as based on the functional marker gene *amoA*. The maximum likelihood tree was inferred by the IQ-tree algorithm [33] using 459 unambiguous alignment positions of the bacterial *amoA* gene. Branch support was tested with the Shimodaira–Hasegawa approximate likelihood-ratio test (SH-aLRT; 1000 replicates) and ultrafast bootstraps (1000 replicates) within IQ-

tree. Branch support was set as significant at  $\geq$ 80% for SH-aLRT and  $\geq$ 95% for ultrafast bootstrap values (black semi-circles for significant and white for non-significant). MAGs, clones or species with freshwater-origin are colored blue. *Nitrosococcus watsonii* (NC\_014315) and *Nitrosococcus oceani* (NC\_007484) *amoA* genes were used as outgroup. The scale bar indicates 10% estimated nucleic acid sequence divergence. Accession numbers can be found in Table S3.



**Figure S6.** Phylogeny of MAGs NOB-LC29, NOB-LC32, and COM-LC224 in relation to representatives of *Nitrospira* lineage I and II. MAGs affiliated with either comammox clade A or B were taken from the literature [39, 40]. Clade classification of comammox bacteria is based on their *amoA* gene as proposed by Daims *et al.* [41] and Pjevac *et al.* [42]. The phylogenomic maximum likelihood tree was constructed using the IQ-Tree algorithm [33] on the basis of a concatenated amino acid alignment of 120 translated single copy genes that were established by the GTDB-based taxonomy for phylogenetic inference of bacteria [30, 31]. Branch support was tested with the Shimodaira–Hasegawa approximate likelihood-ratio test (SH-aLRT; 1000 replicates) and ultrafast bootstraps (1000 replicates) within IQ-tree. Branch support was set as significant at ≥80% for SH-aLRT and ≥95% for ultrafast bootstrap values (black semi-circles for significant and white for non-significant). MAGs or

species with freshwater-origin are colored blue. *Leptospirillum ferriphilum* (GCA\_900198525.1) and *Leptospirillum ferrooxidans* (GCA\_000284315.1) were used as outgroup. The scale bar indicates 10% estimated amino acid sequence divergence. Accession numbers can be found in Table S3.

ANI

100	10	10	10	98	98	98	98	98	84	78	78	81	79	79	78	73	75	74	74	74	76	75	Lake Fuxian MAG F1-120-MAGs037	100
10	100	75	74	79	87	88	89	89	80	76	77	77	77	77	89	76	74	74	75	76	75	76	Lake Fuxian MAG F1-120-MAGs035	
10	75	100	84	75	76	77	75	76	78	75	75	75	75	75	77	79	77	76	78	79	76	77	Candidatus Nitrospira nitrosa COMA1	80
10	74	84	100	75	74	74	74	74	74	74	76	74	74	74	78	76	76	75	77	78	75	75	Wastewater biofilm MAG RBC069	
98	/9	75 70	75	100	99	99	99	99	85	79	80	82	80	80	76	75	11	75	76	75 70	70	11	NOB-LC29	60
98	87 88	70	74	99 00		100	99 00	99 00	85 85	79	79	0∠ 81	80	80	70	76	77	70	78	70	77	77	Lake Biwa MAG I BMAG hypo hin1	
90	80	75	74	99	99	99	100	100	85	79	79	81	80	80	76	76	77	76	77	76	77	77	Great Lake MAG MC14 S12 M concoct 129	40
98	89	76	74	90	99	90 00	100	100	85	79	79	82	80	80	76	75	77	76	77	76	77	77	Great Lake MAG MC17 S15 bin 99	
84	80	78	74	85	85	85	85	85	100	79	79	80	79	79	75	75	76	75	77	76	77	76	NOB-LC32	20
78	76	75	74	79	79	79	79	79	79	100	80	83	81	81	76	75	75	75	75	76	77	76	Sand filter MAG CG24F	20
78	77	75	76	80	79	79	79	79	79	80	100	83	83	83	77	75	76	75	76	76	77	76	Sand filter MAG CG24A	
81	77	75	74	82	82	81	81	82	80	83	83	100	84	85	76	75	75	74	77	75	76	77	Peat MAG palsa_1310	
79	77	75	74	80	80	80	80	80	79	81	83	84	100	93	76	75	75	75	75	76	76	76	COM-LC224	
79	77	75	74	80	80	80	80	80	79	81	83	85	93	100	76	75	75	75	75	77	77	77	Sand filter MAG CG24C	
78	89	77	78	76	76	75	76	76	75	76	77	76	76	76	100	77	76	76	78	78	77	77	Tapfilter MAG SG-bin2	
73	76	79	76	75	76	76	76	75	75	75	75	75	75	75	77	100	77	77	76	78	76	77	Nitrospira inopinata ENR4NZ	
75	74	77	76	77	77	77	77	77	76	75	76	75	75	75	76	77	100	78	77	77	77	78	Nitrospira japonica NJ11	
74	75	76	/5	75	76	76	/6	76	/5 77	/5 75	75	74	/5 75	/5 75	76	11	/8	100	/6	/9	11	11	Drinking water MAG KM1	
74	75	78	70	76	78	70	76	76	76	75	76	11	75 76	75	78 70	70	11	70	001	80	11	11	Candidatus Nitrospira krettil comreactor 17	
76	70	79	75	13	70	70	70	/0 77	70	70	70	75	70	77	77	76	77	77	80 77	77	100	70		
70	76	70	75	77	77	77	77	77	76	76	76	77	76	77	77	77	78	77	77	77	78	100	Nitrospira renta BSTU Nitrospira moscoviensis NSP M 1	
15	10	"	15	11	"	"	"		10	10	10	11	10	11	11	"	10	"	"	"	10	100	Thurospira moscoviensis NSF IN-T	
											۵Δι													
100	62	67	67	m	07	6	06	06	02	77	76	90	67	70	60	66	70	70	67	60	71	71	Laka Funiar MAO F4 400 MAO-007	100
63	100	75	7/	90 61	9/ 70	90 70	90 70	90	65 65	70	70	00 71	67	70	00	74	10	70 69	75	76	71	71	Lake Fuxian MAG F1-120-MAGs037	100
67	75	100	74 94	60	70	10	60	60	67	10	66	66	62	67	71	74	66	65	75	70	13	74 60	Lake Fuxian MAG F1-120-MAGs035	
67	75	84	100	69	70	69	69	69	67	65	66	65	50	66	71	69	65	64	70	76	68	68	Candidatus Nitrospira nitrosa COMA'i	90
on on	61	69	69	100	08	08	08	08	85	70	78	80	65	80	70	68	72	72	69	70	72	72		
97	70	70	70	98	100	99	99	99	87	80	80	83	72	81	71	69	73	72	70	70	73	73	Lake Baikal MAG Baikal, G1	80
96	70	69	69	98	99	100	99	99	87	79	79	82	71	81	70	68	72	72	69	70	72	72	Lake Biwa MAG LBMAG hypo bin1	
96	70	69	69	98	99	99	100	100	87	79	79	82	71	81	70	69	72	72	69	70	72	73	Great Lake MAG MC14 S12 M concoct 129	
96	69	69	69	98	99	99	100	100	87	80	79	82	72	81	70	69	72	72	69	70	72	73	Great Lake MAG MC17 S15 bin 99	70
83	65	67	67	85	87	87	87	87	100	76	76	79	67	78	68	66	70	69	67	67	70	69	NOB-LC32	
77	70	66	65	79	80	79	79	80	76	100	79	80	72	81	68	66	68	67	67	67	69	68	Sand filter MAG CG24E	60
76	70	66	66	78	80	79	79	79	76	79	100	80	73	83	68	66	68	68	66	67	68	68	Sand filter MAG CG24A	
80	71	66	65	80	83	82	82	82	79	80	80	100	72	83	67	66	69	68	66	67	69	69	Peat MAG palsa_1310	
6/	62	60	59	65	72	/1	/1	72	67	72	73	72	100	82	61	60	61	60	61	61	62	61	COM-LC224	
78	71	6/ 71	00	80	81	81	70	81	78	81	83	83	82	100	100	67	69	69	6/ 72	08	69 70	09 71	Sand filter MAG CG24C	
66	7/	70	60	68	60	68	60	60	66	66	66	66	60	67	72	100	66	65	73	73	68	70	Nitrospira inopinata ENRANZ	
70	69	66	65	72	73	72	72	72	70	68	68	69	61	69	66	66	100	72	65	66	68	68	Nitrospira inopinata ENICENZ	
70	68	65	64	72	72	72	72	72	69	67	68	68	60	69	65	65	72	100	65	66	67	67	Drinking water MAG KM1	
67	75	76	75	69	70	69	69	69	67	67	66	66	61	67	73	70	65	65	100	77	69	69	Candidatus Nitrospira kreftii comreactor17	
68	76	77	76	70	70	70	70	70	67	67	67	67	61	68	73	70	66	66	77	100	69	70	Candidatus Nitrospira nitrificans COMA2	
71	73	68	68	72	73	72	72	72	70	69	68	69	62	69	70	68	68	67	69	69	100	72	Nitrospira lenta BS10	
71	74	69	68	72	73	72	73	73	69	68	68	69	61	69	71	70	68	67	69	70	72	100	Nitrospira moscoviensis NSP M-1	
لم ا	5	ູດ	≶	z	5	5	۵	۵	z	ŝ	ŝ	P€	2	ŝ	ਙ	S	Ş	Q	ß	ĉ	₹	Z		
ke	ike	anc	ast	B	ke	ke	rea	rea	B	but	anc	eat	Ň	and	βþ	itro	itro	ij	anc	anc	itro	itro		
Ē	Ę	tide	ev.	È	β,	<u>₿</u>	Ê	Ē	È	Ē	f	Ž	Ē	Ē	lter	spi	spi	inc	fide	fida	spi	spi		
Xie	X.	atu	/ate	ស្ល	ik.	wa	ake	ake	8	ēŗ	ter	å	ស្ល	fér	Ś	ra	a	ξ	atu	atu	ľa	ľa		
ñ	n L	s 7	¥,	G	N	₹	Š	Š	N	Ž	Ž	pa	24	ž	AG	ino	jap	ate	۶Z	۶Z	len	mo		
Ă٨	Ă٨	litrc	ŏ		Α	Å	A	A		å	Ĝ	Isa		á	ŝ	pin	Ŋ,	۳	litro	litro	ťa	SC		
۵	G	g	îlm		G	Ē	2	2		0	0	Ľ.		0	Ģ	ata	ica	Ă	dsc	dsc	BS	40		
Ξ.	Ц	ira	Z		3ail	М,	ถุ	ð		32	35	31		32	jn:	Ē	Ę	с) Т	vira	vira	10	ens		
-12	4	ni	AG		a	AG	14	17		ħ	4A	0		đ	N	R	11	ŝ	Ā	nii		sis		
	-02	frog	R		ה'	j	<u>s</u> 1	5								47		-	efti	trifi		N N		
ΑN	Ř	ä	ВС		<u> </u>	od/	2	с С								N			2	cal		Ď		
Gs	ຼົດ	8	ő			<u>b</u>	2	Ę.											m	ns (		Ę		
03.	503	ž	9			7	onc	99											rea	8		-		
7	ΰï	7					ğ												otc	Ň				
							<u>1</u>												Ĕ	ŝ				
							29												7					

**Figure S7.** Pairwise genome-wide average nucleotide identities (ANI) and average amino acid identities (AAI) for lineage II *Nitrospira* including MAGs NOB-LC32, NOB-LC29, and COM-LC224 (shown in bold).

#### Supplementary Information

RH-aLRT ①UF BT		Candidatus Nitrospira defluvii ERS379726	Nitrospira Lineage I				
significant  not sign	ificant	L <i>Nitrospira</i> sp. ND1					
		NOB-LC32	Nitrospira Lineage II				
	e 1	INOB-LC29					
		Lake Baikal MAG Baikal_G1					
		Great Lake MAG MC14 S12 M concoct 129					
		Great Lake MAG MC17 S15 bin 99					
		Lake Biwa MAG LBMAG hypo bin1					
		Nitrospira lenta BS10					
Comammox		Nitrospira inopinata ENR4					
Clade A		Biofilter MAG SG-bin2					
	թե	Sand filter MAG CG24E					
		Candidatus Nitrospira kreftii comreactor17					
		Candidatus Nitrospira nitrificans COMA2					
		———— Candidatus Nitrospira nitrosa COMA1					
Comammox							
Clade B		- COM-LC224					
		- Sand filter MAG CG24C					
	Nitro	ospira moscoviensis M-1 clone 17					
	Niti	rospira moscoviensis NSP M-1					
	L Nitro	ospira moscoviensis nxrB clone 20					
	Was	ste water MAG H3_NOB1					
	Wa:	ste water MAG OLB3					
	Nit	trospira japonica NJ11					
_	Drir	nking water MAG KM1					
	Candic	datus Nitrospira bockiana clone 12	<i>Nitrospira</i> Lineage V				
	Candia	datus Nitrospira bockiana clone 15					
	Candio	datus Nitrospira bockiana clone 17					
ρ	Nitrospira	a calida Ns10 clone 3	Nitrospira Lineage VI				
-    <sup>.</sup> "	Nitrospira	a calida Ns10 clone 1					
	sp. Ecoma	res 2 1 clone 6	Nitrospira Lineage IV				
Nitrospira	sp. Ecoma	ares 2.1 clone 4					
rNitrospira mar	ina ATCC 4	43039 clone 7					
Nitrospira mai	rina ATCC	43039 clone 9					
Nitrospira mar	ina ATCC 4	13039 clone 17					
Nitroenira mar	ina ATCC	13039 done 12					
Nitrospira mar	ina ATCC	43039 clone N4					
Nitrospira marii	na ATCC 4	3039 clone11	0.1				

**Figure S8.** Phylogeny of MAGs NOB-LC29, NOB-LC32, and COM-LC224 as based on the functional marker gene *nxrB*. Classification of the comammox MAGs into clade A or B was transferred from the phylogenomic tree (Fig. S6), as comammox species cannot be distinguished based on their *nxrB* gene from canonical *Nitrospira* species [41, 43]. The maximum likelihood tree was inferred by the IQ-tree algorithm [33] using 1,205 unambiguous alignment positions of the *nxrB* gene. Branch support was tested with the Shimodaira–Hasegawa approximate likelihood-ratio test (SH-aLRT; 1000 replicates)

and ultrafast bootstrap (1000 replicates) within the IQ-tree software package. Branch support was set as significant at  $\geq$ 80% for SH-aLRT and  $\geq$ 95% for ultrafast bootstrap values (black semi-circles for significant and white for non-significant). MAGs, clones or species with freshwater-origin are colored blue. *Hydrogenobaculum* sp. (NC\_011126), *Natronomonas pharaonis* (NC\_007426.1) and *Candidatus* Kuenenia stuttgartiensis (CT573072) *narH*-like genes were used as outgroup. The scale bar indicates 10% estimated nucleic acid sequence divergence. Accession numbers can be found in Table S3.



Figure S9. Phylogeny of COM-LC224 as based on the functional marker gene amoA. Classification into clade A or B was done as proposed by by Daims et al. [41] and Pjevac et al. [42]. The maximum likelihood tree was inferred by the IQ-tree algorithm [33] using 414 unambiguous alignment positions of the amoA gene. Branch support was tested with the Shimodaira–Hasegawa approximate likelihood-ratio test (SH-aLRT; 1000 replicates) and ultrafast bootstrap (1000 replicates) within the IQ-tree software package. Branch support was set as significant at  $\geq$ 80% for SH-aLRT and  $\geq$ 95% for ultrafast bootstrap values (black semi-circles for significant and white for non-significant). MAGs or species with freshwater-origin are colored blue. Nitrosospira multiformis (U91603), Nitrosospira sp. (WP 041514847.1), Nitrosomonas europaea (L08050) and Nitrosomonas communis (WP\_046851395) amoA genes were used as outgroup. The scale bar indicates 10% estimated nucleic acid sequence divergence. Accession numbers can be found in Table S3.

#### Supplementary Information



**Figure S10.** Network analysis of ammonia oxidation-related and co-transcribed AOA-LC4 genes over the yearly cycle. Only genes with a strong and significant (Spearman's  $r \ge 0.8$ , FDR-corrected p-value <0.05) correlation to at least two of the *amoABC* genes and an average expression higher than the median of all transcribed genes were considered for the network analysis. Correlation coefficient's strength is visualized by increasing width of the edges. Detailed gene annotation and transcription values can be found in Table S2.



**Figure S11.** <sup>15</sup>N-nitrite/nitrate production from <sup>15</sup>N-ammonium in incubations of hypolimnetic water taken from 85 m depth. Produced <sup>15</sup>N-nitrite/nitrate is shown in relation to the first sampling time point. Linear regression through the time points was used to infer ammonia oxidation rates. All incubations were performed in biological triplicates at 4°C in the dark over a period of 48 h, except for June with 67 h. Samples were taken in 2019 on June 18<sup>th</sup>, July 29<sup>th</sup>, August 28<sup>th</sup> and November 5<sup>th</sup>. Incubations started within 1–7 h after sampling.