1   ***AmoA*-targeted polymerase chain reaction primers for the specific detection and**

2   **quantification of comammox *Nitrospira* in the environment**

3

4   Petra Pjevac[1], Clemens Schauberger[1§], Lianna Poghosyan[2§], Craig W. Herbold[1], Maartje A.

5   H.J. van Kessel[2], Anne Daebeler[1], Michaela Steinberger[1], Mike S. M. Jetten[2], Sebastian

6   Lücker[2]*, Michael Wagner[1], Holger Daims[1]*

7

8   [1]Department of Microbiology and Ecosystem Science, Division of Microbial Ecology,

9   Research Network 'Chemistry meets Microbiology', University of Vienna, Althanstrasse 14,

10   1090 Vienna Austria

11   [2]Department of Microbiology, IWWR, Radboud University, Heyendaalseweg 135, 6525 AJ

12   Nijmegen, the Netherlands

13   [§]These authors contributed equally to this study.

14

15   ***Corresponding authors:** Holger Daims, Department of Microbiology and Ecosystem

16   Science, Division of Microbial Ecology, Research Network 'Chemistry meets Microbiology',

17   University of Vienna, Althanstrasse 14, 1090 Vienna, Austria. Phone: +43 1 4277 76604. E-

18   mail: daims@microbial-ecology.net

19   Sebastian Lücker, Department of Microbiology, IWWR, Radboud University,

20   Heyendaalseweg 135, 6525 AJ Nijmegen, the Netherlands. Phone: +31 24 36 52564. E-

21   mail: s.luecker@science.ru.nl

22

23   **Keywords:** nitrification, comammox, *Nitrospira*, *amoA*, marker gene, PCR

24

25   **Running title:** Comammox *amoA*-targeted PCR primers

Abstract

Nitrification, the oxidation of ammonia via nitrite to nitrate, has always been considered to be catalyzed by the concerted activity of ammonia- and nitrite-oxidizing microorganisms. Only recently, complete ammonia oxidizers ('comammox'), which oxidize ammonia to nitrate on their own, were identified in the bacterial genus *Nitrospira*, previously known to contain only canonical nitrite oxidizers. *Nitrospira* are widespread in nature, but for assessments of the distribution and functional importance of comammox *Nitrospira* in ecosystems cultivation-independent tools to distinguish comammox from strictly nitrite-oxidizing *Nitrospira* are required. Here we developed new PCR primer sets that specifically target the *amoA* genes coding for subunit A of the distinct ammonia monooxygenase of comammox *Nitrospira*. While existing primers capture only a fraction of the known comammox *amoA* diversity, the new primer sets cover as much as 95% of the comammox *amoA* clade A and 92% of the clade B sequences in a reference database containing 326 comammox *amoA* genes with sequence information at the primer binding sites. Application of the primers to 13 samples from engineered systems (a groundwater well, drinking water treatment and wastewater treatment plants) and other habitats (rice paddy and forest soils, rice rhizosphere, brackish lake sediment and freshwater biofilm) detected comammox *Nitrospira* in all samples and revealed a considerable diversity of comammox in most habitats. Excellent primer specificity for comammox *amoA* was achieved by avoiding the use of highly degenerate primer preparations and by using equimolar mixtures of oligonucleotides that match existing comammox *amoA* genes. Quantitative PCR with these equimolar primer mixtures was highly sensitive and specific, and enabled the efficient quantification of clade A and clade B comammox *amoA* gene copy numbers in environmental samples. Thus, the new comammox *amoA*-targeted primers will enable more encompassing studies of nitrifying microorganisms in diverse ecosystems.

Introduction

Nitrification is an essential process of the global biogeochemical nitrogen cycle and plays a pivotal role in biological wastewater treatment and in drinking water production. The recent discovery of the first complete ammonia oxidizers ('comammox') in the bacterial genus *Nitrospira* (Daims *et al.*, 2015; van Kessel *et al.*, 2015) was highly unexpected since *Nitrospira* were always regarded as canonical nitrite-oxidizing bacteria (NOB) (Watson *et al.*, 1986; Ehrich *et al.*, 1995; Spieck *et al.*, 2006; Lebedeva *et al.*, 2008; 2011). This discovery has raised the important questions of how commonly comammox *Nitrospira* occur in nitrifying microbial communities and how relevant they are for nitrification compared to ammonia-oxidizing bacteria (AOB), archaea (AOA), and NOB.

Members of the genus *Nitrospira* have been assigned to at least six sublineages (Daims *et al.,* 2001; Lebedeva *et al.*, 2008; 2011). They are widespread in virtually all oxic natural and engineered ecosystems, and an impressively high diversity of coexisting uncultured *Nitrospira* strains has been detected in wastewater treatment plants and soils (Pester *et al.*, 2014; Gruber-Dorninger *et al.*, 2015). All known comammox organisms belong to *Nitrospira* sublineage II (Daims *et al.*, 2015; van Kessel *et al.*, 2015; Pinto *et al.,* 2015; Palomo *et al.,* 2016). This sublineage contains also canonical NOB, which lack the genes for ammonia oxidation (Daims *et al.,* 2001, Koch *et al.*, 2015). Moreover, the known comammox *Nitrospira* do not form a monophyletic clade within *Nitrospira* lineage II in phylogenies based on 16S rRNA or *nxrB*, the gene encoding subunit beta of the functional key enzyme nitrite oxidoreductase. Instead, they mingle with the strict NOB in these phylogenetic trees (Daims *et al.*, 2015; van Kessel *et al.*, 2015; Pinto *et al.,* 2015). Finally, it remains unknown whether other *Nitrospira* sublineages (Daims *et al.,* 2001; Lebedeva *et al.*, 2008, 2011) also contain comammox members. Consequently, it is impossible to infer from 16S rRNA gene or *nxrB* phylogenies whether yet uncharacterized *Nitrospira* bacteria are comammox or strict NOB, although such attempts have been published (Gonzalez-Martinez *et al.*, 2016).

Intriguingly, comammox *Nitrospira* possess novel types of ammonia monooxygenase (AMO) and hydroxylamine oxidoreductase (HAO), the key enzymes of aerobic ammonia oxidation. The comammox AMO is phylogenetically distinct from the AMO forms of canonical AOB and

80   AOA (Daims *et al.*, 2015; van Kessel *et al.*, 2015). The *amoA* gene encoding AMO subunit A

81   has become a widely used functional and phylogenetic marker gene for bacterial and

82   archaeal ammonia oxidizers (Rotthauwe *et al.,* 1997; Purkhold *et al.*, 2000; Juniper *et al.*,

83   2008; Pester *et al.*, 2012). Public database mining for sequences related to the unique

84   comammox *amoA* revealed the presence of putative comammox organisms in various

85   environments including soils (paddy rice soils, other agricultural soils, forest soils, grassland

86   soils), freshwater habitats (wetlands, rivers, lakes, groundwater basins), groundwater wells

87   (GGWs), full-scale wastewater treatment plants (WWTPs), and drinking water treatment

88   plants (DWTPs) (Daims *et al.*, 2015; van Kessel *et al.*, 2015). While this provides preliminary

89   evidence of a broad habitat range of comammox organisms, current knowledge about the

90   environmental distribution and abundance of comammox still is very limited.

91   The *amoA* genes of comammox *Nitrospira* form two monophyletic sister clades (Daims *et al.*,

92   2015; van Kessel *et al.*, 2015), which are referred to as clade A and clade B. Clade A also

93   contains some genes that were previously assigned to the methanotroph *Crenothrix*

94   *polyspora* (Stoecker *et al.,* 2006), but this assignment has recently been corrected (Oswald

95   *et al.,* in revision). Established PCR primer sets specifically targeting the *amoA* genes of

96   AOB or AOA (Sinigalliano *et al.*, 1995; Rotthauwe *et al.*, 1997; Juretschko *et al.*, 1998;

97   Stephen *et al.,* 1999; Nold *et al.*, 2000; Norton et al., 2002; Webster *et al.*, 2002; Okano *et

98   al.*, 2004; Francis et al., 2005; Treusch *et al.*, 2005; Mincer *et al.*, 2007; Juniper *et al.*, 2008;

99   Meinhardt *et al.,* 2015) do not amplify any clade A and clade B comammox *amoA*. Co-

100   amplification of some comammox *amoA* genes occurs with a primer set targeting both

101   betaproteobacterial *amoA* and the A subunit of the particulate methane monooxygenase

102   (*pmoA* of pMMO; Holmes *et al.*, 1995; Luesken *et al.,* 2011). These primers, however, only

103   target a fraction of the known comammox *amoA* genes. A recently established two-step PCR

104   protocol (Wang *et al.*, 2016) relies on the forward primer published by Holmes and

105   colleagues (1995) and a new, highly degenerate reverse primer. It detects the bacterial

106   copper-containing monooxygenase (CuMMO) genes including *pmoA*, betaproteobacterial

107   *amoA*, and at least the comammox *amoA* genes that are amplified by the Holmes forward

108   primer. However, because of its broad coverage of the CuMMOs, this primer pair does not

109   allow a specific detection or quantification by quantitative PCR (qPCR) of comammox *amoA*

110   genes.

111   To enable the direct detection and quantification of comammox *amoA* genes in

112   environmental samples, we designed in this study two new *amoA*-targeted primer sets

113   specific for clade A or clade B comammox *amoA* genes, respectively. Subsequently, we

114   applied these new primers to efficiently screen various habitats for the presence of

115   comammox organisms and to rapidly quantify the abundances of comammox in selected

116   samples.

117

118   **Materials and Methods**

119   *Database mining and sequence collection*

120   The amino acid sequences of bacterial AmoA and PmoA were extracted from publicly

121   available metagenomic datasets stored in the Integrated Microbial Genomes databases

122   (IMG-ER and -MER) using the functional profiler tool against a specific bacterial

123   AmoA/PmoA pfam (PF02461). For characterization of comammox AmoA,

124   betaproteobacterial AmoA, and PmoA, sequences collected from the Integrated Microbial

125   Genomes databases were augmented with nearly full length amino acid sequences collected

126   from the Pfam site and from NCBI Genbank as described in Daims *et al.* (2015). Collected

127   sequences were filtered against a hidden Markov model (hmm) using hmmsearch

128   (http://hmmer.janelia.org/) with the AmoA/PmoA hmm (PF02461) requiring an expect value

129   of <0.0001, and clustered at 90% identity using USEARCH (Edgar, 2010). Cluster centroids

130   were aligned using Mafft (Katoh *et al.*, 2002) and phylogenetic affiliation with comammox

131   AmoA or betaproteobacterial AmoA was verified with phylogenetic trees calculated using

132   FastTree (Prince *et al.*, 2009). Cluster centroids were expanded and reclustered at 99%

133   identity with USEARCH and the phylogenetic affiliation with comammox and/or

134   betaproteobacterial AmoA was re-assessed using FastTree.

135

136 *Phylogenetic analyses*

137 For amino acid sequences that clustered with comammox AmoA and a selection of outgroup

138 sequences affiliating with bacterial AmoA and PmoA, the corresponding nucleic acid

139 sequences were recovered and aligned according to their amino acid translations using

140 MUSCLE (Edgar, 2004). These nucleotide sequence alignments were manually corrected

141 for frameshifts and used for calculating maximum likelihood trees with RaxmlHPC

142 (Stamatakis *et al.,* 2005) using the GTRCAT approximation of rate heterogeneity and 1,000

143 bootstrap iterations. The generated reference tree was imported into the ARB software

144 package (Ludwig *et al.,* 2004) and used as reference tree for the subsequent addition of

145 short and nearly full-length comammox *amoA*, betaproteobacterial *amoA* and *pmoA*

146 nucleotide sequences that were recovered from public databases. These sequences were

147 added to the reference tree without changes of the overall tree topology by using the

148 "parsimony interactive" tool of ARB.

149

150 *Primer design and PCR*

151 Two degenerate PCR primer pairs (Table 1) targeting clade A or clade B comammox *amoA*

152 genes, respectively, were designed and evaluated using the *amoA*/*pmoA* reference

153 sequence dataset and the 'probe design' and 'probe match' functions of ARB (Ludwig *et al*.,

154 2004). The oligonucleotide primers were obtained from Biomers (Ulm, Germany).

155 The optimal annealing temperature for the primers was determined by temperature gradient

156 PCR using total genomic DNA extracted by a phenol-chloroform based method (details

157 below) from a GWW sample (Wolfenbüttel, Germany), which contained both clade A and

158 clade B comammox *Nitrospira* (Daims *et al.,* 2015). An annealing temperature range from 42

159 to 52 °C was chosen for experimental evaluation based on the theoretical melting

160 temperature of 48 °C of the designed primers

161 (http://biotools.nubic.northwestern.edu/OligoCalc.html). Thermal cycling was carried out with

162 an initial denaturation step at 94 °C for 5 min, which was followed by 25 cycles of

163 denaturation at 94 °C for 30 s, primer annealing at 42 to 52 °C for 45 s, and elongation at

164 72 °C for 1 min. Cycling was completed by a final elongation step at 72 °C for 10 min. The

165   optimal annealing temperature for both primer pairs was found to be 52 °C. At lower

166   annealing temperatures, unspecific amplification of DNA fragments shorter than the

167   expected amplicon length (415 bp) was observed. Since the amplification efficiency of

168   correctly sized amplicons was lower at 52 °C than at 48-50 °C, temperatures above 52 °C

169   were not evaluated. The PCR reactions were performed in a DreamTaq Green PCR mix with

170   1× Dream Taq Green Buffer containing 2 mM $MgCl_2$, 0.025 U DreamTaq DNA polymerase,

171   0.2 mM dNTPs, 0.5 µM primers and 0.1 mg/mL bovine serum albumin (Fermentas, Thermo

172   Fischer Scientific, Waltham, MA, USA).

173

174   *Screening of environmental samples for comammox* amoA

175   Total nucleic acids were extracted from the environmental samples (Table 2) by using i) the

176   Fast DNA SPIN Kit for soil (MP Biomedicals, Santa Ana, CA, USA), ii) the PowerSoil DNA

177   Isolation Kit (MO BIO Laboratories, Carlsbad, CA, USA), or iii) a phenol-chloroform based

178   total nucleic acid extraction protocol described by Angel *et al.,* (2012). PCRs were performed

179   with either the DreamTaq Green PCR mix as described above, or with the PerfeCTa SYBR

180   green FastMix PCR premix containing 1.5 mM of $MgCl_2$ (Quanta Bioscience, Beverly, MA,

181   USA) and 0.1 µM primers. Amplicons of the expected length (415 bp) were purified by using

182   the QIAquick PCR purification Kit (Qiagen, Hilden, Germany). Alternatively, the QIAquick Gel

183   extraction Kit (Qiagen, Hilden, Germany) was used to purify PCR products of the correct

184   length from agarose gels if additional, unspecific amplification products of different lengths

185   were present. The purified PCR products were cloned in *E. coli* with the TOPO-TA cloning kit

186   (Invitrogen, Karlsruhe, Germany) or the pGEM-T Easy Vector System (Promega, Mannheim,

187   Germany) by following the manufacturer's instructions. PCR products obtained from cloned

188   vector inserts from the TOPO-TA kit were Sanger-sequenced with the M13F primer at

189   Microsynth (Balgach, Switzerland). For clones generated with the pGEM-TEasy Vector

190   SysteM, plasmids were extracted with the GeneJET plasmid miniprep kit (Thermo Fischer

191   Scientific, Waltham, MA, USA) and the inserts were Sanger-sequenced with the M13F

192   primer at Baseclear (Leiden, the Netherlands). The recovered sequences were quality

193   checked and vector trimmed using the Sequencher version 4.6.1 (GeneCodes Corporation,

7

194   Ann Arbor, MI, USA) or the Chromas version 2.6.1 (Technelysium Pty Ltd, Brisbane,

195   Australia) software packages, and were added as described above to the *amoA*/*pmoA*

196   nucleotide sequence reference tree in ARB.

197

198   *Quantitative PCR*

199   Quantification of comammox *amoA* genes was performed with individually prepared

200   equimolar primer mixes for comammox *amoA* clade A and clade B genes (Table 1) on a

201   selection of samples from different habitats (Table 3). Amplification of comammox *amoA*

202   genes was performed with 3 min initial denaturation at 95 °C, followed by 45 cycles of 30

203   sec at 95 °C, 45 sec at 52 °C, and 1 min at 72 °C. Fluorescence was measured at 72 °C for

204   amplicon quantification. After amplification, an amplicon melting curve was recorded in

205   0.5 °C steps between 38 and 96 °C. All qPCR assays were performed on a Bio-Rad C1000

206   CFX96 Real-Time PCR system (Bio-Rad, Hercules, CA, USA) in Bio-Rad iQ SYBR Green

207   Supermix (Bio-Rad, Hercules, CA, USA), containing 50 U/ml iTaq DNA polymerase, 0.4 mM

208   dNTPs, 100 mM KCl, 40 mM Tris-HCl, 6 mM $MgCl_2$, 20 mM fluorescein, and stabilizers. The

209   respective equimolar primer mixtures (Table 1) were added to a final concentration of

210   0.5 µM. Environmental DNA samples were added to each PCR reaction to a final volume of

211   20 µl reaction mix. Different template DNA dilutions (10x to 10,000x) resulting in DNA

212   concentrations from 0.004 to 24 ng were applied to minimize possible PCR inhibition caused

213   by excess DNA or co-extracted substances.

214   All assays were performed in triplicate. For each assay, triplicate standard series were

215   generated by tenfold serial dilutions (10^1-10^8 gene copies/µl) from purified M13-PCR

216   products obtained from cloned vector inserts generated with the TOPO-TA cloning kit as

217   described above. Clones of comammox *amoA* genes used for standard curve generation

218   originated from this study. The correlation coefficient (r^2) for each of the external standard

219   curves was ≥0.998. The amplification efficiency for comammox *amoA* clade A and clade B

220   genes was 88.5% and 87.0%, respectively.

221

Results and Discussion

Design and evaluation of comammox amoA-*specific PCR primers*

A principal goal of this study was to develop a specific *amoA*-targeted PCR assay for the efficient and cultivation-independent detection and quantification of comammox *Nitrospira* in environmental samples. Mining of public databases retrieved only a small number (376) of sequences related to clade A or clade B comammox *amoA*, and only 32 of these sequences cover almost the full length of the comammox *amoA* genes (840-846 bp for clade A, 885-909 bp for clade B according to the available genomic and metagenomic datasets from comammox *Nitrospira*). To achieve a good coverage of the known comammox *amoA* gene diversity, primer design was restricted to the range of aligned nucleotide positions for which sequence information exists in at least 80% of the publicly available comammox *amoA* sequences. Furthermore, availability of sequence information for at least one of the primer binding sites in the great majority of betaproteobacterial *amoA* gene sequences (14,000/14,019) and bacterial *pmoA* gene sequences (9,913/10,289) facilitated the design of highly comammox *amoA*-specific primers. Two degenerate primer pairs (ComaA-244F/659R and ComaB-244F/659R) targeted 95% and 83% of the *amoA* gene sequences from comammox clade A or B, respectively (Table 1). *In silico* analyses confirmed that each primer contained at least 4 mismatches to all sequences within the respective other comammox *amoA* clade. Furthermore, at least 3 mismatches to almost all sequences affiliated with betaproteobacterial *amoA* and with *pmoA* genes were present. Merely 23 out of 14,019 betaproteobacterial *amoA* had only two mismatches to the comammox *amoA* clade A-targeted primers. Further, only 22 and 2 out of 9,913 *pmoA* gene sequences showed only two mismatches to the comammox *amoA* clade A- and clade B-targeted primers, respectively. No betaproteobacterial *amoA* and no *pmoA* gene sequences had less than two mismatches to any of the comammox *amoA*-targeted primer pairs.

The new primer sets were first evaluated with a pasty iron sludge sample from the GWW Wolfenbüttel, which was known to harbor both clade A and clade B comammox *Nitrospira* according to a previous metagenomic analysis (Daims *et al.*, 2015). A single amplicon of the expected length (415 bp) was obtained from this sample with both primer pairs after PCR

9

251   with an annealing temperature of 52 °C. Cloning and sequencing of the PCR products

252   exclusively retrieved sequences of comammox *amoA* from clade A or clade B, respectively

253   (Fig. 1). After application of the two primer sets to 12 additional samples from various

254   environments, comammox *amoA* clade A was detected in 11 samples and clade B in 7

255   samples (see below and Table 2). Single, correctly sized amplicons were obtained from the

256   majority of samples (8 of the clade A-positive samples and 6 of the clade B-positive

257   samples, Table 2). In the remaining cases, unspecific PCR products of different lengths were

258   observed in addition to the expected *amoA* amplicons. These unspecific PCR products

259   appeared irrespectively of the applied DNA extraction method and the number of PCR

260   cycles, but were overcome by either purifying the target *amoA* amplicon by agarose gel

261   excision (for the lake Herrensee sample) or by a PCR-based clone screening for the right

262   insert size after cloning (for the remaining samples with unspecific PCR products). In

263   summary, we retrieved 444 comammox *amoA* gene sequences in this study by PCR, cloning

264   and sequencing from the 13 environmental samples (Table 2). Only two of the 446 CuMMO

265   sequences did not cluster with comammox *amoA* clades, but were classified as *pmoA*.

266   Furthermore, one sequence obtained with the comammox *amoA* clade A-targeted primers

267   turned out to be affiliated with comammox *amoA* clade B. Thus, the degenerate comammox

268   *amoA*-targeted primers offer a straightforward, fast and robust approach for the detection

269   and identification of comammox *Nitrospira* in complex samples. Notably, with an amplicon

270   size of 415 bp the primers are also suitable for high-throughput amplicon sequencing by

271   current Illumina technologies.

272   Since amplification of non-target DNA was observed with some samples, we took into

273   account that primer degeneracy may impair PCR specificity (Linhart and Shamir, 2002).

274   Furthermore, primer degeneracy can cause an uneven amplification of different sequence

275   variants (Polz and Cavanaugh, 1998). In general, less degenerate, defined primer mixtures

276   can reduce unspecific primer binding and yield cleaner amplification products from complex

277   environmental samples (Linhart and Shamir, 2002). This is of particular importance for

278   obtaining correct population size estimates in qPCR assays. Consequently, less degenerate

279   primer mixtures consisting of separately synthesized versions of the respective forward and

280   reverse primer, with each primer version containing only one or no base ambiguity (Table 1),

281   were also tested. All selected primer versions in these mixtures matched real comammox

282   *amoA* gene sequences in our database. In contrast, the original (more degenerate) primer

283   sets inevitably contained also co-synthesized oligonucleotide versions that did not match any

284   known target comammox *amoA* sequence. The less degenerate mixtures offered the same

285   coverage of comammox *amoA* clade A (95%), and by adding one additional oligonucleotide

286   to the forward primer mix we even increased the coverage of comammox *amoA* clade B

287   from 83 to 92%. Considerably improved PCR results were achieved by applying these

288   primer mixtures to the samples prone to unspecific amplification, as only amplicons of the

289   expected size were obtained after PCR (data now shown). Based on their improved

290   specificity and the higher coverage of comammox *amoA* clade B, the use of the defined,

291   manually pooled equimolar primer mixtures (Table 1) is recommended. For comammox

292   *amoA*-targeted qPCR experiments that use the primer sets presented here, application of

293   the less degenerate, equimolar primer mixtures is mandatory.

294

295   *Environmental detection and diversity of comammox* Nitrospira

296   Intriguingly, comammox *amoA* genes were detected in all samples (Table 2) encompassing

297   forest soil, rice paddy soils and rice rhizosphere, a freshwater biofilm and brackish lake

298   sediment, as well as WWTPs and DWTPs. After cloning and sequencing, the obtained *amoA*

299   sequences were clustered in operational taxonomic units (OTUs) based on a sequence

300   identity threshold of 95% (Francis *et al.,* 2003). It should be noted that these OTUs might not

301   delineate species, as an appropriate species-level sequence identity cutoff for comammox

302   *amoA* remains unknown. This would need to be determined by correlating *amoA* sequence

303   identities with 16S rRNA identities or with genome-wide average nucleotide identities

304   (Purkhold *et al.*, 2000; Richter and Roselló-Mora, 2009; Pester *et al.*, 2014) once more

305   genomic data from comammox *Nitrospira* become available. A phylogenetic analysis of the

306   *amoA* gene OTU representatives revealed a substantial diversity of comammox *Nitrospira* in

307   almost all samples (Fig. 1, Table 2, Table S1). Both clade A and clade B comammox

308   *Nitrospira* were detected in seven samples. Only clade A members were found in the

309   WWTPs, the brackish lake sediment, and the river biofilm enrichment, whereas we retrieved

310   only clade B *amoA* gene sequences from the forest soil (Table 2). The widespread

311   occurrence of comammox *Nitrospira* in the analyzed samples is consistent with the

312   previously reported presence of mostly uncultured and uncharacterized *Nitrospira* members

313   in the respective habitat types (e.g. Daims *et al.*, 2001; Martiny *et al.*, 2005; Ke *et al.*, 2013;

314   Pester *et al.*, 2014). Future studies should determine which fraction of these environmental

315   *Nitrospira* are strict NOB or comammox, respectively. The presence of multiple comammox

316   OTUs, and in particular the co-occurrence of both clade A and clade B comammox, in

317   several samples (Table 2) indicates ecological niche partitioning that enables the

318   coexistence of different comammox strains. For co-occurring *Nitrospira* in WWTPs, niche

319   partitioning based on the preferred nitrite concentrations, on the capability to utilize formate

320   as an alternative substrate, and on the tendency to co-aggregate with AOB has already been

321   demonstrated (Maixner *et al.*, 2006; Gruber-Dorninger *et al.*, 2015). It is tempting to

322   speculate that underlying mechanisms of niche differentiation of comammox *Nitrospira* could

323   also be different substrate (ammonia) concentration optima and alternative energy

324   metabolisms, such as the oxidation of formate and of hydrogen (Koch *et al.*, 2014; Koch *et*

325   *al.*, 2015).

326

327   *Quantification of comammox* amoA *genes by quantitative PCR*

328   Since the discovery of AOA (Könneke *et al.*, 2005), numerous studies compared the

329   abundances of AOB and AOA, and tried to assess the contributions of different groups of

330   ammonia oxidizers to nitrification, in various engineered and natural environments (e.g.

331   Leininger *et al.*, 2006; Chen *et al.*, 2008; Jia and Conrad, 2009; Mussmann *et al.*, 2011;

332   Daebeler *et al.*, 2012; Ke *et al.*, 2013; Bollmann *et al.*, 2014). To adequately investigate the

333   niche differentiation and labor partitioning between ammonia oxidizers, it will be necessary to

334   include comammox *Nitrospira* in such comparisons. This will require a robust method for the

335   rapid and accurate quantification of comammox *Nitrospira*. Thus, we established qPCR

336   assays using the equimolar primer mixtures that target comammox *amoA* clade A or clade

337   B, respectively. Both assays had a high efficiency, accuracy and sensitivity, as the

12

338 quantification of as few as ten copies of comammox *amoA* standards was achieved

339 (Supplementary Figure 1).

340 As proof of applicability, we quantified comammox *amoA* clade A and clade B genes in five

341 different sample types: one activated sludge sample (WWTP VetMed, Vienna); paddy rice

342 soil from Vercelli, Italy; beech forest soil from Klausen-Leopoldsdorf, Austria; pasty iron

343 sludge from the riser pipe of GWW Wolfenbüttel, Germany; and the trickling filter sample

344 from the DWTP Friedrichshof, Germany (Table 3). Melting curve analyses confirmed the

345 specificity of both assays (Supplementary Figure 2). Amplification of non-target DNA

346 occurred only in those samples that lacked the respective target organisms of the primers

347 according to the aforementioned results of the end point PCR (i.e., WWTP VetMed for

348 clade B and Klausen-Leopoldsdorf forest soil for clade A comammox) and was easily

349 detectable based on a distinct slope of the melting curves of the unspecific amplicons

350 (Supplementary Figure 2). The abundances of clade A and clade B comammox *amoA* were

351 in the same order of magnitude in the samples containing both groups (Table 3).

352

353 *Summary and Outlook*

354 The newly developed primer sets were designed for maximal coverage of the known

355 comammox *Nitrospira* lineages, high specificity for comammox *amoA* genes, and a broad

356 range of applications including high-throughput amplicon sequencing and qPCR. By using

357 these primer sets, we detected the presence of comammox *Nitrospira* in a variety of

358 habitats. In this methodological study, our primary goal was to evaluate the applicability of

359 the newly designed primers for end point PCR and for qPCR with a selection of different

360 sample types. A much broader census of comammox *Nitrospira* in the environment is

361 needed before general conclusions on their distribution and importance for nitrification can

362 be drawn. However, as was previously shown for AOA in a WWTP (Mussmann *et al.*, 2011),

363 high *amoA* gene copy numbers do not always indicate that the respective population is

364 functionally important for ammonia oxidation. Aside from the possibility that abundant

365 putative nitrifiers grow on alternative substrates, one can also not exclude *amoA* gene

366 quantification biases caused by extracellular DNA, which may for example be secreted to a

13

367 different extent by the phylogenetically diverse nitrifiers. Moreover, at least in AOA a basal

368 *amoA* transcription can occur even in the absence of detectable ammonia-oxidizing activity

369 (Mussmann *et al.*, 2011; Ke *et al.*, 2013). In contrast, pronounced shifts in the *amoA*

370 transcription levels upon certain environmental stimuli, such as a changing availability of

371 ammonium, likely are useful activity indicators for AOA, AOB, and comammox *Nitrospira*.

372 Although not tested yet in this context, the new comammox *amoA*-specific primers should in

373 principle be suitable for reverse transcription-qPCR of *amoA* and should thus allow the

374 quantification of comammox *amoA* transcripts in environmental samples and during

375 incubation experiments. Thus, the primers can serve as a tool for a broad range of study

376 designs to elucidate the ecological significance of comammox *Nitrospira* in comparison to

377 other ammonia oxidizers.

378

392

393

Figures and Tables

Table 1. Comammox *amoA*-targeted primers designed in this study, and *in silico* evaluation of the primer coverage of the target clades in our comammox *amoA* reference sequence database.

Table 2. Environmental samples analyzed in this study, PCR amplification cycles, and numbers of cloned comammox *amoA* sequences and OTUs retrieved from each sample.

Table 3. Copy numbers of comammox *amoA* genes in environmental samples as determined by qPCR analysis.

Figure 1. Maximum likelihood tree showing the phylogenetic affiliation of the comammox *amoA* genes obtained in this study (printed in boldface) to reference *amoA* sequences from comammox clades A and B. One representative sequence from each OTU is shown. Identifiers of clones (in parentheses) and of OTUs are indicated. Previously described comammox Nitrospira strains are shown in red. The outgroup consisted of selected betaproteobacterial *amoA* and proteobacterial *pmoA* genes. The scale bar indicates estimated change per nucleotide.

References

Angel R, Claus P, Conrad R. Methanogenic archaea are globally ubiquitous in aerated soils and become active under wet anoxic conditions. ISME J. 2012; 6: 847-862.

Bollmann A, Bullerjahn GS, McKay RM. Abundance and diversity of ammonia-oxidizing archaea and bacteria in sediments of trophic end members of the Laurentian Great Lakes, Erie and Superior. PLoS One. 2014; 9: e97068.

Chen X., Zhu YG, Xia Y, Shen JP, He JZ. Ammonia-oxidizing archaea: important players in paddy rhizosphere soil?. Environ Microbiol. 2008; 10: 1978-1987.

422   Daebeler A, Abell GC, Bodelier PL, Bodrossy L, Frampton DM, Hefting MM, et al. Archaeal
423   dominated ammonia-oxidizing communities in Icelandic grassland soils are moderately
424   affected by long-term N fertilization and geothermal heating. Front Microbiol. 2012; 3: 352.
425   Doi: 10.3389/fmicb.2012.00352

426   Daims H, Nielsen JL, Nielsen PH, Schleifer KH, Wagner M. In Situ Characterization of
427   *Nitrospira*-Like Nitrite-Oxidizing Bacteria Active in Wastewater Treatment Plants. Appl
428   Environ Microbiol. 2001; 67: 5273-5284.

429   Daims H, Lebedeva EV, Pjevac P, Han P, Herbold C, Albertsen M, et al. Complete
430   nitrification by *Nitrospira* bacteria. Nature. 2015; 528: 504-509.

431   Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput.
432   Nucleic Acids Res. 2004; 32: 1792-1797.

433   Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics.
434   2010; 26: 2460-2461.

435   Ehrich S, Behrens D, Lebedeva E, Ludwig W, Bock E. A new obligately
436   chemolithoautotrophic, nitrite-oxidizing bacterium, *Nitrospira moscoviensis* sp. nov. and its
437   phylogenetic relationship. Arch Microbiol. 1995; 164: 16-23.

438   Francis CA, O'Mullan GD, Ward BB. Diversity of ammonia monooxygenase (*amoA*) genes
439   across environmental gradients in Chesapeake Bay sediments. Geobiology. 2003; 1: 129-
440   140.

441   Francis CA, Roberts KJ, Beman JM, Santoro AE, Oakley BB. Ubiquity and diversity of
442   ammonia-oxidizing archaea in water columns and sediments of the ocean. Proc Nat Acad
443   Sci USA. 2005; 102: 14683-14688.

444   Gruber-Dorninger C, Pester M, Kitzinger K, Savio DF, Loy A, Rattei, T., et al. Functionally
445   relevant diversity of closely related Nitrospira in activated sludge. ISMEJ. 2015; 9: 643-655.

446 Gonzalez-Martinez A, Rodriguez-Sanchez A, van Loosdrecht MM, Gonzalez-Lopez J,
447 Vahala R. Detection of comammox bacteria in full-scale wastewater treatment bioreactors
448 using tag-454-pyrosequencing. Environ Sci Pollut R. 2016; 23: 25501-25511.

449 Holmes AJ, Costello A, Lidstrom ME, Murrell JC. Evidence that participate methane
450 monooxygenase and ammonia monooxygenase may be evolutionarily related. FEMS
451 Microbiol Lett. 1995; 132: 203-208.

452 Jia Z, Conrad R. Bacteria rather than Archaea dominate microbial ammonia oxidation in an
453 agricultural soil. Environ Microbiol. 2009; 11: 1658-1671.

454 Junier P, Kim OS, Molina V, Limburg P, Junier T, Imhoff JF, et al. Comparative in silico
455 analysis of PCR primers suited for diagnostics and cloning of ammonia monooxygenase
456 genes from ammonia-oxidizing bacteria. FEMS Microbiol Ecol. 2008; 64: 141-152.

457 Juretschko S, Timmermann G, Schmid M, Schleifer KH, Pommerening-Röser A, Koops HP,
458 et al. Combined molecular and conventional analyses of nitrifying bacterium diversity in
459 activated sludge: *Nitrosococcus mobilis* and *Nitrospira*-like bacteria as dominant populations.
460 Appl Environ Microbiol. 1998; 64: 3042-3051.

461 Katoh K, Misawa K, Kuma KI, Miyata T. MAFFT: a novel method for rapid multiple sequence
462 alignment based on fast Fourier transform. Nucleic Acids Res. 2002; 30: 3059-3066.

463 Ke X, Angel R, Lu Y, Conrad R. Niche differentiation of ammonia oxidizers and nitrite
464 oxidizers in rice paddy soil. Environ Microbiol. 2013; 15: 2275-2292.

465 van Kessel MA, Speth DR, Albertsen M, Nielsen PH, den Camp HJO, Kartal B, et al.
466 Complete nitrification by a single microorganism. Nature. 2015; 528: 555-559.

467 Koch H, Galushko A, Albertsen M, Schintlmeister A, Gruber-Dorninger C, Lücker S, et al.
468 Growth of nitrite-oxidizing bacteria by aerobic hydrogen oxidation. Science. 2014; 345: 1052-
469 1054.

470 Koch H, Lücker S, Albertsen M, Kitzinger K, Herbold C, Spieck E, et al. Expanded metabolic
471 versatility of ubiquitous nitrite-oxidizing bacteria from the genus *Nitrospira*. Proc Nat Aca Sci
472 USA. 2015; 112: 11371-11376.

473 Könneke M, Bernhard AE, José R, Walker CB, Waterbury JB, Stahl DA. Isolation of an
474 autotrophic ammonia-oxidizing marine archaeon. Nature. 2005; 437: 543-546.

475 Lebedeva EV, Alawi M, Maixner F, Jozsa PG, Daims H, Spieck E. Physiological and
476 phylogenetic characterization of a novel lithoautotrophic nitrite-oxidizing bacterium,
477 '*Candidatus* Nitrospira bockiana'. IJSEM. 2008; 58: 242-250.

478 Lebedeva EV, Off S, Zumbrägel S, Kruse M, Shagzhina A, Lücker S, et al. Isolation and
479 characterization of a moderately thermophilic nitrite-oxidizing bacterium from a geothermal
480 spring. FEMS Microbiol Ecol. 2011; 75: 195-204.

481 Leininger S, Urich T, Schloter M, Schwark L, Qi J, Nicol GW, et al. Archaea predominate
482 among ammonia-oxidizing prokaryotes in soils. Nature. 2006; 442: 806-809.

483 Linhart C, Shamir R. The degenerate primer design problem. Bioinformatics. 2002; 18:
484 S172-S181.

485 Ludwig W, Strunk O, Westram R, Richter L, Meier H, Buchner A, et al. ARB: a software
486 environment for sequence data. Nucleic Acids Res. 2004; 32: 1363-1371.

487 Luesken FA, van Alen TA, van der Biezen E, Frijters C, Toonen G, Kampman C, et al.
488 Diversity and enrichment of nitrite-dependent anaerobic methane oxidizing bacteria from
489 wastewater sludge. Appl Microbiol Biot. 2011; 92: 845-854.

490 Maixner F, Noguera DR, Anneser B, Stoecker K, Wegl G, Wagner M, et al. Nitrite
491 concentration influences the population structure of *Nitrospira*-like bacteria. Environ
492 Microbiol. 2006; 8: 1487-1495.

493 Martiny AC, Albrechtsen HJ, Arvin E, Molin S. Identification of bacteria in biofilm and bulk
494 water samples from a nonchlorinated model drinking water distribution system: detection of a

18

495   large nitrite-oxidizing population associated with *Nitrospira* spp. Appl Environ Microbiol.

496   2005; 71: 8611-8617.

497   Meinhardt KA, Bertagnolli A, Pannu MW, Strand SE, Brown SL, et al. Evaluation of revised

498   polymerase chain reaction primers for more inclusive quantification of ammonia-oxidizing

499   archaea and bacteria. Environ Microbiol Rep. 2015; 7: 354-363.

500   Mincer TJ, Church MJ, Taylor LT, Preston C, Karl DM, DeLong EF. Quantitative distribution

501   of presumptive archaeal and bacterial nitrifiers in Monterey Bay and the North Pacific

502   Subtropical Gyre. Environ Microbiol. 2007; 9: 1162-1175.

503   Mußmann M, Brito I, Pitcher A, Damsté JSS, Hatzenpichler R, Richter A, et al.

504   Thaumarchaeotes abundant in refinery nitrifying sludges express *amoA* but are not obligate

505   autotrophic ammonia oxidizers. Proc Nat Acad Sci USA. 2011; 108: 16771-16776.

506   Nold SC, Zhou J, Devol AH, Tiedje JM. Pacific Northwest Marine Sediments Contain

507   Ammonia-Oxidizing Bacteria in the β Subdivision of the *Proteobacteria*. Appl Environ

508   Microbiol. 20001; 66: 4532-4535.

509   Norton JM, Alzerreca JJ, Suwa Y, Klotz MG. Diversity of ammonia monooxygenase operon

510   in autotrophic ammonia-oxidizing bacteria. Arch Microbiol. 2002; 177: 139-149.

511   Okano Y, Hristova KR, Leutenegger CM, Jackson LE, Denison RF, Gebreyesus B, et al.

512   Application of real-time PCR to study effects of ammonium on population size of ammonia-

513   oxidizing bacteria in soil. Appl Environ Microbiol. 2004; 70: 1008-1016.

514   Oswald K, Graf JS, Littmann S, Tieken D, Brand A, Wehrli B, et al. *Crenothrix* are major

515   methane consumers in stratified lakes. Under revision in ISMEJ.

516   Palomo A, Fowler SJ, Gülay A, Rasmussen S, Sicheritz-Ponten T, Smets BF. Metagenomic

517   analysis of rapid gravity sand filter microbial communities suggests novel physiology of

518   *Nitrospira* spp. ISMEJ. 2016; 10:2569-258.

19

519 Pester M, Rattei T, Flechl S, Gröngröft A, Richter A, Overmann J, et al. *amoA*-based

520 consensus phylogeny of ammonia-oxidizing archaea and deep sequencing of *amoA* genes

521 from soils of four different geographic regions. Environ Microbiol. 2012; 14: 525-539.

522 Pester M, Maixner F, Berry D, Rattei T, Koch H, Lücker S, et al. NxrB encoding the beta

523 subunit of nitrite oxidoreductase as functional and phylogenetic marker for nitrite-oxidizing

524 *Nitrospira*. Environ Microbiol. 2014; 16: 3055-3071.

525 Pinto AJ, Marcus DN, Ijaz UZ, Bautista-de lose Santos QM, Dick GJ, Raskin L.

526 Metagenomic evidence for the presence of comammox *Nitrospira*-like bacteria in a drinking

527 water system. mSphere. 2016; 1: e00054-15.

528 Polz MF, Cavanaugh CM. Bias in template-to-product ratios in multitemplate PCR. Appl

529 Environ Microbiol. 1998; 64: 3724-3730.

530 Price MN, Dehal PS, Arkin AP. FastTree: computing large minimum evolution trees with

531 profiles instead of a distance matrix. Mol Biol Evol. 2009; 26: 1641-1650.

532 Purkhold U, Pommerening-Röser A, Juretschko S, Schmid MC, Koops HP, Wagner, M.

533 Phylogeny of all recognized species of ammonia oxidizers based on comparative 16S rRNA

534 and *amoA* sequence analysis: implications for molecular diversity surveys. Appl Environ

535 Microbiol. 2000; 66: 5368-5382.

536 Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species

537 definition. Proc Nat Acad Sci USA. 2009; 106: 19126-19131.

538 Rotthauwe JH, Witzel KP, Liesack W. The ammonia monooxygenase structural gene *amoA*

539 as a functional marker: molecular fine-scale analysis of natural ammonia-oxidizing

540 populations. Appl Environ Microbiol. 1997; 63: 4704-4712.

541 Sinigalliano CD, Kuhn DN, Jones RD. Amplification of the *amoA* gene from diverse species

542 of ammonium-oxidizing bacteria and from an indigenous bacterial population from seawater.

543 Appl Environ Microbiol. 1995; 61: 2702-2706.

544   Spieck E, Hartwig C, McCormack I, Maixner F, Wagner M, Lipski A, et al. Selective

545   enrichment and molecular characterization of a previously uncultured *Nitrospira*-like

546   bacterium from activated sludge. Environ Microbiol. 2006; 8: 405-415.

547   Stamatakis A, Ludwig T, Meier H. RAxML-III: a fast program for maximum likelihood-based

548   inference of large phylogenetic trees. Bioinformatics. 2005; 21: 456-463.

549   Stephen JR, Chang YJ, Macnaughton SJ, Kowalchuk GA, Leung KT, Flemming CA, White

550   DC. Effect of toxic metals on indigenous soil β-subgroup proteobacterium ammonia oxidizer

551   community structure and protection against toxicity by inoculated metal-resistant bacteria.

552   Appl Environ Microbiol. 1999; 65: 95-101.

553   Stoecker K, Bendinger B, Schöning B, Nielsen PH, Nielsen JL, Baranyi C, et al. Cohn's

554   *Crenothrix* is a filamentous methane oxidizer with an unusual methane monooxygenase.

555   Proc Nat Acad Sci USA. 2006; 103: 2363-2367.

556   Treusch AH, Leininger S, Kletzin A, Schuster SC, Klenk HP, Schleper C. Novel genes for

557   nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic

558   crenarchaeota in nitrogen cycling. Environ Microbiol. 2005; 7: 1985-1995.

559   Wang JG, Xia F, Zeleke J, Zou B, Rhee SK, Quan ZX. An improved protocol with a highly

560   degenerate primer targeting copper-containing membrane-bound monooxygenase (CuMMO)

561   genes for community analysis of methane-and ammonia-oxidizing bacteria. FEMS Microbiol

562   Ecol. 2016; doi: 10.1093/femsec/fiw244

563   Watson SW, Bock E, Valois FW, Waterbury JB, Schlosser U. *Nitrospira marina* gen. nov. sp.

564   nov.: a chemolithotrophic nitrite-oxidizing bacterium. Arch Microbiol. 1986; 144: 1-7.

565   Webster G, Embley TM, Prosser JI. Grassland management regimens reduce small-scale

566   heterogeneity and species diversity of β-proteobacterial ammonia oxidizer populations. Appl

567   Environ Microbiol. 2002; 68: 20-30.

568 **Table 1.** Comammox *amoA*-targeted primers designed in this study, and *in silico* evaluation

569 of the primer coverage of the target clades in our comammox *amoA* reference sequence

570 database.

Target gene	Primer name	Primer sequence (5' – 3')	Amplicon size	Target group coverage[1]
		Degenerate primer pairs		
Comammox *Nitrospira* clade A *amoA*	comaA-244F	TAYAAYTGGGTSAAYTA	415	93/98 (95%)
	comaA-659R	ARATCATSGTGCTRTG		
Comammox *Nitrospira* clade B *amoA*	comaB-244F	TAYTTCTGGACRTTYTA	415	188/228 (83%)
	comaB-659R	ARATCCARACDGTGTG		
		Single oligonucleotides for equimolar primer mixtures		
Comammox *Nitrospira* clade A *amoA*	comaA-244f_a	TACAACTGGGTGAACTA	415	93/98 (95%)
	comaA-244f_b	TATAACTGGGTGAACTA		
	comaA-244f_c	TACAATTGGGTGAACTA		
	comaA-244f_d	TACAACTGGGTCAACTA		
	comaA-244f_e	TACAACTGGGTCAATTA		
	comaA-244f_f	TATAACTGGGTCAATTA		
	comaA-659r_a	AGATCATGGTGCTATG		
	comaA-659r_b	AAATCATGGTGCTATG		
	comaA-659r_c	AGATCATGGTGCTGTG		
	comaA-659r_d	AAATCATGGTGCTGTG		
	comaA-659r_e	AGATCATCGTGCTGTG		
	comaA-659r_f	AAATCATCGTGCTGTG		
Comammox *Nitrospira* clade B *amoA*	comaB-244f_a	TAYTTCTGGACGTTCTA	415	209/228 (92%)
	comaB-244f_b	TAYTTCTGGACATTCTA		
	comaB-244f_c	TACTTCTGGACTTTCTA		
	comaB-244f_d	TAYTTCTGGACGTTTTA		
	comaB-244f_e	TAYTTCTGGACATTTTA		
	comaB-244f_f	TACTTCTGGACCTTCTA		
	comaB-659r_a	ARATCCAGACGGTGTG		
	comaB-659r_b	ARATCCAAACGGTGTG		
	comaB-659r_c	ARATCCAGACAGTGTG		
	comaB-659r_d	ARATCCAAACAGTGTG		
	comaB-659r_e	AGATCCAGACTGTGTG		
	comaB-659r_f	AGATCCAAACAGTGTG		

571 [1] Indicated are the number of comammox *amoA* sequences in the respective target clade that fully

572 match the primers, the total number of sequences in this target clade, and the primer coverage in per

573 cent.

574

22

575 **Table 2.** Environmental samples analyzed in this study, PCR amplification cycles, and

576 numbers of cloned comammox *amoA* sequences and OTUs retrieved from each sample.

Sample name	ID	Habitat type	Gene amplified	Amplification cycles	N_{clones}	N_{OTUs}
Wolfenbüttel	WB	Engineered (GWW)	Clade A *amoA*	25	24	5
			Clade B *amoA*	25	28[3]	3
Friedrichshof	FH	Engineered (DWTP)	Clade A *amoA*	25	23	4
			Clade B *amoA*	25	22[4]	3
Spiekeroog	SP	Engineered (DWTP)	Clade A *amoA*	25	22	1
			Clade B *amoA*	25	23	2
Breehei[1]	BH	Engineered (DWTP)	Clade A *amoA*	30	10	3
			Clade B *amoA*	30	10	3
Lieshout	LS	Engineered (WWTP)	Clade A *amoA*	30	8	2
VetMed	VM	Engineered (WWTP)	Clade A *amoA*	30	20[3]	4
Ingolstadt	IS	Engineered (WWTP)	Clade A *amoA*	30	5	1
Recirculating aquaculture system[1]	RAS	Engineered	Clade A amoA	30	7	4
			Clade B amoA	30	8	3
Vercelli rice soil[2]	VS	Rice paddy soil	Clade A *amoA*	35	26	7
			Clade B *amoA*	35	41	5
Vercelli rice rhizosphere	VR	Rhizosphere	Clade A *amoA*	35	31	8
			Clade B *amoA*	35	29	6
Klausen-Leopoldsdorf	KLD	Forest soil	Clade B *amoA*	35	60	8
Herrensee[1]	HS	Brackish lake sediment	Clade A *amoA*	30	10	3
River Schwarza enrichment (inoculated with river biofilm and fed with ammonium)	RS	Freshwater biofilm	Clade A *amoA*	35	37	2

577 [1] Unspecific PCR products were obtained with the degenerate comammox *amoA* clade A targeted
578 primers.
579 [2] Unspecific PCR products were obtained with the degenerate comammox *amoA* clade B targeted
580 primers.
581 [3] One of the obtained sequences was classified as *pmoA* gene by phylogenetic analysis.
582 [4] One of the obtained sequences was classified as comammox *amoA* clade A gene by phylogenetic
583 analysis.

23

584   **Table 3.** Copy numbers of comammox *amoA* genes in environmental samples as

585   determined by qPCR analysis.

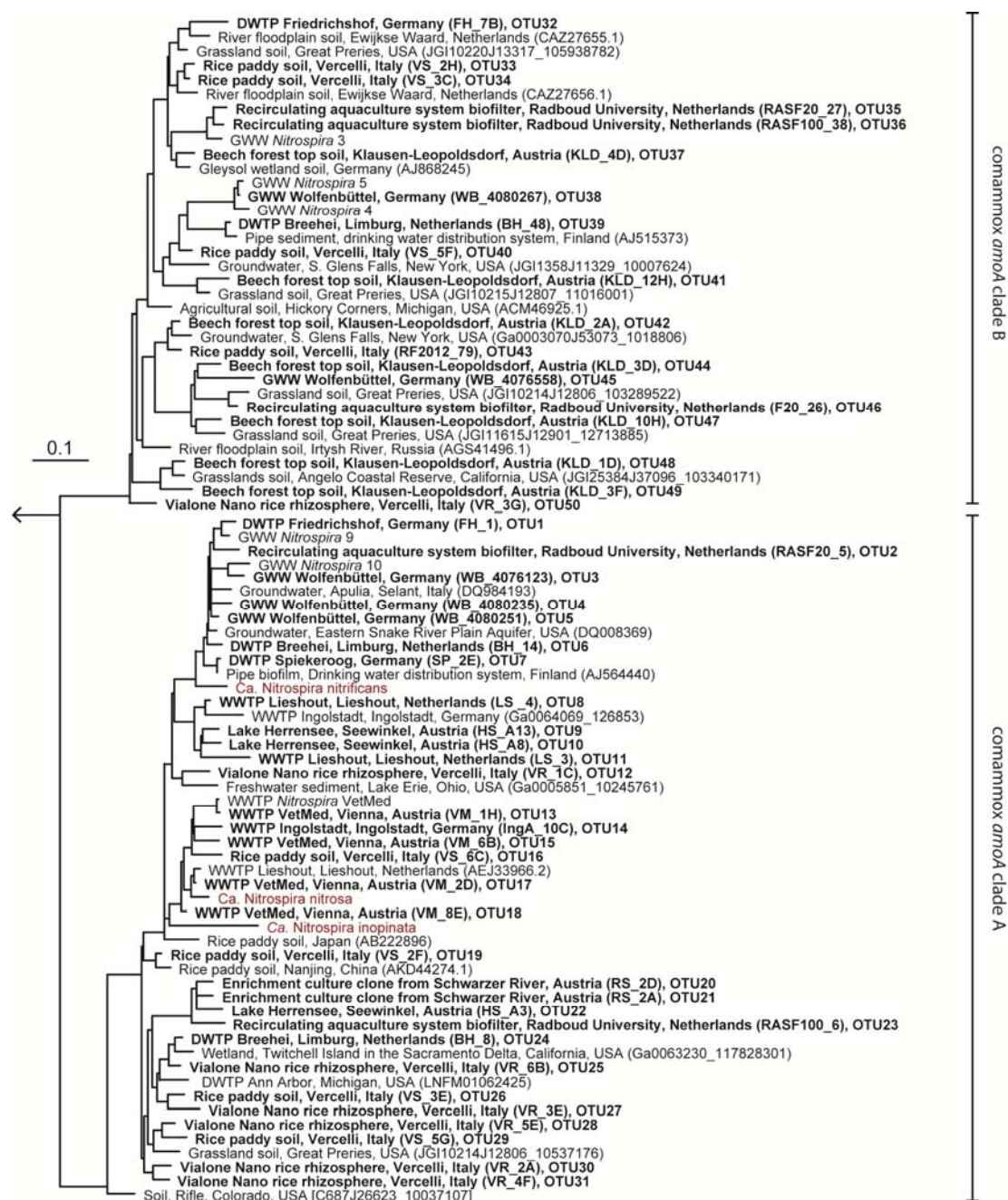| Sample name	Habitat type	Gene amplified	Copies / ng DNA
Wolfenbüttel	Engineered (GWW)	Clade A *amoA*	$1.1\pm0.2 \times10^4$
		Clade B *amoA*	$2.5\pm0.2 \times10^4$
Friedrichshof	Engineered (DWTP)	Clade A *amoA*	$1.7\pm0.3 \times10^3$
		Clade B *amoA*	$3.7\pm0.6 \times10^3$
VetMed	Engineered (WWTP)	Clade A *amoA*	$5.1\pm1.7 \times10^2$
Vercelli rice soil	Rice paddy soil	Clade A *amoA*	$4.6\pm1.0 \times10^2$
		Clade B *amoA*	$4.0\pm0.5 \times10^2$
Klausen-Leopoldsdorf	Forest soil	Clade B *amoA*	$3.9\pm1.0 \times10^2$

586

24

Figure 1. Maximum likelihood tree showing the phylogenetic affiliation of the comammox *amoA* genes obtained in this study (printed in boldface) to reference *amoA* sequences from comammox clades A and B. One representative sequence from each OTU is shown. Identifiers of clones (in parentheses) and of OTUs are indicated. Previously described comammox *Nitrospira* strains are shown in red. The outgroup consisted of selected betaproteobacterial *amoA* and proteobacterial *pmoA* genes. The scale bar indicates estimated change per nucleotide.