A new primer set for Clade I *nosZ* that recovers genes from a broader range of taxa

Bangzhou Zhang[1,2], C. Ryan Penton[3,4], Zhenhua Yu[1,5], Chao Xue[1,6], Qiongyun Chen[7], Zhangran Chen[1,7], Changsheng Yan[7], Qiang Zhang[7], Mengxin Zhao[1,8], John F. Quensen[1], James M. Tiedje[1]*

1 Center for Microbial Ecology, Michigan State University, East Lansing, MI, USA
2 Department of Gastroenterology, Zhongshan Hospital, Xiamen University, Xiamen, China
3 Center for Fundamental and Applied Microbiomics, Biodesign Institute, Arizona State University, Tempe, AZ, USA
4 College of Integrative Sciences and Arts, Faculty of Science and Mathematics, Arizona State University, Mesa, AZ, USA
5 Key Laboratory of Mollisols Agroecology, Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Harbin, China
6 Jiangsu Provincial Key Lab for Solid Organic Waste Utilization, National Engineering Research Center for Organic-based Fertilizers, Jiangsu Collaborative Innovation Center for Solid Organic Waste Resource Utilization, Nanjing Agricultural University, Nanjing, China
7 Institute for Microbial Ecology, School of Medicine, Xiamen University, Xiamen, China
8 State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China

*Corresponding author: James M. Tiedje, tiedjej@msu.edu

Acknowledgments
This work was funded in part by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research (Awards DE-SC0018409 and DE-FC02-07ER64494), by the National Science Foundation Long-term Ecological Research Program (DEB 1637653) at the Kellogg Biological Station, by Michigan State University AgBioResearch, and by National Natural Science Foundation of China (81800517)

Abstract: Denitrification is an important global N cycle process. The gene encoding NosZ that converts nitrous oxide (N_2O) to N_2 has been widely used as a biomarker to study denitrifying communities. However, conventional PCR primers target a limited range of the genetically diverse Clade I *nosZ*, and the amplicons are too long for sequencing on current NGS platforms. To address these issues, we developed a new PCR primer set that amplifies a 355-bp region of Clade I *nosZ* and captures broader taxonomic coverage than conventional primers in *in-silico* tests. When compared with the widely-used nosZF_nosZR_Rich_2003 set using the same soil samples and same sequencing depth, the new set retrieved genes from four times more unique species, with consistently higher general diversity-based metrics. The new primer set performed well with different sequencing platforms (Ion Torrent and Illumina), and among a wide variety of soils from polar to tropical, desert to agricultural, and surface to a very low biomass subsoil, with significant differences in denitrifying community diversity and composition. This new primer set for Clade I together with the primers recently reported for Clade II (Chee-Sanford et al. 2020) provide a more comprehensive assessment of denitrifier gene hosts, their ecological patterns and the degree of novelty in retrieved gene sequences.

Keywords: nosZ, Clade I, primer design, NGS, higher coverage, soil type

Introduction

Nitrous oxide (N_2O) is an intermediate product of biological denitrification. Although there are multiple sources of planetary N_2O, terrestrial ecosystems account for the majority of N_2O emissions, which includes the loss of fixed nitrogen (e.g. fertilizers) from agricultural fields (Hallin et al. 2018; IPCC 2019). Moreover, atmospheric N_2O is an important contributor to global warming with approximately 300 times more warming potential per mole than CO_2 (IPCC 2019), and accelerates the depletion of the stratospheric ozone layer (Ravishankara et al. 2009). N_2O can be converted into inert nitrogen gas (N_2) by the microbial N_2O reductase (NosZ), located either in the periplasm or membrane-associated (Zumft and Kroneck 2007). Microbial reduction of N_2O to N_2 is the only known sink of N_2O in the biosphere, and can be affected by various factors, including pH and Cu availability (Samad et al. 2016; Shen et al. 2020; Thomson et al. 2012). Hence, particular attention has been paid to N_2O-reducting microbial populations in order to understand their role in N_2O dynamics. This is often accomplished by using the microbial *nosZ* gene as a biomarker for the process.

Well-known active denitrifiers are largely Alpha-, Beta-, and Gammaproteobacteria, as well as some halophilic Euryarchaeota and carry Clade I (typical) *nosZ* (Hallin et al. 2018). Recent pioneering work based on genome analyses has identified Clade II (atypical) *nosZ*, another group of abundant N_2O-reducing microbes with a distant phylogeny to Clade I. These have higher gene sequence and organism diversity as assessed by metagenomics and qPCR (Jones et al. 2013; Orellana et al. 2014; Samad et al. 2016; Sanford et al. 2012; Tsiknia et al. 2015), and were

proposed to play more important roles in reducing N_2O to N_2 in terrestrial habitats. Current research has focused on studying their niche differentiation. For example, in a pot experiment with agricultural soils, Clade II relative abundances were observed to be more dominant in bulk soil, while Clade I abundances were more abundant in the rhizosphere (Graf et al. 2016). Clade I may also be more important in N_2O reduction in aquatic systems (coastal sediments, salt marshes, constructed wetlands, bioreactors) (Hallin et al. 2018). Studies on kinetics revealed that Clade I-carrying strains have significantly higher half-saturation constants than Clade II-carrying strains (Yoon et al. 2016), which implies that Clade II performs better at low N_2O concentrations while Clade I likely processes more N_2O at higher concentrations, for example in fertilized agricultural soils.

Studies of functional gene-harboring microbial communities and their variations in different conditions or environments are often based on amplicon metagenomics (Harter et al. 2016; Penton et al. 2015; Wang et al. 2019; Zhao et al. 2018). This relies on a primer set(s) that performs reliably in challenging sample types, that is sensitive yet specific, and comprehensively covers a diversity of the targeted gene (Schöler et al. 2017). However, primers to effectively detect microbial guilds possessing *nosZ* have been a challenge due to the high genetic diversity of these sequences (Hallin et al. 2018; Ma et al. 2019). In fact, the more dominant and diverse Clade II sequences remained unknown until they were revealed in genomic and metagenomic sequences (Jones et al. 2013; Orellana et al. 2014; Sanford et al. 2012). Recently, a suite of new primer sets were designed to specifically target 7 of 10 subclades of Clade II for amplicon sequencing (Chee-Sanford et al. 2020). As for Clade I, many existing primer sets do not produce amplicons with lengths (Table S1) suitable for popular high-throughput sequencing platforms (Ma et al. 2019). As reference databases have grown, and technologies have advanced, it is important to re-evaluate the performances of potential primers under a broad range of soil types and conditions and with different sequencing platforms.

This study revisits existing primer sets for Clade I using larger reference data sets in order to determine if improvements in primer design and coverage are possible. We did this by collecting an extensive set of high quality Clade I *nosZ* references and developing a new primer set that, *in-silico*, effectively targets a broader range of the Clade I populations while generating amplicons appropriate for high-throughput targeted amplicon sequencing. Secondly, we evaluated the fidelity and experimental performance of this new set in terms of coverage and sequence novelty by comparing them with a widely-used primer pair on identical soil DNA samples. Finally, this new primer set was used to amplify Clade I *nosZ* from diverse soil types and tested with both Ion Torrent PGM and Illumina MiSeq platforms to validate the general usability for providing insight into a wider denitrifier gene diversity.

Materials and methods

Clade I *nosZ* subclades and primer design

A total of 344 nucleotide sequences of Clade I *nosZ* used for references were obtained from the RDP Functional Gene Repository (FunGene) (Fish et al. 2013) with filters of HMM coverage of 80%, length of 340 amino acids, and a score of 900 (all minimum values). All of these sequences were aligned using Clustal X (Larkin et al. 2007). Potential conserved regions with length of 18–22 base pairs and no more than three of the same consecutive nucleotides were picked in MEGA 6 (Tamura et al. 2013). Degenerate primers specific to conserved regions were designed with the general criteria of a degeneracy less than 60, a primer melting temperature in the range of 60-63°C where possible, and an amplicon length suitable for use on next generation high-throughput sequencing platforms.

In-silico comparison of primer sets

To explore the improvement of the new primer set, we compared its coverage *in-silico* by FunGene Probe Match (Fish et al. 2013) with 11 previously reported primer sets (primer designation, bp covered): nosZ1F_nosZ1R_Henry_2006 (247 bp) (Henry et al. 2006); nosZ2F_nosZ2R_Henry_2006 (267 bp) (Henry et al. 2006); nosZF_nosZR_Rich_2003 (701 bp) (Rich et al. 2003); nosZF_nosZ1622R_Throback_2004 (454 bp) (Throback et al. 2004); nosZF_nos1773R_Throback_2004 (247 bp) (Throback et al. 2004); nosZF-nosZR_Rosch_2002 (701 bp) (Rosch et al. 2002); Nos1527F_Nos1773R_Scala_1998 (247 bp) (Scala and Kerkhof 1998); nosLb_nosRb_Chenby_1998 (302 bp) (Cheneby et al. 1998); nosZ1126F_nosZ1184R_Chen_2012 (759 bp) (Chen et al. 2012); and nosZ1F_nosZ2F_Philippot-2013 (458 bp) (Philippot et al. 2013).These lengths are relative to the 1938 bp *nosZ* of *R. palustris* CGA009. By allowing a maximum of two mismatches for each primer sequence, all primer sets were tested against four datasets, including the one originally used for primer design, the Clade I references (324 sequences) used in Xander (Wang et al. 2015), the Clade I *nosZ* (241 sequences) assembled from 21 rhizosphere soil samples by Jiarong Guo (personal communication), and Clade I sequences (1645 sequences) downloaded from FunGene with a minimum length of 688 amino acids on Oct. 6th 2018. Sequences targeted by both forward and reverse primers were counted as effective targets and recorded (Table S1). The targets by each primer set were visualized in a phylogenetic tree based on the Clade I references utilized in Xander by the Interactive Tree of Life (iTOL) (Letunic and Bork 2016).

Comparison in the experimental performance

To test the experimental performance of the new primers, total DNA of four soil samples from an experimental soil warming treatment established in Oklahoma were amplified and sequenced with the new primer set by the Ion Torrent PGM 400bp kit at the Research Technology Support Facility at Michigan State University. These soil samples were coarse-silty and well drained with neutral pH (Table 1) and previously studied with nosZF_nosZR_Rich_2003 primers (Penton et al. 2015),. The fidelity of the primers was tested on six common denitrifiers: *R. palustris* CGA009, *Marinobacter aquaeolei* VT8 (ATCC 700491), *Alicycliphilus denitrificans* K601 (DSMZ 14773), *Ochrobactrum anthropi* (ATCC 49188), *Brucella ovis* (ATCC 25840), and *Pseudomonas stutzeri*

(ATCC 17588) and seven non-denitrifiers: *Sphingomonas echinoides*, *Enterococcus gallinarum*, *Escherichia fergusoni*, *Klebsiiella variicola*, *Streptococcus mutans*, *Cytobacillus kochii*, and *Shigella sonnei*. The library preparation was similar to the prior procedure with some minor modification (Zhang et al. 2015). Briefly, a 20 μL PCR mixture contained 1× green buffer (PROMEGA), 2.38 mM $MgCl_2$ (Promega), 0.25 mM each deoxynucleoside triphosphate (Promega), 500 nM each primer (IDT), 0.1 mg/ml bovine serum albumin (NEB), 0.5 U of Taq polymerase (REF M8295, Promega), and 1 ng/μL template DNA (final concentration). PCR cycling conditions were 2 min at 95°C, 35 cycles of 45 s at 95°C, 45s at 53°C, and 1 min at 72°C, and an extra 7 min at 72°C for a final extension.

Application in different soil types and sequencing with different platforms

To verify primer performance in soils with diverse characteristics, the new primer set was used to amplify *nosZ* genes in triplicate surface (0-25 cm) soil samples collected from eight USA states (e.g. Alaska to Florida, Hawaii to Michigan), covering agricultural corn fields, various forests (tropical to boreal taiga), a tall grass prairie and a desert shrubland (soil properties in Table S2). Soils were mixed , and the DNA was extracted with the Power Soil DNA isolation kit (MO BIO) according to the manufacturer's protocol, then amplified and sequenced by PGM using above PCR and sequencing conditions, respectively. To test primer performance under low biomass conditions, soil samples were collected from six replicates at different soil depths (0-10 cm, 10-25 cm, 25-50 cm) under two biofuel crops (continuous corn and switchgrass) established for ten years at MSU's Kellogg Biological Station (soil properties in Table S3), following a sampling protocol described previously (Zhang et al. 2017). The *nosZ* genes were amplified with the same thermal cycling conditions mentioned above and sequenced using Illumina MiSeq platform with PE 250 reagents.

Sequence analysis

Raw sequence reads were trimmed by the Initial Process tool in the FunGene pipeline (Fish et al. 2013) before downstream analyses. The filters for data from the new primer set were a forward primer maximum edit distance of 2, a reverse primer maximum edit distance of 1, a maximum number of N's of 0, a minimum sequence length of 300 (excluding primers), and a minimum read Q score of 22 (Zhang et al. 2015). The filters for raw reads from the old primer set were similar, except a minimum sequence length of 260 (excluding primers), a minimum read Q score of 20, and without requiring the reverse primer. The filtered reads were subjected to chimera deletion in *de novo* mode by USEARCH (Edgar et al. 2011), and then translated and frameshift corrected against the Clade I-Xander references with FrameBot (Wang et al. 2013) using default settings. The frameshift-corrected amino acid sequences were randomly resampled to 1,759 sequences per sample, aligned with the FunGene HMMER3 Aligner, and finally clustered to OTUs at 3% amino acid dissimilarity by RDP mcClust with the complete linkage algorithm. OTU representative sequences were further evaluated by BLASTp against Clade I-Xander references to assign taxonomy. Alpha diversity indexes, including observed richness (Obs), Shannon diversity (H), and

Pielou's evenness (J), were computed based on the OTU table using the R package vegan (Oksanen et al. 2015), and plotted using the package ggplot2 (Wickham et al. 2017). The NMDS ordination and the significance test among treatments by PERMANOVA with 999 permutations were performance based on Bray-Curtis distance using vegan package (Oksanen et al. 2015). SIMPROF was analyzed using clustsig package (Whitaker and Christman 2014).

Results and discussion

Primer design and evaluation *in silico*

Based on the conserved regions of 344 high quality *nosZ* reference sequences from FunGene, we designed one primer pair with a degeneracy of 54 for the forward primer (5'-GGCAARCTVTCDCCVAC-3') and a degeneracy of 36 for the reverse (5'-AVCGGTCYTTVGAGAAYTT-3'. This new primer set covers a 355-bp fragment (nosZ1039F - nosZ1393R) of *Rhodopseudomonas palustris* CGA009 (Fig. S1), which is within the range of current sequencing abilities of both Illumina and Ion Torrent platforms. By allowing a maximum of two mismatches of each primer sequence to the Clade I *nosZ* references in Xander (Wang et al. 2015), this new primer set had the highest coverage, compared with the previously reported 11 primer pairs (Fig. 1, Table S1). Many *nosZ*-harboring taxa, including *Photobacterium* and *Vibrio* in Gammaproteobacteria, *Neisseria* and *Kingella* in Betaproteobacteria, and *Haloferax* and *Haloarcula* in Archaea were only targeted by the new primer set. This coverage improvement was also observed in additional reference sets, including the one originally used for primer design , the Clade I *nosZ* assembled from 21 rhizosphere soil samples, and the Clade I *nosZ* recently downloaded from Functional Gene Repository (Table S1). Together, these *in-silico* tests indicated that the newly designed primer set improved Clade I *nosZ* coverage and had a suitable degeneracy and amplicon length for next generation sequencing.

Experimental evaluation showing wider coverage and novel targets

The fidelity of the new primer set was verified by PCR amplification and sequencing against DNA of six confirmed denitrifiers and no amplification from seven non-denitrifiers. Total DNA of four soil samples from an experimental soil warming treatment previously studied (Penton et al. 2015) with nosZF_nosZR_Rich_2003 primers were re-analyzed with the new primer set for performance comparison. There were no significant differences in α-diversity indexes between these two primer sets at the OTU level. However, significantly higher richness (observed species and Chao1, $P<0.01$) were recovered by the new set at the species-level, based on BLASTp hits (Table 2), along with a relatively smaller variation among replicates. When evaluated by nearest taxa, approximately four times more unique species and genera, as well as 11 more microbial families, were recovered by the new primer set (Table 3). Since these two primer sets target different regions of *nosZ* and produce different amplicon lengths, it is more meaningful to focus on taxonomic levels from BLASTp for comparison.

The composition of the Clade I *nosZ*-harboring communities were significantly different (PERMANOVA, P<0.05) between the two primer datasets (Table S4). As for detailed taxonomic composition, the majority of targets matched to Proteobacteria. In the top 30 hits that accounted for 95% of the total reads, more Alphaproteobacteria (*Bradyrhizobium*, *Rhodopseudomonas*), Gammaproteobacteria (Candidatus *Competibacter*) and uncultured bacteria, but less Betaproteobacteria (*Castellaniella*) were retrieved by the new set (Fig. S2). More importantly, targets of the new primer set exhibited much lower identities to the reference sequences than the old primer set within the Alphaproteobacteria and Gammaproteobacteria, while a similar identity distribution was observed to the old set within the Betaproteobacteria (Fig. 2). These lower-identity distribution patterns indicated more novel Clade I *nosZ*-harboring bacteria were recovered by the new primer set. It is noteworthy that SIMPROF analysis showed consistent clustering for samples amplified by each primer set (Fig. S3) and similar rank orders for α-diversity indexes except richness (Table 1), which together suggests that use of the two primer sets should result in the same general diversity-derived biological conclusions.

Clade I *nosZ*-harboring communities in diverse soil types and with soil depth

To verify whether the primer set works well across diverse soil samples, we used it to amply *nosZ* genes in samples with various soil textures, organic carbon content, pH, microbial biomass, and land uses including agricultural fields, tropical and boreal forest, tall grass prairie, and desert shrubland (Table S2), as well as soils from different soil depths (Table S3). *NosZ* was readily amplified from all soils with the new primer set and successfully sequenced by PGM or Illumina platforms. We observed significant differences in α-diversity indexes (Fig. 3A) and in overall community composition (Fig. 3B, PERMANOVA, F=3.673, P=0.001) among the five land use types. Furthermore, there were remarkable differences in the composition of the closest matching dominant genera (mean abundance >1%) among different samples (Fig. S4).). Less variation within corn soils was also observed in α-diversity indexes (Fig. 3A), overall community composition (insert in Fig. 3B, PERMANOVA, F=2.586, P=0.002), and among the dominant genera (Fig. S4). This is likely due to the more similar soil and climate conditions for Midwest corn production areas (IA, KS, MI, and WI) as well as homogenization of community selection due to shared agricultural practices such as cultivation and fertilization (Ji et al. 2020).

As expected, we found significantly lower α-diversity indexes (richness, Shannon, and evenness) in deeper soils under biofuel crops of corn and switchgrass, especially in the lower biomass 25-50 cm layer (Figs. 3C and S5), where lower microbial diversity and biomass were reported (Castellano-Hinojosa et al. 2018; Zhang et al. 2017). Significant differences in overall community composition (PERMANOVA, P=0.001) were also observed among soil depths and between crops (Fig. 3D), as well as a significant interaction between depth and crop type (Fig 3E, PERMANOVA, P<0.01), indicating that the new primer set worked equally well for sequencing on the Illumina platform.

Conclusions

The new primer set we developed exhibited higher *in-silico* reference database coverage than existing primer sets, retrieved more *nosZ*-sequences similar to those in diverse species and captured additional novel sequences in tests using soil samples already analyzed by another commonly used primer set. We also verified that the new primer set performed well across a wide range of soil characteristics and could distinguish the denitrifying community composition and diversity among closely as well as distantly related soils. Since the amplicon length of this primer set is 355bp, this new primer set functions well with the sequencing ability of Ion Torrent and Illumina platforms. The combination of this novel primer set along with the suite of new primers reported for the more diverse Clade II (Chee-Sanford et al. 2020) result in a high coverage of known sequence diversity among the *nosZ*-harboring microbial community, thus allowing for greater insight into how the catalysts for a globally important process, N_2O reduction, respond to perturbation across a variety of environments.

Funding

This work was funded in part by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research (Awards DE-SC0018409 and DE-FC02-07ER64494), by the National Science Foundation Long-term Ecological Research Program (DEB 1637653) at the Kellogg Biological Station, by Michigan State University AgBioResearch, and by National Natural Science Foundation of China (81800517).

Conflicts of interest/Competing interests

Not applicable

Availability of data and material

Raw sequencing data generated with Ion Torrent PGM and Illumina MiSeq were deposited in the NCBI Sequence Read Archive (SRA) database under the study accession PRJNA640470 and PRJNA640229, respectively.

References

Castellano-Hinojosa A, González-López J, Bedmar EJ (2018) Distinct effect of nitrogen fertilisation and soil depth on nitrous oxide emissions and nitrifiers and denitrifiers abundance. Biol Fertil Soils 54:829-840

Chee-Sanford JC, Connor L, Krichels A, Yang WH, Sanford RA (2020) Hierarchical detection of diverse Clade II (atypical) nosZ genes using new primer sets for classical- and multiplex PCR array applications. J Microbiol Meth 172:105908

Chen Z, Liu J, Wu M, Xie X, Wu J, Wei W (2012) Differentiated response of denitrifying communities to fertilization regime in paddy soil. Microb Ecol 63:446-459

Cheneby D, Hartmann A, Hénault C, Topp E, Germon J (1998) Diversity of denitrifying microflora and ability to reduce N2O in two soils. Biol Fertil Soils 28:19-26

Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27:2194-2200

Fish JA, Chai B, Wang Q, Sun Y, Brown CT, Tiedje JM, Cole JR (2013) FunGene: the functional gene pipeline and repository. Front Microbiol 4:291

Graf DRH, Zhao M, Jones CM, Hallin S (2016) Soil type overrides plant effect on genetic and enzymatic N_2O production potential in arable soils. Soil Biol Biochem 100:125-128

Hallin S, Philippot L, Löffler FE, Sanford RA, Jones CM (2018) Genomics and Ecology of Novel N2O-Reducing Microorganisms. Trends Microbiol 26:43-55

Harter J, Weigold P, El-Hadidi M, Huson DH, Kappler A, Behrens S (2016) Soil biochar amendment shapes the composition of N2O-reducing microbial communities. Sci Total Environ 562:379-390

Henry S, Bru D, Stres B, Hallet S, Philippot L (2006) Quantitative detection of the nosZ gene, encoding nitrous oxide reductase, and comparison of the abundances of 16S rRNA, narG, nirK, and nosZ genes in soils. Appl Environ Microbiol 72:5181-5189

IPCC (2019) Climate Change and Land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems. www.ipcc.ch/srccl/

Ji Y, Conrad R, Xu H (2020) Responses of archaeal, bacterial, and functional microbial communities to growth season and nitrogen fertilization in rice fields. Biol Fertil Soils 56:81-95

Jones CM, Graf DR, Bru D, Philippot L, Hallin S (2013) The unaccounted yet abundant nitrous oxide-reducing microbial community: a potential nitrous oxide sink. ISME J 7:417-426

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23:2947-2948

Letunic I, Bork P (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res 44:W242-245

Ma Y, Zilles JL, Kent AD (2019) An evaluation of primers for detecting denitrifiers via their functional genes. Environ Microbiol 21:1196-1210

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H (2015) vegan: community ecology package.

Orellana LH, Rodriguez-R LM, Higgins S, Chee-Sanford JC, Sanford RA, Ritalahti KM, Loffler FE, Konstantinidis KT (2014) Detecting nitrous oxide reductase (nosZ) genes in soil metagenomes: method development and implications for the nitrogen cycle. mBio 5:e01193-01114

Penton CR, St Louis D, Pham A, Cole JR, Wu L, Luo Y, Schuur EA, Zhou J, Tiedje JM (2015) Denitrifying and diazotrophic community responses to artificial warming in permafrost and tallgrass prairie soils. Front Microbiol 6:746
Philippot L, Spor A, Henault C, Bru D, Bizouard F, Jones CM, Sarr A, Maron PA (2013) Loss in microbial diversity affects nitrogen cycling in soil. ISME J 7:1609-1619
Ravishankara AR, Daniel JS, Portmann RW (2009) Nitrous oxide (N_2O): the dominant ozone-depleting substance emitted in the 21st century. Science 326:123-125
Rich JJ, Heichen RS, Bottomley PJ, Cromack K, Jr., Myrold DD (2003) Community composition and functioning of denitrifying bacteria from adjacent meadow and forest soils. Appl Environ Microbiol 69:5974-5982
Rosch C, Mergel A, Bothe H (2002) Biodiversity of denitrifying and dinitrogen-fixing bacteria in an acid forest soil. Appl Environ Microbiol 68:3818-3829
Samad MS, Biswas A, Bakken LR, Clough TJ, de Klein CAM, Richards KG, Lanigan GJ, Morales SE (2016) Phylogenetic and functional potential links pH and N_2O emissions in pasture soils. Sci Rep-Uk 6:35990-35990
Sanford RA, Wagner DD, Wu Q, Chee-Sanford JC, Thomas SH, Cruz-Garcia C, Rodriguez G, Massol-Deya A, Krishnani KK, Ritalahti KM, Nissen S, Konstantinidis KT, Loffler FE (2012) Unexpected nondenitrifier nitrous oxide reductase gene diversity and abundance in soils. Proc Natl Acad Sci 109:19709-19714
Scala DJ, Kerkhof LJ (1998) Nitrous oxide reductase (nosZ) gene-specific PCR primers for detection of denitrifiers and three nosZ genes from marine sediments. Fems Microbiol Lett 162:61-68
Schöler A, Jacquiod S, Vestergaard G, Schulz S, Schloter M (2017) Analysis of soil microbial communities based on amplicon sequencing of marker genes. Biol Fertil Soils 53:485-489
Shen W, Xue H, Gao N, Shiratori Y, Kamiya T, Fujiwara T, Isobe K, Senoo K (2020) Effects of copper on nitrous oxide (N_2O) reduction in denitrifiers and N_2O emissions from agricultural soils. Biol Fertil Soils 56:39-51
Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30:2725-2729
Thomson AJ, Giannopoulos G, Pretty J, Baggs EM, Richardson DJ (2012) Biological sources and sinks of nitrous oxide and strategies to mitigate emissions. Philos Trans R Soc Lond Ser B Biol Sci 367:1157-1168
Throback IN, Enwall K, Jarvis A, Hallin S (2004) Reassessing PCR primers targeting nirS, nirK and nosZ genes for community surveys of denitrifying bacteria with DGGE. FEMS Microbiol Ecol 49:401-417
Tsiknia M, Paranychianakis NV, Varouchakis EA, Nikolaidis NP (2015) Environmental drivers of the distribution of nitrogen functional genes at a watershed scale. FEMS Microbiol Ecol 91:fiv052
Wang H, Han J, Deng N, An S (2019) Effects of the natural restoration time of abandoned farmland in a semiarid region on the soil denitrification rates and abundance and community structure of denitrifying bacteria. Appl Microbiol Biotechnol 103:1939-1951
Wang Q, Fish JA, Gilman M, Sun Y, Brown CT, Tiedje JM, Cole JR (2015) Xander: employing a novel method for efficient gene-targeted metagenomic assembly. Microbiome 3:32
Wang Q, Quensen JF, Fish JA, Lee TK, Sun Y, Tiedje JM, Cole JR (2013) Ecological patterns of nifH genes in four terrestrial climatic zones explored with targeted metagenomics using FrameBot, a new informatics tool. mBio 4:e00592-00513
Whitaker D, Christman M (2014) clustsig: Significant Cluster Analysis.

Wickham H, Chang W, RStudio (2017) ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. Springer, New York

Yoon S, Nissen S, Park D, Sanford RA, Loffler FE (2016) Nitrous Oxide Reduction Kinetics Distinguish Bacteria Harboring Clade I NosZ from Those Harboring Clade II NosZ. Appl Environ Microbiol 82:3793-3800

Zhang B, Penton CR, Xue C, Quensen JF, Roley SS, Guo J, Garoutte A, Zheng T, Tiedje JM (2017) Soil depth and crop determinants of bacterial communities under ten biofuel cropping systems. Soil Biol Biochem 112:140-152

Zhang B, Penton CR, Xue C, Wang Q, Zheng T, Tiedje JM (2015) Evaluation of the Ion Torrent Personal Genome Machine for gene-targeted studies using amplicons of the nitrogenase gene *nifH*. Appl Environ Microbiol 81:4536-4545

Zhao S, Wang Q, Zhou J, Yuan D, Zhu G (2018) Linking abundance and community of microbial NO-producers and NO-reducers with enzymatic NO production potential in a riparian zone. Sci Total Environ 642:1090-1099

Zumft WG, Kroneck PM (2007) Respiratory transformation of nitrous oxide (N2O) to dinitrogen by Bacteria and Archaea. Adv Microbiol Physiol 52:107-227

Figure captions

Fig. 1 Coverage comparison of the new primer set with eleven other primer pairs against references consisting of 324 sequences from Xander (Wang et al. 2015), visualized by iTOL (Letunic and Bork 2016). These eleven primer pairs corresponding to legend keys (covering length relative to 1938 bp *nosZ* of *R. palustris* CGA009) were nosZ1F_nosZ1R_Henry_2006 (247 bp) (Henry et al. 2006), nosZ2F_nosZ2R_Henry_2006 (267 bp) (Henry et al. 2006), nosZF_nosZR_Rich_2003 (701 bp) (Rich et al. 2003), nosZF_nosZ1622R_Throback_2004 (454 bp) (Throback et al. 2004), nosZF_nos1773R_Throback_2004 (247 bp) (Throback et al. 2004), nosZF-nosZR_ Rosch_2002 (701 bp) (Rosch et al. 2002), Nos1527F_Nos1773R_Scala_1998 (247 bp) (Scala and Kerkhof 1998), nosLb_nosRb_Chenby_1998 (302 bp) (Cheneby et al. 1998), nosZ1126F-nosZ1184R_Chen_2012 (759 bp) (Chen et al. 2012), and nosZ1F_nosZ2F_Philippot_2013 (458 bp) (Philippot et al. 2013). More details about these primers are listed in Table S1.

Fig. 2 BLASTp identity distribution of recovered hits by the newly designed nosZ1039F_nosZ1393R_this study primer set (labeled this study) and the widely-used nosZF_nosZR_Rich_2003 set (labeled Rich_2003 (Rich et al. 2003)) against the public references consisting of 324 sequences from Xander (Wang et al. 2015). A. Distribution based on Alpha- and Betaproteobacteria. B. Distribution based on Gamaproteobacteria and other classes.

Fig. 3 Differences in denitrifying communities among soils from eight USA states, representing soil, climate and vegetation diversity (A and B), and with soil depth and between crops (C-E). A. Difference (ANOVA, $P<0.001$) in Shannon diversity, with letters indicating the grouping. B. Differences (PERMANOVA, $F=3.673$, $P=0.001$) in overall community structures, with cornfields from four states more similar to each other. The insert shows the significant differences (PERMANOVA, $F=2.586$, $P=0.002$) among the cornfields. BorealF, boreal forest; SubtrDryF, subtropical dry forest; TropMoistF, tropical moist forest; GrassShrub, grassland-shrubland. C. Difference (Kruskal-Wallis, $P<0.01$) in α-diversity, with letters indicating the grouping. D-E. Differences in overall community structures among soil depths, crops, and their interactions by PERMANOVA.