Journal of Microbiological Methods 160 (2019) 49–59

Contents lists available at ScienceDirect

Journal of Microbiological Methods

journal homepage: www.elsevier.com/locate/jmicmeth

Optimization of PCR primers to detect phylogenetically diverse *nrfA* genes associated with nitrite ammonification

Jordan Cannon[a], Robert A. Sanford[b], Lynn Connor[c], Wendy H. Yang[a,b], Joanne Chee-Sanford[c,*]

[a] Dept. of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA
[b] Dept. of Geology, University of Illinois at Urbana-Champaign, Urbana, IL, USA
[c] USDA-ARS, Urbana, IL, USA

ARTICLE INFO

Keywords:
nrfA
PCR primers
Nitrite ammonification
DNRA

ABSTRACT

Dissimilatory nitrate reduction to ammonium (DNRA) is now known to be a more prevalent process in terrestrial ecosystems than previously thought. The key enzyme, a pentaheme cytochrome *c* nitrite reductase NrfA associated with respiratory nitrite ammonification, is encoded by the *nrfA* gene in a broad phylogeny of bacteria. The lack of reliable and comprehensive molecular tools to detect diverse *nrfA* from environmental samples has hampered efforts to meaningfully characterize the genetic potential for DNRA in environmental systems. In this study, modifications were made to optimize the amplification efficiency of previously-designed PCR primers, targeting the diagnostic region of NrfA between the conserved third- and fourth heme binding domains, and to increase coverage to include detection of environmentally relevant Geobacteraceae-like *nrfA*. Using an alignment of the primers to > 270 bacterial *nrfA* genes affiliated with 18 distinct clades, modifications to the primer sequences improved coverage, minimized amplification artifacts, and yielded the predicted product sizes from reference-, soil-, and groundwater DNA. Illumina sequencing of amplicons showed the successful recovery of *nrfA* gene fragments from environmental DNA based on alignments of the translated sequences. The new primers developed in this study are more efficient in PCR reactions, although gene targets with high GC content affect efficiency. Furthermore, the primers have a broader spectrum of detection and were validated rigorously for use in detecting *nrfA* from natural environments. These are suitable for conventional PCR, qPCR, and use in PCR access array technologies that allow multiplex gene amplification for downstream high throughput sequencing platforms.

1. Introduction

Phylogenetically diverse microorganisms mediate the reduction of nitrite to ammonium, a metabolic reaction catalyzed by the pentaheme c-type cytochrome nitrite reductase NrfA (Einsle et al., 2011; Simon and Klotz, 2013). NrfA catalyzes the second step in the two-step process commonly referred to as dissimilatory nitrate reduction to ammonium (DNRA) (Rutting et al., 2011; Tiedje, 1988). This key reaction, namely nitrite ammonification, occurs at a critical branching point in the nitrogen (N) cycle where nitrite serves as an intermediate for both oxidation and reduction biochemical reactions (Baggs, 2011; Kuypers et al., 2018; Stein and Klotz, 2011; Stein and Klotz, 2016; Welsh et al., 2014). In contrast to denitrification where the terminal products are gaseous N-oxides, nitrite ammonification retains nitrogen in the system, with implications for potential N conservation in many ecosystems (Bhowmik et al., 2017, Chen et al., 2015a, 2015b, Giblin et al., 2013, Rutting et al., 2011, Silver et al., 2001, Smith et al., 2015, Templer et al., 2008). Recent studies have shown that DNRA is more prevalent in some ecosystems (e.g. soil) than previously thought, but the potential for DNRA to occur across diverse environments is still poorly characterized (Rutting et al., 2011; Yang et al., 2017). The distinction of NrfA as a key enzyme in nitrite ammonification has consequently allowed *nrfA* to serve as the bacterial genetic marker to investigate DNRA in these environments. Although some molecular methods have been developed in recent years to detect *nrfA* genes, the current PCR based tools available are not optimized to detect the diversity of NrfA contributing to DNRA activity in soil and other environmental systems.

DNRA has been demonstrated to be an important nitrate reduction pathway in aquatic and estuarine ecosystems (Decleyre et al., 2015; Gardner et al., 2006; Jensen et al., 2011; Koop-Jakobsen and Giblin, 2010; Kuypers et al., 2018; Lam et al., 2009; Roberts et al., 2014; Song et al., 2014) and wetlands (Bernard et al., 2015; Gao et al., 2017;

* Corresponding author.
E-mail address: Joanne.Cheesanford@ars.usda.gov (J. Chee-Sanford).

https://doi.org/10.1016/j.mimet.2019.03.020
Received 4 March 2019; Received in revised form 20 March 2019; Accepted 20 March 2019
Available online 21 March 2019

Jahangir et al., 2017). While nitrite ammonification activity has traditionally been thought to occur where the C: N ratio is high and redox potential is low (Tiedje, 1988), DNRA has also been found to occur across diverse upland terrestrial ecosystems where these conditions do not always apply, including forests (Sotta et al., 2008; Minick et al., 2016; Rutting and Muller, 2008; Silver et al., 2005; Templer et al., 2008), grasslands (Chen et al., 2015b), deserts (Yang et al., 2017), and agricultural soils (Chen et al., 2015a, Pandey et al., 2018, Shan et al., 2016, Zhang et al., 2015). A recent metagenome study revealed DNRA is a high frequency trait among widely diverse bacterial taxa in a range of different soil habitats (Nelson et al., 2016), indicating the large genetic potential for nitrite ammonification in these systems. This suggests that DNRA is a more widespread process in the environment than previously thought and that our understanding of controls for nitrite ammonifiers and DNRA is incomplete.

Ecosystem analysis of DNRA is complicated by NrfA activity occurring among ecophysiologically disparate bacteria, with 18 NrfA clades previously designated based on a cutoff of 30% amino acid sequence identity divergence (Welsh et al., 2014). Differences among clades, such as different proteins associated with delivering electrons to NrfA (NrfB vs. NrfH), could have implications for controls on DNRA rates. The emerging view on the wide-ranging types of habitats and conditions where DNRA can occur taken together with the high diversity of bacterial NrfA highlights the need for updated molecular tools that can provide comprehensive coverage able to span multiple types of ecosystems to further advance studies that can track dynamic changes in NrfA communities. While the genetic diversity of *nrfA* is apparent, apart from a few taxonomic groups that have been the focus of some studies (Giacomucci et al., 2012; Mania et al., 2014; Su et al., 2012; van den Berg et al., 2016; van den Berg et al., 2017; Yin et al., 2002; Yoon et al., 2015), little is known about the majority of taxa harboring *nrfA*. The significance of *nrfA* diversity within a specified environment is not yet fully understood, however the extent to which DNRA occurs in natural environments due to the presence of a pentaheme cytochrome *c* nitrite reductase depends on active members of the extant NrfA community. Broad detection of *nrfA* will reveal both the clade-level community structure and diversity of nitrite ammonifiers and thus, DNRA potential in a system.

In one of the first studies to develop DNA-based methodology to detect the *nrfA* gene, Mohan et al. (Mohan et al., 2004) designed PCR primer sets that were based on a limited set (< 10) of genomic reference sequences. Ten years later, Welsh et al. (2014) performed an extensive phylogenetic analysis of non-redundant *nrfA* from 272 available reference genomes and identified a primer set targeting coding regions of highly conserved NrfA heme binding domains in 16 of 18 diverse clades that could effectively expand detection of *nrfA* in natural environments, soils included. These initial primers were predicated on conserved nucleotide sequences that corresponded with the third and fourth heme-binding domains of the NrfA pentaheme protein. This region of NrfA contains conserved amino acid residues with key diagnostic features specific for nitrite ammonification that are easily analyzed with translated PCR amplicon sequences. While the primer set did yield correct *nrfA* gene amplicons from soil, primer design features precluded universality in PCR conditions that could apply broadly to different environmental samples and an acknowledged inability to detect environmentally relevant Geobacteraceae-like *nrfA* genes (Clade I) made their use suboptimal.

One of the problematic design features of the primer set designed by Welsh et al. (2014) are the low primer melting temperatures (~47 °C) that could prevent efficient binding to the target regions of most gene phylotypes. This leads to potential amplification biases that would result in an inaccurate accounting of *nrfA* genes of some taxa within the different clades, particularly problematic if DNA from natural environments are composed of widely varying abundances of bacterial populations harboring *nrfA* genes. A number of studies have used the previously published primers to detect *nrfA* in wide-ranging environments (Bu et al., 2017; Lam et al., 2009; Lindemann et al., 2016; Pandey et al., 2018; Smith et al., 2007; Song et al., 2014; Stief et al., 2018; Takeuchi, 2006; Tatti et al., 2017; Yu et al., 2018). Some of these studies have attempted to compensate for low detection thresholds and problematic PCR issues by increasing the number of cycles used in PCR. Unfortunately, this approach likely exchanges one problem for another; excessive PCR cycles (> 35) are known to lead to non-target amplicon products (Qiu et al., 2001; Sipos et al., 2007; Wu et al., 2010). Any inaccuracies in *nrfA* detection resulting in missed coverage or inefficient amplification would also affect meaningful interpretation of results, particularly true in qPCR analyses. Thus, previous *nrfA*-targeting primer sets risk erroneous accounting of the relative contributions of taxa from different NrfA clades to the DNRA activity in an environmental sample and even miss important clades (e.g. Clade I). Consequently, this prompted a necessary reevaluation of primer design.

The objective of this study was to improve the breadth of coverage and the efficiency of PCR amplification by *nrfA* primers. To do this, we modified the primers designed by Welsh et al. (2014) to increase the primer melting temperatures and improve primer pair compatibility, thus allowing optimization of PCR annealing temperatures and universality of reaction conditions that can effectively target wide-ranging *nrfA* G + C contents. These optimized primer features served to reduce the erroneous need for using excessive PCR cycles when applied to environmental samples and thus minimizes likely amplification artifacts. Amplicons generated from these primers targeted regions that yield products mainly between 236 bp and 278 bp. This amplicon size range is suitable for multiple methodologies including conventional PCR, qPCR, RT-qPCR, and multiplex amplification arrays (e.g. Fluidigm Access Array™) used for amplicon-based sequencing platforms (e.g. Illumina). The resulting two new forward primers, required to extend coverage to Geobacteraceae (Clade I), coupled to a new reverse primer, yielded PCR products that retain all the conserved genetic features of the third- and fourth heme-binding domains encoded by *nrfA*. This improves diagnostic detection across the broad diversity of NrfA in a variety of environments.

2. Materials and methods

2.1. NrfA sequence selection

A previous phylogenetic analysis of 272 full-length NrfA protein sequences, based on Bayesian inference, distinguished 18 clades possessing conserved features diagnostic of pentaheme NrfA proteins (Welsh et al., 2014). Using the nucleotide alignment of *nrfA* from these reference sequences, between 2 and 11 sequences (median = 5) were manually chosen from each clade to represent the sequence diversity within the clade for the purpose of primer design. Clade membership size was also taken into consideration in choosing the number of representatives. Generally, a greater number of representatives were selected as clade size increased, resulting in 99 reference sequences chosen for the initial primer design in silico evaluation. The resulting final sets of new primers were ultimately tested in silico against a library of 271 *nrfA* sequences assembled here (see Fig. 2 in (Cannon et al., 2019)) and previously from Welsh et al. (2014). NrfA sequences from three metagenome-assembled genomes (European Nucleotide Archive # PRJEB20068) belonging to Clades K and N and derived from the agricultural soils used in this study (described below) were also included for this analysis (Orellana et al., 2017).

2.2. Sequence alignment and primer design

A highly conserved region of the *nrfA* gene encoding the 3rd heme-binding domain was previously used to design forward primer nrfAF2aw (Welsh et al., 2014) (Table 1) to be paired with reverse primer nrfAR1 that targeted a region in the conserved 4th heme-binding domain (Mohan et al., 2004). These regions were the starting

Table 1
Primer sequences and properties of *nrfA* primers used in this study.

Primer Name	T_m^{primer} (°C)[a]	Sequence (5′→3′)	Reference
Forward			
nrfAF2aw	47.3	CAR TGY CAY GTB GAR TA	(Welsh et al., 2014)
nrfAF2awMOD	54.9	GSI CAR TGY CAY GTI GAR TA	This study
nrfAF2awMODgeo	54.2	GSI CAR TGY CAY GTI ASB TA	This study
Reverse			
nrfAR1	47.6	TWN GGC ATR TGR CAR TC	(Mohan et al., 2004)
nrfAR1MOD	53.4	GGC ATR TGR CAR TCI RYR CA	This study

[a] Melting temperature of primer to itself is calculated using MacVector software v. 16.0.8 which follows the thermodynamic parameters described by SantaLucia Jr. (1998). T_m values are calculated using N in place of inosine (I) substitutions.

basis for modifications that would raise primer melting temperatures and retain temperature compatibility between the forward and reverse primers, raise the primer:target annealing temperature, and further improve coverage of *nrfA* diversity with four criteria in accordance with general principles of primer design considered (Dieffenbach et al., 1993) (< https://www.biocompare.com/Bench-Tips/133581-Primers-by-Design-Tips-for-Optimal-DNA-Primer-Design/ >): 1) primer melting temperatures (T_m) in the range 52–58 °C for more efficient annealing and better compatibility with certain qPCR and amplicon sequencing platforms, 2) primer length of 18–22 base pairs, 3) %GC content between 40 and 60%, and 4) ΔT_m < 5 °C between primers and their corresponding target sequences. All sequence alignments, mismatch identification, and analyses of temperature characteristics were made in silico using tools in MacVector software (v. 16.0.8, MacVector, Inc.). The resulting primer sequences were further analyzed for consensus alignment in silico against reference sequences grouped by clades, and graphically represented as sequence logos generated using WebLogo (v. 2.8.2 (2005-09-08), < https://weblogo.berkeley.edu/ >) (Crooks et al., 2004). Due to the increase in primer degeneracies introduced in the new designs, inosine substitutions were made for all positions with an "N" and subsequently used for all PCR optimization tests.

2.3. Validation of primers

DNA extracts from reference strains originating from a variety of environments that represent 11 of the 18 total *nrfA* clades were used to test new primer pair candidates. The reference DNA included *Serratia fonticola* strain HAc5 (Clade A) (Genbank #JX293824.1), *E. coli* K12 (Clade A), *Shewanella oneidensis* MR-1 (Clade C), *Shewanella loihica* (Clades C and D), *Wolinella succinogenes* (Clade F), *Geobacter bemidjiensis* Bem (Clade I), *Anaeromyxobacter dehalogenans* st. 2CP-1 (Clades J and K), *Desulfitobacterium hafniense* DCB-2 (Clade M and N), *Desulfovibrio vulgaris* str. Hildenborough (Clade O), and *Bacillus* sp. strain UAAc7 (Clade P) (Genbank #JX293830.1). Full *nrfA* sequences were obtained from the Functional Gene Pipeline and Repository (FUNGENE) (http://fungene.cme.msu.edu/) database, version 9.5 (February 2018). *S. fonticola* strain HAc-5 and *Bacillus* sp. strain UAAc-7 were previously isolated from agricultural soils and draft genomes were obtained from each of them (Chee-Sanford, unpublished).

DNA was extracted from reference cultures and soil as described previously (Welsh et al., 2014) using a phenol: chloroform extraction method (Tsai and Olson, 1991) modified by the addition of glycogen (20 mg/ml) to enhance the recovery of DNA during precipitation. Soil DNA samples consisted of equal volumes of DNA pooled accordingly from extracts of soil taken in April 2012 and November 2012 from depths of 0–5 cm, 5–20 cm, and 20–30 cm at agricultural sites near Havana, Illinois (HW) and Urbana, Illinois (UM). DNA from additional soil and groundwater samples used specifically for amplicon sequencing were extracted using an abbreviated phenol:chloroform protocol (Griffiths et al., 2000) and then followed by glycogen-enhanced recovery as described above. Final DNA concentrations (~8–10 ng/μl) were measured using Qubit 2.0 fluorometry (Invitrogen) and DNA band intensities estimated against quantitative DNA ladders following gel electrophoresis.

2.4. Optimization of PCR

All primers were HPLC-purified and obtained from IDT (Integrated DNA Technologies, Skokie, IL, USA). Stock concentrations (100 μM) of each primer were made by adding Invitrogen™ UltraPure™ DNase/RNase-Free Distilled Water (Thermo Fisher Scientific Waltham, MA, USA) and subsequently diluted for use in PCR. PCR reactions were performed in 25 μl volumes using the Takara ExTaq PCR kit (Clontech) and a MJ Research PTC-200 Gradient Thermal Cycler. The optimized reaction mixture was the following: 1× PCR buffer, 0.2 mM each deoxynucleoside triphosphate (dNTPs), 0.025 U/μl TaKaRa *Ex Taq* DNA polymerase, 3.2 μM each forward and reverse primers, and ~1 ng reference template DNA. Thermocycling conditions were the following: initial denaturation step at 95 °C for 5 min, followed by 25 or 30 cycles of [95 °C for 30 s, 56 °C for 30 s, and 72 °C for 30 s] and a final extension of 72 °C for 10 min. The optimal annealing temperature of 56 °C was empirically determined using a temperature gradient program for amplification of DNA from *E. coli* K12, *S. fonticola* HAc-5, and *A. dehalogenans* 2CP-1. Due to the degenerate nature of the primers, optimal concentrations per reaction using 1.6-, 3.2-, and 6.4 μM of each primer were independently tested against DNA from these same organisms and the final concentration was based on visualizing maximum product yields by gel analysis. DNA from ten *nrfA* containing organisms served as positive controls to test the efficacy of the primers (Table 2). PCR products were resolved by gel electrophoresis using 2.5% High Resolution Agarose (fragments < 1 kb) (Gold Biotechnology, Olivette, MO, USA) in 1 × TBE buffer on a HU13 Midi horizontal gel unit (Scieplas Ltd., Cambridge, UK) at 4 V/cm for 80 min. Amplicon product sizes varied depending on the reference *nrfA* (Table 2). DNA ladders consisted of 1 μl of Low Molecular Weight DNA Ladder (New England Biolabs Inc., Ipswich, MA, USA) and 5 μl of Quick-Load Purple 2-Log DNA Ladder (0.1–10.0 kb).

PCR temperature cycling conditions were evaluated using soil DNA under fixed annealing temperature (56 °C) and varying the total number of cycles from 25 to 35. Qualitative yields of products were assessed by gel electrophoresis as described above.

To further corroborate the specificity of the forward primers nrfAF2awMOD and nrfAF2awMODgeo when paired with the reverse primer nrfAR1Mod, primer combinations (individual pairs (F and R)) were used to evaluate amplicons generated from a mixed DNA pool (1 ng each) of reference DNA from *S. fonticola* HAc-5, *S. oneidensis*, *A. dehalogenans* 2CP-1, and *G. bemidjiensis*. A combined pool of both forward primers with the reverse primer was also tested against the same reference DNA to assess any inhibition that could result from competing reactions.

2.5. Amplified fragment length polymorphism (AFLP) analysis

Amplified fragment length polymorphism (AFLP) analysis was used

Table 2
Characteristics of primers and expected products analyzed in silico against ten reference organisms representing eleven of eighteen NrfA clades.

Organism	Clade[a]	PCR product size (bp)	Primer pair[b]	Forward primer Tm to target (°C)	Reverse primer Tm to target (°C)	ΔT_m of pair to targets (°C)	Optimal annealing temperature (°C)	Product GC content (%)	Product T_m (°C)
E. coli K12	A	236	nrfAF2awMOD, nrfAR1MOD	59.4	57.5	1.8	56.2	48.3	76.8
Serratia fonticola[c]	A	236	nrfAF2awMOD, nrfAR1MOD	62.7	57.5	5.2	56.5	49.6	77.4
Shewanella oneidensis MR-1	C	245	nrfAF2awMOD, nrfAR1MOD	58.2	58.6	0.4	54.4	41.2	74.0
Shewanella loihica PV4 2nd	C	245	nrfAF2awMOD, nrfAR1MOD	56.1	57.1	1.0	55.4	46.9	76.4
Shewanella loihica PV4 1st	D	239	nrfAF2awMODgeo, nrfAR1MOD	62.2	63.3	1.1	59.7	55.6	79.9
Wolinella succinogenes	F	269	nrfAF2awMOD, nrfAR1MOD	62.2	61.5	0.7	58.0	49.8	77.8
Geobacter bemidjiensis Bem 1st	I	266	nrfAF2awMODgeo, nrfAR1MOD	62.2	66.0	3.8	60.4	57.5	80.9
Geobacter bemidjiensis Bem 2nd	I	266	nrfAF2awMODgeo, nrfAR1MOD	61.4	63.2	1.9	61.7	62.8	83.1
Anaeromyxobacter dehalogenans 2CP-1 2nd	J	248	nrfAF2awMOD, nrfAR1MOD	64.6	68.8	4.2	63.7	66.9	84.6
Anaeromyxobacter dehalogenans 2CP-1 1st	K	236	nrfAF2awMOD, nrfAR1MOD	64.6	68.8	4.2	63.6	66.9	84.5
Desulfitobacterium hafniense DCB-2 2nd	M	236	nrfAF2awMOD, nrfAR1MOD	56.2	63.2	7.1	56.1	49.6	77.4
Desulfitobacterium hafniense DCB-2 1st	N	236	nrfAF2awMOD, nrfAR1MOD	54.9	60.7	5.8	55.5	48.7	77.0
Desulfovibrio vulgaris str. Hildenborough	O	278	nrfAF2awMOD, nrfAR1MOD	61.8	65.8	4.0	62.1	63.7	83.6
Bacillus UAAc-7	P	251	nrfAF2awMOD, nrfAR1MOD	60	61.5	1.4	55.5	43	74.9
AVERAGE				60.5	62.4	3.0	58.5		

[a] NrfA clade designations were based on phylogenetic analysis of 272 reference sequences (Welsh et al., 2014).
[b] Primer pairs were tested using N-substitutions in place of inosine bases to calculate T_m values using tools available in MacVector version 16.0.8 software.
[c] *Serratia fonticola* (FUNGENE CP011254).

to assess the amplification efficiencies from a pool of different reference *nrfA* and to demonstrate the potential utility of this method in creating a community fingerprint of the amplicons obtained. Reference DNA covering a range of *nrfA* %GC content (41.2% to 66.9%) of the yielded products with predicted amplicon lengths ranging from 236 bp to 278 bp, facilitating separation in the fragment analysis. To test AFLP for profiling a *nrfA* community, a mixed pool containing 1 ng DNA from each of seven reference organisms that included *D. hafniense* strain DCB-2, *Bacillus* UAAc-7, *D. vulgaris* strain Hildenborough, *S. oneidensis* MR-1, *W. succinogenes*, *S. fonticola* HAc-5, and *A. dehalogenans* 2CP-1 was used in 50 µl PCR reactions under conditions as described above, except the forward primer was labeled at the 5′ end with 6-FAM (Integrated DNA Technologies, Skokie, IL, USA). Additionally, separate 25 µl PCR reactions were conducted with each individual reference (1 ng/reaction). The PCR products from individual reactions were pooled in equal volumes following PCR amplification before submitting for fragment size analysis (Roy J. Carver Biotechnology Center, University of Illinois, Urbana, IL) and the resulting AFLP was compared to the AFLP pattern generated following amplification of the DNA mixed pool. To test the application of AFLP to an environmental sample, soil DNA were also amplified for 30 cycles using ~8–10 ng of DNA in individual reactions and PCR reaction mixtures were modified with the addition of 25 µg/ml T4 gene 32 protein (Roche Applied Science, Indianapolis, IN, USA) and spiked with 1 ng DNA from *D. vulgaris* as an internal standard.

3. Results

3.1. Primer characteristics and nrfA clade coverage

An in silico assessment of the alignment of our previous primers, nrfAF2aw and nrfAR1, to 99 reference *nrfA* sequences from all 18 clades revealed flanking regions suitable for modifications that met most of the desired design features for optimal PCR (see Materials and Methods). Cutting three bases off the 5′ end of the reverse primer and adding six bases to the 3′ end increased the length to 20 bases and resulted in higher theoretical primer melting temperatures (T_m^{primer} = 53.4 °C) (Table 1). This reverse primer (nrfAR1MOD) targets region of the gene that are highly conserved for all clades with the exception of Clades D and R (Fig. 1; see Fig. 2 (Cannon et al., 2019)). Translation of the primer target region revealed that the mismatches from the expected amino acid sequence in this region for all clades was reduced from 19.5% to 11.6% for nrfAR1 and nrfAR1MOD, respectively. A large portion of improved coverage of this primer was associated with Clade I. An analysis of 271 reference DNA sequences showed that 90.4% matched the new reverse primer allowing for up to one mismatch (Fig. 1; see Table 1 in (Cannon et al., 2019)). Allowing two mismatches covered 97% of the sequences across all clades, with only Clade R receiving moderate coverage (58.3%) (see Table 1 in (Cannon et al., 2019)). The original reverse primer (nrfAR1) (Mohan et al., 2004) also had additional mismatches with Clade I and only provided coverage for about 80% of the reference sequences from all clades.

An extension of the initial forward primer (nrfAF2aw) by three bases at the 5′ end increased the length to 20 bases and included an additional degeneracy that further improved coverage and increased the T_m^{primer} to 54.9 °C (Table 1). The coverage for each of the clades ranged from 0%–100% (average = 63.8%) allowing for one mismatch, with two highly conserved bases (TA) maintained at the 3′ end of the primer (Fig. 1; see Fig. 1 and Fig. 2 in (Cannon et al., 2019)). Clades E and I are entirely missed by the new forward primer nrfAF2awMOD due to multiple mismatches at different positions in the target region (Fig. 1). The graphical representation (i.e. WebLogo) of target:primer matches further showed key mismatches at positions 4 and 6 and positions 16–18 across the majority of members for Clades E and I, respectively (Fig. 1, see Fig. 2 in (Cannon et al., 2019)). Fractional coverage (36.4% - 50%) by nrfAF2awMOD to *nrfA* genes in Clades D, M, N, Q, and R (Fig. 1) are due to sequence divergences that are not covered by the consensus sequence of the forward primer (see Fig. 1 and Fig. 2 in (Cannon et al., 2019)). Since many members of Clade D (e.g. *Shewanella loihica* PV4) have multiple copies of *nrfA* belonging to another clade, the use of the modified forward primer nrfAF2awMOD would still allow detection of these taxa as part of the *nrfA* genes detected in DNA sample (Fig. 2). For Clades D, M, N, Q, and R the implicit awareness of partial clade coverage by the primer must be considered in data interpretation, although if two mismatches are allowed in evaluating nrfAF2awMOD, the average coverage of all clades increased to ~84%, providing a substantial improvement for Clades E, M, N and R (see Table 1 in (Cannon et al., 2019)).

Genes associated with NrfA Clade I, which included Geobacteraceae, shared among all its members mismatches corresponding to nucleotides 16–18 in the redesigned forward primer nrfAF2awMOD (Fig. 2; see Fig. 1 and Fig. 2 in (Cannon et al., 2019)). To address this issue and to target this environmentally relevant clade, modifications were made to bases at these positions yielding the forward primer nrfAF2awMODgeo. This clade-specific primer matched 87% of the reference sequences of Clade I allowing for one mismatch and also improved coverage for members of Clades D, M, N and R (Fig. 1, see Fig. 3 in (Cannon et al., 2019)). Most of Clade D, however, possess *nrfA* genes that do not match the new reverse primer and therefore would not yield amplicons paired with this Clade I-targeted forward primer. By allowing up to two mismatches this nrfAF2awMODgeo primer would provide 100% coverage of all members within Clade I (see Fig. 3 and Table 1 in (Cannon et al., 2019)). The predicted combined coverage for the two new forward primers increased to 78.2% of all the reference genes evaluated with one mismatch allowed.

3.2. Amplification of nrfA from reference organisms and soil

Extending the length of primers from 17 to 20 bases resulted in commensurate increases in primer T_m from ~47–48 °C to 53–55 °C and $\Delta T_m \leq 1.5$ °C (Table 1). The T_m characteristics for the reference targets to their corresponding forward primers nrfAF2awMOD or nrfAF2awMODgeo and reverse primer nrfAR1MOD ranged from 54.9 °C to 64.6 °C and 57.1 °C to 68.8 °C, respectively (Table 2). The calculated optimal annealing temperatures ranged from 55.4 °C to 63.7 °C, with 56 °C empirically determined to be the optimal for amplification. Percent GC contents ranged between 41.2% and 66.9% for all PCR products. Under the optimized PCR conditions, the amplified *nrfA* genes from seven reference organisms yielded single amplicons of the expected sizes using the nrfAF2awMOD/nrfAR1MOD primer pair (Fig. 3). In a mixed pool of selected reference genomic DNA, primer pair nrfAF2awMODgeo/nrfAR1MOD only amplified *G. bemidjiensis* BEM *nrfA* producing the expected size of 266 bp (Fig. 4A). An additional product (~360 bp) was notably amplified from *G. bemidjiensis* in all reactions with the Clade I-inclusive primer set, but its efficiency of amplification appeared to be lower than the intended *nrfA* target (Fig. 4A; see Fig. 4 in (Cannon et al., 2019)). Inspection of the *G. bemidjiensis* genome indicated the potential for primers binding with mismatches to heme-binding domains within another c-type cytochrome protein unrelated to NrfA that would yield the larger PCR product (data not shown). Neither product from *G. bemidjiensis* was amplified with the broad coverage primer set nrfAF2awMOD/nrfAR1MOD. When both primer sets were compared at an annealing temperature of 56 °C using increasing numbers of PCR amplification cycles (i.e. 25, 30 and 35) with soil DNA, correct sized products were visible by gel electrophoresis with as few as 25 cycles with only the new broad coverage primer set (Fig. 4B). The Clade I-inclusive primer set (nrfAF2awMODgeo/nrfAR1MOD) resulted in correct *nrfA* products soil DNA after 30 or 35 cycles, confirming their presence and suggesting their possible lower abundance in this environment.

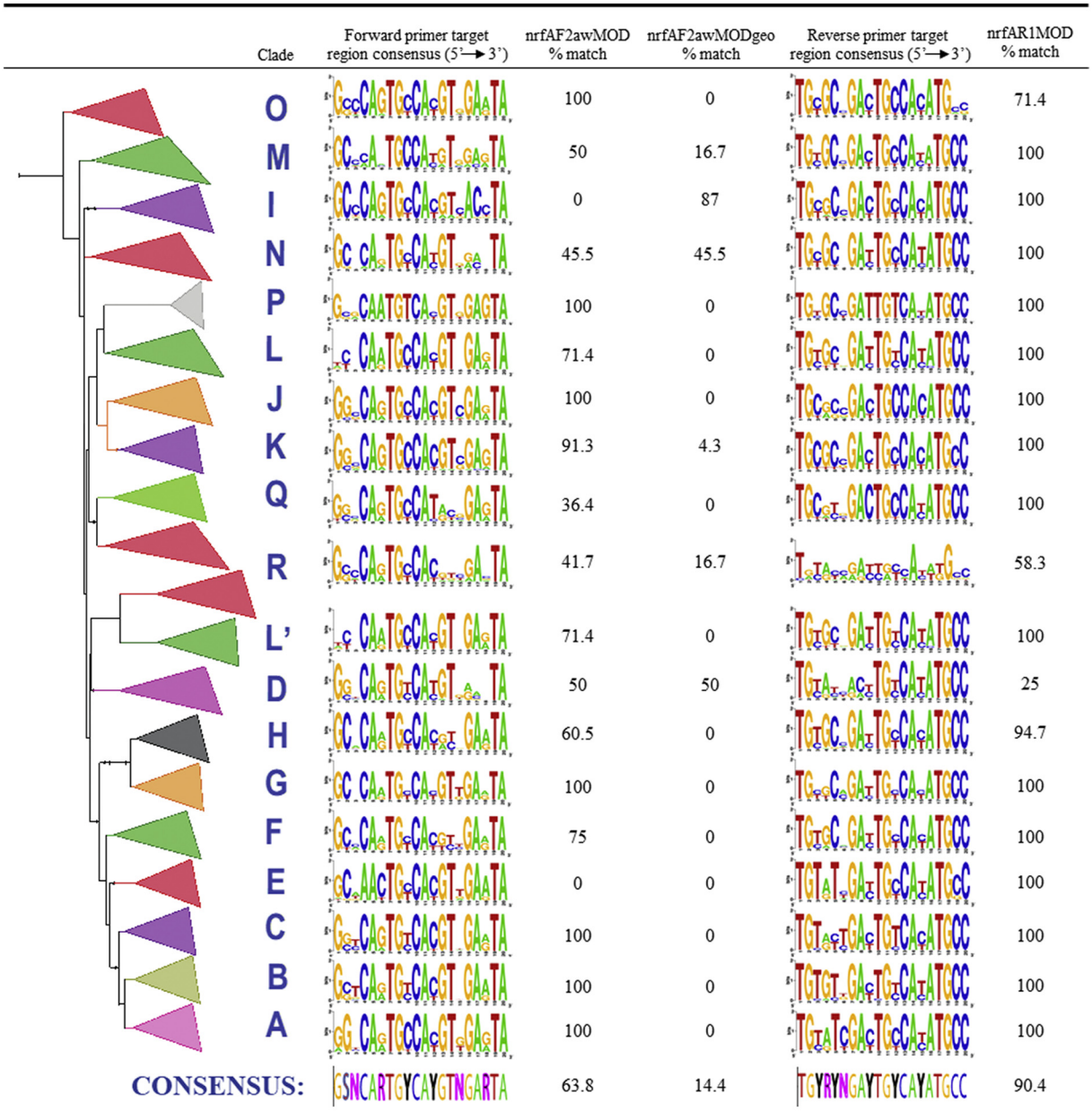

Fig. 1. WebLogos depicting stacks of nucleotides (one stack for each position in the sequence) corresponding to regions targeted by the forward primers nrfAF2awMOD or nrfAF2awMODgeo, and reverse primer nrfAR1MOD, respectively. The overall height of the stack indicates the sequence conservation at that position, while the height of symbols within the stack indicates the relative frequency of each nucleic acid at that position. Each WebLogo represents the match to the corresponding membership within each of the 18 clades delineated by a phylogenetic tree showing the relationship among 271 reference *nrfA* sequences. The consensus sequence of the reverse primer target region depicts the reverse and complement of the target sequence. Percent (%) matches represent the coverage of reference sequences within each clade or overall for all sequences with the corresponding primer allowing one mismatch.

Amplified fragment length polymorphism (AFLP) analysis of a mixed pool of reference *nrfA* containing seven reference genomes (*A. dehalogenans* 2CP-1 Clades J & K; *S. fonticola* HAc-5 Clade A; *D. hafniense* DCB-2 Clades M&N; *S. oneidensis* MR-1 Clade C; *Bacillus* st. UAAc-7 Clade P; *W. succinogenes* Clade F; *D. vulgaris* st. Hildenborough Clade O) demonstrated the recovery of all expected products when using the nrfAF2awMOD - nrfAR1MOD primer pair (Fig. 5A). The mixed pool served as a proxy for diverse *nrfA* that could be potentially present in an environmental sample. The relative amount of amplified product (i.e. fragment peak height), when using identical template DNA amounts, for each member in a mixed DNA pool compared to corresponding products generated in individual PCR reactions indicated amplification efficiency differences that were dependent on the reference taxa. For example, the relative amplification efficiencies of *A. dehalogenans* and *Bacillus* UAAc-7 appeared to be lower when in competition with other references compared to being amplified individually (Fig. 5A and Fig. 5B). In addition, amplicons with different %G + C content with the same size from different taxa were separated in the fragment analysis. For example, *A. dehalogenans, S. fonticola* HAc5 and *D. hafniense* all have *nrfA* AFLP peaks corresponding to a fragment size of 236, however the *A. dehalogenans* gene fragment represented as peak 1a migrates faster and is separated from the other two of the same size. (Fig. 5). This suggested that AFLP analysis may not accurately reflect the true size of the amplicon and is most likely associated with G + C content. To evaluate the utility of AFLP profiling to provide a relatively simple methodological tool for *nrfA* community comparisons in environmental

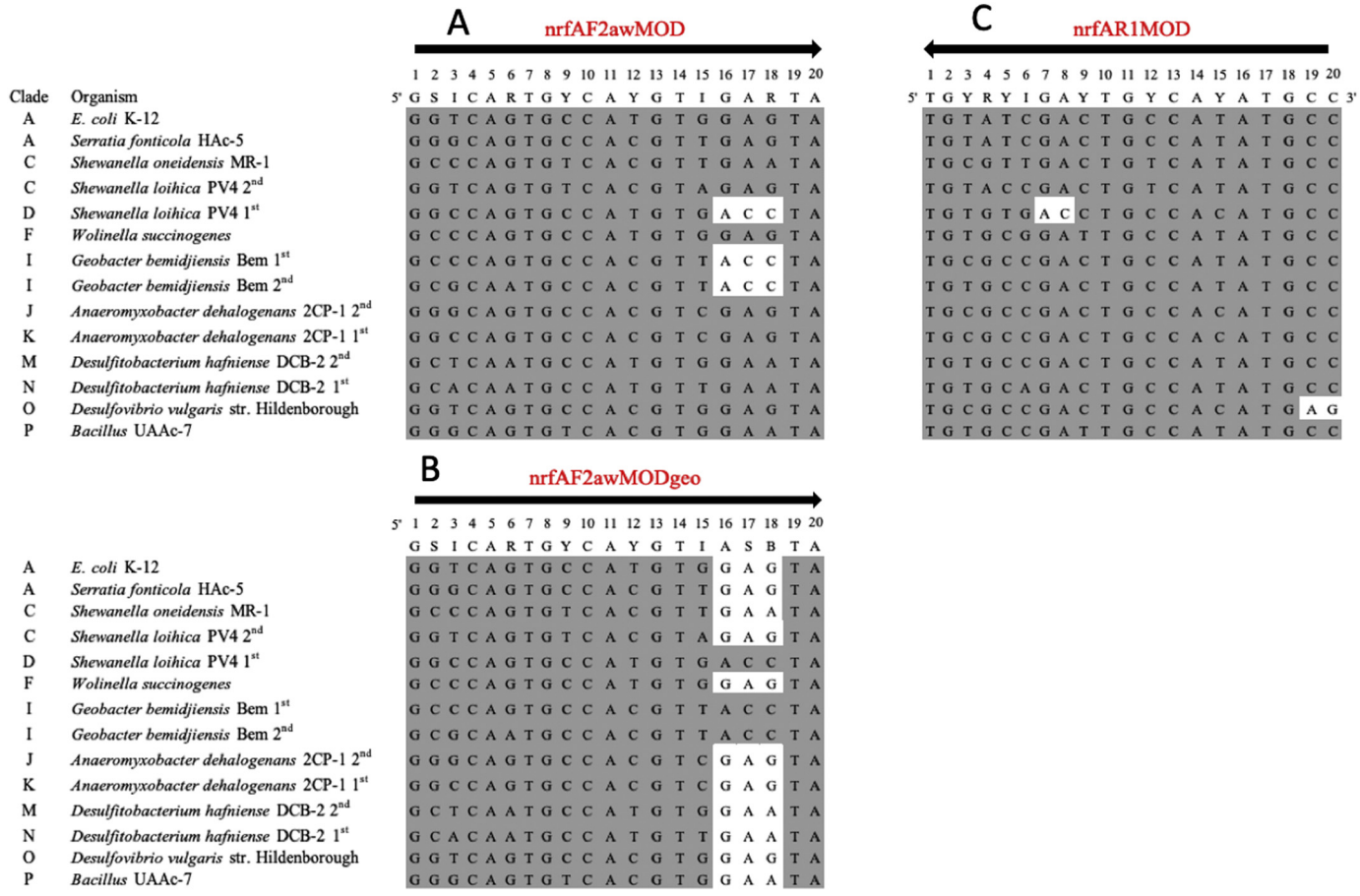

Fig. 2. Nucleotide alignment of the target regions of *nrfA* from ten reference organisms representing eleven of the eighteen clades corresponding to forward primers (A) nrfAF2awMOD and (B) nrfAF2awMODgeo, and reverse primer (C) nrfAR1MOD. The reverse and complement sequence is shown for nrfAR1MOD. Shaded nucleotides match corresponding bases in the primer. Unshaded nucleotides are mismatches between primer and target site.

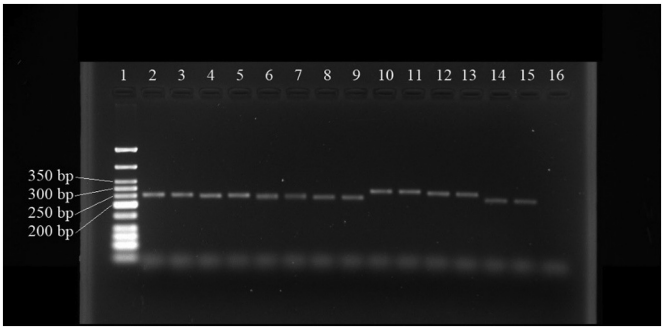

Fig. 3. Gel electrophoresis image of PCR products amplified from DNA of seven reference organisms using primer pair nrfAF2awMOD / nrfAR1MOD. Lane 1: DNA ladder, Lanes 2–3: *Bacillus* UAAc-7, Lanes 4–5: *S. oneidensis* MR-1, Lanes 6–7: *A. dehalogenans* 2CP-1, Lanes 8–9: *D. hafniense* DCB-2, Lanes 10–11: *D. vulgaris* str. Hildenborough, Lanes 12–13: *W. succinogenes*, Lanes 14–15: *S. fonticola* HAc-5, and Lane 16: ultrapure water.

studies, both primer sets were applied to different soil DNA (Fig. 5C; see Fig. 5 in (Cannon et al., 2019)). Fragments corresponding to known amplicon sizes (e.g. ~236 bp) were detected in several test soils. The occurrence of several additional fragments in the correct size range, some matching reference taxa, demonstrated the presence of diverse *nrfA* genes in the soil (Fig. 5C). Variation in AFLP fragment patterns was noted with different soil DNA and may be used to compare the differences in the community composition of *nrfA*- containing microbial populations.

4. Discussion

The redesign of primer sets to target the *nrfA* gene, necessitated based on a critical assessment of the primer set (nrfAF2aw/nrfAR1), yields compatible primer sets with improved primer melting temperature characteristics and primer length features that will more reliably yield efficient PCR amplification when applied to environmental samples. Optimizing primer features recognizes PCR as an indispensable method for functional gene studies in microbial ecology, but the effectiveness of this approach relies on primers and PCR conditions that are optimized for target detection and rigorously tested for universality in application. This is especially relevant for N-cycling genes and high apparent diversity in light of the growing number of sequence deposits made from genomic-, metagenomic- and amplicon-based sequencing studies (Bonilla-Rosso et al., 2016; Jones et al., 2014; Orellana et al., 2014; Sanford et al., 2012; Welsh et al., 2014). Unlike the polyphyletic nature of genes encoding nitrous oxide reductase (NosZ) (Jones et al., 2013; Sanford et al., 2012) and nitrite reductase (NirS, NirK) (Bonilla-Rosso et al., 2016) which makes universal primer sets difficult to design, the two primer sets for NrfA developed in this study provide broad coverage that include 17 of 18 clades now further inclusive of Clade I members applicable to a variety of environments. Of additional significance is the suitability of the new primer sets for multiple types of PCR applications (e.g. amplicon sequencing and qPCR).

The evaluation of the new primers in our study also served to underscore the importance of attention paid to PCR application standards that can be commonly overlooked by practitioners due to the perceived ease and simplicity often associated with PCR methods. Recent

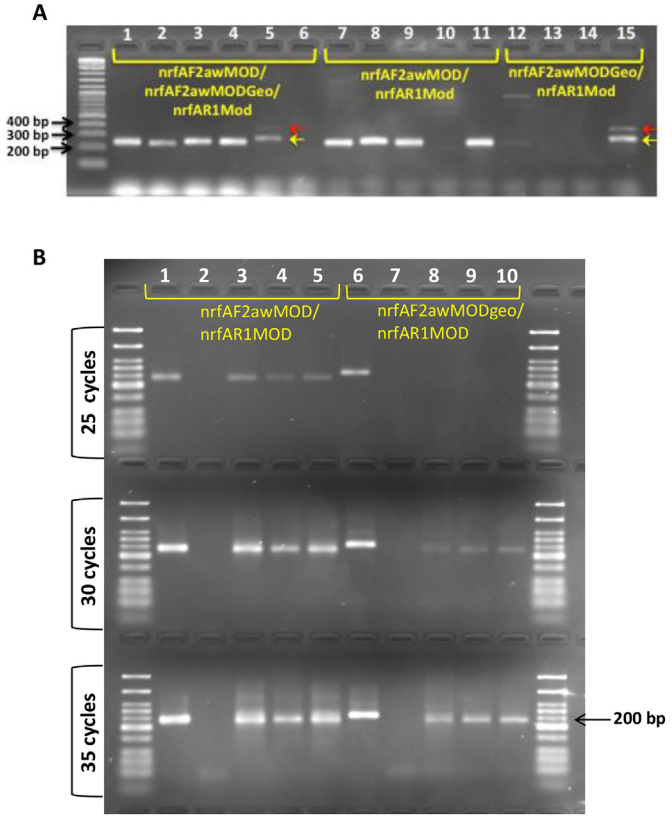

Fig. 4. (A) Gel electrophoresis image of PCR products amplified from reference DNA showing specificity of primer pairs in Lanes 1–6: combined forward primers nrfAF2awMOD and nrfAF2awMODgeo paired with the reverse primer nrfAR1MOD; Lane 7–11: forward primer nrfAF2awMOD paired with reverse primer nrfAR1MOD; and Lanes 12–15: forward primer nrfAF2awMODgeo paired with reverse primer nrfAR1MOD. Lanes 1,11: seven pooled reference DNA, Lanes 2, 7, 12: *S. fonticola* HAc-5, Lanes 3, 8, 13: *S. oneidensis* MR-1, Lanes 4, 9, 14: *A. dehalogenans* 2CP-1, Lanes 5, 10, 15: *G. bemidjiensis* Bem, Lane 6: ultrapure water. Lower yellow arrow indicates the expected *nrfA* amplicon product. Upper red arrow shows larger amplicon with *G. bemidjiensis* attributed to ~360 bp product that is not associated with *nrfA*. (B) Gel electrophoresis of *nrfA* amplicons generated with primer sets nrfAF2awMOD/nrfAR1MOD (Lanes 1–5) and nrfAF2awMODgeo/nrfAR1MOD (Lanes 6–10) comparing 25, 30, and 35 cycles of PCR with reference controls and soil DNA as templates. Lane 1: *A. dehalogenans* 2CP-1, Lanes 2 & 7: ultrapure water, Lanes 3 & 8: soil DNA Hw 4/2012, Lanes 4 & 9: soil DNA UM 4/2012, Lanes 5 & 10: soil DNA HW 11/2012, Lane 6: *G. bemidjiensis.* All PCR reactions were performed with annealing temperature 56 °C. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

environmental studies that used the previous primer set nrfAF2aw/nrfAR1 employed upwards of 50 cycles of amplification (Bu et al., 2017; Pandey et al., 2018; Song et al., 2014), presumed out of necessity due to low copy numbers of the *nrfA* gene in contrast to 16S rRNA genes. Despite the theoretical and methodological issues of using excessive cycles for PCR (> 35 cycles), it has been common practice in *nrfA* studies to do just that (Lam et al., 2009; Lindemann et al., 2016; Mohan et al., 2004; Smith et al., 2007; Takeuchi, 2006). This less than optimal approach leads to potential consequences for non-target amplicon formation that can be further compounded by insufficient coverage of gene targets by the primers. For example, excessive cycles during PCR can diminish polymerase fidelity, lower polymerase processivity, and propagate non-specific products or point mutations (Qiu et al., 2001; Sipos et al., 2007; Wu et al., 2010). We observe larger non-target products when amplifying *nrfA* from soil DNA even when the number of PCR cycles was increased to just 35 cycles (Fig. 4B). This highlights the importance of considering the fundamental principles of PCR. Apart from the more conventional factors that can affect PCR, additional application to multi-template PCR involves additional considerations that could also cause undue influences on the production of artifacts and biases (for review, see (Kalle et al., 2014)). These include minimizing bias in detection of different target gene homologs that may span a wide magnitude range in environmental samples and reducing the negative effects on amplification efficiency due to co-extracted PCR inhibitors present in the environmental matrix. As the number of cycles above the optimum increases, the negative impacts, such as the increased number of non-target sequences or amplification artifacts that do not represent *nrfA*, will likely contribute more to the resulting amplicon pool. This would clearly lead to misinterpretation of any PCR-based analyses, particularly when applied to SYBR green-based qPCR, which would lead to overestimates of *nrfA* gene abundance. In using the primers from this study, the need for > 35 cycles of PCR to detect amplicons from an environmental sample would most likely yield a pool primarily resulting in non-target artifacts. Because the new primer sets here yield more efficient amplification of *nrfA*, we do not recommend their use with > 35 cycles.

Our design process allowed both the forward and reverse primer sequences to be modified by increasing their lengths and changing the base composition to match commonly specified PCR standards. For example, due to the increased primer length, the melting temperatures for the modified primer set nrfAF2awMOD and nrfAR1MOD are now higher (e.g. ~56 °C) but still allow broad coverage in 16 of the 18 NrfA clades (Fig. 1). The changes in primer length and base composition allow diverse *nrfA* to be effectively targeted using a single optimized PCR annealing temperature of 56 °C. The additional degeneracies incorporated into the new primers extend coverage for highly variable bases among homologous *nrfA* genes beyond that covered with the previously published primer set. In addition, substituting inosine (I) for "N" bases at the wobble positions in both the forward and reverse primer sequences reduces the overall degree of degeneracy while facilitating more universality in PCR reaction conditions (Ben-Dov and Kushmaro, 2015). When considering both the amplicon sizes generated and the fixed annealing temperature for PCR with the new primer sets, the benefit further extends their compatibility for use with fixed-temperature thermocycler platforms in multiplex array technology (e.g. Fluidgm Access Array™). These platforms allow multiple target amplifications and are designed for generating products to use in downstream amplicon-based sequencing or gene quantification with a large number of samples.

The use of the nrfAF2awMOD/nrfAR1MOD primer pair enables broad detection of all clades except Clades E and I. With two mismatches allowed, however, the coverage improves significantly in Clade E (see Table 1 in (Cannon et al., 2019)). Seven of the 18 clades show 100% coverage of the reference sequences evaluated in silico. The amino acid translation of the amplicon sequence provides a means to observe motifs diagnostic for a functional NrfA between the third and fourth heme-binding sites. Specifically this region contains a KXRH or KXQH motif in all NrfA proteins (Welsh et al., 2014). To further verify the specificity of the new broad coverage primer pair nrfAF2awMOD/nrfAR1MOD and its applicability to multiplex PCR, the primer set was included in a Fluidigm access array platform to generate amplicons followed by Illumina sequencing. We confirmed the recovery of several clades of NrfA from representative amplicons generated from groundwater and soil DNA samples using a Fluidigm access array system (see Fig. 6 and Fig. 7 in (Cannon et al., 2019)). In this sequencing effort, multiple and diverse sequences were found to be associated with the different clades recovered (see Fig. 6 in (Cannon et al., 2019)), with > 99% of the retrieved sequences corresponding to verified NrfA based on a translated alignment of the amplified sequences showing the expected conserved amino acid residues between the 3rd and 4th heme-binding domains (Welsh et al., 2014). Most of the sequences obtained from soil DNA corresponded to NrfA belonging to Clades J and K, while the representative sequences from groundwater DNA contained diverse

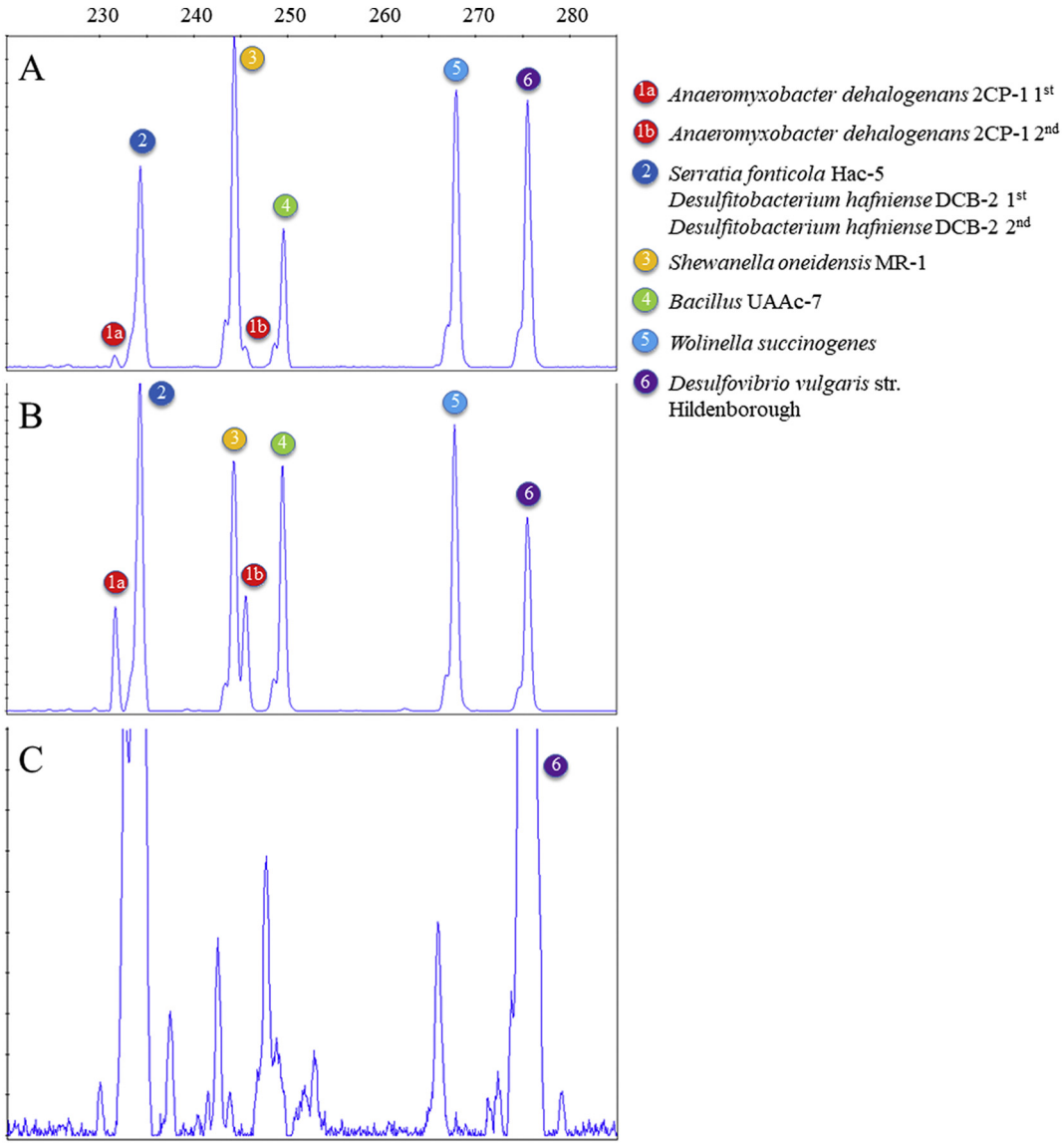

Fig. 5. Amplified fragment length polymorphism (AFLP) profiles following PCR amplification of (A) DNA pooled from reference organisms *A. dehalogenans* 2CP-1 (1st copy Clade K, 2nd copy Clade J), *S. fonticola* HAc-5 (Clade A), *D. hafniense* DCB-2 (Clades M&N), *S. oneidensis* MR-1 (Clade C), *Bacillus* UAAc7 (Clade P), *W. succinogenes* (Clade F), and *D. vulgaris* (Clade O) compared to (B) PCR products of reference organisms individually amplified and pooled post-PCR. The AFLP profile for (C) soil sample UM2 4/2012 shows amplicons within the range of expected sizes for diverse *nrfA* along with the expected product of *D. vulgaris* added as an internal reference. All reactions were performed using primer pair 6-FAM- nrfAF2awMOD/nrfAR1MOD.

NrfA belonging to Clades G, H, K, L, and M. Remarkably, the sequencing results showed very low occurrence of non-*nrfA* gene sequences (< 5%), which we have observed with other functional gene amplicon sequencing like *nosZ* (data not shown).

Since the forward primer nrfAF2awMOD lacked consensus with all members of Clade I, the primer pair nrfAF2awMODgeo and nrfAR1MOD was newly designed to cover this clade (Fig. 1, Fig. 2; see Fig. 2 in (Cannon et al., 2019)). Clade I members include the family Geobacteraceae, known for their abundance in many environments and importance in metal cycling (Aklujkar et al., 2009; Berthrong et al., 2014; Bu et al., 2017; Roling, 2014; van den Berg et al., 2016). A recent collective survey of 365 soil metagenomes showed a high proportion of *Geobacter* with the trait for DNRA, suggesting their potential significance in soil biogeochemical N-cycling (Nelson et al., 2016). The primer set nrfAF2awMODgeo and nrfAR1MOD also extends coverage of *nrfA* to allow detection of additional diversity within other clades that are missed by the broad range primer set, e.g. Clade N (see Fig. 3 in (Cannon et al., 2019)). Other reference genes in Clades D and R are also now detected along with one soil metagenomic-derived sequence (e.g. Havana Deep 038 Clade K). These *nrfA* would otherwise be missed using just the broad-coverage primer set (see Fig. 2 in (Cannon et al., 2019)). The design approach we used makes it possible to combine both new forward primers with the reverse primer in a single PCR reaction that would then provide comprehensive coverage that includes nearly all the designated *nrfA* clades.

The application of AFLP using fluorescent-labeled forward primers provides a means to evaluate amplification differences in a pool of different reference genomes containing *nrfA* when combined or mixed together after individual PCR reactions. It also provides a method to obtain a community fingerprint of *nrfA* genes amplified from DNA extracted from environmental samples (Fig. 5). This approach successfully takes advantage of predicted size differences among the amplicons generated from *nrfA* of disparate taxa (Table 2). Recovery of all the expected sequences in a defined pool of references were seen in both AFLP profiles (Fig. 5; see Fig. 4 in (Cannon et al., 2019)), however the results suggest lower amplification efficiency for DNA (i.e. relative peak

height) from some reference genomic DNA when in the presence of competing templates compared correspondingly to individual reactions. This lower efficiency occurs in organisms like *A. dehalogenans* with high genomic G + C content (74.9%), a feature that is known to impact amplification. As such, caution should be used when attempting to report relative phylogenetic distribution based on the size of the AFLP peaks as an accurate representation of *nrfA* community structure. It is important to note this effect of %G + C content would also be applicable to amplicon sequencing and the common practice of reporting relative number of unique sequences (OTUs) retrieved. Additional optimization of PCR thermocycler conditions could be evaluated with these new primer sets to address this issue. For example, increasing the initial denaturing conditions and extending subsequent melt cycles has been shown to significantly reduce amplification bias associated with high %G + C content targets (Aird et al., 2011; Head et al., 2014). We also observed an effect of %G + C content during AFLP analysis where fragment migration differed between *nrfA* amplicons with high G + C content *A. dehalogenans* (Clade K) and lower G + C content *S. fonticola* (Clade A) and *D. hafniense* (Clades M & N) that should yield a single collective peak due to their common amplicon size (236 bp) (Table 2, Fig. 5A & 5B). While AFLP does not afford concomitant sequence information, it is worthwhile to note that such a profiling approach can be a useful tool for comparative analysis to assess changes in a large proportion of the extant *nrfA* DNA (and cDNA) community structure and can potentially inform sample selection strategies for further amplicon-based sequencing efforts. Taken together with the high specificity to *nrfA* demonstrated for the primer sets and amplicon size resolution capabilities offered by fragment analysis and considering the caveats above, AFLP further provides a useful tool for *nrfA* community comparisons.

5. Conclusions

Specific functional traits and functional genes, rather than phylogenetic markers, are thought to serve as better proxies for studying the relationship between microbial diversity and function (Bernhard and Kelly, 2016). The use of PCR methods remains a vital tool for detection of functional genes in natural environments over a wide dynamic range of phylotypes and abundances. Updating the design of primer sets that are informed continuously with available new sequences from genomes and environmental studies ensures better coverage of diversity and selection of molecular tools more pertinent to a particular environment. With one mismatch allowed per primer, the primer sets designed in this study provide 100% coverage of NrfA diversity in six of the eighteen clades (Clades A, B, C, G, J and P) and extensive coverage of members within an additional ten clades that now includes the environmentally relevant Clade I (i.e. Geobacteraceae). When two mismatches are considered, the coverage extends to 17 of 18 clades (all except Clade D) that together target representative taxa common to many terrestrial, aquatic, and gut environments. Interestingly, NrfA sequenced from Clades J and K appeared to dominate in soils (unpublished data) using both the previously published *nrfA* primers set nrfAF2aw/nrfAR1 (Welsh et al., 2014) and the new ones reported here. This is further corroborated in metagenome-assembled *nrfA* contigs and genomes from soil libraries (Orellana et al., 2017), demonstrating the utility of the new primer set designed here that provides nearly full coverage for these clades (Fig. 1).

The new primer sets here are better optimized for single annealing temperature conditions requisite in multiplex platforms that work in parallel with other PCR primers to generate DNA pools of target genes that can be subsequently sequenced and translated into protein fragments (see Fig. 7 in (Cannon et al., 2019)). The new reverse primer nrfAR1MOD can also be anticipated for better use to generate a targeted cDNA pool in a single reverse transcription reaction, allowing for potentially more optimized *nrfA* transcript studies. The PCR primers are also suitable candidates for use in conventional- and high throughput qPCR platforms due to their improved efficiency for amplification of *nrfA* and product sizes. The emergent diversity of *nrfA* as revealed by genomics and a growing database of sequences retrieved from natural systems challenges existing paradigms for microbial populations previously thought to dominate nitrate and nitrite reduction in natural environments, and more recently, also highlights the role of nitrite ammonifiers as significant sources of N_2O emissions in terrestrial systems (Mania et al., 2014; Rutting et al., 2011; Stremińska et al., 2012). The new primers here should benefit studies that continue to elucidate the potential prominence of nitrite ammonifiers in N-cycling and provide new insight into patterns of soil N retention and loss in a variety of terrestrial and aquatic environments.

Acknowledgments

This work was funded in part by the U.S. National Science Foundation grant DEB-1656027 to WHY and RS, and by the U.S. Department of Agriculture NIFA grant 2016-67030-25211 to WHY. We wish to thank Alex Krichels for soil DNA samples used for sequencing analysis. Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

References

Aird, D., Ross, M.G., Chen, W.S., Danielsson, M., Fennell, T., Russ, C., Jaffe, D.B., Nusbaum, C., Gnirke, A., 2011. Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. Genome Biol. 12, R18.

Aklujkar, M., Krushkal, J., DiBartolo, G., Lapidus, A., Land, M.L., Lovley, D.R., 2009. The genome sequence of *Geobacter metallireducens*: features of metabolism, physiology and regulation common and dissimilar to Geobacter sulfurreducens. BMC Microbiol. 9, 109.

Baggs, E., 2011. Soil microbial sources of nitrous oxide: recent advances in knowledge, emerging challenges and future direction. Curr. Opin. Environ. Sustain. 3, 321–327.

Ben-Dov, E., Kushmaro, A., 2015. Inosine at different primer positions to study structure and diversity of prokaryotic populations. Curr. Issues. Mol. Biol. 17, 53–56.

Bernard, R.J., Mortazavi, B., Kleinhuizen, A.A., 2015. Dissimilatory nitrate reduction to ammonium (DNRA) seasonally dominates NO_3^- reduction pathways in an anthropogenically impacted sub-tropical coastal lagoon. Biogeochemistry. 125, 47–64.

Bernhard, A., Kelly, J., 2016. Editorial: linking ecosystem function to microbial diversity. Front. Microbiol. 7.

Berthrong, S.T., Yeager, C.M., Gallegos-Graves, L., Steven, B., Eichorst, S.A., Jackson, R.B., Kuske, C.R., 2014. Nitrogen fertilization has a stronger effect on soil nitrogen-fixing bacterial communities than elevated atmospheric CO2. Appl. Environ. Microbiol. 80, 3103–3112.

Bhowmik, A., Cloutier, M., Ball, E., Bruns, M.A., 2017. Underexplored microbial metabolisms for enhanced nutrient recycling in agricultural soils. AIMS Microbiol. 3, 826–845.

Bonilla-Rosso, G., Wittorf, L., Jones, C., Hallin, S., 2016. Design and evaluation of primers targeting genes encoding NO-forming nitrite reductases: implications for ecological inference of denitrifying communities. Sci. Rep. 6.

Bu, C., Wang, Y., Ge, C., Ahmad, H.A., Gao, B., Ni, S.Q., 2017. Dissimilatory nitrate reduction to ammonium in the yellow river estuary: rates, abundance, and community diversity. Sci. Rep. 7, 6830.

Cannon, J., Sanford, R., Connor, L., Yang, W., Chee-Sanford, J., 2019. Sequence alignments and validation of PCR primers used to detect phylogenetically diverse *nrfA* genes associated with dissimilatory nitrate reduction to ammonium (DNRA). Data in Brief (submitted).

Chen, Z., Ding, W., Xu, Y., Muller, C., Rutting, T., Yu, H., Fan, J., Zhang, J., Zhu, T., 2015a. Importance of heterotrophic nitrification and dissimilatory nitrate reduction to ammonium in a cropland soil: evidences from a N-15 tracing study to literature synthesis. Soil Biol. Biochem. 91, 65–75.

Chen, Z., Wang, C., Gschwendtner, S., Willibald, G., Unteregelsbacher, S., Lu, H., Kolar, A., Schloter, M., Butterbach-Bahl, K., Dannenmann, M., 2015b. Relationships between denitrification gene expression, dissimilatory nitrate reduction to ammonium and nitrous oxide and dinitrogen production in montane grassland soils. Soil Biol. Biochem. 87, 67–77.

Crooks, G.E., Hon, G., Chandonia, J.M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190.

Decleyre, H., Heylen, K., Van Colen, C., Willems, A., 2015. Dissimilatory nitrogen reduction in intertidal sediments of a temperate estuary: small scale heterogeneity and novel nitrate-to-ammonium reducers. Front. Microbiol. 6, 1124.

Dieffenbach, C.W., Lowe, T.M., Dveksler, G.S., 1993. General concepts for PCR primer design. PCR Methods Appl. 3, S30–S37.

Einsle, O., Klotz, M., Stein, L., 2011. Structure and function of formate-dependent cytochrome c nitrite reductase. In: NrfA. Methods in Enzymology, Vol 46: Research on Nitrification and Related Processes, Pt B. vol. 496. pp. 399–422.

Gao, D., Li, X., Lin, X., Wu, D., Jin, B., Huang, Y., Liu, M., Chen, X., 2017. Soil dissimilatory nitrate reduction processes in the Spartina alterniflora invasion chronosequences of a coastal wetland of southeastern China: dynamics and environmental implications. Plant Soil 421, 383–399.
Gardner, W.S., McCarthy, M.J., An, S., Sobolev, D., Sell, K.S., Brock, D., 2006. Nitrogen fixation and dissimilatory nitrate reduction to ammonium (DNRA) support nitrogen dynamics in Texas estuaries. Limnol. Oceanogr. 51, 558–568.
Giacomucci, L., Purdy, K.J., Zanardini, E., Polo, A., Cappitelli, F., 2012. A new non-degenerate primer pair for the specific detection of the nitrite reductase gene nrfA in the genus *Desulfovibrio*. J. Mol. Microbiol. Biotechnol. 22, 345–351.
Giblin, A., Tobias, C., Song, B., Weston, N., Banta, G., Rivera-Monroy, V., 2013. The importance of dissimilatory nitrate reduction to ammonium (DNRA) in the nitrogen cycle of coastal ecosystems. Oceanography 26, 124–131.
Griffiths, R.I., Whiteley, A.S., O'Donnell, A.G., Bailey, M.J., 2000. Rapid method for co-extraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. Appl. Environ. Microbiol. 66, 5488–5491.
Head, S.R., Komori, H.K., LaMere, S.A., Whisenant, T., Van Nieuwerburgh, F., Salomon, D.R., Ordoukhanian, P., 2014. Library construction for next-generation sequencing: overviews and challenges. Biotechniques 56 61–64, 66, 68, passim.
Jahangir, M.M.R., Fenton, O., Muller, C., Harrington, R., Johnston, P., Richards, K.G., 2017. In situ denitrification and DNRA rates in groundwater beneath an integrated constructed wetland. Water Res. 111, 254–264.
Jensen, M.M., Lam, P., Revsbech, N.P., Nagel, B., Gaye, B., Jetten, M.S., Kuypers, M.M., 2011. Intensive nitrogen loss over the Omani shelf due to anammox coupled with dissimilatory nitrite reduction to ammonium. Isme J. 5, 1660–1670.
Jones, C.M., Graf, D.R.H., 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.
Jones, C., Spor, A., Brennan, F., Breuil, M., Bru, D., Lemanceau, P., Griffiths, B., Hallin, S., Philippot, L., 2014. Recently identified microbial guild mediates soil N_2O sink capacity. Nat. Clim. Chang. 4, 801–805.
Kalle, E., Kubista, M., Rensing, C., 2014. Multi-template polymerase chain reaction. Biomol. Detect Quantif. 2, 11–29.
Koop-Jakobsen, K., Giblin, A.E., 2010. The effect of increased nitrate loading on nitrate reduction via denitrification and DNRA in salt marsh sediments. Limnol. Oceanogr. 55, 789–802.
Kuypers, M.M.M., Marchant, H.K., Kartal, B., 2018. The microbial nitrogen-cycling network. Nat. Rev. Microbiol. 16, 263–276.
Lam, P., Lavik, G., Jensen, M.M., van de Vossenberg, J., Schmid, M., Woebken, D., Gutiérrez, D., Amann, R., Jetten, M.S., Kuypers, M.M., 2009. Revising the nitrogen cycle in the Peruvian oxygen minimum zone. Proc. Natl. Acad. Sci. U. S. A. 106, 4752–4757.
Lindemann, S., Zarnoch, C., Castignetti, D., Hoellein, T., 2016. Effect of eastern oysters (Crassostrea virginica) and seasonality on nitrite reductase gene abundance (*nirS, nirK, nrfA*) in an urban estuary. Estuar. Coasts 39, 218–232.
Mania, D., Heylen, K., van Spanning, R., Frostegard, A., 2014. The nitrate-ammonifying and nosZ-carrying bacterium *Bacillus vireti* is a potent source and sink for nitric and nitrous oxide under high nitrate conditions. Environ. Microbiol. 16, 3196–3210.
Minick, K.J., Pandey, C.B., Fox, T.R., Subedi, S., 2016. Dissimilatory nitrate reduction to ammonium and N_2O flux: effect of soil redox potential and N fertilization in loblolly pine forests. Biol. Fertil. Soils 52, 601–614.
Mohan, S., Schmid, M., Jetten, M., Cole, J., 2004. Detection and widespread distribution of the *nrfA* gene encoding nitrite reduction to ammonia, a short circuit in the biological nitrogen cycle that competes with denitrification. FEMS Microbiol. Ecol. 49, 433–443.
Nelson, M., Martiny, A., Martiny, J., 2016. Global biogeography of microbial nitrogen-cycling traits in soil. Proc. Natl. Acad. Sci. U. S. A. 113, 8033–8040.
Orellana, L., Rodriguez-R, L., Higgins, S., Chee-Sanford, J., Sanford, R., Ritalahti, K., Loffler, F., Konstantinidis, K., 2014. Detecting nitrous oxide reductase (nosZ) genes in soil metagenomes: method development and implications for the nitrogen cycle. Mbio. 5.
Orellana, L.H., Chee-Sanford, J.C., Sanford, R.A., Loffler, F.E., Konstantinidis, K.T., 2017. Year-round shotgun metagenomes reveal stable microbial communities in agricultural soils and novel ammonia oxidizers responding to fertilization. Appl. Environ. Microbiol. 84.
Pandey, A., Suter, H., He, J.Z., Hu, H.W., Chen, D., 2018. Nitrogen addition decreases dissimilatory nitrate reduction to ammonium in rice paddies. Appl. Environ. Microbiol. 84.
Qiu, X., Wu, L., Huang, H., McDonel, P.E., Palumbo, A.V., Tiedje, J.M., Zhou, J., 2001. Evaluation of PCR-generated chimeras, mutations, and heteroduplexes with 16S rRNA gene-based cloning. Appl. Environ. Microbiol. 67, 880–887.
Roberts, K.L., Kessler, A.J., Grace, M.R., Cook, P.L.M., 2014. Increased rates of dissimilatory nitrate reduction to ammonium (DNRA) under oxic conditions in a periodically hypoxic estuary. Geochim. Cosmochim. Acta 133, 313–324.
Roling, W.F.M., 2014. The family Geobacteraceae. In: Rosenberg, E., Delong, E.F., Lory, S., Stackebrand, E., Thompson, F. (Eds.), *The Prokaryotes - Deltaproteobacteria* and *Epsilonproteobacteria*. Springer-Verlag, Berlin Heidelberg, pp. 157–172.
Rutting, T., Muller, C., 2008. Process-specific analysis of nitrite dynamics in a permanent grassland soil by using a Monte Carlo sampling technique. Eur. J. Soil Sci. 59, 208–215.
Rutting, T., Boeckx, P., Muller, C., Klemedtsson, L., 2011. Assessment of the importance of dissimilatory nitrate reduction to ammonium for the terrestrial nitrogen cycle. Biogeosciences. 8, 1779–1791.
Sanford, R., Wagner, D., Wu, Q., Chee-Sanford, J., Thomas, S., Cruz-Garcia, C., Rodriguez, G., Massol-Deya, A., Krishnani, K., Ritalahti, K., Nissen, S., Konstantinidis, K., Loffler, F., 2012. Unexpected nondenitrifier nitrous oxide reductase gene diversity and abundance in soils. Proc. Natl. Acad. Sci. U. S. A. 109, 19709–19714.
SantaLucia Jr., J., 1998. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc. Natl. Acad. Sci. U. S. A. 95, 1460–1465.
Shan, J., Zhao, X., Sheng, R., Xia, Y., Ti, C., Quan, X., Wang, S., Wei, W., Yan, X., 2016. Dissimilatory nitrate reduction processes in typical Chinese Paddy soils: rates, relative contributions, and influencing factors. Environ. Sci. Technol. 50, 9972–9980.
Silver, W., Herman, D., Firestone, M., 2001. Dissimilatory nitrate reduction to ammonium in upland tropical forest soils. Ecology 82, 2410–2416.
Silver, W., Thompson, A., Reich, A., Ewel, J., Firestone, M., 2005. Nitrogen cycling in tropical plantation forests: potential controls on nitrogen retention. Ecol. Appl. 15, 1604–1614.
Simon, J., Klotz, M., 2013. Diversity and evolution of bioenergetic systems involved in microbial nitrogen compound transformations. BBA-Bioenergetics 1827, 114–135.
Sipos, R., Szekely, A.J., Palatinszky, M., Revesz, S., Marialigeti, K., Nikolausz, M., 2007. Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targetting bacterial community analysis. FEMS Microbiol. Ecol. 60, 341–350.
Smith, C., Nedwell, D., Dong, L., Osborn, A., 2007. Diversity and abundance of nitrate reductase genes (*narG* and *napA*), nitrite reductase genes (*nirS* and *nrfA*), and their transcripts in estuarine sediments. Appl. Environ. Microbiol. 73, 3612–3622.
Smith, C., Dong, L., Wilson, J., Stott, A., Osborn, A., Nedwell, D., 2015. Seasonal variation in denitrification and dissimilatory nitrate reduction to ammonia process rates and corresponding key functional genes along an estuarine nitrate gradient. Front. Microbiol. 6.
Song, B., Lisa, J., Tobias, C., 2014. Linking DNRA community structure and activity in a shallow lagoonal estuarine system. Front. Microbiol. 5.
Sotta, E.D., Corre, M.D., Veldkamp, E., 2008. Differing N status and N retention processes of soils under old-growth lowland forest in eastern Amazonia, Caxiuanã, Brazil. Soil Biol. Biochem. 40, 740–750.
Stein, L., Klotz, M., 2011. Surveying N_2O-producing pathways in bacteria. Methods in enzymology: research on nitrification and related processes, Vol 486. Part A 486, 131–152.
Stein, L., Klotz, M., 2016. The nitrogen cycle. Curr. Biol. 26, R94–R98.
Stief, P., Lundgaard, A.S.B., Treusch, A.H., Thamdrup, B., Grossart, H.P., Glud, R.N., 2018. Freshwater copepod carcasses as pelagic microsites of dissimilatory nitrate reduction to ammonium. FEMS Microbiol. Ecol. 94.
Stremińska, M.A., Felgate, H., Rowley, G., Richardson, D.J., Baggs, E.M., 2012. Nitrous oxide production in soil isolates of nitrate-ammonifying bacteria. Environ. Microbiol. Rep. 4, 66–71.
Su, W., Zhang, L., Li, D., Zhan, G., Qian, J., Tao, Y., 2012. Dissimilatory nitrate reduction by *Pseudomonas alcaliphila* with an electrode as the sole electron donor. Biotechnol. Bioeng. 109, 2904–2910.
Takeuchi, J., 2006. Habitat segregation of a functional gene encoding nitrate ammonification in estuarine sediments. Geomicrobiol J. 23, 75–87.
Tatti, E., Goyer, C., Zebarth, B., Wertz, S., Burton, D., Chantigny, M., Filion, M., Zeng, J., 2017. Over-winter dynamics of soil bacterial denitrifiers and nitrite ammonifiers influenced by crop residues with different carbon to nitrogen ratios. Appl. Soil Ecol. 110, 53–64.
Templer, P., Silver, W., Pett-Ridge, J., DeAngelis, K., Firestone, M., 2008. Plant and microbial controls on nitrogen retention and loss in a humid tropical forest. Ecology 89, 3030–3040.
Tiedje, J.M., 1988. Ecology of Denitrification and Dissimilatory Nitrate Reduction to Ammonium. Biology of Anaerobic Microorganisms. John Wiley and Sons, New York, NY, pp. 179–244.
Tsai, Y., Olson, B., 1991. Rapid method for direct extraction of DNA from soil and sediments. Appl. Environ. Microbiol. 57, 1070–1074.
van den Berg, E., Boleij, M., Kuenen, J., Kleerebezem, R., van Loosdrecht, M., 2016. DNRA and denitrification coexist over a broad range of acetate/N-NO_3^- ratios, in a chemostat enrichment culture. Front. Microbiol. 7.
van den Berg, E., Rombouts, J., Kuenen, J., Kleerebezem, R., van Loosdrecht, M., 2017. Role of nitrite in the competition between denitrification and DNRA in a chemostat enrichment culture. AMB Express 7.
Welsh, A., Chee-Sanford, J.C., Connor, L.M., Löffler, F.E., Sanford, R.A., 2014. Refined NrfA phylogeny improves PCR-based nrfA gene detection. Appl. Environ. Microbiol. 80, 2110–2119.
Wu, J.Y., Jiang, X.T., Jiang, Y.X., Lu, S.Y., Zou, F., Zhou, H.W., 2010. Effects of polymerase, template dilution and cycle number on PCR based 16 S rRNA diversity analysis using the deep sequencing method. BMC Microbiol. 10, 255.
Yang, W.H., Ryals, R.A., Cusack, D.F., Silver, W.L., 2017. Cross-biome assessment of gross soil nitrogen cycling in California ecosystems. Soil Biol. Biochem. 107, 144–155.
Yin, S., Chen, D., Chen, L., Edis, R., 2002. Dissimilatory nitrate reduction to ammonium and responsible microorganisms in two Chinese and Australian paddy soils. Soil Biol. Biochem. 34, 1131–1137.
Yoon, S., Sanford, R., Loffler, F., 2015. Nitrite control over dissimilatory nitrate/nitrite reduction pathways in *Shewanella loihica* strain PV-4. Appl. Environ. Microbiol. 81, 3510–3517.
Yu, T.T., Li, M., Niu, M.Y., Fan, X.B., Liang, W.Y., Wang, F.P., 2018. Difference of nitrogen-cycling microbes between shallow bay and deep-sea sediments in the South China Sea. Appl. Microbiol. Biotechnol. 102, 447–459.
Zhang, J., Lan, T., Muller, C., Cai, Z., 2015. Dissimilatory nitrate reduction to ammonium (DNRA) plays an important role in soil nitrogen conservation in neutral and alkaline but not acidic rice soil. J. Soils Sediments 15, 523–531.