Four PCR primers necessary for the detection of periplasmic nitrate reductase genes in all groups of Proteobacteria and in environmental DNA

Tobias Klatte, Laura Evans, Rebekah N. Whitehead and Jeffrey A. Cole

School of Biosciences, University of Birmingham, Birmingham B15 2TT, U.K.

Abstract

Generic primers are available for detecting bacterial genes required for almost every reaction of the biological nitrogen cycle, the one notable exception being napA (gene for the molybdoprotein of the periplasmic nitrate reductase) encoding periplasmic nitrate reductases. Using an iterative approach, we report the first successful design of three forward oligonucleotide primers and one reverse primer that, in three separate PCRs, can amplify napA DNA from all five groups of Proteobacteria. All 140 napA sequences currently listed in the NCBI (National Center for Biotechnology Information) database are predicted to be amplified by one or more of these primer pairs. We demonstrate that two pairs of these primers also amplify PCR products of the predicted sizes from DNA isolated from human faeces, confirming their ability to direct the amplification of napA fragments from mixed populations. Analysis of the resulting amplicons by high-throughput sequencing will enable a good estimate to be made of both the range and relative abundance of nitrate-reducing bacteria in any community, subject only to any unavoidable bias inherent in a PCR approach to molecular characterization of a highly diverse target.
wider range of α-, β- and γ-proteobacteria and have used them to estimate the balance between different groups of nitrate-reducing bacteria in community DNA from 18 different environments [17,18]. However, none of the primers currently available are suitable for detecting all known napA sequences represented in the DNA sequence databases. This review describes a strategy to overcome this limitation, its successful application and its limitations.

A standardized PCR protocol to detect periplasmic nitrate gene products

The PCR protocol was developed using as templates chromosomal DNA isolated from E. coli K-12, D. desulfuricans 27774, Campylobacter jejuni, Haemophilus influenzae, Habella chejuensis, Wolinella succinogenes, Campylobacter lari, Campylobacter hominis, Campylobacter curvus, Campylobacter concisus, Bortetella bronchiseptica, Bradyrhizobium japonicum and Paracoccus pantotrophus. DNA sequences of species that were known to contain the napA gene were downloaded from the NCBI (National Center for Biotechnology Information) database. They were aligned with the ClustalX tool and displayed by the BoxShade tool in order to detect conserved regions. Degenerate primers were designed based on these alignments and produced by Alta Biosciences.

After many preliminary experiments that included trials with hot start, touchdown and gradient PCR, a standardized touchdown protocol was developed that included a 5 min denaturation step at 95°C followed by 32 touchdown cycles in which the annealing temperature was decreased by 0.5°C per cycle from 55°C down to 39°C. Samples were held at the annealing temperature for 1 min. In each subsequent cycle, the sample was denatured for 30 s at 95°C and extended for 150 s at 72°C. The amount of product was amplified in 30 steps in which the denaturation and extension steps were exactly as in the touchdown cycles, but primers were annealed for 1 min at 39°C. After a final 10 min extension step at 72°C the sample was stored at 4°C. Products were checked by electrophoresis in a 100 ml 1% agarose gel.

Limited applicability of previously described primers

In preliminary experiments, various primers described in the literature or developed in-house were tested for their ability to direct the synthesis of a limited range of napA target genes. The design of primers V16 and V17 was based on the known sequences of napA genes from a limited number of bacteria that included E. coli and P. denitrificans [15] (Figure 1). The ability of these primers to recognize these templates, but not napA from D. desulfuricans strain 27774, was confirmed. Similar results were obtained with related primers described by Bru et al. [17]. Conversely, primers DdV16 and DdV67, which were used in the determination of the sequence of D. desulfuricans nap operon [16], were unable to recognize E. coli or P. denitrificans templates.

Design of primers that recognize both D. desulfuricans and E. coli napA templates

We then attempted to design universal primers that would amplify napA DNA from all groups of proteobacteria by comparing the sequences of highly divergent napA genes from 3α-, 3β-, 6γ-, 25- and 46-proteobacteria (Figure 1). Although very few regions of sequence identity were found, two primers resulting from this analysis, 1173b and R2, were predicted to amplify napA fragments from both E. coli and D. desulfuricans. This prediction was confirmed. However, none of the primer combinations available recognized napA genes in C. jejuni or C. lari.

Design of primers that recognize napA templates from all groups of proteobacteria

Further analysis revealed that this limited range of sequences could be split into two subgroups within which were more regions of sequence similarity. Two pairs of degenerate primers based on these regions of similarity were designed. Primers SF1173d and SR2294 successfully amplified the expected napA fragments of both E. coli and D. desulfuricans DNA, but failed to recognize C. jejuni DNA (Figures 1 and 2 and Table 1). Primers LF716 and LR1837 amplified napA fragments from both E. coli and C. jejuni DNA, but not from D. desulfuricans DNA. When primer SF1173d was combined with LR1837, the expected 0.5 kb fragments were generated, but no product was obtained with D. desulfuricans DNA as template (Figure 2). The primer pair LF716 and SR2294 generated the expected 1.5 and 1.8 kb products with E. coli and C. jejuni DNA templates (Figure 2). However, not every expected product was generated, and some templates directed the synthesis of more than one product.

The sequence analysis was then extended to all 140 napA sequences from 70 bacterial species in the NCBI database in February 2010. Each of the four primers was checked against the database sequences to determine whether there was a mismatch in the last six bases at the 3’-end of each primer sequence that might prevent effective amplification of the required fragment (Table 1). This analysis revealed only three sequences, those from Geobacter lovleyi, C. lari, and C. hominis, that failed to meet the stringent criteria set with the forward primers SF1173d and LF716. One of the reverse primers, SR2294, was predicted to recognize every napA sequence in the database, and was therefore potentially truly generic. Furthermore, this primer allowed problems predicted when using primer LR1837 to be avoided (Table 1).

In a final attempt to design a universal forward primer, alignments were first generated for all five subdivisions of proteobacteria, with a further alignment of sequences that posed the greatest problems. This resulted in the design of
Figure 1 | Oligonucleotide primers used to detect napA genes in all groups of proteobacteria

The top five segments of the Figure show the sequences of five of the primers used to develop the generic protocol for the detection of napA in various groups of proteobacteria above the typical sequences they were designed to recognize. The lower diagrams show where within the napA genes of Escherichia coli, D. desulfuricans and C. jejuni the various primers mentioned in the text are predicted to bind.

Table 1 | Overview of napA sequences that might not be amplified using each of the four primers

| Class/primer | LF716 | LR1837 | SF1173d | SR2294 |
|--------------|-------|--------|---------|--------|
| α-Proteobacteria (eight sequences) | – | B. japonicum, Rhodobacterales bacterium, Sinorhizobium meliloti, Agrobacterium tumefaciens, Rhodospirillum centrum, Pseudovibrio sp. | – | – |
| β-Proteobacteria (11 sequences) | Candidatus, ‘Accumulibacter phosphatis’ clade, Cupriavidus tawananensis, Burkholderia xenovorans | – | – | – |
| γ-Proteobacteria (34 sequences) | Aggregatibacter actinomycetemcomitans, Calwellia psychrophila | Aeromonas salmonicida, Shigella sonnei, Haemophilus ducreyi, Haemophilus somnus, Vibrio parahaemolyticus, H. chejuensis | – | – |
| δ-Proteobacteria (four sequences) | G. lovleyi, D. desulfuricans | – | G. lovleyi | – |
| ε-Proteobacteria (13 sequences) | C. lari, C. hominis, Nautilia profundicola | Campylobacter fetus, C. hominis, Sulfospirillum deleyianum, Arcobacter butzleri | C. jejuni, C. lari, C. hominis, W. succinogenes, Helioctebacter hepaticus, Arcobacter butzleri | – |
Figure 2 | Sequences used to design the forward primer, SF1173d and the alternative reverse primer, LR1837

The gel photographs on the right show PCR products obtained using various combinations of the two forward primers, LF716 or SF1173d in combination with the reverse primers LR1837 and SR2294. The five DNA templates tested were *E. coli* plasmid pRS460 and chromosomal DNA (lanes 1 and 2 respectively); *D. desulfuricans* 27774 (lanes 3) and *C. jejuni* (lanes 4). The template for products in lanes 5 was a heavily contaminated sample of *Salmonella enterica* DNA: purified *S. enterica* serovar Typhimurium DNA gave PCR products similar to those from *E. coli*, as would be expected.

primer OF640 that was predicted to recognize templates such as *C. lari*, *C. hominis* and *G. lovleyi* (Figure 2B). This forward primer in combination with the universal reverse primer SR2294 directed the synthesis of the predicted 1.8 kb fragment on the two *Campylobacter* templates (*G. lovleyi* DNA was unavailable). Control experiments showed that the LF716 forward primer with either SR2294 or LR1837 also generated the expected fragments on these templates, but multiple products were obtained even with pure DNA templates. For this reason, use of the third forward primer, OF640, is recommended for the detection, amplification and quantification of difficult target *napA* genes from ε-proteobacteria.

In summary, all known *napA* genes can now be amplified using one of three forward primers, LF716, SF1173d or OF640, in combination with the single reverse primer, SR2294. Although they are all degenerate primers that also recognize sequences unrelated to *napA* on some templates, the limited degeneracy restricts this problem to an acceptable degree. However, due to the generation of multiple products in some reactions, the three forward primers should not be used together in a single multiplex PCR.

Amplification of *napA* sequences using community DNA as a template

We recently described the use of community DNA from the faeces of patients suffering from the gastric diseases ulcerative colitis, irritable bowel syndrome and Crohn’s disease to compare the abundance of *Faecalibacterium prausnitzii* before and after successful treatment with those of a group of healthy controls [18]. Sulfate-reducing bacteria were abundant in all of these samples, as were enteric bacteria. Virtually all of the samples yielded PCR products of the expected size range with two of the three pairs of primers (the primer pair OF640+SR2294 was not tested in these experiments). However, as expected when degenerate primers
are used with community DNA as a template, multiple PCR products were obtained from most samples. Nevertheless, as discussed below, these PCRs can now be combined with high-throughput DNA sequencing that is increasingly widely available to determine the diversity of napA sequences in any bacterial community.

**High-throughput DNA sequencing avoids a requirement for highly specific primers**

Before the availability of high-throughput DNA sequencing, analysis of the variety of species capable of catalysing major environmental changes required the cloning and sequencing of individual PCR products one at a time, a process that was both time consuming and extremely expensive. Consequently, it was essential to design oligonucleotide primers that were not only sufficiently degenerate to capture all of the relevant target genes in community DNA but also sufficiently specific to avoid the generation of PCR products irrelevant to the investigation. The ability to generate up to 10^6 DNA sequences in a single experiment, for example using Roche 454 DNA sequencing technology, renders the second problem largely irrelevant because species that represent less than 0.1% of the population can be detected even in a mixture of relevant and irrelevant amplicons [19]. Primer design can now focus on ensuring the amplification of the widest range of individual PCR products one at a time, a process that was both time consuming and extremely expensive.

Consequently, it was essential to design oligonucleotide primers that were not only sufficiently degenerate to capture all of the relevant target genes in community DNA but also sufficiently specific to avoid the generation of PCR products irrelevant to the investigation. The ability to generate up to 10^6 DNA sequences in a single experiment, for example using Roche 454 DNA sequencing technology, renders the second problem largely irrelevant because species that represent less than 0.1% of the population can be detected even in a mixture of relevant and irrelevant amplicons [19]. Primer design can now focus on ensuring the amplification of the widest range of individual PCR products one at a time, a process that was both time consuming and extremely expensive.

Consequently, it was essential to design oligonucleotide primers that were not only sufficiently degenerate to capture all of the relevant target genes in community DNA but also sufficiently specific to avoid the generation of PCR products irrelevant to the investigation. The ability to generate up to 10^6 DNA sequences in a single experiment, for example using Roche 454 DNA sequencing technology, renders the second problem largely irrelevant because species that represent less than 0.1% of the population can be detected even in a mixture of relevant and irrelevant amplicons [19]. Primer design can now focus on ensuring the amplification of the widest range of individual PCR products one at a time, a process that was both time consuming and extremely expensive.

Consequently, it was essential to design oligonucleotide primers that were not only sufficiently degenerate to capture all of the relevant target genes in community DNA but also sufficiently specific to avoid the generation of PCR products irrelevant to the investigation. The ability to generate up to 10^6 DNA sequences in a single experiment, for example using Roche 454 DNA sequencing technology, renders the second problem largely irrelevant because species that represent less than 0.1% of the population can be detected even in a mixture of relevant and irrelevant amplicons [19]. Primer design can now focus on ensuring the amplification of the widest range of individual PCR products one at a time, a process that was both time consuming and extremely expensive.

Consequently, it was essential to design oligonucleotide primers that were not only sufficiently degenerate to capture all of the relevant target genes in community DNA but also sufficiently specific to avoid the generation of PCR products irrelevant to the investigation. The ability to generate up to 10^6 DNA sequences in a single experiment, for example using Roche 454 DNA sequencing technology, renders the second problem largely irrelevant because species that represent less than 0.1% of the population can be detected even in a mixture of relevant and irrelevant amplicons [19]. Primer design can now focus on ensuring the amplification of the widest range of individual PCR products one at a time, a process that was both time consuming and extremely expensive.

Consequently, it was essential to design oligonucleotide primers that were not only sufficiently degenerate to capture all of the relevant target genes in community DNA but also sufficiently specific to avoid the generation of PCR products irrelevant to the investigation. The ability to generate up to 10^6 DNA sequences in a single experiment, for example using Roche 454 DNA sequencing technology, renders the second problem largely irrelevant because species that represent less than 0.1% of the population can be detected even in a mixture of relevant and irrelevant amplicons [19]. Primer design can now focus on ensuring the amplification of the widest range of individual PCR products one at a time, a process that was both time consuming and extremely expensive.

Consequently, it was essential to design oligonucleotide primers that were not only sufficiently degenerate to capture all of the relevant target genes in community DNA but also sufficiently specific to avoid the generation of PCR products irrelevant to the investigation. The ability to generate up to 10^6 DNA sequences in a single experiment, for example using Roche 454 DNA sequencing technology, renders the second problem largely irrelevant because species that represent less than 0.1% of the population can be detected even in a mixture of relevant and irrelevant amplicons [19]. Primer design can now focus on ensuring the amplification of the widest range of individual PCR products one at a time, a process that was both time consuming and extremely expensive.

Consequently, it was essential to design oligonucleotide primers that were not only sufficiently degenerate to capture all of the relevant target genes in community DNA but also sufficiently specific to avoid the generation of PCR products irrelevant to the investigation. The ability to generate up to 10^6 DNA sequences in a single experiment, for example using Roche 454 DNA sequencing technology, renders the second problem largely irrelevant because species that represent less than 0.1% of the population can be detected even in a mixture of relevant and irrelevant amplicons [19]. Primer design can now focus on ensuring the amplification of the widest range of individual PCR products one at a time, a process that was both time consuming and extremely expensive.

Consequently, it was essential to design oligonucleotide primers that were not only sufficiently degenerate to capture all of the relevant target genes in community DNA but also sufficiently specific to avoid the generation of PCR products irrelevant to the investigation. The ability to generate up to 10^6 DNA sequences in a single experiment, for example using Roche 454 DNA sequencing technology, renders the second problem largely irrelevant because species that represent less than 0.1% of the population can be detected even in a mixture of relevant and irrelevant amplicons [19]. Primer design can now focus on ensuring the amplification of the widest range of individual PCR products one at a time, a process that was both time consuming and extremely expensive.

Consequently, it was essential to design oligonucleotide primers that were not only sufficiently degenerate to capture all of the relevant target genes in community DNA but also sufficiently specific to avoid the generation of PCR products irrelevant to the investigation. The ability to generate up to 10^6 DNA sequences in a single experiment, for example using Roche 454 DNA sequencing technology, renders the second problem largely irrelevant because species that represent less than 0.1% of the population can be detected even in a mixture of relevant and irrelevant amplicons [19]. Primer design can now focus on ensuring the amplification of the widest range of individual PCR products one at a time, a process that was both time consuming and extremely expensive.
15 Flanagan, D.A., Gregory, L.G., Carter, J.P., Karakas-Sen, A., Richardson, D.J. and Spiro, S. (1999) Detection of genes for periplasmic nitrate reductase in nitrate respiring bacteria and in community DNA. FEMS Microbiol. Lett. 177, 263–270
16 Marietou, A., Richardson, D., Cole, J. and Mohan, S. (2005) Nitrate reduction by *Desulfovibrio desulfuricans*: a periplasmic nitrate reductase system that lacks NapB, but includes a unique tetraheme c-type cytochrome, NapM. FEMS Microbiol. Lett. 248, 217–225
17 Bru, D., Sarr, A. and Philippot, L. (2007) Relative abundances of proteobacterial membrane-bound and periplasmic nitrate reductases in selected environments. Appl. Environ. Microbiol. 73, 5971–5974
18 Jia, W., Whitehead, R.N., Griffiths, L., Dawson, C., Waring, R.H., Ramsden, D.B., Hunter, J.O. and Cole, J.A. (2010) Is the abundance of *Faecalibacterium prausnitzii* relevant to Crohn’s disease? FEMS Microbiol. Lett. 310, 138–144
19 Demanèche, S., Philippot, L., David, M.M., Navarro, E., Vogel, T.M. and Simonet, P. (2009) Characterisation of denitrification gene clusters of soil bacteria via a metagenomic approach. Appl. Environ. Microbiol. 75, 534–537

Received 12 October 2010
doi:10.1042/BST0390321