



Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton

Amy Apprill^{1,*}, Sean McNally^{1,2}, Rachel Parsons², Laura Weber¹

¹Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA

²Bermuda Institute of Ocean Sciences, Ferry Reach, St. George's GE01, Bermuda

ABSTRACT: High-throughput sequencing of small subunit ribosomal RNA (SSU rRNA) genes from marine environments is a widely applied method used to uncover the composition of microbial communities. We conducted an analysis of surface ocean waters with the commonly employed hypervariable 4 region SSU rRNA gene primers 515F and 806R, and found that bacteria belonging to the SAR11 clade of *Alphaproteobacteria*, a group typically making up 20 to 40% of the bacterioplankton in this environment, were greatly underrepresented and comprised <4% of the total community. Using the SILVA reference database, we found a single nucleotide mismatch to nearly all SAR11 subclades, and revised the 806R primer so that it increased the detection of SAR11 clade sequences in the database from 2.6 to 96.7%. We then compared the performance of the original and revised 806R primers in surface seawater samples, and found that SAR11 comprised 0.3 to 3.9% of sequences with the original primers and 17.5 to 30.5% of the sequences with the revised 806R primer. Furthermore, an investigation of seawater obtained from aquaria revealed that SAR11 sequences acquired with the revised 806R primer were more similar to natural cellular abundances of SAR11 detected using fluorescence *in situ* hybridization counts. Collectively, these results demonstrate that a minor adjustment to the 806R primer will greatly increase detection of the globally abundant SAR11 clade in marine and lake environments, and enable inclusion of this important bacterial lineage in experimental and environmental-based studies.

KEY WORDS: SSU rRNA gene · 16S · SAR11 · *Bacteria* · Fluorescence *in situ* hybridization · FISH

Resale or republication not permitted without written consent of the publisher

INTRODUCTION

The ability to deeply sequence microbial small subunit ribosomal RNA (SSU rRNA) genes has provided considerable insight into the structure, stability, composition and dynamics of microbial populations associated with aquatic environments (e.g. Sogin et al. 2006, Huber et al. 2007, Gilbert et al. 2009, Nelson et al. 2011, Vergin et al. 2013a). Primers that target the hypervariable region 4 (V4) of the SSU rRNA gene, especially 515F and 806R, are now commonly employed in studies that investigate and compare the taxonomic diversity of microbial communities

(Caporaso et al. 2012, Kozich et al. 2013), including the Earth Microbiome Project's exploration of the global microbiome (Gilbert et al. 2014). The V4 primers are popular because they target *Bacteria* and *Archaea*, and produce an appropriate amplicon size for next-generation sequencing. Additionally, there is now a growing abundance of V4 sequence data, allowing for meaningful comparisons among and across environments.

Our application of these commonly used V4 SSU rRNA gene primers to surface seawater samples revealed a surprisingly low detection of sequences belonging to the SAR11 clade of *Alphaproteobacte-*

ria (<4% of the community). SAR11 constitutes about 1 in 3 microbial cells in the surface ocean, and can comprise up to 1 in 5 microbial cells in the mesopelagic ocean and some lakes (Morris et al. 2002, Carlson et al. 2009, Salcher et al. 2011, Treusch et al. 2012, Vergin et al. 2013b). The low detection of this dominant bacterioplankton using the V4 SSU rRNA gene primers could bias interpretation of experimental and environmental microbial dynamics, particularly for studies focused on understanding microbial processes in surface seawater.

In any PCR- and primer-based taxonomic investigation, members of a microbial community may be omitted, distorted, and/or misrepresented, typically due to primer mismatches or PCR biases (Acinas et al. 2005, Hong et al. 2009, Lee et al. 2012, Pinto & Raskin 2012, Logares et al. 2014). In fact, even a single base mismatch within a primer can result in a significant lack of detection of community members (Bru et al. 2008, Mao et al. 2012); thus, primers are frequently and continually altered to better account for the targeted community. In this study, we present evidence of the 806R primer bias against most subgroups of the SAR11 SSU rDNA using an *in silico* analysis of the V4 primers in relation to the recognized SAR11 subgroups in the SILVA reference database. We also compare the performance of the original 806R reverse primer and a modified 806RB primer that was designed to enhance the number of SAR11 sequences using a head-to-head sequencing study of the same surface seawater communities. In addition, we compare the retrieval of SAR11 sequences with the 2 primer sets to a PCR-independent survey of SAR11 cellular abundances in aquaria seawater samples. Our results demonstrate that a minor adjustment to the existing 806R primer will enhance the recovery of SAR11 populations in aquatic environments.

MATERIALS AND METHODS

Primer coverage evaluation and re-design

The primers 515F (5'-GTG CCA GCM GCC GCG GTA A-3') and 806R (5'-GGA CTA CHV GGG TWT CTA AT-3') (Caporaso et al. 2011), targeting the V4 region of the SSU rRNA gene, were examined in relation to SSU rRNA gene sequences publicly available in the non-redundant SILVA SSU Ref database (v.115, released August 2013) (Quast et al. 2013) using the ARB software (Ludwig et al. 2004). A total of 3600 SAR11 sequences were specifically evaluated in relation to the V4 primers. ARB software was

also utilized for the re-design of the 806R primer. In addition, primer coverage of the original reverse primer and the revised reverse primer was evaluated using SILVA TestPrime 1.0 (Klindworth et al. 2013) with version 117 of the SILVA SSU Ref database with no mismatches allowed. Melting temperatures of the primer sets were examined using OligoCalc (Kibbe 2007), and primer dimers were evaluated using OligoAnalyzer 3.1 (Integrated DNA Technologies; www.idtdna.com/calc/analyzer).

Seawater sample collection

Nucleic acids were collected from surface waters along the coastline of the southeastern Red Sea and from coral reefs surrounding the islands and atolls in the Federated States of Micronesia (see Table S1 in Supplement 1 at www.int-res.com/articles/suppl/a075p129_supp.pdf). Samples from an aquaria-based experiment housed at the Bermuda Institute of Ocean Sciences (BIOS) were also utilized in this study. Seawater from Ferry Reach, Bermuda (32.37035° N, 64.69545° W) was sampled from a direct inflow line. In addition, this water was also held in separate aerated 30 l aquaria (similar to de Putron et al. 2011) where it was sampled for nucleic acids (500 ml) and fluorescence *in situ* hybridization (FISH) analyses (50 ml) over the course of 12 d.

DNA analysis

Seawater collected from the Red Sea (20 l) and Micronesia (2 l) was filtered onto 142 mm, 0.22 µm Durapore membrane filters (Millipore) and 25 mm, 0.2 µm Supor polyethersulfone membranes (Pall), respectively, and immediately frozen in liquid nitrogen. Total genomic DNA was extracted from these samples using an extraction method employing bead beating for 10 min with sucrose-EDTA lysis buffer (0.75 M sucrose, 20 mM EDTA, 400 mM NaCl, 50 mM Tris) and 100 ml of 10% sodium dodecyl sulfate, followed by a Proteinase K digestion for 4 h at 55°C, and finally purification with the DNeasy kit purification (Qiagen) (Santoro et al. 2010). Similarly, microbial biomass originating from the BIOS seawater inflow line and aquaria seawater was concentrated onto 0.2 µm Supor polyethersulfone membranes using a 47 mm support filter and a gelman rig under gentle vacuum (~100 mm Hg). Each filter was stored in 1 ml of sterile sucrose lysis buffer (20 mM EDTA, 400 mM NaCl, 0.75 M sucrose, 50 mM Tris-HCl) at

–80°C. DNA was extracted by incubating the filter and buffer in 1 mg ml⁻¹ of lysozyme at 37°C for 20 min, followed by the addition of sodium dodecyl sulfate to 0.5% and an incubation with 160 µg ml⁻¹ proteinase K for 2 h at 37°C, and a phenol-chloroform purification (Giovannoni et al. 1990). Primers targeting the V4 region of the SSU rRNA gene (515F and 806R, presented above), were utilized for PCR amplification using unique barcoded primer combinations for each sample. In addition, the same DNA samples were amplified separately with the 515F primer and a modified 806RB primer using the identical barcoding approach. The modified reverse primer replaces the 'H' degeneracy in the original 806R primer with a 'N', and is designed to enhance SAR11 targets (revised primer 806RB; 5'-GGA CTA CNV GGG TWT CTA AT-3'). The primers were designed after Kozich et al. (2013) and were each equipped with a unique 8 bp barcode, 10 bp pad and 2 bp link that followed the above-mentioned primers (see 'Materials and methods' in Supplement 1). Triplicate 25 µl PCR reactions were conducted for each sample, and each reaction contained 1.25 U of *GoTaq* Flexi DNA Polymerase (Promega), 5× Colorless *GoTaq* Flexi Buffer, 2.5 mM MgCl₂, 200 µM dNTP mix (Promega), 200 nM of each barcoded primer, and 1 to 4 ng of genomic template. The reaction conditions consisted of an initial denaturation step at 95°C for 2 min, followed by 27 to 38 cycles of 95°C for 20 s, 55°C for 15 s, and 72°C for 5 min, concluding with an extension step at 72°C for 10 min. The reactions were carried out in a Bio-Rad thermocycler (Bio-Rad Laboratories). Reaction products (5 µl) were screened on a 1% agarose/TBE gel. The HyperLadder 50 bp DNA ladder (generally 5 ng µl⁻¹) (Bioline) was used to confirm appropriate amplicon size. The number of PCR cycles varied between samples in order to produce similar, minimal yields, but each sample was subjected to nearly identical PCR cycles with both primer sets. The 3 replicate reactions were pooled and subsequently purified using the QIAquick Purification Kit (Qiagen), and quantified using the Qubit 2.0 Fluorometer with the dsDNA High Sensitivity Assay (Life Technologies). For each primer set, barcoded amplicons were pooled into equimolar ratios. These amplicon pools were then shipped to the University of Illinois W.M. Keck Center for Comparative and Functional Genomics where they were used for construction of 2 separate libraries that were subsequently sequenced using 2 × 250 bp paired-end MiSeq (Illumina), as detailed previously (similar to Kozich et al. 2013). Control samples included sterile water (negative controls) in which PCR did not yield any

detectable amplification with either primer set. A mock community sample (positive control; obtained through BEI Resources, NIAID and NIH as part of the Human Microbiome Project: Genomic DNA from Microbial Mock Community B [even, low concentration], v5.1L, HM-782D) was amplified with the 515F/806RB primers and sequenced to assess amplification bias and sequencing error rate.

Sequence analysis

Data analyses were conducted using mothur v.1.33.3 (Schloss et al. 2009) and included contig construction of the paired ends, quality filtering, amplicon size selection (253 bp median size) and alignment to the SSU rRNA gene. Chimera detection was conducted via UCHIME (Edgar et al. 2011) using mothur, and chimeric sequences were removed. Taxonomic classification of sequences was conducted in mothur with the SILVA SSU Ref database (v.117) using the k-nearest neighbor algorithm on sequences sub-sampled to the same depth with each primer pair (10 000, 12 000, or 17 500 sequences sample⁻¹ for BIOS aquaria, Micronesia, and Red Sea sample sets, respectively). The sequencing error analysis conducted on the mock community sequences amplified with the 515F/806RB primers revealed a sequencing error rate of 0.0012%. Data are accessible in NCBI's Short Read Archive under BioProject ID PRJNA279146.

Microbial abundances and FISH

To determine microbial abundances from the BIOS inflow and aquaria samples, the seawater was fixed with formalin (10% final concentration) and stored at –80°C. Upon analysis, the samples were thawed and filtered onto 0.2 µm filters pre-stained with Irgalan black (0.2 g in 2% acetic acid) under gentle vacuum (~100 mm Hg) and post-stained with 4,6-diamidino-2-phenyl dihydrochloride (DAPI; 5 µg ml⁻¹, SIGMA-Aldrich) (Porter & Feig 1980). Slides were then enumerated using an AX70 epifluorescent microscope (Olympus) under UV excitation at 100× magnification as previously described (Parsons et al. 2014). At least 500 cells filter⁻¹ (12 fields) were counted.

Fluorescence *in situ* hybridization (FISH) was used to quantify the abundance of SAR11 in the BIOS aquaria samples using a probe suite (152R-Cy3, 441R-Cy3, 542R-Cy3, 732R-Cy3) and was conducted as previously described (Morris et al. 2002, Parsons et al.

Table 1. *In silico* analysis of the percentage of SAR11 sequences recovered using the original hypervariable region 4 (V4) primers and revised 806RB primer with no mismatches allowed

SAR11 subgroup ^a	No. of sequences in database ^b	—SAR11 sequences recovered— Original V4 primers (%)	515F/806RB (%)
Surface 1	2195	0.6	96.0
Surface 2	271	0	96.7
Surface 3	12	0	100.0
Surface 4	36	91.7	91.7
Deep 1	94	0	98.5
Chesapeake– Delaware Bay	106	0.9	99.1
LD12 freshwater	405	5.3	97.9

^aSubgroups currently defined by SILVA SSU rRNA project
^bSILVA SSU reference 117 non-redundant database

2012). Image analysis coupled with epifluorescence microscopy was used to process FISH slides excited with Cy3 (550 nm) and UV wavelengths. The images were captured using a Retiga Exi CCD digital camera with QCapture software (v.2.0; QImaging) and processed with Image Pro software (v.7.0; Media Cybernetics) as previously described (Parsons et al. 2014). SAR11 percentages were calculated as SAR11 FISH abundances compared to total cellular abundances.

RESULTS AND DISCUSSION

In silico analysis and modification of the 806R primer

Evaluation of the V4 region 515F and 806R primers in ARB revealed that most (3579 of 3600) SAR11 sequences in the SILVA non-redundant SSU 115 database bind to the 515F primer, but the majority (3509 of 3600) of these sequences have a single internal mismatch to the 806R primer. Specifically, there is a mismatch at the 'D' position of the reverse complement orientation of the primer (position 799 of the *Escherichia coli* SSU rRNA molecule; see Fig. S1 in Supplement 1 at www.int-res.com/articles/suppl/a075p129_supp.pdf). SAR11 Surface 1, 2, and 3 subgroups as well as the Deep 1, Chesapeake–Delaware Bay, and LD12-freshwater subgroups contain a 'C' basepair at this mismatched position, and the 'D' degeneracy will anneal with all bases except 'C', potentially leading to insufficient detection of SAR11 SSU rDNA during PCR (Fig. S1). Only the SAR11 Surface 4 subgroup with a 'G' in the 799 position will

theoretically bind to 806R. This primer mismatch was verified using TestPrime and an updated version of the SILVA non-redundant database (SSU 117, containing 3659 SAR11 sequences); analysis revealed that the original V4 primer set targets 2.6% of the SAR11, the main taxonomic group being the Surface 4 subgroup (Table 1). Replacing the 'D' basepair in the reverse complement of the original 806R primer to a 'N' basepair (reverse complement now reads as 3'-ATT AGA WAC CCB NGT AGT CC-5') theoretically allows the 'C' base in the remaining SAR11 sequences to anneal during PCR (as well as all other nucleotide possibilities at this position). *In silico* analysis of this revised primer (806RB 5'-GGA CTA CNV GGG TWT CTA AT-3') with Test-

Prime and the SSU 117 database revealed an increase in the SAR11 sequence targets to 96.6%, which includes nearly full recovery of the Chesapeake–Delaware Bay, Deep 1, LD12-freshwater, and Surface 1, 2 and 3 subgroups (Table 1). This analysis included sequences recognized as belonging to finer resolution groupings within the subgroups (subclades) (Vergin et al. 2013b), but the taxonomy presented here conforms to that currently recognized by the SILVA rRNA database project. The minor modification made to the original 806R primer did not alter the theoretical melting temperature of the primer beyond the range of the original barcoded primers (all between 63.8 and 67°C, salt adjusted). Heterodimer analysis revealed that the delta G decreased with the revised primer (from $-9.85 \text{ kcal mol}^{-1}$ in the original 515F/806R primers to $-13.42 \text{ kcal mol}^{-1}$ in the new 515F/806RB primers), making a 5-base primer dimer more likely. Even so, primer dimers were observed using both primer sets when higher-than-optimal PCR cycles were utilized, and our methodology optimized PCR cycles per sample to avoid formation of these products.

Comparison of SAR11 recovery with V4, modified primers and FISH counts

When seawater samples from the Red Sea and Micronesia were amplified with the original V4 primers, SAR11 sequences were found to comprise 0.3 to 3.9% of the bacterial and archaeal communities, with the Surface 1 and 4 subgroups most represented. In contrast, SAR11 composed 17.5 to 30.5%

of the communities in the same samples when samples were amplified with the revised reverse primer, and the Surface 1 subgroup increased about 10-fold in these communities, with the Surface 2 and unclassified SAR11 subgroups (which were previously nearly undetectable) now represented in all communities (Fig. 1, Table S2 in Supplement 1). These data suggest that although the original 806R primer mismatch is internal, substitution of the mismatched base clearly increases detection of SAR11. This performance of an internal priming mismatch is consistent with previous observations (Sipos et al. 2007). As expected, the representation of non-SAR11 lineages decreased within the community when the revised primer set was used (Fig. 2).

Cellular abundances of SAR11 examined using FISH in the BIOS aquaria samples revealed that 16.7 to 29.4% of the bacterioplankton community was comprised of SAR11. These cellular abundances were substantially greater than the sequence estimates of SAR11 using the original V4 primers (0.2 to 2.0%) would suggest, and were more comparable to sequence estimates recovered using the revised primer (11.4 to 28.7%) (Table 2). It was expected that SAR11 abundances would be more variable in the aquaria compared to ocean waters, and these results will be examined in more detail in relation to a coral-based experiment (S. McNally et al. unpubl. data). The discrepancy between SAR11 sequences and FISH counts are within the range reported in another

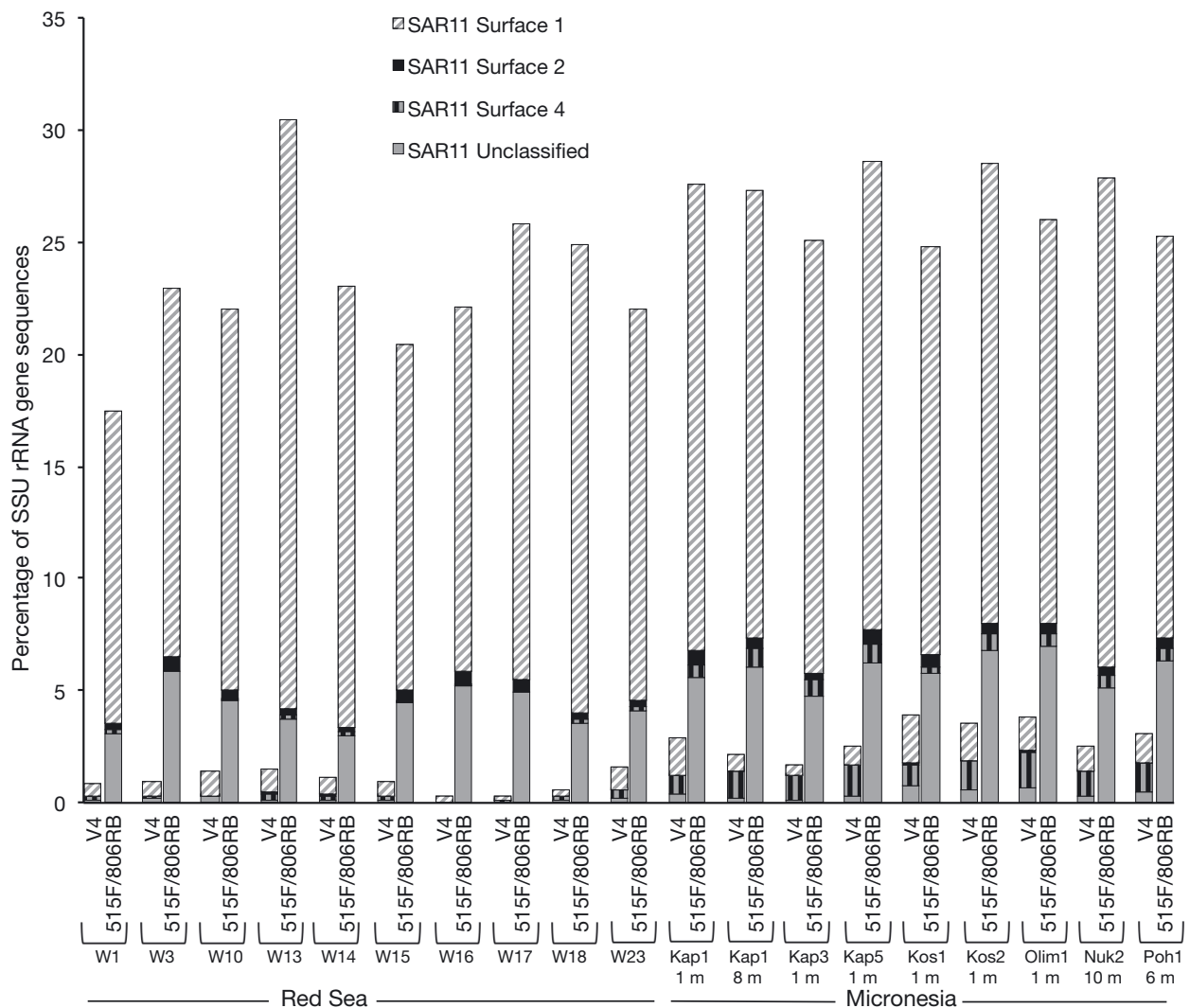


Fig. 1. Taxonomic comparison of sequences belonging to the SAR11 subgroups recovered from seawater samples using the original hypervariable region 4 (V4) and 515F/806RB primers from coral reef waters in the Red Sea (17 500 sequences sample⁻¹) and Federated States of Micronesia (12 000 sequences sample⁻¹). W, Kap, Kos, Olim, Nuk and Poh refer to the sample names (see Table S1 in Supplement 1 at www.int-res.com/articles/suppl/a075p129_supp.pdf)

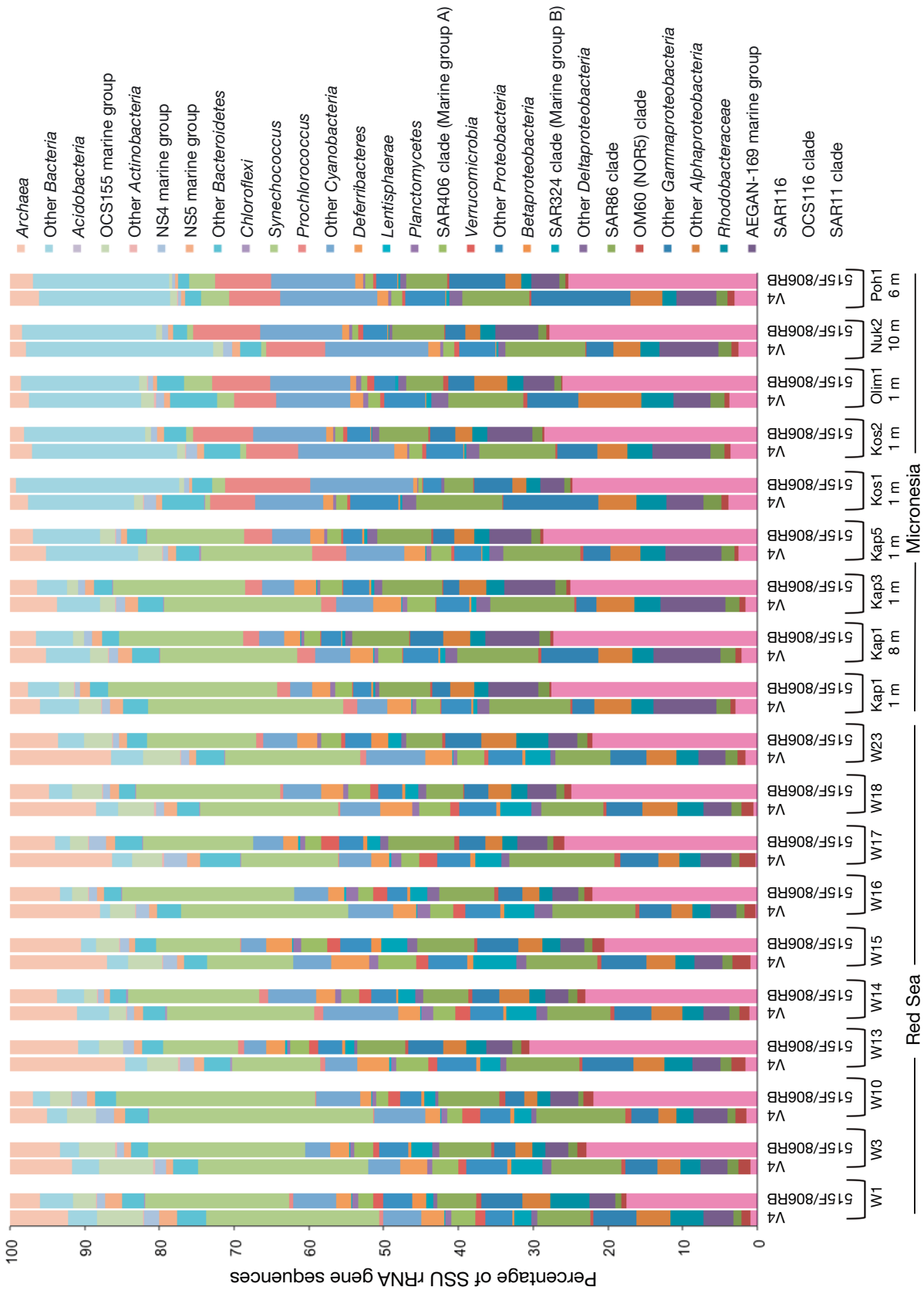


Fig. 2. Comparison of the major taxonomic groups of *Bacteria* and *Archaea* recovered from seawater samples examined with the hypervariable region 4 (V4) and 515F/806RB primers. Samples are from coral reef waters in the Red Sea (17 500 sequences sample⁻¹) and Federated States of Micronesia (12,000 sequences sample⁻¹). W, Kap, Kos, Olim, Nuk and Poh refer to the sample names (see Table S1 in Supplement 1 at www.int-res.com/articles/suppl/a075p129_supp.pdf).

study examining the coastal waters of Bermuda (Parsons et al. 2014). SAR11 cells typically contain a single ribosomal RNA operon (Giovannoni et al. 2005, Grote et al. 2012). As a result, natural abundances of SAR11 could be underestimated in comparison to abundances of other bacterioplankton containing multiple copies of this gene operon if only PCR-based surveys are used to profile seawater microbial communities.

Minor impacts on the recovery of other surface seawater taxa using the revised primer

The contribution of SAR11 to the Red Sea and Micronesian seawater bacterioplankton communities was entirely removed in order to compare the amplification of non-SAR11 taxonomic groups between the original V4 and 515/806RB primers. This analysis revealed that a majority (27 of 29) of the taxonomic groups were not significantly affected by the primer modification (see Supplement 1: Table S3, 2 sample *t*-tests, $p > 0.05$ as well as full taxonomic comparisons provided in Fig. S2 at www.int-res.com/articles/suppl/a075p129_supp.pdf; see also Table S4 in Supplement 2 at www.int-res.com/articles/suppl/a075p129_supp.xlsx). Two taxonomic groups, SAR86 (*Gammaproteobacteria*) and other *Proteobacteria* (comprised of sequences not belonging to the Alpha, Beta, Delta and Gamma lineages), were significantly impacted by the primer alteration ($t = 2.33$, $p = 0.025$ for SAR86; $t = 3.12$, $p = 0.004$ for other *Proteobacteria*). In general, there was less recovery of SAR86 sequences and

other *Proteobacteria*-affiliated sequences with the modified primers, but the inconsistency in the recovery of these sequences in seawater appeared limited to several samples (Fig. S3A,B & Table S3). These results only represent the performance of the primers in a limited marine environment (surface seawater); a more thorough examination of non-marine taxa is warranted before applying the revised primer to other environments.

CONCLUSIONS

The *in silico* and sequencing data presented here provide support for a minor revision to the V4 primer 806R that would result in an increased detection of SAR11 bacterioplankton without incurring a large bias in the detection of other surface bacterioplankton taxa. Several published studies have applied the original V4 primers to surface marine and lake samples, and thus may have underestimated SAR11 abundances in these waters (Paver et al. 2013, Taylor et al. 2014, Orsi et al. 2015). In some marine samples, however, underestimation of SAR11 with the V4 primers may not impact the findings of the study (i.e. for sponges and corals, see Cuvelier et al. 2014, Meyer et al. 2014). The Earth Microbiome Project (EMP) has employed the V4 primers for studies exploring Earth's microbial environment, and they have recently acknowledged the limitation of these primers for detecting SAR11 (Gilbert et al. 2014). The performance of the 806RB primer proposed and evaluated in this study should be examined in tandem

Table 2. Abundance of SAR11 compared to the entire bacterioplankton community recovered using the original V4 primers, the revised 806RB primer, and SAR11 cellular abundances determined using fluorescence *in situ* hybridization (FISH) from aquaria containing seawater

Sample	Percent of bacterioplankton		
	V4 primers 515F/806R (SAR11 sequences ^a)	Revised 515F/806RB (SAR11 sequences ^a)	SAR11 FISH counts (SD)
Inflow to aquaria	1.9 (192)	28.7 (2867)	29.4 (4.9)
Seawater aquaria 7, Day 2	0.6 (54)	11.4 (1136)	24.0 (2.6)
Seawater aquaria 9, Day 2	1.1 (105)	14.9 (1488)	29.0 (4.6)
Seawater aquaria 7, Day 4	1.4 (133)	16.5 (1646)	20.2 (3.5)
Seawater aquaria 9, Day 4	0.2 (15)	16.0 (1599)	22.0 (3.3)
Seawater aquaria 7, Day 6	1.3 (133)	14.7 (1474)	20.0 (6.3)
Seawater aquaria 9, Day 6	1.6 (157)	12.6 (1255)	21.2 (4.9)
Seawater aquaria 7, Day 8	1.7 (168)	16.5 (1651)	22.7 (3.0)
Seawater aquaria 8, Day 8	1.8 (184)	26.6 (2160)	22.4 (3.6)
Seawater aquaria 7, Day 10	0.9 (92)	12.0 (1193)	16.7 (9.5)
Seawater aquaria 8, Day 10	2.0 (203)	23.4 (2338)	22.0 (1.8)
Seawater aquaria 8, Day 12	0.5 (53)	14.3 (1428)	19.5 (9.5)

^aSequence data subsampled to 10 000 sequences sample⁻¹

using the EMP methodologies (which are different from those presented here), to understand whether this primer revision is suitable for the characterization of microbial communities in other environments. While no primer set perfectly captures the diversity of the *Bacteria* and *Archaea* residing in Earth's diverse environments, the proposed revision to the existing 806R primer will enhance recognition of the globally abundant SAR11 clade in aquatic environments.

Acknowledgements. This project was supported by NSF award OCE-1233612 to A.A. with contributions from BIOS Grant in aid award to S.McN. and NSF Oceanic Microbial Observatory OCE-0801991 subcontract to BIOS managed by R.P. We thank Justin Ossolinski and Konrad Hughen for collecting the Red Sea seawater samples, Alyson Santoro and Matthew Neave for assistance with the Micronesian seawater samples, Chris Wright and the University of Illinois W. M. Keck Center for Comparative and Functional Genomics for sequencing and Rob Knight, Jack Gilbert and Jed Furhman for helpful discussions.

LITERATURE CITED

- Acinas SG, Sarma-Rupavtarm R, Klepac-Ceraj V, Polz MF (2005) PCR-induced sequence artifacts and bias: insights from comparison of two 16S rRNA clone libraries constructed from the same sample. *Appl Environ Microbiol* 71:8966–8969
- Bru D, Martin-Laurent F, Philippot L (2008) Quantification of the detrimental effect of a single primer-template mismatch by real-time PCR using the 16S rRNA gene as an example. *Appl Environ Microbiol* 74:1660–1663
- Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D and others (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA* 108:4516–4522
- Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D and others (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6:1621–1624
- Carlson CA, Morris R, Parsons R, Giovannoni SJ, Vergin K (2009) Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the northwestern Sargasso Sea. *ISME J* 3:283–295
- Cuvellier ML, Blake E, Mulheron R, McCarthy PJ, Blackwelder P, Vega Thurber R, Lopez JV (2014) Two distinct microbial communities revealed in the sponge *Cinachyrella*. *Front Microbiol* 5:581
- de Putron SJ, cCorkle DC, Cohen AL, Dillon AB (2011) The impact of seawater saturation state and bicarbonate ion concentration on coral calcification. *Coral Reefs* 30: 321–328
- Edgar RC, Haas B, Clemente J, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27:2194–2200
- Gilbert JA, Field D, Swift P, Newbold L and others (2009) The seasonal structure of microbial communities in the Western English Channel. *Environ Microbiol* 11: 3132–3139
- Gilbert JA, Jansson JK, Knight R (2014) The Earth Microbiome Project: successes and aspirations. *BMC Biol* 12:69
- Giovannoni SJ, DeLong EF, Schmidt TM, Pace NR (1990) Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. *Appl Environ Microbiol* 56:2572–2575
- Giovannoni SJ, Tripp HJ, Givan S, Podar M and others (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* 309:1242–1245
- Grote J, Thrash JC, Huggett MJ, Landry ZC, Carini P, Giovannoni SJ, Rappé MS (2012) Streamlining and core genome conservation among highly divergent members of the SAR11 clade. *MBio* 3:e00252-12
- Hong S, Bunge J, Leslin C, Jeon S, Epstein SS (2009) Polymerase chain reaction primers miss half of rRNA microbial diversity. *ISME J* 3:1365–1373
- Huber JA, Mark Welch DB, Morrison HG, Huse SM, Neal PR, Butterfield DA, Sogin ML (2007) Microbial population structures in the deep marine biosphere. *Science* 318:97–100
- Kibbe WA (2007) OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res* 35(Suppl 2):W43–W46
- Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 41:e1
- Kozich JJ, Westcott SL, Baxter NT, Highlander S, Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 79:5112–5120
- Lee CK, Herbold CW, Polson SW, Wommack KE, Williamson SJ, McDonald IR, Cary SC (2012) Groundtruthing next-gen sequencing for microbial ecology-biases and errors in community structure estimates from PCR amplicon pyrosequencing. *PLoS ONE* 7:e44224
- Logares R, Sunagawa S, Salazar G, Cornejo-Castillo FM and others (2014) Metagenomic 16S rDNA Illumina tags are a powerful alternative to amplicon sequencing to explore diversity and structure of microbial communities. *Environ Microbiol* 16:2659–2671
- Ludwig W, Strunk O, Westram R, Richter L and others (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* 32:1363–1371
- Mao DP, Zhou Q, Chen CY, Quan ZX (2012) Coverage evaluation of universal bacterial primers using the metagenomic datasets. *BMC Microbiol* 12:66
- Meyer JL, Paul VJ, Teplitski M (2014) Community shifts in the surface microbiomes of the coral *Porites astreoides* with unusual lesions. *PLoS ONE* 9:e100316
- Morris RM, Rappé MS, Connon SA, Vergin KL, Siebold WA, Carlson CA, Giovannoni SJ (2002) SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* 420:806–810
- Nelson CE, Alldredge AL, McCliment EA, Amaral-Zettler LA, Carlson CA (2011) Depleted dissolved organic carbon and distinct bacterial communities in the water column of a rapid-flushing coral reef ecosystem. *ISME J* 5:1374–1387
- Orsi WD, Smith JM, Wilcox HM, Swalwell JE, Carini P, Worden AZ, Santoro AE (2015) Ecophysiology of uncultivated marine euryarchaea is linked to particulate organic matter. *ISME J*, doi:10.1038/ismej.2014.260
- Parsons RJ, Breitbart M, Lomas MW, Carlson CA (2012) Ocean time-series reveals recurring seasonal patterns of

- virio plankton dynamics in the northwestern Sargasso Sea. *ISME J* 6:273–284
- Parsons RJ, Nelson CE, Carlson CA, Denman CC and others (2014) Marine bacterioplankton community turnover within seasonally hypoxic waters of a subtropical sound: Devil's Hole, Bermuda. *Environ Microbiol*, doi:10.1111/1462-2920.12445
 - Paver SF, Hayek KR, Gano KA, Fagen JR and others (2013) Interactions between specific phytoplankton and bacteria affect lake bacterial community succession. *Environ Microbiol* 15:2489–2504
 - Pinto AJ, Raskin L (2012) PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets. *PLoS ONE* 7:e43093
 - Porter KG, Feig YS (1980) The use of DAPI for identifying and counting aquatic microflora. *Limnol Oceanogr* 25: 943–948
 - Quast C, Pruesse E, Yilmaz P, Gerken J and others (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:D590–D596
 - Salcher MM, Pernthaler J, Posch T (2011) Seasonal bloom dynamics and ecophysiology of the freshwater sister clade of SAR11 bacteria 'that rule the waves' (LD12). *ISME J* 5:1242–1252
 - Santoro AE, Casciotti KL, Francis CA (2010) Activity, abundance and diversity of nitrifying archaea and bacteria in the central California Current. *Environ Microbiol* 12: 1989–2006
 - Schloss PD, Westcott SL, Ryabin T, Hall JR and others (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75:7537–7541
 - Sipos R, Székely AJ, Palatinszky M, Révész S, Márialigeti K, Nikolausz M (2007) Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targeting bacterial community analysis. *FEMS Microbiol Ecol* 60:341–350
 - Sogin ML, Morrison HG, Huber JA, Welch DM and others (2006) Microbial diversity in the deep sea and the underexplored 'rare biosphere'. *Proc Natl Acad Sci USA* 103: 12115–12120
 - Taylor JD, Cottingham SD, Billinge J, Cunliffe M (2014) Seasonal microbial community dynamics correlate with phytoplankton-derived polysaccharides in surface coastal waters. *ISME J* 8:245–248
 - Treusch AH, Demir-Hilton E, Vergin KL, Worden AZ and others (2012) Phytoplankton distribution patterns in the northwestern Sargasso Sea revealed by small subunit rRNA genes from plastids. *ISME J* 6:481–492
 - Vergin KL, Done B, Carlson CA, Giovannoni SJ (2013a) Spatiotemporal distributions of rare bacterioplankton populations indicate adaptive strategies in the oligotrophic ocean. *Aquat Microb Ecol* 71:1–13
 - Vergin KL, Beszteri B, Monier A, Cameron Thrash J and others (2013b) High-resolution SAR11 ecotype dynamics at the Bermuda Atlantic Time-series Study site by phylogenetic placement of pyrosequences. *ISME J* 7: 1322–1332

*Editorial responsibility: Jed Fuhrman,
Los Angeles, California, USA*

*Submitted: July 3, 2014; Accepted: April 1, 2015
Proofs received from author(s): May 29, 2015*