Journal of Visualized Experiments

Next-generation sequencing of 16S ribosomal RNA gene amplicons

--Manuscript Draft--

Manuscript Number:	JoVE51709R2
Full Title:	Next-generation sequencing of 16S ribosomal RNA gene amplicons
Article Type:	Invited Methods Article - JoVE Produced Video
Keywords:	Metagenomics; Bacteria; 16S rRNA gene; Amplicon sequencing; Next-generation sequencing; benchtop sequencers
Manuscript Classifications:	2.3: Bacteria
Corresponding Author:	Etienne Yergeau National Research Council Canada Montreal, Quebec CANADA
Corresponding Author Secondary Information:	
Corresponding Author E-Mail:	etienne.yergeau@cnrc-nrc.gc.ca
Corresponding Author's Institution:	National Research Council Canada
Corresponding Author's Secondary Institution:	
First Author:	Sylvie Sanschagrin
First Author Secondary Information:	
Other Authors:	Sylvie Sanschagrin
Order of Authors Secondary Information:	
Abstract:	One of the major questions in microbial ecology is "who is there?" This question can be answered using various tools, but one of the long-lasting gold standards is to sequence 16S ribosomal RNA gene amplicons generated by domain-level PCR reactions amplifying from genomic DNA. Traditionally, this was performed by cloning and Sanger (capillary electrophoresis) sequencing of PCR amplicons. The advent of next-generation sequencing has tremendously simplified and increased the sequencing depth for 16S rRNA gene sequencing. The introduction of benchtop sequencers now allows small labs to perform their 16S rRNA sequencing in-house in a matter of days. Here, an approach for 16S rRNA gene amplicon sequencing using a benchtop next-generation sequencer is detailed. The environmental DNA is first amplified by PCR using primers that contain sequencing adapters and barcodes. They are then coupled to spherical particles via emulsion PCR. The particles are loaded on a disposable chip and the chip is inserted in the sequencing machine after which the sequencing is performed. The sequences are retrieved in fastq format, filtered and the barcodes are used to establish the sample membership of the reads. The filtered and binned reads are then further analysed using publically available tools. An example analysis where the reads were classified with a taxonomy-finding algorithm within the software package Mothur is given. The method outlined here is simple, inexpensive and straightforward and should help smaller labs to take advantage from the ongoing genomic revolution.
Author Comments:	
Additional Information:	
Question	**Response**

TITLE: Next-generation sequencing of 16S ribosomal RNA gene amplicons

AUTHORS: Sylvie Sanschagrin and Etienne Yergeau

AFFILIATIONS:
Sylvie Sanschagrin
Energy, Mining and Environment
National Research Council Canada
Montreal, Canada
Sylvie.Sanschagrin@cnrc-nrc.gc.ca

Etienne Yergeau
Energy, Mining and Environment
National Research Council Canada
Montreal, Canada
Etienne.Yergeau@cnrc-nrc.gc.ca

CORRESPONDING AUTHOR: Etienne Yergeau

KEYWORDS: Metagenomics, Bacteria, 16S ribosomal RNA gene, Amplicon sequencing, Next-generation sequencing, benchtop sequencers.

SHORT ABSTRACT:
Characterizing microbial community has been a longstanding goal in environmental microbiology. Next-generation sequencing methods now allow for the characterization of microbial communities at an unprecedented depth with minimal cost and labor. We detail here our approach to sequence bacterial 16S ribosomal RNA genes using a benchtop sequencer.

LONG ABSTRACT:
One of the major questions in microbial ecology is “who is there?” This question can be answered using various tools, but one of the long-lasting gold standards is to sequence 16S ribosomal RNA (rRNA) gene amplicons generated by domain-level PCR reactions amplifying from genomic DNA. Traditionally, this was performed by cloning and Sanger (capillary electrophoresis) sequencing of PCR amplicons. The advent of next-generation sequencing has tremendously simplified and increased the sequencing depth for 16S rRNA gene sequencing. The introduction of benchtop sequencers now allows small labs to perform their 16S rRNA sequencing in-house in a matter of days. Here, an approach for 16S rRNA gene amplicon sequencing using a benchtop next-generation sequencer is detailed. The environmental DNA is first amplified by PCR using primers that contain sequencing adapters and barcodes. They are then coupled to spherical particles via emulsion PCR. The particles are loaded on a disposable chip and the chip is inserted in the sequencing machine after which the sequencing is performed. The sequences are retrieved in fastq format, filtered and the barcodes are used to establish the sample membership of the reads. The filtered and binned reads are then further analysed using publically available tools. An example analysis where the reads were classified

with a taxonomy-finding algorithm within the software package Mothur is given. The method outlined here is simple, inexpensive and straightforward and should help smaller labs to take advantage from the ongoing genomic revolution.

INTRODUCTION:

Metagenomic sequencing is a very powerful technology as it targets the entirety of the genetic information contained in an environmental sample. There are different flavors of metagenomic sequencing, including shotgun sequencing, large-insert libraries and amplicon sequencing. Amplicon sequencing offers the advantage of being relatively inexpensive, fast and able to produce reads from a single genomic region that can be generally aligned. In addition, the data analysis workflow for amplicon sequencing is mostly standardized. However, since it is based on PCR, it has all the biases related to incomplete specificity, incomplete coverage and primer biases [1,2], which makes this approach semi-quantitative at best. Several genomic regions can be targeted for amplicon sequencing including functional genes, but the most popular options are to use marker genes such as the 16S rRNA gene to generate a community profile. Traditionally, 16S rRNA gene amplicon sequencing was carried out using labour-intensive techniques that included cloning in *E. coli*, colony picking and plasmid extraction followed by Sanger sequencing on the isolated plasmids, and, consequently, most studies analyzed fewer than 100 clones per sample. Next-generation sequencing brought two major advances: massive parallelization of the sequencing reactions and, most importantly, clonal separation of templates without the need to insert gene fragments in a host. This has simplified tremendously the sequencing of 16S rRNA gene amplicons, which is now back as a routine feature of many environmental microbiology studies, resulting in a “renaissance” for 16S rRNA gene amplicon sequencing [3].

Since the advent of Roche 454 sequencing in 2005[4], several other next-generation sequencing technologies have appeared on the market (e.g. Illumina, Solid, PacBio). More recently, the introduction of bench-top sequencers brought to small labs the sequencing capacity once exclusive to large sequencing centers. Five benchtop machines are currently available: the 454 GS Junior, the Ion Torrent Personal Genome Machine (PGM) and Proton, and the Illumina MiSeq and NextSeq 500. While all these sequencers offer less reads per run and fewer bases per dollar than most full-scale sequencers they are more flexible, rapid and their low acquisition and run costs makes them affordable for small academic laboratories. Benchtop sequencers are particularly well suited for amplicon, small genome and low-complexity metagenome sequencing in environmental microbiology studies, because this type of studies generally does not require an extreme depth of sequencing. For example, it is generally agreed that for 16S rRNA gene sequencing studies the number of reads per sample is not paramount, as ~1,000 reads can generate the same patterns as multi-million reads datasets[5]. Having said that, benchtop next-generation sequencers still generate large amounts of sequence data, with maximal yields of ~35Mbp (454 GS Junior), ~2Gbp (Ion Torrent PGM), ~10-15Gbp (Ion Torrent Proton), ~10Gbp (Illumina MiSeq) and ~100Gbp (Illumina Next Seq 500), which is more than enough for most environmental microbiology studies.

Next-generation sequencing of 16S rRNA amplicons using benchtop sequencers has been recently applied to a wide variety of environments. For example, the Ion Torrent PGM has been used for community analyses of uranium mine tailings that had particularly high pH and low permeability[6], of recirculating aquaculture systems[7], of hydrocarbon-contaminated Arctic soils [8,9], of oil sands mining affected sediments and biofilms from the Athabasca River [10,11], of the rhizosphere of willows planted in contaminated soils [12], of the human and animal bodies [13-16] and of anaerobic digesters [17].

In this contribution we detail our approach to sequence 16S rRNA gene amplicons in-house using a benchtop next-generation sequencer (the Ion Torrent PGM). After DNA extraction, 16S rRNA genes are amplified using domain-level bacterial primers that contain sequencing adapters and unique, sample-specific sequences (barcodes). The amplicons are purified, quantified and pooled at an equimolar ratio. The pooled samples are then clonally amplified in an emulsion PCR and sequenced. Resulting sequences are analysed using publicly available bioinformatics tools (e.g. Mothur).

PROTOCOL:

1. 16S rRNA gene amplicon library preparation by the fusion method

1.1) Thaw the primers (Forward, F343-IonA-MIDXX: 5'-**CCA TCT CAT CCC TGC GTG TCT CCG AC***T CAG* XXX XXX XXX XTA CGG RAG GCA GCA G-3'; reverse, R533-IonP1: 5'- **CCT CTC TAT GGG CAG TCG GTG AT**A TTA CCG CGG CTG CTG GC-3'; bold: Ion Torrent specific adapters; italics: sequencing key for calibrating signal intensity at the beginning of the sequencing run; regular font: template-specific primers; X…X: 10 bp barcode, see Table 1 for some barcode sequences), the 2X HotStartTaq Plus master mix and the samples. Mix thoroughly before use, except for the master mix that should be mixed gently.

1.2) Prepare a PCR mix (20 µl reactions) with 0.5 µM of the reverse primer, 0.4 mg/ml of Bovine Serum Albumine (BSA) and 1X HotStartTaq Plus Master mix. Add 18.5 µL of the master mix to each PCR tube.

1.3) Add a forward primer (0.5 µl at 20mM) with a different barcode for each of the samples that will be sequenced together. Add 1 µL of sample (1-10ng/µl) to the PCR tube.

1.4) Place the tubes in a PCR machine and run the following program: initial denaturation at 95°C for 5 min; 25 cycles of 95°C for 30s, 55°C for 30s and 72°C for 45s; final elongation at 72°C for 10 min. PCR products can be kept at 4°C overnight or frozen until used.

2. Amplicon purification, quantification and pooling

2.1) Prepare a 2% agarose gel using the largest combs available. Add 5 µL of 6X gel loading dye to the reaction and load on the gel, separating each sample by an empty well. Run the gel at 70V for 50 min.

2.2) Take a picture of the gel and cut the bands (expected size around 250 bp: 190 bp product plus 63 bp adapters and barcodes) using a clean scalpel.

2.3) Weigh the excised bands and proceed with the gel extraction and purification of the PCR products (with e.g. Qiaquick Gel Extraction kit).

2.4) Quantify each PCR product with PicoGreen and make dilutions to obtain a stock solution at $5x10^9$ molecule/µL (approximately 1ng/µl). If some samples show lower amplification, adjust the dilutions to a lower concentration, keeping in mind that the lowest concentration suitable for subsequent procedures is $1.57x10^7$ molecule/µl (26pM).

2.5) Pool 10 µL of each diluted PCR product to obtain an equimolar pool of samples. Prepare one separate pool for each of the planned sequencing run. The pooled PCR products can be kept in the freezer until used.

3. Emulsion PCR and sequencing

3.1) Perform emulsion PCR

3.1.1) Dilute the pooled PCR products to 26 pM in a volume of 25 µl. Thaw reagents from an Ion PGM template kit.

3.1.2) Prepare the emulsion PCR mix (reagents provided in the kit) in a PCR hood and add the pooled PCR products and the sphere particles (provided) on the bench. Mix thoroughly and insert the mixture in a filter cartridge and carefully top with emulsion oil (provided).

3.1.3) Slowly reverse the filter cartridge and load on the automated emulsion PCR apparatus (e.g. Ion One Touch 2 instrument). Select the appropriate program and start the procedure.

3.2) After the automated emulsion PCR has completed, remove the supernatant from the collection tubes and retrieve the sphere particles at the bottom of the tubes. Re-suspend the spheres in 100µl of wash solution (provided in the kit) and take a 2.0 µl aliquot for library quantification.

3.3) Perform the library quantification.

3.3.1) Mix 100µl of SSPE buffer (150 mM NaCl, 10mM $NaPO_4$, 1mM Na_2-EDTA) with 2µl of Adapter B'-Fam (5'-*FAM*-CTG AGA CTG CCA AGG CAC ACA GGG GAT AGG-3') probe and 2µl of Adapter A-Cy5 (5'-*Cy5*-CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG-3') probe. Add 52 µl of this mixture to the 2.0µl aliquot taken in step 3.2 and add 2µl of wash solution (from the template kit) to the remaining 52µl as a blank for quantitation.

3.3.2) Incubate at 95°C for 2 minutes, then at 37°C for 2 minutes. Transfer the mixtures in 1.5ml tubes.

3.3.3) Add 1.0 ml of TEX buffer (10mM Tris, 1mM EDTA, 0.01% Triton X-100, pH 8.0), mix and centrifuge at 15,500 X g for 3 minutes at room temperature. Remove the supernatant and leave 20µl in each tube. Repeat this washing procedure two times.

3.3.4) Add 180µl of TEX buffer, resuspend the pelleted particles and transfer into 500µl Qubit tubes.

3.3.5) Measure fluorescence for FAM and Cy5 using a Qubit apparatus and calculate the ratio of positive spheres using the calculator available on the Ion Community website.

3.4) Enrich the remainder of the sphere particles for positive spheres (containing amplified DNA) using an automated enrichment system (e.g. Ion One Touch Enrichment System).

3.4.1) Pipet reagents in the appropriate wells of the provided 8-well strip: Well 1: sphere particles from step 3.2; Well 2: pre-washed MyOne Streptavidin C1 beads (13 µl); Well 3, 4 and 5: wash solution (provided in the template kit); Well 6: empty; Well 7: melt-off solution (125mM NaOH, 0.1% Tween 20); Well 8: empty.

3.4.2) Install a tip on the pipetting arm and a 0.2ml tube for sample collection. Start the instrument. At the end of the enrichment, add 10µl of neutralizing solution (provided). The enriched beads can be kept a 4°C for up to 15 days.

3.5) Initialize the sequencing instrument

3.5.1) Connect to the sequencing instrument and prepare a sequencing run plan, specifying the details of the sequencing run.

3.5.2) Install the wash solution #2 (provided in the Ion PGM sequencing kit), the wash solution #1 (350µl 100mM NaOH) and the auto-pH buffer solution (provided).

3.5.3) Start the initialization procedure and when prompted add the dATP, dGTP dCTP and dTTP nucleotides (provided) in their respective 50 ml tubes. Screw the tubes onto the sequencing machine and continue the initialization procedure. When the initialization is completed, start the run procedure immediately.

3.6) Prepare samples for sequencing and load the chip

3.6.1) Collect the enriched spheres from step 3.4 at the bottom of the tube by centrifuging at 15,500 X g for 1.5 minutes. Remove the supernatant leaving exactly 3 µl.

3.6.2) Add 3 µl of sequencing primer (provided in the sequencing kit), mix thoroughly by pipetting up and down to resuspend the particles. Incubate at 95°C for 2 minutes, then at 37°C for 2 minutes.

3.6.3) Add 1µl of polymerase (provided), mix well and incubate at room temperature for 5 minutes. Proceed to load the enzyme-sphere mixture on the sequencing chip within 30 minutes.

3.6.4) During incubation, load the sequencing chip on the sequencer and perform a chip quality check. Remove the liquid from the chip by centrifugation and load the sample in the chip by slowly pipeting down the enzyme-sphere mix (7µl) and avoiding the introduction of bubbles in the chip.

3.6.5) Centrifuge the chip three times for 1 min in a micro-centrifuge. Change the orientation of the chip at each centrifugation and mix by pipetting up and down in between each run.

3.6.6) Load the chip in the instrument and start the run.

4. Basic sequence data analysis

4.1) Connect to the sequencing data server and download the fastq file. Create a tab-delimited "oligos" file containing the primer and barcode information (see example in Table 1). Download the Greengenes reference files from the Mothur website (http://www.mothur.org/wiki/Taxonomy_outline).

4.2) Launch Mothur and perform analyses

4.2.1) Convert the fastq file to a fasta and a quality file using the following command: *fastq.info(fastq=test.fastq)*

4.2.2) Determine the group membership of each sequence and trim the sequences using the following command: *trim.seqs(fasta=test.fasta, oligos=barcode.txt, qfile=test.qual, qwindowaverage=20, qwindowsize=50, minlength=150, keepforward=T)*
NOTE: The various options can be modified according to the stringency desired

4.2.3) Classify the sequences in the greengenes taxonomy using the following command: *classify.seqs(fasta=test.trim.fasta, method=wang, group=test.groups, template=gg_99_otus.fasta, taxonomy=gg_99_otus.tax, cutoff=50)*

REPRESENTATIVE RESULTS:
After purification on gel, with 25 cycles of PCR amplification, the amplification products are usually at a concentration of 0.2-10.0 ng in 50µl of water. This may vary widely depending on the starting DNA concentration, the type of sample and the purification kit used. It is recommended to keep the number of PCR cycles to the lowest possible to avoid chimera

formation and decrease amplification biases, keeping in mind that all samples should be amplified using the same number of cycles. To minimize the number of polyclonal reads and empty spheres and maximize the number of reads, the Qubit ratio should be between 0.1 and 0.3 and the FAM fluorescence should be above 200. Using a 314 chip on an Ion Torrent PGM, the average output is around 0.3-0.5M good quality reads after filtering of the results in Mothur. Table 2 shows a typical breakdown of the number of reads after each step of the procedure for a run containing 36 multiplexed environmental samples amplified with primers targeting the V3-4 region of the 16S and analysed using Mothur. In Mothur, the trim.seqs procedure generate a "*.trim.fasta" file containing the sequences that passed the quality filters and a "*.scrap.fasta" that contains the sequences that did not pass the quality filters along with the reason for rejection in the sequence header. When supplied with barcodes in the "oligos" file, this command will also generate a "*.groups" file that contains the group membership of every sequence based on the barcode sequence. The classify.seqs procedure generates a ".tax.summary" that can be opened in Excel. This file contains the summary of the taxonomic affiliation (in lines) for each of the samples (in columns). This file can be used for downstream statistical analyses and to visualize community composition at various taxonomic levels. The ".taxonomy" file contains the detailed taxonomic affiliation for each sequence. The average community composition at the phylum/class level across all 36 samples is shown in Fig. 1.

Figure Legends:
Figure 1: Average community composition at the phylum/class level across all samples.

Description of tables:
Table 1: Example "oligos" file for use in Mothur
Table 2: Number of reads produced from a typical run for 36 environmental samples multiplexed on one Ion Torrent 314 chip.

DISCUSSION:

The method presented here is straightforward and inexpensive, and should allow many laboratories to access the power of metagenomic sequencing. Although it varies depending on the sequencing platform used, once the libraries are constructed very little hands-on time is required, with most of the process being automatized. For the sequencing platform used here (Ion Torrent PGM), the complete procedure can be performed within two days of work. At the moment of writing (September 2013), the reagent costs related to the example detailed above were as follows: PCR amplification of 36 samples: $25, gel purification and PicoGreen DNA quantitation of 36 samples: $125, emulsion PCR for one pooled amplicon sample: $150 and sequencing reagents: $250, for a total of $550 or $15 per sample or $0.0015 per quality-filtered read. This price does not include instrument service contract, instrument depreciation, technician salary and laboratory space usage.

One of the most important steps is to pool all the products in an equimolar ratio, in order to retrieve similar number of reads for each of the samples. PicoGreen quantification was used here, but other methods might be suitable, though less accurate (e.g. UV quantification, gel-based quantification). Even by doing the most accurate quantification and pooling, there is some variability in the number of reads per sample, and in the typical run detailed in Table 2, it ranges from 4,380 to 32,750 reads, with an average of 10,338 reads. If processing large number of samples (more than 40-50), single-column gel-purification can be replaced by gel-purification in plate or purification using beads with a stringent size cutoff (e.g. AMPure beads).

To date, the most used next-generation sequencing technology for the 16S rRNA gene is 454. The Ion Torrent sequencing technology used in this protocol is conceptually very similar to 454 and both technologies are prone to the same type of sequencing errors. Not surprisingly, it was shown that Ion Torrent sequencing resulted in sequencing results very similar to 454 sequencing[10]. Recently, many researchers have explored the use of Illumina technology for 16S rRNA gene amplicon sequencing [18,19]. In any case, it would be easy to adapt the current protocol for other benchtop sequencers like the Illumina MiSeq or the 454 GS Junior by changing the fusion primer sequences to match the adapters and barcodes needed for these sequencing technologies, like in the method recently described for the Illumina MiSeq [19]. Alternatively, researchers could follow steps 1 and 2 of the protocol detailed here and send the pooled amplicons to a sequencing center where the emulsion PCR and sequencing would be performed.

The 16S rRNA gene reads were trimmed and classified using Mothur, but many other analyses can be performed on 16S rRNA gene amplicons. For instance, beta diversity can be evaluated by calculating the Unifrac distances between each sample pair using the procedure outlined at http://unifrac.colorado.edu/ [20]. Alpha diversity indices and number of operational taxonomical units of each sample can be calculated using tools within QIIME like AmpliconNoise [21] or using the procedure outlined by Huse et al.[22] and available within Mothur.

The primers used here amplified the variable regions 3 and 4 from the 16S rRNA gene, but many other regions could be targeted. In present study, 16S rRNA genes were amplified from

plant material and the choice of primer was made to avoid amplification of chloroplast 16S rRNA gene [23,24]. There are a wide variety of other primers available that vary in term of the product length, taxonomic power and usefulness[25,26]. However, in all cases 200-400 bp reads of the 16S rRNA gene cannot be reliably classified at the species level, and analyses are limited to the genus and higher taxonomical levels. Other genes could be more appropriate if species level information is needed, like the cpn60 and rpoB genes[27,28]. Future drastic drops in the cost of sequencing and increases in the power of analytical tools might make it feasible to replace 16S rRNA gene sequencing by shotgun metagenomics, but until then 16S rRNA gene sequencing remains the gold standard of environmental microbiology.

ACKNOWLEDGMENTS:
Development of the method presented here has been carried out with various sources of funding, including Genome Canada and Genome Quebec, Environment Canada STAGE program and internal NRC funds.

DISCLOSURES:
The authors have nothing to disclose.

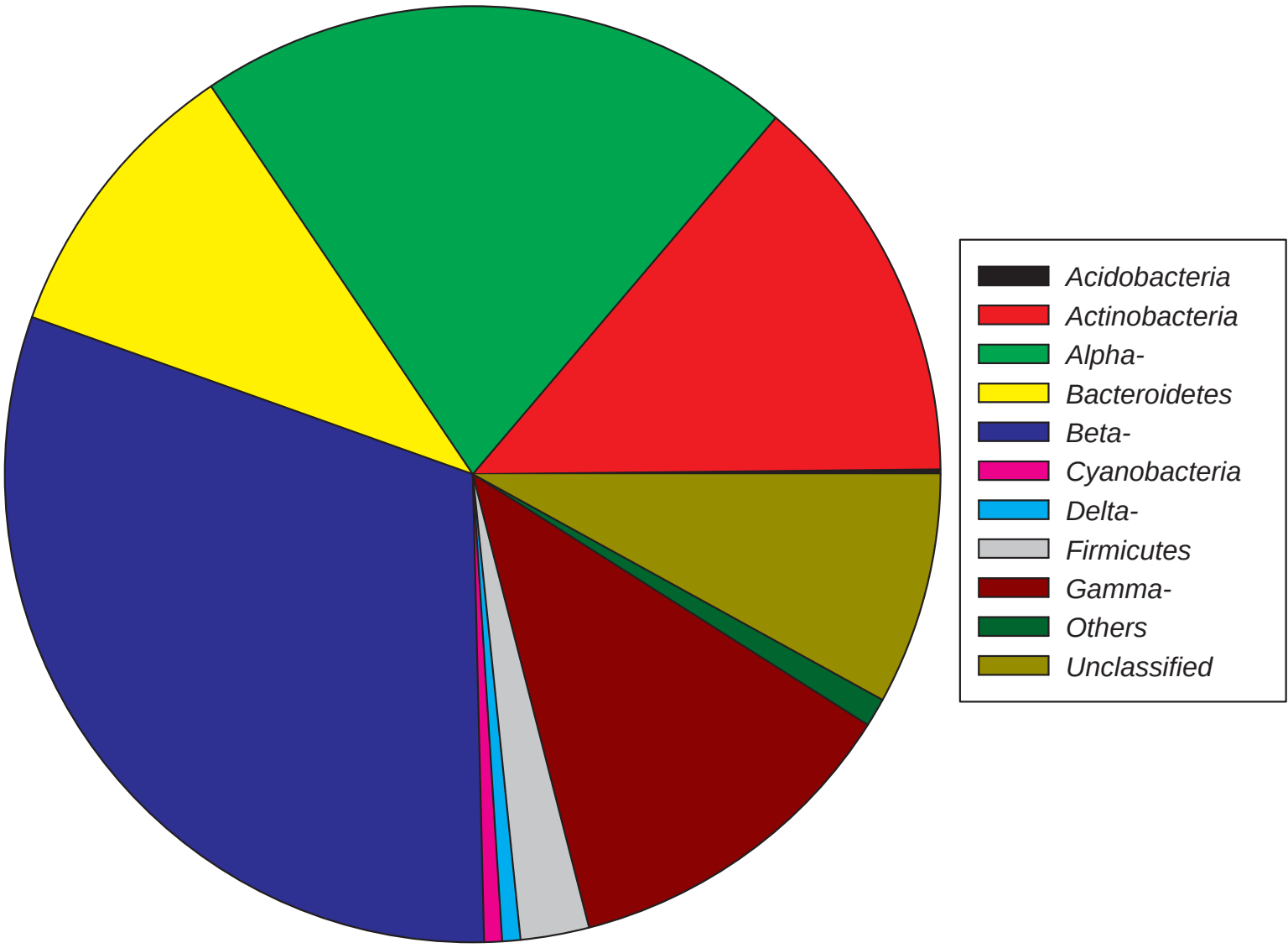
REFERENCES

1 Sipos, R., Szekely, A., Revesz, S. & Marialigeti, K. Addressing PCR biases in environmental microbiology studies. *Methods in molecular biology* **599**, 37-58, doi:10.1007/978-1-60761-439-5_3 (2010).
2 Schloss, P. D., Gevers, D. & Westcott, S. L. Reducing the Effects of PCR Amplification and Sequencing Artifacts on 16S rRNA-Based Studies. *PLoS ONE* **6**, e27310, doi:10.1371/journal.pone.0027310 (2011).
3 Tringe, S. G. & Hugenholtz, P. A renaissance for the pioneering 16S rRNA gene. *Current Opinion in Microbiology* **11**, 442–446, doi:10.1016/j.mib.2008.09.011 (2008).
4 Margulies, M. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376–380, doi:10.1038/nature03959 (2005).
5 Kuczynski, J. *et al.* Microbial community resemblance methods differ in their ability to detect biologically relevant patterns. *Nature Methods* **7**, 813-819, doi:10.1038/nmeth.1499 (2010).
6 Bondici, V. F. *et al.* Microbial communities in low-permeability, high pH uranium mine tailings: characterization and potential effects. *Journal of Applied Microbiology* **113**, 1671-1686, doi:10.1111/jam.12180 (2013).
7 Auffret, M. *et al.* Impact of water quality on the bacterial populations and off-flavours in recirculating aquaculture systems. *FEMS Microbiology Ecology* **84**, 235-247, doi:10.1111/1574-6941.12053 (2013).
8 Bell, T. H., Yergeau, E., Juck, D., Whyte, L. G. & Greer, C. W. Alteration of microbial community structure affects diesel biodegradation in an Arctic soil. *FEMS Microbiology Ecology* **85**, 51-61, doi:10.1111/1574-6941.12102 (2013).
9 Bell, T. H. *et al.* Predictable bacterial composition and hydrocarbon degradation in Arctic soils following diesel and nutrient disturbance. *ISME J.* **7**, 1200-1210, doi:10.1038/ismej.2013.1 (2013).

10 Yergeau, E. *et al.* Next-generation sequencing of microbial communities in the Athabasca River and its tributaries in relation to oil sands mining activities. *Applied and Environmental Microbiology* **78**, 7626-7637, doi:10.1128/AEM.02036-12 (2012).
11 Yergeau, E. *et al.* Aerobic biofilms grown from Athabasca watershed sediments are inhibited by increasing bituminous compounds concentrations. *Applied and Environmental Microbiology* **79**, 7398-7412, doi:10.1128/AEM.02216-13 (2013).
12 Yergeau, E., Sanschagrin, S., Maynard, C., St-Arnaud, M. & Greer, C. W. Microbial expression profiles in the rhizosphere of willows depend on soil contamination. *ISME J.* **8**, 344-358, doi:10.1038/ismej.2013.163 (2013).
13 Deagle, B. E., Thomas, A. C., Shaffer, A. K., Trites, A. W. & Jarman, S. N. Quantifying sequence proportions in a DNA-based diet study using Ion Torrent amplicon sequencing: which counts count? *Molecular ecology resources* **13**, 620-633, doi:10.1111/1755-0998.12103 (2013).
14 Milani, C. *et al.* Assessing the Fecal Microbiota: An Optimized Ion Torrent 16S rRNA Gene-Based Analysis Protocol. *PLoS ONE* **8**, e68739, doi:10.1371/journal.pone.0068739 (2013).
15 Jünemann, S. *et al.* Bacterial community shift in treated periodontitis patients revealed by Ion Torrent 16S rRNA gene amplicon sequencing. *PLoS ONE* **7**, e41606, doi:10.1371/journal.pone.0041606 (2012).
16 Petrof, E. *et al.* Stool substitute transplant therapy for the eradication of Clostridium difficile infection: 'RePOOPulating' the gut. *Microbiome* **1**, 3, doi:10.1186/2049-2618-1-3 (2013).
17 Whiteley, A. S. *et al.* Microbial 16S rRNA Ion Tag and community metagenome sequencing using the Ion Torrent (PGM) Platform. *J. Microbiol. Meth.* **91**, 80-88, doi:10.1016/j.mimet.2012.07.008 (2012).
18 Caporaso, J. G. *et al.* Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **6**, 1621-1624, doi:10.1038/ismej.2012.8 (2012).
19 Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*, doi:10.1128/AEM.01043-13 (2013).
20 Hamady, M., Lozupone, C. & Knight, R. Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. *ISME J.* **4**, 17-27, doi:10.1038/ismej.2009.97 (2010).
21 Quince, C., Lanzen, A., Davenport, R. J. & Turnbaugh, P. J. Removing noise from pyrosequenced amplicons. *BMC Bioinformatics* **12**, 38, doi:10.1186/1471-2105-12-38 (2011).
22 Huse, S. M., Welch, D. M., Morrison, H. G. & Sogin, M. L. Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environmental Microbiology* **12**, 1889-1898, doi:doi: 10.1111/j.1462-2920.2010.02193.x (2010).
23 Edwards, J. E., Huws, S. A., Kim, E. J. & Kingston-Smith, A. H. Characterization of the dynamics of initial bacterial colonization of nonconserved forage in the bovine rumen. *FEMS Microbiology Ecology* **62**, 323-335, doi:10.1111/j.1574-6941.2007.00392.x (2007).
24 Rastogi, G., Tech, J. J., Coaker, G. L. & Leveau, J. H. J. A PCR-based toolbox for the culture-independent quantification of total bacterial abundances in plant environments. *J. Microbiol. Meth.* **83**, 127-132, doi:10.1016/j.mimet.2010.08.006 (2010).
25 Baker, G. C., Smith, J. J. & Cowan, D. A. Review and re-analysis of domain-specific 16S primers. *J. Microbiol. Meth.* **55**, 541–555, doi:10.1016/j.mimet.2003.08.009 (2003).
26 Schloss, P. D. The effects of alignment quality, distance calculation method, sequence filtering, and region on the analysis of 16S rRNA gene-based studies. *PLoS computational biology* **6**, e1000844, doi:10.1371/journal.pcbi.1000844 (2010).

27 Links, M. G., Dumonceaux, T. J., Hemmingsen, S. M. & Hill, J. E. The Chaperonin-60 Universal Target Is a Barcode for Bacteria That Enables *De Novo* Assembly of Metagenomic Sequence Data. *PLoS ONE* **7**, e49755, doi:10.1371/journal.pone.0049755 (2012).
28 Vos, M., Quince, C., Pijl, A. S., de Hollander, M. & Kowalchuk, G. A. A Comparison of *rpoB* and 16S rRNA as Markers in Pyrosequencing Studies of Bacterial Diversity. *PLoS ONE* **7**, e30600, doi:10.1371/journal.pone.0030600 (2012).

Click here to download Figure: PieChart.EPS



Click here to download Table: Table1.xlsx

forward	TACGGRAGGCAGCAG	
barcode	CTAAGGTAAC	Sample01
barcode	TAAGGAGAAC	Sample02
barcode	AAGAGGATTC	Sample03
barcode	TACCAAGATC	Sample04
barcode	CAGAAGGAAC	Sample05
barcode	CTGCAAGTTC	Sample06
barcode	TTCGTGATTC	Sample07
barcode	TTCCGATAAC	Sample08
barcode	TGAGCGGAAC	Sample09
barcode	CTGACCGAAC	Sample10

Click here to download Table: Table2.xlsx

	# of reads	% of previous step
Number of wells	1,262,519	-
Wells with beads	1,114,108	88.20%
Beads with templates	1,112,746	99.90%
Monoclonal beads	826,805	74.30%
Good quality reads (Output from the sequencer)	782,204	94.60%
Pass Mothur filters (min. avg. quality score of 20 over a 50bp window, min. length of 150bp)	372,168	47.60%
Classified at the phylum level in GreenGenes (50% confidence threshold)	342,171	91.90%
Classified at the family level in GreenGenes (50% confidence threshold)	316,512	92.50%
Classified at the genus level in GreenGenes (50% confidence threshold)	289,899	91.60%

Avg. per sample
35,070
30,947
30,910
22,967
21,728
10,338
9,505
8,792
8,053

*Excel Spreadsheet- Table of Materials/Equipment

Click here to download Excel Spreadsheet- Table of Materials/Equipment: 20131001-Reagent_list_JoVE.xlsx

Reagent	Company	Catalog Number
Ion 314 Chip Kit v2	Life Technologies	4482261
Ion PGM Sequencing 200 Kit v2	Life Technologies	4482006
Ion PGM Template OT2 200 Kit	Life Technologies	4480974
HotStarTaq Plus Master Mix Kit	Qiagen	203646
Primers and probes	IDT	NA
Qiaquick Gel Extraction Kit	Qiagen	28704
BSA 20 mg/ml	Roche	10,711,454,001
Dynabeads MyOne Streptavidin C1	Life Technologies	65001

National Research Council Canada | Conseil national de recherches Canada

Licence to Publish (Crown Copyright)

Manuscript number _____

No changes are to be made to this licence without prior consent of NRC Legal Services.

To: MyJoVE Corp.
One Alewife Center, Suite 200
Cambridge, MA 02140
USA

(contact name and address of publisher)

This refers to the manuscript entitled (hereinafter the "manuscript"):

Next-generation sequencing of 16S ribosomal RNA gene amplicons

written by the author(s): ☐ Check if you need more lines to enter all the author names of this manuscript.

Sylvie Sanschagrin
Etienne Yergeau

to be published in:

Journal of Visualized Experiments

1. Definitions

Accepted Manuscript:
The version of a manuscript that has been accepted for publication following peer review. Content and layout follow publisher's submission requirements.

Version of Record:
A fixed version of an Accepted Manuscript that has been made available by a publisher. This includes any "early release" article that is formally identified as being published even before the compilation of a volume issue and assignment of associated metadata, as long as it is citable via some permanent identifier(s). This does not include any "early release" article that has not yet been "fixed" by processes that are still to be applied, such as copy-editing, proof corrections, layout, and typesetting.

2. The authors (except those listed in this paragraph) contributed to the manuscript on behalf of the National Research Council of Canada ("NRC"), thereby establishing a copyright belonging to the Crown in Right of Canada, that is, to the Government of Canada. NRC is not able to deal with the rights relating to the contribution of any co-authors listed below, and you should seek their permission separately [*may say* none]. ☐ Check if you need more lines to enter all the author names of this manuscript.

None

(list author names as well as their respective organization, email address and phone number, if available)

3. With respect to the Crown's copyright, NRC grants you permission to publish the Version of Record in the publication named above, on the condition that clear attribution is given to the authors and the National Research Council of Canada. After that publication, you may re-publish the Version of Record in any form or medium, with the same condition about attribution.

4. You are hereby [*specify* hereby *or* not] granted the right to license others to copy or publish the Version of Record. (If this right is granted, it is subject to the same condition about attribution).

5. NRC, for itself and those authors not listed in section 2, warrants that:
   a. The manuscript is the original work of the named author(s).
   b. A Version of Record has not been published elsewhere and will not be permitted to be published elsewhere, if you accept the manuscript for publication.
   c. The manuscript contains no infringing, libellous or other unlawful statements.

6. NRC retains the following rights, provided that when reproducing the manuscript or extracts from it, the publisher is acknowledged:
   a. To post a copy of the Accepted Manuscript on the NRC's website, the author(s)' funding body's designated archive and any other open access digital repository, subject to the embargo period specified herewith: [*specify* none, 3 months, 6 months *or* 12 months], none . For the purposes of clarity, where the embargo period is left blank, NRC will assume that a copy of the Accepted Manuscript may be posted immediately upon the date of acceptance of the manuscript for publication. A copy of the Version of Record will only be posted if so directed by the publisher. For this article, the publisher's preference is that we [*specify:* not replace *or* replace] the Accepted Manuscript with the Version of Record once the Version of Record becomes available.
   b. To reproduce the Accepted Manuscript for non-commercial purposes including other formats and other forms of expression.
   c. To reproduce the Accepted Manuscript for the purpose of education or research by the Government of Canada and to permit other institutions with which the author(s) is affiliated to reproduce the Accepted Manuscript for the purpose of education or research.
   d. To reuse figures, tables or an abstract created by the author(s).
   e. To authorize others to make any non-commercial use of the Accepted Manuscript so long as the author(s) receives credit as author(s) and the publication in which the Version of Record has been published is linked to the source of publication of the Version of Record (link to the publisher's or journal site).

7. It is believed that this authorization will provide you with all the scope of authority you require from the Government of Canada, but it does NOT transfer the copyright to you.

Signed at on

(City and Province/State) (Date)

NATIONAL RESEARCH COUNCIL OF CANADA

Per: A. G. Rosenthal.

(Director General)

(Institute or Branch)

Cover letter and rebuttal Comments
Click here to download Rebuttal Comments: 20140117-JoVeCoverLetter-v3.docx

Friday, January 24, 2014

Sephorah Zaman, Ph.D.
Science Editor
JoVE
17 Sellers St., Cambridge, MA 02139

Subject: Manuscript re-submission

Dear Dr. Zaman,

Following reviewer's comments, we have substantially modified our manuscript JoVE 51709 (see point by point revision below) and we would like to re-submit this manuscript entitled "Next-generation sequencing of 16S ribosomal RNA gene amplicons" and co-authored by Sylvie Sanschagrin and Etienne Yergeau for consideration in the Journal of Visualized Experiments. Ms. Sanschagrin optimized the protocols in the laboratory and wrote the protocol section, while I supervised her in the laboratory, conceptualized the protocols and wrote the remaining of the manuscript.

This protocol is ideally suited for JoVE's unique multimedia format as it is a seemingly straightforward protocol, but with a lot of tricky steps. Most researchers also never have the opportunity to visualize the sequencing process itself as it is often carried out by an external sequencing center. In the manuscript we also mention that 16S rRNA sequencing is the gold standard for determining community composition in environmental microbiology, which should make our protocol interesting to a wide audience.

Most sincerely,

Etienne Yergeau, PhD
National Research Council of Canada,
6100 Royalmount Ave.
Montréal (Qc)
H4P 2R2, Canada
514-255-5584
Etienne.Yergeau@cnrc-nrc.gc.ca

Editorial comments:

1) All of your previous revisions have been incorporated into the most recent version of the manuscript. Please download this version of the Microsoft word document from the "file inventory" to use for any subsequent changes.

Done.

2) Please do not underline any text in the protocol. Bold lettering, italics or a different font size may be used instead.

Corrected.

3) Editor removed the company weblink from step 3.3.5 since JoVE is unable to publish manuscripts containing commercial sounding language. The calculator may instead be listed in the Table of Reagents/Equipment with the weblink (http://ioncommunity.lifetechnologies.com/) included in the "Comments" column of the table.

Ok.

4) Please specify the centrifugation speed in steps 3.6.4 and 3.6.5.

This centrifugation is performed in a micro-centrifuge in which it is impossible to set the centrifugation speed.

5) Please simplify step 3.1 of the protocol so that individual steps contain 2-3 actions per step and no more than 4 sentences per step. We recommend that you split your longer steps into sub-steps numbered accordingly, i.e. step 1 is followed by sub-step 1.1 then 1.1.1 as necessary.

Done.

6) After adding spaces between steps, the protocol length is over our 3 page limit for filmable content. Please use yellow highlighting to identify a total of 2.75 pages of protocol text (which includes headings and spaces) to identify which portions of the procedure are most important to include in the video, i.e. which steps should be visualized to tell the most cohesive story of your protocol steps. See JoVE's instructions for authors for more clarification and remember that the non-highlighted protocol steps will remain in the manuscript and therefore will still be available to the reader.

Done.

7) Please update the title of your article in the "License to Publish" form.

Done.

8) Please disregard the comment below if all of your figures are original.
If you are re-using figures from a previous publication, you must obtain explicit permission to re-use the figure from the previous publisher (this can be in the form of a letter from an editor or a link to the editorial policies that allows you to re-publish the figure). Please upload the text of the re-print

permission (may be copied and pasted from an email/website) as a Word document to the Editorial Manager site in the "Supplemental files (as requested by JoVE)" section. Please also cite the figure appropriately in the figure legend, i.e. "This figure has been modified from [citation]."

All our figures are original.

Reviewer #1:

Manuscript Summary:
N/A

Major Concerns:
Most of the methods are not provided in sufficient detail to replicate, and most of the method that is described is already provided by the instrument manufacturer.

We disagree. We re-read carefully the methods and we find that most steps are sufficiently detailed. We could not reproduce protocols from the manufacturers and summarized most steps with enough details to be reproduced.

The choice of primer selection and bioinformatics methods are not adequately explained.

Our goal was to give an example of the analytical methods used and provide the reader with potential alternatives for primers, bioinformatic pipelines and even sequencing platforms. As such we don't feel the need to justify extensively the choice of the methods, but we added references in the discussion for the readers to make their own choice of primer, sequencing technology and analysis pipeline.

Minor Concerns:
N/A

Reviewer #2:
Manuscript Summary:
The manuscript presents a workflow for analyzing bacterial community composition using 16S rRNA gene sequencing on the Ion Torrent DNA sequencing platform.

Major Concerns:
1. In the introduction, it would be helpful to include a few references that show some of the issues with PCR-based analysis of communities. This would be particularly helpful for researchers who are not yet familiar with the approach and the caveats.

We added two recent references related to PCR issues.

2. In the introduction, the authors suggest that depth of sequencing is not important for bacterial community analysis. This may be accurate if one is only interested in the most abundant taxonomic groups, but the I am sure the authors are aware that there is an exceptionally high diversity of less abundant species in all ecosystems and that deep sequencing has provided much insights into the composition of this "rare biosphere". The authors must add a couple references here (at least Sogin's original paper) and perhaps a more recent example (ie Gibbons et al. PNAS, Evidence for a persistent

seed bank throughout the global ocean.

We already mentioned that environmental microbiology studies GENERALLY don't need an extreme depth of sequencing. The rare biosphere is an interesting concept, but is far from the preoccupation of most microbial ecology studies, as it is not well defined what the role of these rare organisms might be. We therefore maintain our statement that extreme depth of sequencing is generally not required.

We prefer not referring to the Sogin paper as the rare biosphere detected there was subsequently shown to be caused by sequencing artefacts, see:

Kunin V, Engelbrektson A, Ochman H, Hugenholtz P (2010). Wrinkles in the rare biosphere: pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ Microbiol* **12:** 118-123.
Huse SM, Welch DM, Morrison HG, Sogin ML (2010). Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environ Microbiol* **12:** 1889-1898.

3. It would be helpful if the authors explained which region of the 16S rRNA gene they are targeting, and why this region was selected. Okay, I see that this issue is addressed in the discussion, but perhaps it should be moved to the introduction and a reference should be added that compares the usefulness of the different regions of the 16S. Pat Schloss has a reference that may be useful here.

We added the suggested reference and several sentences in this paragraph, but decided to keep this section in the discussion.

4. Other than the PCR, the protocol is based on kits. It would be helpful if the authors were clear about which kit was being used when. For example, 3.3 could say "Perfrom the library quantification using Qubit and DNA quantification kit XXX" Also, I don't see these reagents listed at the end of the manuscript! They should be included.

[**Editorial Comment:** Please keep JoVE's protocol requirements in mind as you address the above comment as well as some of Reviewer 2 comments. No commercial sounding language should be added to the protocol, including company and brand names. All commercial products should be sufficiently referenced in the table of materials/equipment.]

The Qubit quantification was performed using custom-made oligos. As for the rest of the protocol, it is only the emulsion PCR and sequencing steps that rely on kits, the PCR amplification, purification quantification and data analysis are customized. All the kits used are mentioned in the reagent table.

Minor Concerns:
na

Additional Comments to Authors:
na

Reviewer #3:
Review of Sanschagrin and Yergeau - "Next-generation sequencing of 16S rRNA gene amplicons"

This chapter is a straightforward description of the workflow for generating 16S rRNA gene amplicon sequence libraries using the Ion Torrent Personal Genome Machine (PGM) platform. Details are

provided for primer design, PCR amplification and sequencing. Some discussion of the 16S rRNA gene amplicon conceptual approach for microbial community characterization is also provided. The manuscript is acceptable, but not particularly novel; there are many such manuscripts describing this style of workflow. Specific comments are provided below.

We agree that similar worflows are detailed elsewhere. The novelty resides in the use of video for demonstrating this particular method.

L25: "...methods now ALLOW for the characterization..."

Corrected.

L27: Spell out ribosomal RNA the first time; also probably in the title.

Done.

L31-32: "...standards is to sequence 16S rRNA gene amplicons generated by domain-level PCR reactions amplifying from genomic DNA. Traditionally, this was performed by cloning and Sanger (capillary electrophoresis) sequencing of PCR amplicons."

Corrected.

L36: Perhaps remove personal pronouns from the text

Corrected.

L43: "An example analysis where the reads were classified with a taxonomy-finding algorithm within the software package Mothur...

Corrected.

L49: I'm a bit picky about the use of the word metagenomic for domain-level amplicon sequencing (i.e. I don't use it).

Recent reviews on metagenomics include amplicon sequencing.

L52: Not sure why amplicon sequencing is more straightforward that shotgun metagenomics. It's faster, for sure. Also, another advantage of amplicon sequencing (particularly 16S) is that the data analysis workflow is also really standardized.

We replaced straightforward by fast and added a sentence about standardization of methods.

L58: "...traditionally, 16S rRNA gene AMPLICON sequencing was..."

Corrected.

L60: "..sequencing on isolated plasmids; consequently, most studies analyzed FEWER than 100 clones

per sample."

Corrected.

L68: "...Roche 454 pyrosequencing..."

Corrected.

L73: replace "less" with "fewer". Also "bases" instead of "bp"

Corrected.

L76-77: I think this is quite debatable; certain applications in microbial ecology demand very high sequencing output - for example, shotgun metagenomics. Here, I think small benchscale sequencers cannot match the output and price from the larger sequencers. 5-10 Gb per sample might easily be used for environmental metagenome.

We re-phrased that to specifically refer to amplicon, small genomes and low diversity metagenome sequencing.

L81-82: Refer to Ion Torrent PGM and Ion Torrent Proton. Proton is now generating 10-15 Gb with the PI chip v2; MiSeq is probably closer to 10 Gb, assuming 50 M reads of 200 bases (realistically, the 300 base reads generate 200 bases of quality data).

Corrected.

L86: I think "recently" is not the right word; this has been done for years already.

We disagree: benchtop sequencers are on the market for only 2-3 years.

L86: "a wide variety of environments from deep ocean sediments to human microbiome (REFs). More recently, the Ion Torrent PGM has also been used for community analyses of diverse environments as well."

We modified this sentence.

L101: "PGM"

We added the name of the sequencer used for the protocol.

L101: "After DNA extraction, 16S rRNA genes are amplified using domain-level bacterial primers that contain sequencing adapters and unique, sample-specific sequences (barcodes)."

Modified.

L103: "The DNA fragments in the pooled samples are then clonally amplified in an emulsion PCR, and then sequenced, Resulting sequences are then analyzed using publically available BIOINFORMATIC

tools..."

Modified.

L111: Should indicate that this approach is called the fusion PCR method by Ion Torrent.

We modified the title of Step 1 to indicate this fact.

L114: Should explain what the sequencing key is. Also, reference is need for HotStartTaq; is this a proofreading polymerase?

Modified.

L111-116: Mention table of barcodes. I would recommend if using Ion Torrent barcodes to using those of the same length (if possible). Ion Torrent has 10,11 and 12 base barcodes, and this can be annoying for some methods to demultiplex. You may also want to recommend Golay-12 barcodes as recommended by Caporaso. Also, any barcode can be entered into the PGM or Proton system, and demultiplexing can be performed on instrument. Also, indicate what type of purification is used for the primers and approximate cost. What is the target group of the primer set? Do they pick up Archaea? Are there any known taxa that are definitely missed? Is there a reference for this primer set?

The different length of the barcode doesn't cause problem in Mothur. We don't recommend the primers of Caporaso for the Ion Torrent as they were not designed to be different in the Ion Torrent flow space as were the Ion Torrent recommended barcodes. We added some information in the discussion about the primers used. The cost for primer is negligible, as we get enough primer for ~1000 reactions at a cost of approximately $10. It is included in the PCR cost mentioned in the discussion.

L131: I personally think this is an insane approach - that is - to do gel purification for each sample. This might be tolerable for 10 samples, but what if you have 96? Or 500? So, not really feasible when you try to increase the throughput to take full advantage of the PGM using a 318 chip, for example. AMPure beads (or similar), for example, can be used in 96-well throughput. Particularly if the size of the PCR product is larger, then different concentrations of AMPure beads can be used to remove all fragments below 300 bp (0.6X buffer), for example. In addition, you might use other size selection of all samples pooled together in roughly equal quantities and then do gel purification together. This will lead to some variability in the number of reads per sample, but this appears to be unavoidable in any case.

We agree with the reviewer, and in fact we are using bead purification in the laboratory when sample numbers exceed 48. We added a sentence in the discussion to mention this alternative.

L141-144: We have found that the recommended concentrations of libraries are excessively high, and that we get very high polyclonals when running at the recommended concentrations. Our best runs have had much lower concentrations (at least ½ or less). We also typically use qPCR to accurately quantitate the library before emulsion PCR using KAPA kits.

In our hands, the concentrations indicated give adequate results.

L165: Need to indicate which lifetech kit is used; this isn't in the reagent table.

This is not a kit, we use custom made oligos.

L247-255: Should really define what all the parameters are.

The parameters are already defined in steps 4.2.1 to 4.2.3.

L289-L297: My feeling is that this is an underestimate of the cost of associated with this method. For one, the service contract on the instrument should be added. In addition, the price for primer synthesis (probably around $1/sample) should be added. Costs should be broken down more clearly - including the emulsion PCR, sequencing reagents, library quantification, etc. Also, if you use the 400 bp kit (for example - 515F/806R primer set) this will increase cost.

Several factors were left out of the equation to calculate costs: service contract, technician salary, laboratory space usage, instrument depreciation, etc. We wanted to indicate the cost in term of reagents, not in term of total cost, as this will vary widely from institution to institution. The price for primers is included in the price for PCR and we can get enough primer for 100-1000 reactions for around $10, which makes it negligible. We changes “library preparation” for emulsion PCR and added “sequencing reagents” to clarify the cost breakdown.

L299-308: I would indicate that actually Illumina MiSeq is the leader for amplicon sequencing at the moment. This appears to be the ability to handle hundreds of samples simultaneously, separate index read that is integrated with QIIME, increasing read length and paired-end reads.

We added some text related to this point in the discussion. In term of numbers of paper published, 454 is clearly the leader, but we agree that MiSeq might be a dominant player in the near future.

L317-329: This method should also be acknowledged to really only provide genus-level and higher classification; species is basically out.

This is now part of the discussion. We also mention that other genes might be better for species level information.

Table 2: Maybe more useful than any details about a single run, you might put the ranges you have gotten from multiple runs.

We choose this run because it was a good average run. We mentioned in the representative results section that these are typical results that we get out of 16S rRNA gene amplicon runs.

Other thoughts:
(1) Since this is specific to Ion Torrent, perhaps the title should be "Ion Torrent sequencing of 16S rRNA gene amplicons"; either that or include details of how to adapt this fusion method for Illumina and Roche 454 pyrosequencing approaches.

The title was modified following previous editorial comments. We added details about adapting this approach for Illumina and 454 in the discussion.

Manuscript with track changes

Click here to download Supplemental File (as requested by JoVE): 20140117-JoVE-v5-with-changes.docx