Instituto de Ingeniería del Agua y Medio Ambiente

SOP intra-FLA bacteria METAGENOMICS

CONTENTS

1. Purpose
2. Procedure description
   - 2.1. Equipment and Materials
   - 2.2. Amplicon primers
   - 2.3. Procedure
3. References

1. Purpose

The purpose of this procedure is to describe a method to analyze the 16S ribosomal RNA gene in order to study the biodiversity among the pathogenic bacteria from inside Free Living Amoeba (FLA) isolated from water.

2. Procedure description

The workflow of the procedure is:

1. Amplicon PCR using the designed primers
2. Library preparation
3. Sequencing on MiSeq
4. Data analysis of the sequences obtained

2.1 Equipment and Materials

EQUIPMENT:

- Standard benchtop microcentrifuge containing a rotor for 2.0 ml reaction tubes
- Fluorometer. QUBIT™ (Invitrogen)

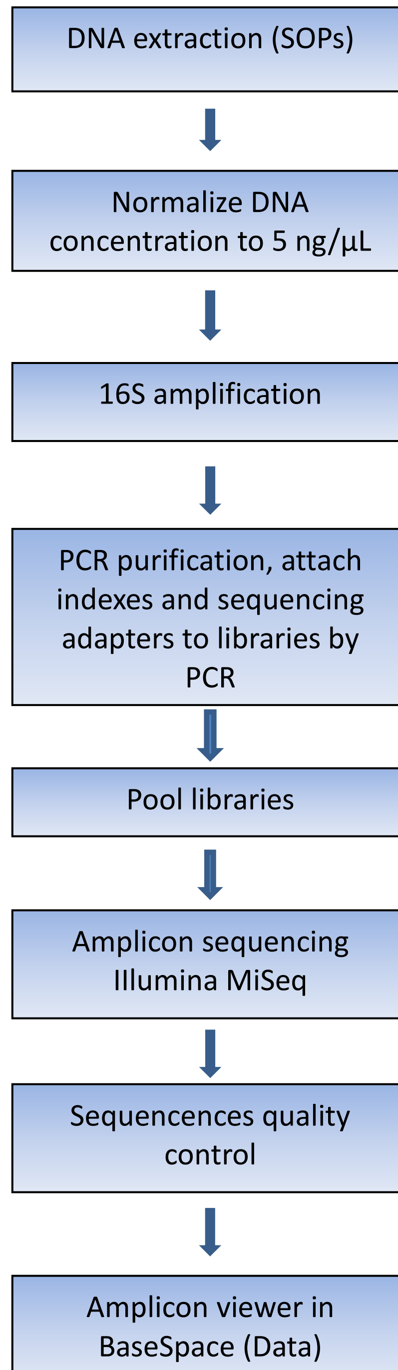
MATERIALS:

- Micropipettes and tips with aerosol resistant filter
- Sterilized microtubes (0.2-1.5 µL)
- Molecular grade water, nuclease-free
- Quant-iT ™ assays kit for use with Qubit.
- KAPA Hifi HotStart ready mix (2X)

METAWATER

Instituto de Ingeniería del Agua y Medio Ambiente

WORKFLOW 16S Metagenomic Sequencing preparation



2.2. Amplicon Primers

The specific sequences of the primers used in this protocol target the V3-V4 16S rRNA region. Specificity of the primers were performed *in silico* and *in vitro* with a battery of reference DNAs from different bacteria (preparing data for publication).

16S Amplicon primers:

FORWARD PRIMER
341F: 5´-CCT ACG GGN GGC WGC AG-3' (Herlemann et al., 2011)
REVERSE PRIMER
802R: 5´- TAC NVG GGT ATC TAA TCC-3' (Claesson et al., 2009)

The Illumina overhang adapter sequences to be added to primer- specific sequences are:

Forward overhang :
5'-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG-3' [primer-specific sequence]

Reverse overhang:
5'- GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA G-3' [primer-specific sequence]

2.3. Procedure

a. PCR

PCR was performed as described by Illumina guide in the documentation provided by Illumina (http://www.illumina.com/content/dam/illumina- support/documents/documentation /chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf).

Amplification reaction:

PCR reaction Mix	Volume
Microbial DNA (5 ng/µl)	2.5 µl
Forward Primer 1 µM	5 µl
Reverse Primer 1 µM	5 µl
2x KAPA HiFi HotStart ReadyMix	12.5 µl
TOTAL	25 µl

Mix and perform PCR in a thermal cycler using the following program:

• 95°C for 3 minutes

• 25 cycles of:
— 95°C for 30 seconds
— 55°C for 30 seconds
— 72°C for 30 seconds

• 72°C for 5 minutes
• Hold at 4°C

b. Steps 2 and 3 (library preparation and sequencing on MiSeq) were performed as described by Illumina guide in the documentation provided by Illumina (http://www.illumina.com/content/dam/illumina- support/documents/documentation /chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf)

c. Bioinformatics data analysis: DNA sequencing data was processed using the QIIME open – source bioinformatics pipeline (Caporaso et al., 2010). OTUs were picked at 97% similarity against the Greengenes database (De Santis et al., 2006) or SILVA database (www.arb-silva.de; Quast el al., 2013)

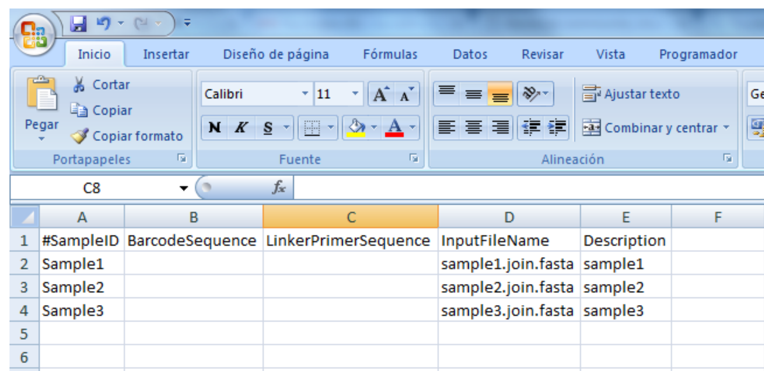
METAWATER

Instituto de Ingeniería del Agua y Medio Ambiente

QIIME WORKFLOW

Data were analyzed using the QIIME 1.8.0 software package (http://qiime.org/). The workflow followed consists of the following steps:

Create mapping file in Excel



	A	B	C	D	E	F
1	#SampleID	BarcodeSequence	LinkerPrimerSequence	InputFileName	Description	
2	Sample1			sample1.join.fasta	sample1	
3	Sample2			sample2.join.fasta	sample2	
4	Sample3			sample3.join.fasta	sample3	
5						
6						

The SampleID, BarcodeSequence, LinkerPrimerSequence, and Description headers need to be present.
Save the document in .txt format

a) Validate mapping file

validate_mapping_file.py -m map.txt -o validate_mapping_file_output -b -p -j Description
In the Linux terminal should appear the message: “No errors or warnings were found in mapping file”

b) Generate a single .fna archive including all your samples

add_qiime_labels.py -m map.txt -i FASTA -c InputFileName
Resulting archive will be combined_seqs.fna

c) Count number of sequences and get the average length of the sequences

count_seqs.py -i combined_seqs.fna -o counted_sequences.txt

d) Pick OTUs and assign taxonomy

pick_open_reference_otus.py -o otusopen/ -i combined_seqs.fna -r 97_otus.fasta -a -O 2 -p uc_fast_params.txt

The archive 97_otus.fasta comes from the greengenes database

(http://greengenes.secondgenome.com/downloads)

Previously, create the document uc_fast_params.txt. It should only include the following line:

pick_otus:enable_rev_strand_match True

A .biom (biological observation matrix) archive is created. It contains OTUs and taxonomy information.

e) Statistics of .biom table

biom summarize-table -i otusopen/otu_table_mc2_w_tax_no_pynast_failures.biom -o otusopen/statistics.txt

f) Convert .biom table into a .txt document

biom convert -i otusopen/otu_table_mc2_w_tax_no_pynast_failures.biom -o otusopen/table_from_biom.txt -b --header-key taxonomy

Now the table_from_biom.txt document can be opened in Excel

3. References

1. Herlemann DP1, Labrenz M, Jürgens K, Bertilsson S, Waniek JJ, Andersson AF. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. ISME J. 2011 Oct;5(10):1571-9. doi: 10.1038/ismej.2011.41.

2. Claesson MJ1, O'Sullivan O, Wang Q, Nikkilä J, Marchesi JR, Smidt H, de Vos WM, Ross RP, O'Toole PW. Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. PLoS One. 2009 Aug 20;4(8):e6669. doi: 10.1371/journal.pone.0006669.

3. Caporaso, J. G. et al. QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7, 335–336 (2010).

4. DeSantis, T. Z. et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl. Environ. Microbiol. 72, 5069–5072 (2006).

5. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Opens external link in new window. Nucl. Acids Res. 41 (D1): D590-D596.

6. Laure Guillou, Dipankar Bachar, Stéphane Audic, David Bass, Cédric Berney, Lucie Bittner, Christophe Boutte, Gaétan Burgaud, Colomban de Vargas, Johan Decelle, Javier del Campo, John R. Dolan, Micah Dunthorn, Bente Edvardsen, Maria Holzmann, Wiebe H.C.F. Kooistra, Enrique Lara, Noan Le Bescot, Ramiro Logares, Frédéric Mahé, Ramon Massana, Marina Montresor, Raphael Morard, Fabrice Not, Jan Pawlowski, Ian Probert, Anne-Laure Sauvadet, Raffaele Siano, Thorsten Stoeck, Daniel Vaulot, Pascal Zimmermann, and Richard Christen The Protist Ribosomal Reference database (PR2): a catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucl. Acids Res. (1 January 2013) 41 (D1): D597-D604 first published online November 27, 2012 doi:10.1093/nar/gks1160