A Bioinformatics Analysis workflow for 16S rRNA Amplicon Sequencing data

In 3 collections

Nov 23, 2020

Lilan Hao[1]

[1]BGI-shenzhen

1 Works for me dx.doi.org/10.17504/protocols.io.bntpmemn

Lilan Hao

ABSTRACT

Deal with the 16S rRNA amplicon sequencing reads, which sequenced by Ion PGM™ Sequencer with Ion 318™Chip v2 with a read length of 400bp, and takes the OTU table(the measured abundance profile of detected operational taxonomic units (OTUs)).

DOI

dx.doi.org/10.17504/protocols.io.bntpmemn

PROTOCOL CITATION

Lilan Hao 2020. A Bioinformatics Analysis workflow for 16S rRNA Amplicon Sequencing data. **protocols.io** https://dx.doi.org/10.17504/protocols.io.bntpmemn

COLLECTIONS

The Female Urinary Microbiota Protocols Collection

Protocols for the female urinary microbiota in relation to the reproductive tract microbiota

Protocols for "The female urinary microbiota in relation to the reproductive tract microbiota."

CREATED

Oct 23, 2020

LAST MODIFIED

Nov 23, 2020

PROTOCOL INTEGER ID

43599

PARENT PROTOCOLS

Part of collection

The Female Urinary Microbiota Protocols Collection

Protocols for the female urinary microbiota in relation to the reproductive tract microbiota

Protocols for "The female urinary microbiota in relation to the reproductive tract microbiota."

ABSTRACT

Deal with the 16S rRNA amplicon sequencing reads, which sequenced by Ion PGM™ Sequencer with Ion

318^{TM}Chip v2 with a read length of 400bp, and takes the OTU table(the measured abundance profile of detected operational taxonomic units (OTUs)).

1 **## Convert the bam files to fastq files**
../bam2fastq-1.1.0/bam2fastq -o vagina/01.trim.data/16_s1.fq ../rawdata/RFEMpxjMAADGAAPGM-47.bam

2 **## Delete the primer**
perl ../iontorrent/re-primers.pl ../v4-v5_primers.fa 2 vagina/01.trim.data/16_s1.fq

3 **## Reverse fastq to fasta and quality**
perl ../fq2qual.pl vagina/01.trim.data/16_s1.fq vagina/01.trim.data/16_s1

4 **## Trim the poly-A/T, multi-N and low-quality sequences.**
mothur "#trim.seqs(fasta= vagina/01.trim.data/16_s1.fa, qfile= vagina/01.trim.data/16_s1.qual, minlength=200,maxambig=1,maxhomop=10,qwindowaverage=25,qwindowsize=50,processors=8)"

mothur "#reverse.seqs(fasta= vagina/01.trim.data/16_s1.trim.fasta, qfile=vagina/01.trim.data/16_s1.trim.qual)"

perl ../creat.group.pl vagina/01.trim.data/16_s1.trim.rc.fasta 16_s1 >> /vagina/02.processing.data/vagina.groups

cat vagina/01.trim.data/16_s1.trim.rc.fasta >> vagina/02.processing.data/vagina.trim.fasta
cat vagina/01.trim.data/16_s1.trim.rc.qual >> vagina/02.processing.data/vagina.trim.qual

5 **## Pick the representative sequences.**
mothur"#unique.seqs(fasta=vagina/02.processing.data/vagina.trim.fasta)"

6 **## Assign to the SAILVA**
mothur "#align.seqs(fasta=vagina/02.processing.data/vagina.trim.unique.fasta, reference=../bin/16s/database/silva.bacteria.fasta, processors=8)"

mothur "#summary.seqs(fasta=vagina/02.processing.data/vagina.trim.unique.align, name=vagina/02.processing.data/vagina.trim.names)" > vagina/02.processing.data/vagina.trim.unique.align.summary.log

7 **## Pick the high-scoring sequences**
mothur "#screen.seqs(fasta=vagina/02.processing.data/vagina.trim.unique.align, name=vagina/02.processing.data/vagina.trim.names,group=vagina/02.processing.data/vagina.groups, optimize=start-end, criteria=90,processors=8)"

mothur "#filter.seqs(fasta=vagina/02.processing.data/vagina.trim.unique.good.align, vertical=T,trump=.,processors=8)"

mothur "#unique.seqs(fasta=vagina/02.processing.data/vagina.trim.unique.good.filter.fasta, name=vagina/02.processing.data/vagina.trim.good.names)"

rm vagina/02.processing.data/vagina.trim.unique.good.align vagina/02.processing.data/vagina.trim.unique.align

mothur "#summary.seqs(fasta=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.fasta, name=vagina/02.processing.data/vagina.trim.unique.good.filter.names)" >

Citation: Lilan Hao (11/23/2020). A Bioinformatics Analysis workflow for 16S rRNA Amplicon Sequencing data. https://dx.doi.org/10.17504/protocols.io.bntpmemn

vagina/02.processing.data/vagina.trim.unique.good.filter.unique.fasta.summary.log

8 **##Check chimera and delete chimeric sequences**

mothur "#pre.cluster(fasta=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.fasta,
name=vagina/02.processing.data/vagina.trim.unique.good.filter.names,group=vagina/02.processing.data/vagina.good.groups,diffs=2)"

mothur "#chimera.uchime(fasta=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.fasta,
name=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.names,group=vagina/02.processing.data/vagina.good.groups,processors=8)"

mothur "#remove.seqs(
accnos=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.devono.uchime.accnos,
fasta=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.fasta,
name=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.names,group=vagina/02.processing.data/vagina.good.groups)"

mothur
"#summary.seqs(fasta=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.pick.fasta,
name=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.pick.names)" >
vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.pick.fasta.summary.log

9 **## Classify according to the Greengene reference sequences (gg_13_8_otus)**

ln -s ../database/greengenes/gg_13_5_99.fasta vagina/02.processing.data/gg_13_5_99.fasta
ln -s ../database/greengenes/gg_13_5_99.gg.tax vagina/02.processing.data/gg_13_5_99.gg.tax

mothur "#classify.seqs(fasta=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.pick.fasta,
template=vagina/02.processing.data/gg_13_5_99.fasta,taxonomy=vagina/02.processing.data/gg_13_5_99.gg.tax,cutoff=50,processors=8)"

10 **## Delete reads of Mitochondria and Chloroplast**

mothur "#remove.lineage(
fasta=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.pick.fasta,
name=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.pick.names,group=vagina/02.processing.data/vagina.good.pick.groups,taxonomy=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.pick.gg.wang.taxonomy,taxon=Mitochondria-Cyanobacteria_Chloroplast-unknown)"

11 **## Make the otu table**

python ../bin/pick_otus.py -i vagina/03.otu.data/vagina.final.uclu.fasta -o vagina/03.otu.data

perl ../bin/OTU_format_trans.pl vagina/03.otu.data/vagina.final.uclu_otus.txt
vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.pick.pick.names 0.03 >
vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.pick.pick.an.list

perl ../bin/tags_taxonomy.stat.v2.pl
vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.pick.pick.names
vagina/02.processing.data/vagina.good.pick.pick.groups
vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.pick.gg.wang.pick.taxonomy.pick
vagina/03.otu.data/vagina.raw.data.info

mothur "#make.shared(
list=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.pick.pick.an.list,
group=vagina/02.processing.data/vagina.good.pick.pick.groups,label=0.03)"

Citation: Lilan Hao (11/23/2020). A Bioinformatics Analysis workflow for 16S rRNA Amplicon Sequencing data. https://dx.doi.org/10.17504/protocols.io.bntpmemn

```
mothur
"#classify.otu(list=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.pick.pick.an.list,
name=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.pick.pick.names,taxonomy=vagina
/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.pick.gg.wang.pick.taxonomy.pick,label=0.03)"

mothur "#summary.single(
shared=vagina/02.processing.data/vagina.trim.unique.good.filter.unique.precluster.pick.pick.an.shared,groupmode=
t)"
```


Citation: Lilan Hao (11/23/2020). A Bioinformatics Analysis workflow for 16S rRNA Amplicon Sequencing data. https://dx.doi.org/10.17504/protocols.io.bntpmemn