Package 'etec16s'

November 10, 2026

Maintainer Joseph N. Paulson <jpaulson@jimmy.harvard.edu>

Author Joseph N. Paulson

Version 1.40.0

License Artistic-2.0

Title Individual-specific changes in the human gut microbiota after challenge with enterotoxigenic Escherichia coli and subsequent ciprofloxacin treatment

Description 16S rRNA gene sequencing data to study changes in the faecal microbiota of 12 volunteers during a human challenge study with ETEC (H10407) and subsequent treatment with ciprofloxacin.

LazyData yes

Depends R (>= 3.3), Biobase, metagenomeSeq(>= 1.12.0)

biocViews ExperimentData, SequencingData, MicrobiomeData

NeedsCompilation no

git_url https://git.bioconductor.org/packages/etec16s

git_branch RELEASE_3_23

git_last_commit f495dec

git_last_commit_date 2026-04-28

Repository Bioconductor 3.23

Date/Publication 2026-11-10

Contents

etec16s-package 2
etec16s 2

Index **5**

etec16s-package

Dataset of 16S samples from 12 subjects challenged with enterotoxigenic Escherichia coli.

Description

Data used in 'Individual-specific changes in the human gut microbiota after challenge with enterotoxigenic Escherichia coli and subsequent ciprofloxacin treatment'. Measurements are number of reads clustered into OTUs (operatonal taxanomic units) by DNAClust.

Author(s)

Joseph N. Paulson

References

Individual-specific changes in the human gut microbiota after challenge with enterotoxigenic Escherichia coli and subsequent ciprofloxacin treatment.

etec16s

Curated dataset of many healthy and moderate-to-severe diarretic gut 16s samples on the 454 FLEX platform.

Description

Data used in 'Individual-specific changes in the human gut microbiota after challenge with enterotoxigenic Escherichia coli and subsequent ciprofloxacin treatment'.

DNA was amplified using 'universal' primers targeting the V1-V2 region of the 16S rRNA gene (small subunit of the ribosome) in bacteria - 338R (5'- CATGCTGCCTCCCGTAGGAGT -3') and 27F (5'- AGAGTTTGATCCTGGCTCAG -3'). Both forward and reverse primers had a 5' portion specific for use with 454 GS-FLX Titanium sequencing technology and the forward primers contained a barcode between the Titanium and gene specific region, so that samples could be pooled to a multiplex level of 132 samples per instrument run.

16S rRNA gene sequencing was performed for all available stool samples. After sequencing, 124 samples passed quality controls, corresponding to data from 5 volunteers with the most unambiguous cases of diarrhea during the study (54 samples) and 7 volunteers without diarrheal symptoms who had the most stool samples (78 samples). The raw data have been submitted to NCBI under project ID: PRJNA298336.

Briefly, the sequenced reads were filtered for quality and were clustered with DNAclust36 to 99 percent identity clusters that were further annotated with Ribosomal Database (RDP) (rdp.cme.msu.edu, release 10.4). Sequences without a nearly identical match to RDP (> 100 bp perfect match and > 97 percent identity, as defined by BLAST) were marked as having "no genus match" and assigned an operational taxonomic unit (OTU) identifier. Chimera checking was done with Perseus/UCHIME.

Selected OTUs (e.g., OTUs belonging to the pre-infection biomarker) were aligned with Mega BLAST against the NCBI nt database. Manual inspection of the results based on percent identity and query sequence coverage was used to refine the automatically determined taxonomic classification. If the best BLAST hit differed by more than 3 percent from the query sequence, or covered less than 95 percent of the query sequence, classification was done at the genus level only based on the taxonomy labels of the top hits.

format

Data is stored as an `MRexperiment-class` object. Using `MRcounts` one can obtain the 16S count matrix produced using using DNAclust (http://dnaclust.sourceforge.net/). The `pData` function accesses a data frame with the following columns:

`SubjectID`: Subject ID

`Dose`: Quantity of ETEC given

`Day`: Relative Day

`AnyDayDiarrhea`: Did the subject ever get diarrhea during the study

`Diarrhea`: Did subject have diarrhea on day of student

`AntiGiven`: Was antibiotic given on specific day

`AntibPrev`: Was antibiotic given previously

The `fData` function accesses a data frame with the following columns:

`OTU.ID`: OTU cluster id

`Center`: Sequencing read center ID

`Phylum`: phylum

`Class`: class

`Order`: order

`Family`: family

`Genus`: genus

`Strain`: Strain

`clusterCenter`: The OTU cluster's representative sequence

Author(s)

Joseph N. Paulson

References

'Individual-specific changes in the human gut microbiota after challenge with enterotoxigenic Escherichia coli and subsequent ciprofloxacin treatment'

See Also

`MRexperiment-class` for the class definition, `cumNorm` to normalize the counts.

Examples

```
data(etec16s)
head(pData(etec16s))
head(fData(etec16s))
```

Index

∗ **datasets**
etec16s, 2

cumNorm, *3*

etec16s, 2
etec16s-package, 2