

# QIIME Tutorial

## Metadata File

sample-id	barcode-sequence	body-site	year	month	day	subject	reported-antibiotic-usage	days-since-experiment-start
#q21types	categorical	categorical	numeric	numeric	numeric	categorical	categorical	numeric
L1S8	AGCTGACTAGTC	gut	2008	10	28	subject-1	Yes	0
L1S57	ACACACTATGGC	gut	2009	1	20	subject-1	No	84
L1S76	ACTACGTGTGGT	gut	2009	2	17	subject-1	No	112
L1S105	AGTGCATGCGT	gut	2009	3	17	subject-1	No	140
L2S155	ACGATGCGACCA	left palm	2009	1	20	subject-1	No	84
L2S175	AGCTATCCACGA	left palm	2009	2	17	subject-1	No	112
L2S204	ATGCAGCTCAGT	left palm	2009	3	17	subject-1	No	140
L2S222	CACGTGACATGT	left palm	2009	4	14	subject-1	No	168
L3S242	ACAGTTGCGCGA	right palm	2008	10	28	subject-1	Yes	0
L3S294	CACGACAGGCTA	right palm	2009	1	20	subject-1	No	84
L3S313	AGTGTACAGGTG	right palm	2009	2	17	subject-1	No	112
L3S341	CAAGTGAGAGAG	right palm	2009	3	17	subject-1	No	140
L3S368	CATCGTATCAAC	right palm	2009	4	14	subject-1	No	168
L5S104	CAGTGCAGGAC	tongue	2008	10	28	subject-1	Yes	0
L5S155	ATCTTAACTGCG	tongue	2009	1	20	subject-1	No	84
L5S174	CAGACATTGCGT	tongue	2009	2	17	subject-1	No	112
L5S203	CGATGCACCAGA	tongue	2009	3	17	subject-1	No	140
L5S222	CTAGAGACTCTT	tongue	2009	4	14	subject-1	No	168
L1S140	ATGGCAGCTCTA	gut	2008	10	28	subject-2	Yes	0
L1S208	CTGAGATACGGG	gut	2009	1	20	subject-2	No	84
L1S257	CCGACTGAGATG	gut	2009	3	17	subject-2	No	140
L1S281	CCTCTCGTGATC	gut	2009	4	14	subject-2	No	168
L2S240	CATATCGCAGTT	left palm	2008	10	28	subject-2	Yes	0
L2S309	CGTGCATTATCA	left palm	2009	1	20	subject-2	No	84
L2S357	CTAACGCACTCA	left palm	2009	3	17	subject-2	No	140
L2S382	CTCAATGACTCA	left palm	2009	4	14	subject-2	No	168
L3S378	ATCGATCTGTGG	right palm	2008	10	28	subject-2	Yes	0
L4S63	CTCGTGGAGTAG	right palm	2009	1	20	subject-2	No	84
L4S112	GGGTACACACA	right palm	2009	3	17	subject-2	No	140
L4S137	GAACGTATCTCT	right palm	2009	4	14	subject-2	No	168
L5S240	CTGGACTCATAG	tongue	2008	10	28	subject-2	Yes	0
L6S20	GAGGCTCATCAT	tongue	2009	1	20	subject-2	No	84
L6S68	GATACGTCCTGA	tongue	2009	3	17	subject-2	No	140
L6S93	GATTAGCACTCT	tongue	2009	4	14	subject-2	No	168

**Figure 1:** Metadata file containing sample information such as: sample ID, body site, and antibiotic usage.

## Barcode File

```
@HWI-EAS440_0386:1:23:17547:1423#0/1
ATGCAGCTCAGT
+
IIIIIIIIIIII
@HWI-EAS440_0386:1:23:14818:1533#0/1
CCCTCAGCGGC
+
DDD@D?@B<<+/
@HWI-EAS440_0386:1:23:14401:1629#0/1
GACGAGTCAGTC
+
GGECDGGGGGGD
@HWI-EAS440_0386:1:23:15259:1649#0/1
AGCAGTCGGAT
+
IIIIIIIIIIII
@HWI-EAS440_0386:1:23:13748:2482#0/1
AGCACACCTACA
+
GGGGBGGEEGGD
@HWI-EAS440_0386:1:23:6532:3028#0/1
GAGAGAATGATC
+
HIIIIIIIIIIII
@HWI-EAS440_0386:1:23:8677:3027#0/1
CACAGTGGACGT
+
FHHHHHHHHHHH
@HWI-EAS440_0386:1:23:5678:3052#0/1
ATAGCTCCATAC
+
IIIIIIIIIIII
@HWI-EAS440_0386:1:23:11889:3171#0/1
ACGTTAGCACAC
+
IIIIIIIIIIII
@HWI-EAS440_0386:1:23:2112:3374#0/1
GAGAGAATGATC
+
FEEBBCEEEEDG
@HWI-EAS440_0386:1:23:5900:3393#0/1
AGGTGTGATCGC
```

**Figure 2:** This file contains the barcodes used during demultiplexing. Each barcode corresponds to a different sample

# QIIME Tutorial

## Citations: .qza File

Details of gut-table.qza

name: "gut-table.qza"

uuid: "b7bfebb8-42a0-414d-af55-037160c98fc2"

type: "FeatureTable[Frequency]"

format: "BIOMV210DirFmt"

Citations

Citation Format: BibTex

[Download](#)

```
@article{frameworkqiime2:2024.10.1|0,
  author = {Bolyen, Evan and Rideout, Jai Ram and Dillon, Matthew R. and Bokulich, Nicholas A. and Abnet, Christian C. and Al-Ghalith, Gabriel A. and Alexa
  doi = {10.1038/s41587-019-0209-9},
  issn = {1546-1696},
  journal = {Nature Biotechnology},
  number = {8},
  pages = {852-857},
  title = {Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2},
  url = {https://doi.org/10.1038/s41587-019-0209-9},
  volume = {37},
  year = {2019}
}

@article{view|types:2024.10.0|BIOMV210DirFmt|0,
  author = {McDonald, Daniel and Clemente, Jose C and Kuczynski, Justin and Rideout, Jai Ram and Stombaugh, Jesse and Wendel, Doug and Wilke, Andreas and H
  doi = {10.1186/2047-217X-1-7},
  journal = {GigaScience},
  number = {1},
  year = {2012}
}
```

**Figure 3:** Screenshot of a .qza file showing gut table information such as citation.

## Provenance: .qza File

qiime2view

File: gut-table.qza

Citations **Provenance** Metadata

Search Provenance

type: ("FeatureData" OR "SampleData")

< 0/0 > Center on Result Clear Search

```
graph TD
  seqs --> demux
  demux --> demultiplexed_seqs
  demultiplexed_seqs --> table
  table --> final(( ))
```

Result Details

```
{ "uuid": "b7bfebb8-42a0-414d-af55-037160...", "type": "FeatureTable[Frequency]"
  , "format": "BIOMV210DirFmt" }
{ "uuid": "b7bfebb8-42a0-414d-af55-037160c98fc2"
  , "type": "FeatureTable[Frequency]"
  , "format": "BIOMV210DirFmt" }
```

**Figure 4:** Screenshot of the provenance of gut-table.qza. It includes steps, commands, and metadata used.

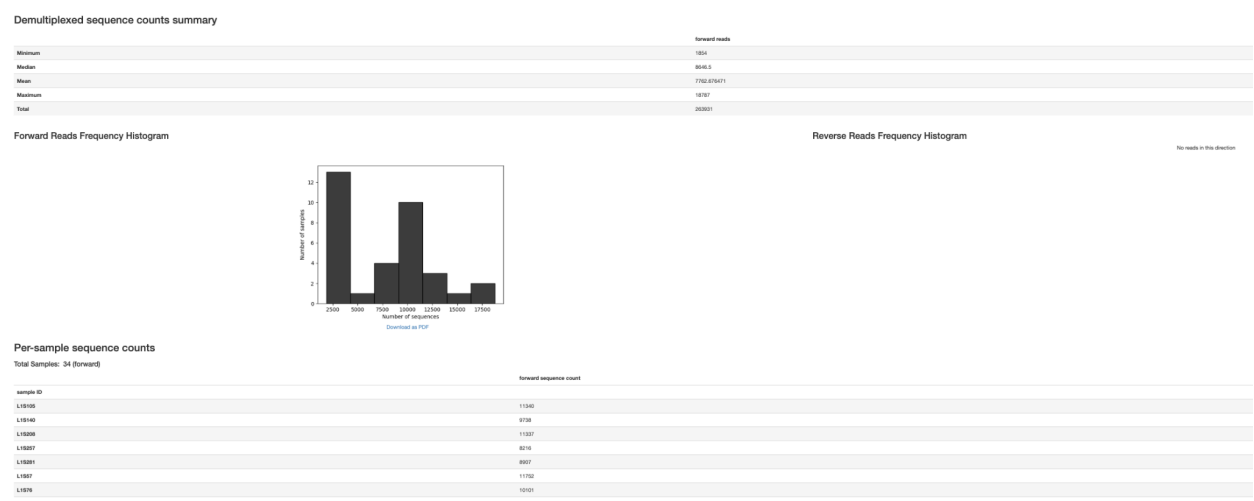
# QIIME Tutorial

## Rep-seq File

```
Rep-seqs — vim dna-sequences.fasta — 113x35
>cb2fe0146e2fbc0101050edb996a0ee2
TACGTAGGTGGCAAGCGTTATCCGGAATTATTGGGCGTAAAGCGCGGTAGGCGGTTTTTAAAGTCTGATGTGAAAGCCACGGCTCAACCGTGGAGGGTCATTGGAAACTGG
AAACTT
>ad41fe8f8be5b01c96549309937e3b14
TACGTAGGGGGCTAGCGTTGTCCGGAATCACTGGGCGTAAAGGGTTCGAGCGGAAATGCAAGTCAGATGTAAAGGCAGTAGCTTAACCTACTGTAAGCATTGAAACTGCA
TATCTTG
>d29fe3c70564fc0f69f2c03e0d1e5561
TACGTAGGTCCCGAGCGTTGTCCGGAATTATTGGGCGTAAAGCGAGCGCAGCGGTTAGATAAGTCTGAAGTTAAAGGCTGTGGCTTAACCATAGTACGCTTTGAAACTGTT
TAACTTG
>fcd4f95c05b868060121ff709085bf21
TACGTATGGAGCGAGCGTTGTCCGGAATTATTGGGCGTAAAGGGTACGAGCGCGGTTAATAAGTCGAATGTTAAAGATCGGGGCTCAACCCCGTAAAGCATTGAAACTGAT
AAACTTG
>1685b25b43eb070cd7655de8c2d8c0a4
TACGTATGTCGCGAGCGTTATCCGGAATTATTGGGCTTAAAGGGCATCTAGGCGGCGAACAAGTTGAAGGTGAAAAGCTGTGGCTCAACCATAGTCTTGCCTACAAAAGTGT
AATGCTA
>fc5b641a0b0408d99ddf2a5ba64da59
TACAGAGGATGCAAGCGTTATCCGGAATGATTGGGCGTAAAGCGCTGTAGTGGCTTTTTAAGTCCGCCGTCAAATCCCAGGGCTCAACCCCTGGACAGCGGTTGAAACTAC
CAAGCTT
>7e598ad34909a3fb8ff1623f1ede11a6
TACGTAGGGTGGAGCGTTGTCCGGAATTACTGGGCGTAAAGGGCTCGTAGTGGTTTTGTCGCGTCTGTGTGAAATTCCGGGGCTTAACCTCCGGGCGTGCAGGCGATACGGG
CATAACT
>76174505aef45f134d8068d5ad2c42aa
TACGTATGTCGCAAGCGTTATCCGGAATTATTGGGCGTAAAGCGCGCTAGGCGGTTTTGTAAGTCTGATGTGAAAATGCGGGGCTCAACCCCGTATTGCGTTGAAACTGCT
TAACTAG
>636c27859c1ac0a7994ca8ed145d03fe
TACGTAGGTCCCGAGCGTTGTCCGGAATTATTGGGCGTAAAGCGAGCGCAGGTGGTTTTAAGTCTGGTGTAAAGGCAGTGGCTCAACCATTTGTATGCATTGAAACTGGT
AGACTTG
>469e4ff2129066178b74b112829fc03c
TACGTAGGGTGGAGCGTTGTCCGGAATTACTGGGCGTAAAGGGCTCGTAGTGGTTTTGTCGCGTCTGTGTGAAATTCTGGGGCTTAACCTCCGGGCGTGCAGGCGATACGGG
CATAACT
>a049763053c277b16c2a318f41eb23b4
TACGTAGGGTGGAGCGTTGTCCGGAATTACTGGGCGTAAAGAGCTCGTAGGCGGTTTTGTCACGTCGTCTGTGAAATCCTAGGGCTTAACCCCTGGACGTGCAGGCGATACGGG
CTGACTT
>684721318896ed01048f3502333c214d
```

**Figure 5:** Screenshot of Rep-seq file output representing the sequences as a fasta file. These sequences get denoised by DADA2.

## Demux.qvz



**Figure 6:** Result summary of the demultiplexed raw sequences. It includes a histogram showing the distribution of forward read counts. No reverse reads, indicating a single-end sequence run.

# QIIME Tutorial

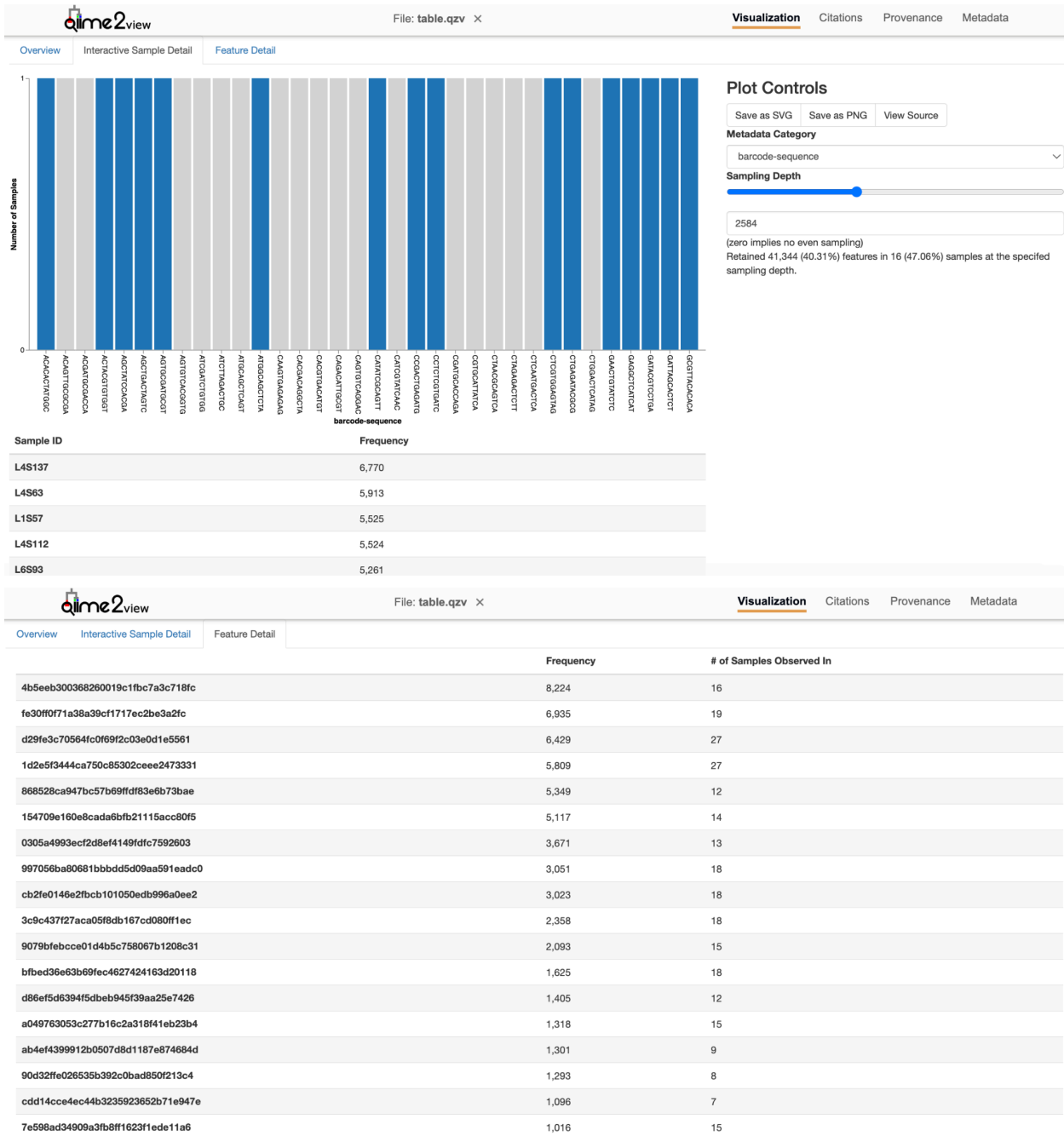
## Stats-data2 File

QIIME2view File: stats-data2.qzv × Visualization Citations Provenance Metadata							
sample-id	input	filtered	percentage of input passed filter	denoised	non-chimeric	percentage of input non-chimeric	
#q2:types	numeric	numeric	numeric	numeric	numeric	numeric	
.1S105	11340	8571	75.58	8499	7780	68.61	
.1S140	9738	7677	78.84	7605	7163	73.56	
.1S208	11337	9261	81.69	9152	8152	71.91	
.1S257	8216	6705	81.61	6627	6388	77.75	
.1S281	8907	7067	79.34	6976	6615	74.27	
.1S57	11752	9299	79.13	9260	8702	74.05	
.1S76	10101	8395	83.11	8337	7867	77.88	
.1S8	12388	7663	61.86	7624	7033	56.77	
.2S155	9263	4112	44.39	3932	3932	42.45	
.2S175	10692	4546	42.52	4386	4386	41.02	
.2S204	7299	3379	46.29	3199	3158	43.27	
.2S222	8386	3485	41.56	3187	3187	38	
.2S240	11986	5183	43.24	5094	5061	42.22	
.2S309	3836	1550	40.41	1419	1419	36.99	
.2S357	6254	2526	40.39	2373	2373	37.94	
.2S382	9301	4279	46.01	4202	4089	43.96	
.3S242	2055	970	47.2	951	897	43.65	
.3S294	2318	1313	56.64	1225	1225	52.85	
.3S313	1996	1191	59.67	1103	1103	55.26	
.3S341	1854	1109	59.82	962	962	51.89	

**Figure 7:** Stats-data2 file containing summary statistics of the DADA2 denoising step. It contains information on read input, filtered, merged, retained.

# QIIME Tutorial

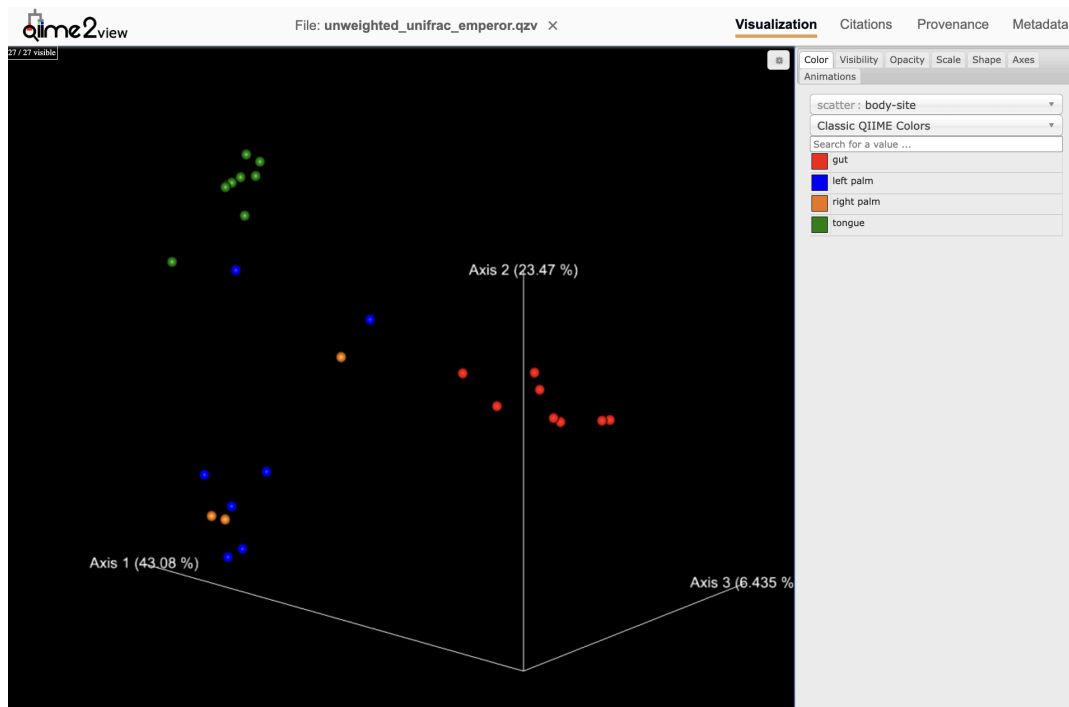
Table.qzv



**Figure 8:** Top figure shows interactive sample detail with adjusted refracted depth. Bottom figure shows feature detail and information of observed representatives per sample.

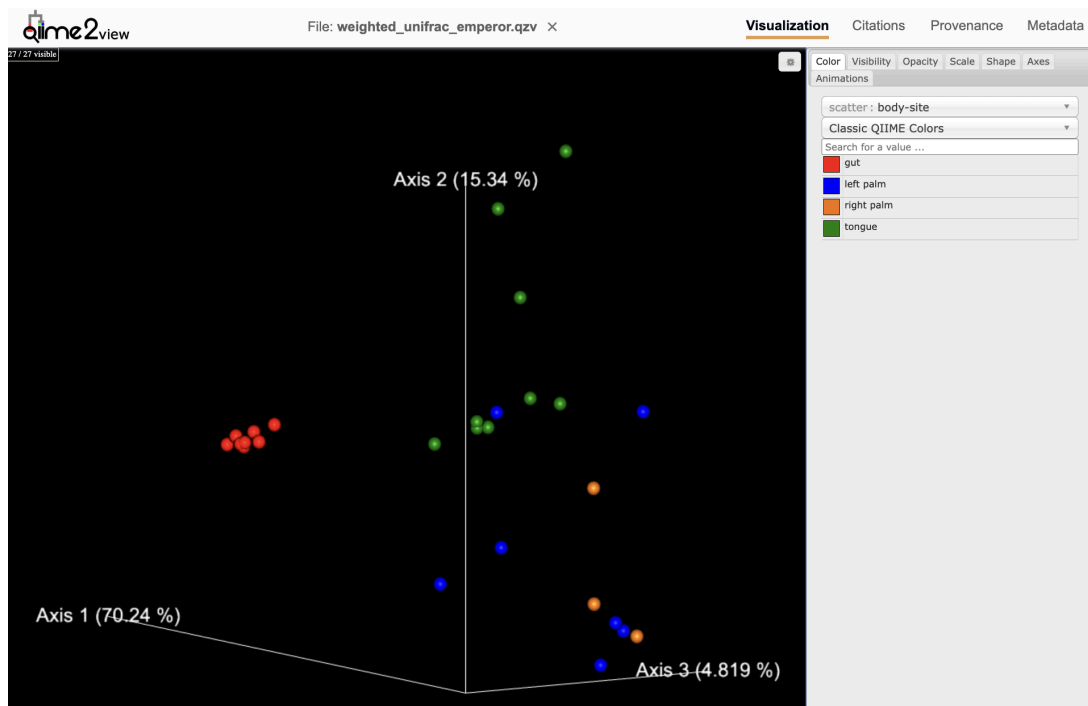
# QIIME Tutorial

## Beta Diversity - Unweighted Unifrac



**Figure 9:** 3D plot of unweighted unifrac (presence or absence) clustered by body-site

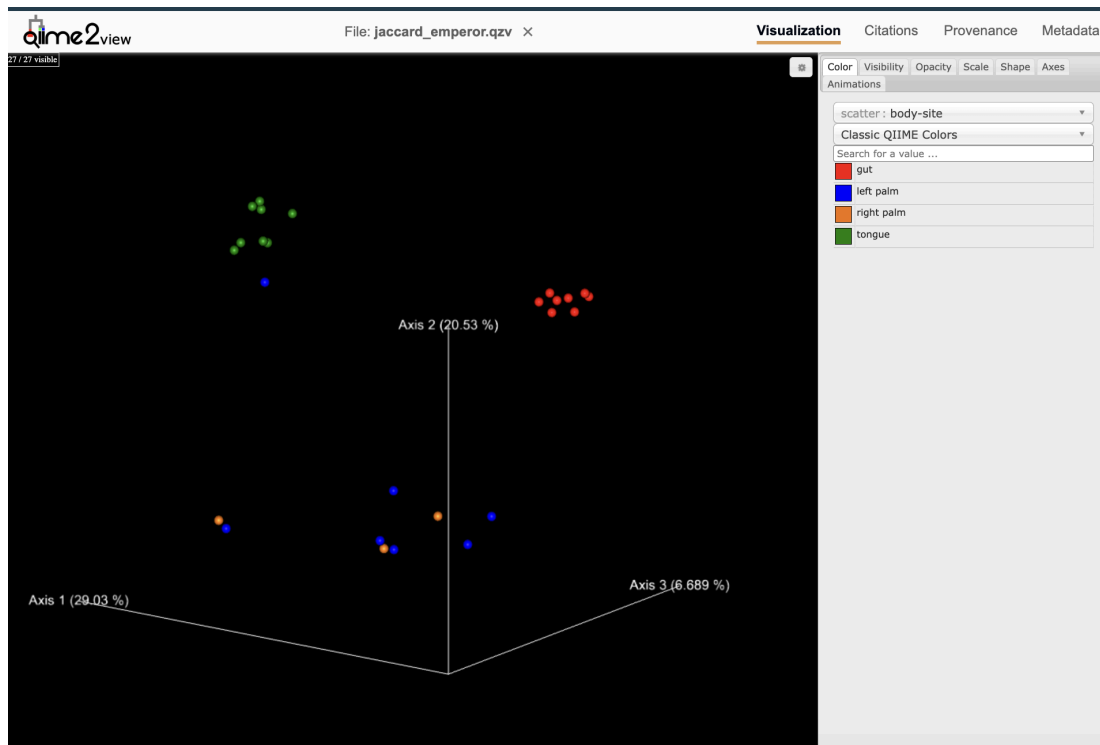
## Beta Diversity - Weighted Unifrac



**Figure 10:** 3D plot of weighted unifrac (presence and relative abundance) clustered by body-site

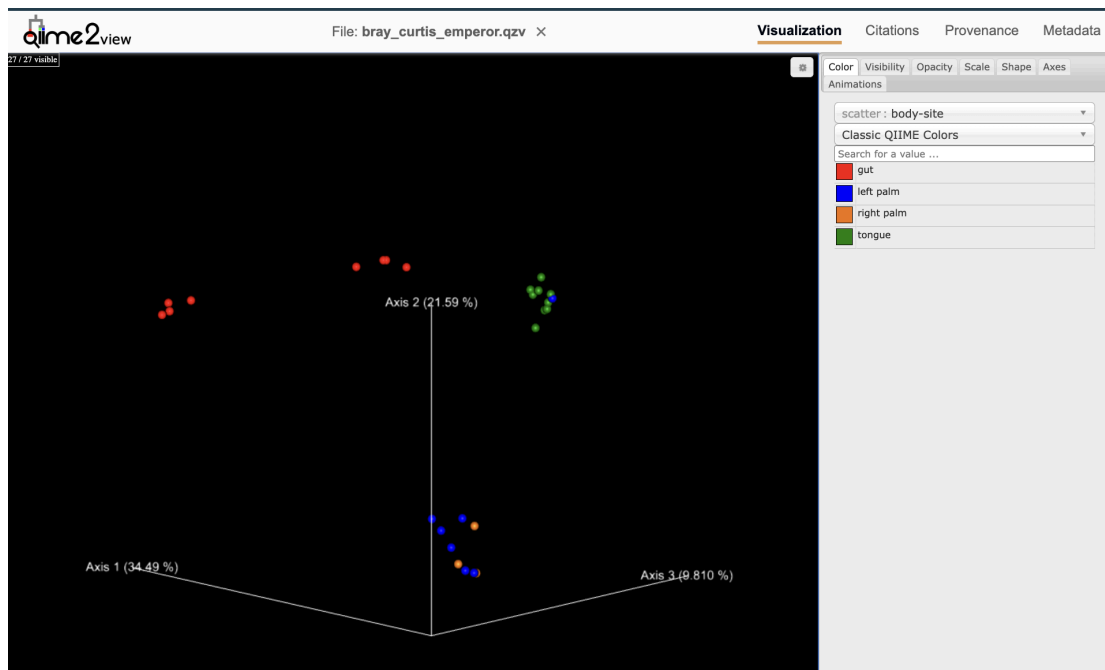
# QIIME Tutorial

## Beta Diversity - Jaccard Distance



**Figure 11:** PCoA plot of jaccard distance (presence or absence) clustered by body-site.

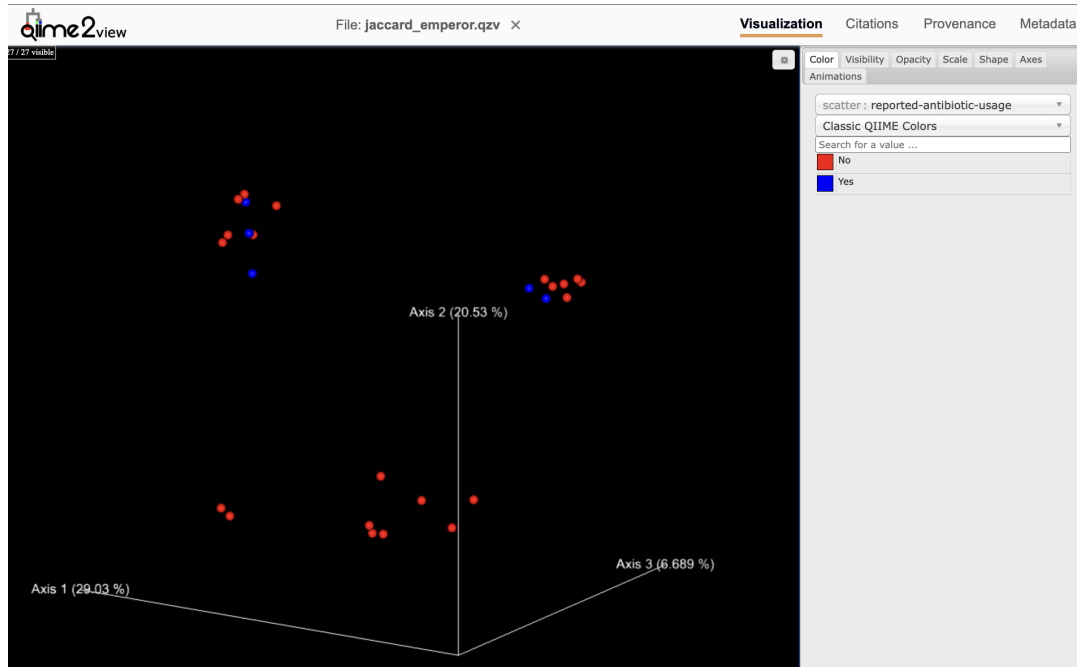
## Beta Diversity - Bray-Curtis



**Figure 12:** 3D plot representing Bray-Curtis distance clustered by body-site. It demonstrates abundant community dissimilarity.

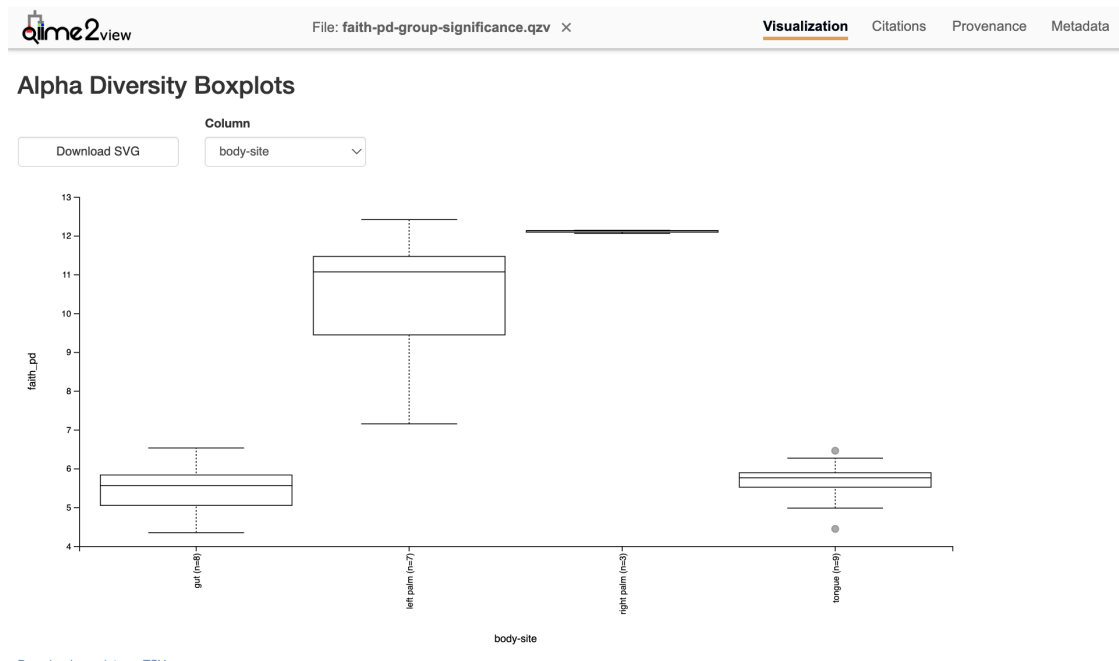
# QIIME Tutorial

## Beta Diversity - Jaccard Distance



**Figure 13:** 3D plot of jaccard distance clustered by reported-antibiotic usage.

## Faith PDs

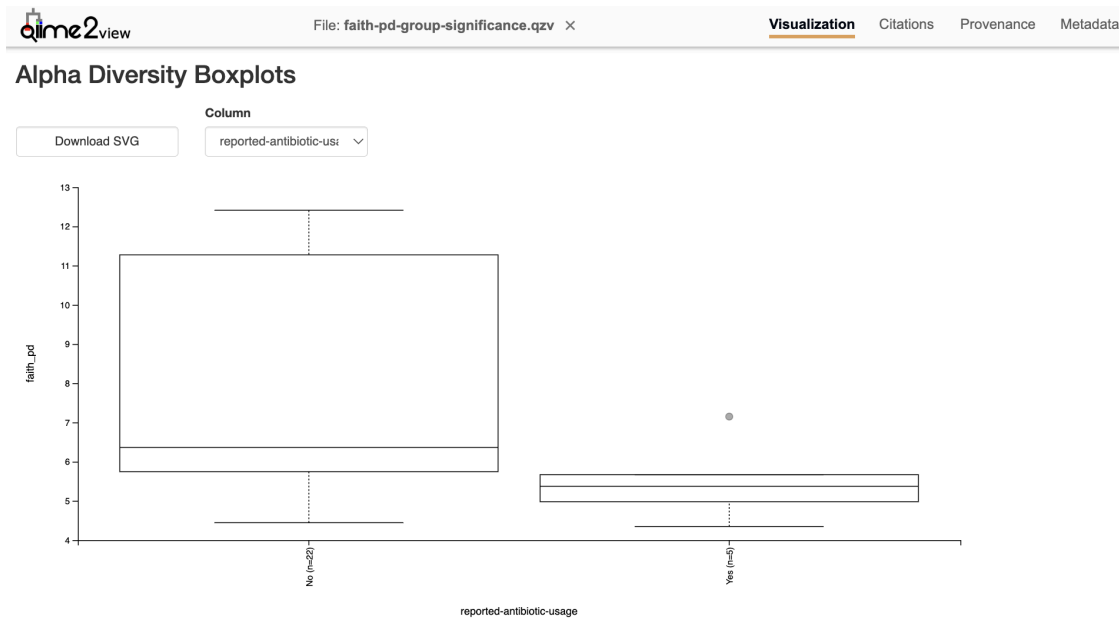


**Figure 14:** Boxplot showing significance of Faith PD (phylogenetic diversity) by body site. A Kruskal-Wallis test was performed to test if differences were statistically significant. A p-value of  $<0.01$  was used to show whether diversity differs meaningfully across groups. Left palm had the greatest diversity compared to the other groups.



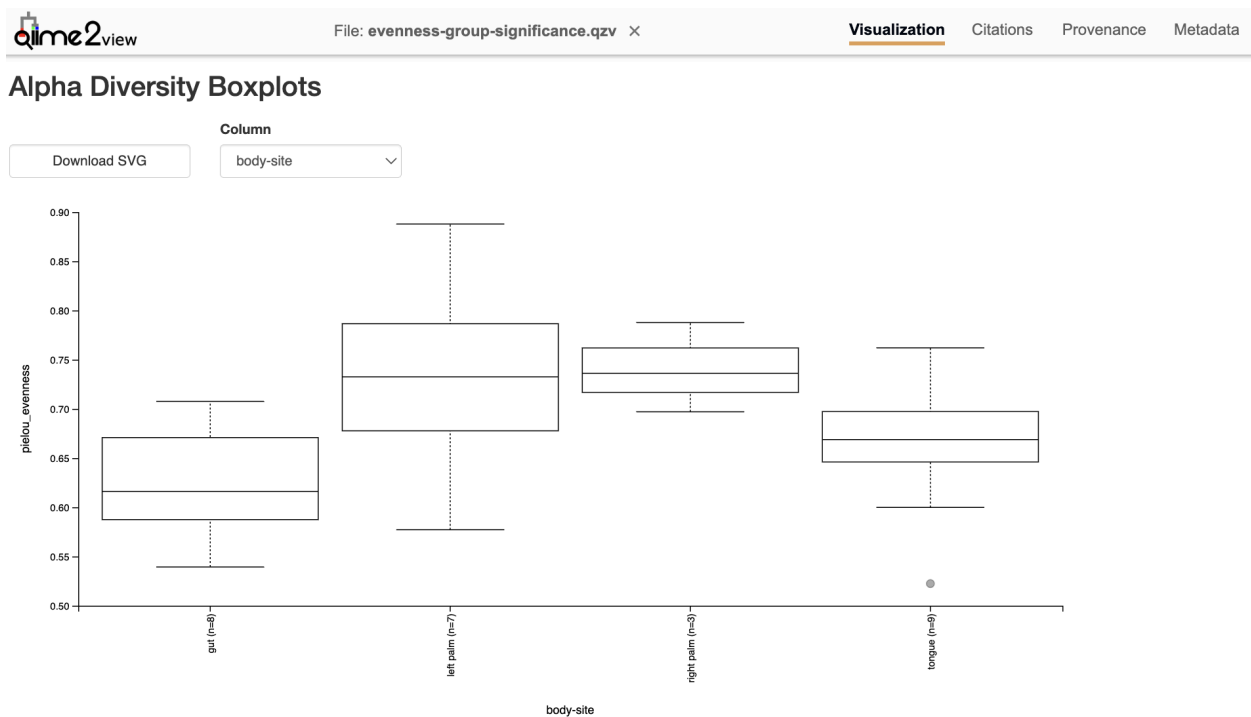
# QIIME Tutorial

## Alpha Diversity - Faith PDs



**Figure 15:** Boxplot representing alpha diversity using faith PD to show richness across samples. The group with no antibiotic usage had the greatest richness.

## Alpha Diversity - Evenness by Body Site



**Figure 16:** Boxplot representing alpha diversity by evenness, how even taxa is distributed by body site. Left and right palm had the greatest evenness taxa from the rest of the group which indicates uniform distribution among that specific microbial taxa.

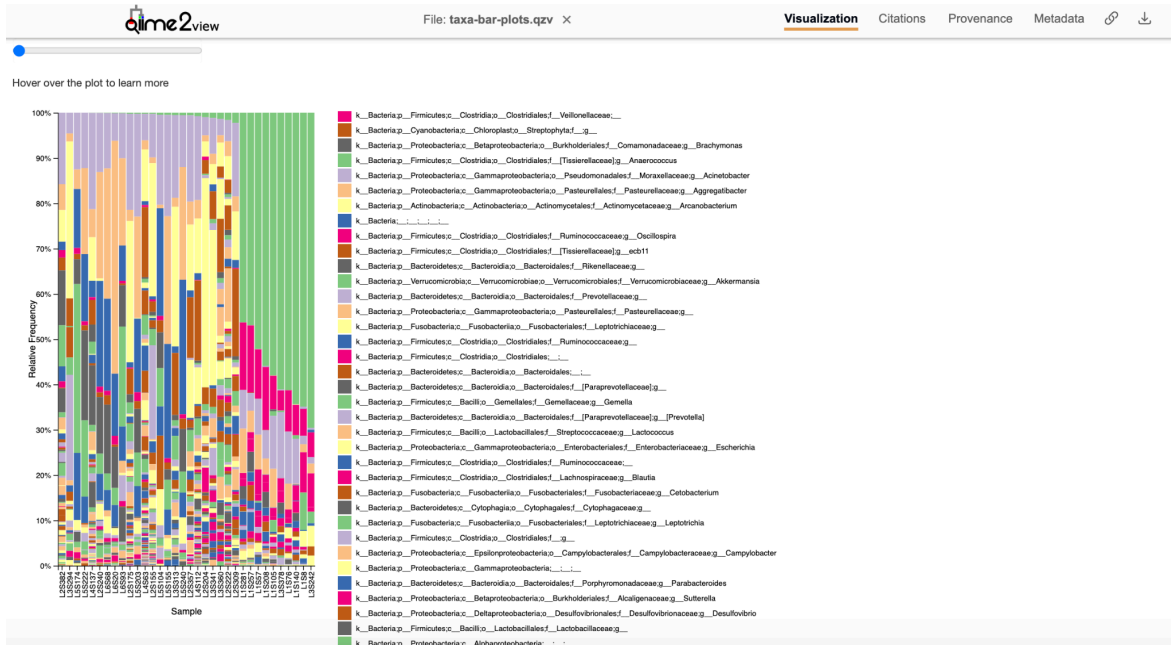
# QIIME Tutorial

## Alpha Diversity - Evenness by Antibiotic Usage



**Figure 17:** Alpha diversity microbial evenness by antibiotic usage. The group with no antibiotic usage had a greater microbial evenness distribution, this group dominates.

## Taxonomic Analysis



**Figure 18:** Taxonomic barplot demonstrating microbial composition at genus level. Large green group represents the Bacteroides genus. Commonly found in the human gut microbiome.