

1. ¿Cuántas jaulas distintas se utilizaron en el estudio? ¿Cuántos ratones había en cada una de ellas? ¿Cuántos posibles valores tiene la variable categórica genotype_and_donor_status y cómo se llaman? Explica cada una de estas variables.

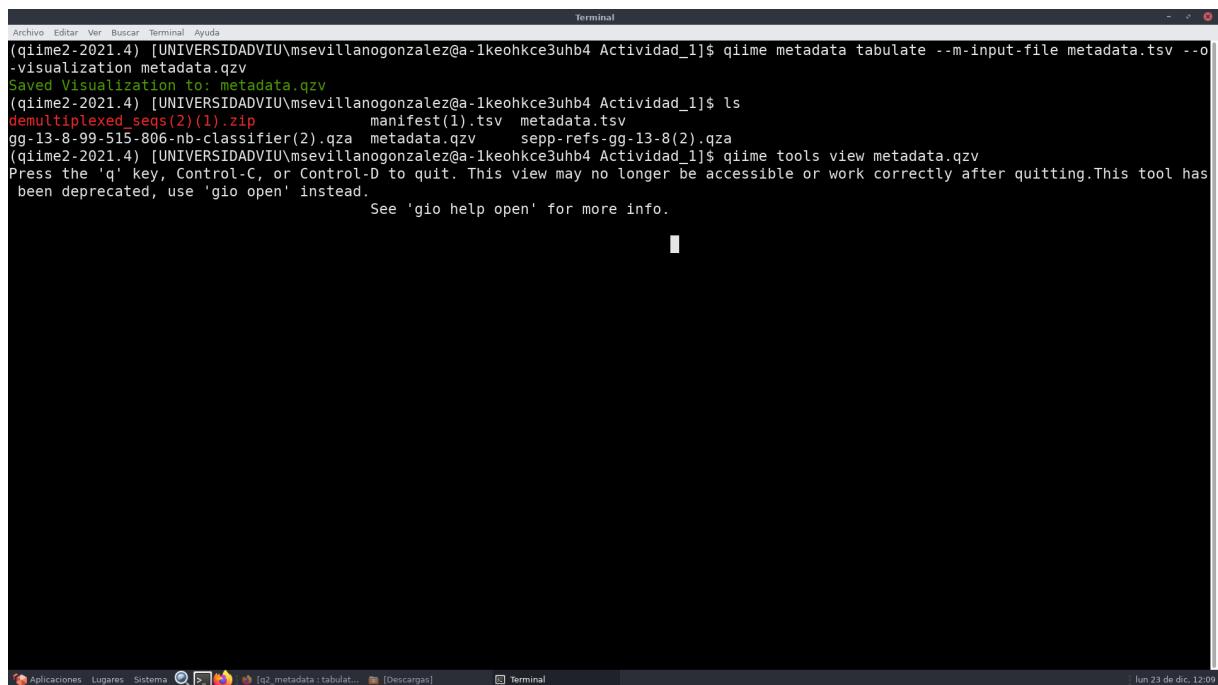
Se utilizaron 6 jaulas distintas, con 2 ratones en cada una de ellas.

La variable categórica genotype_and_donor_status tiene 4 posibles valores:

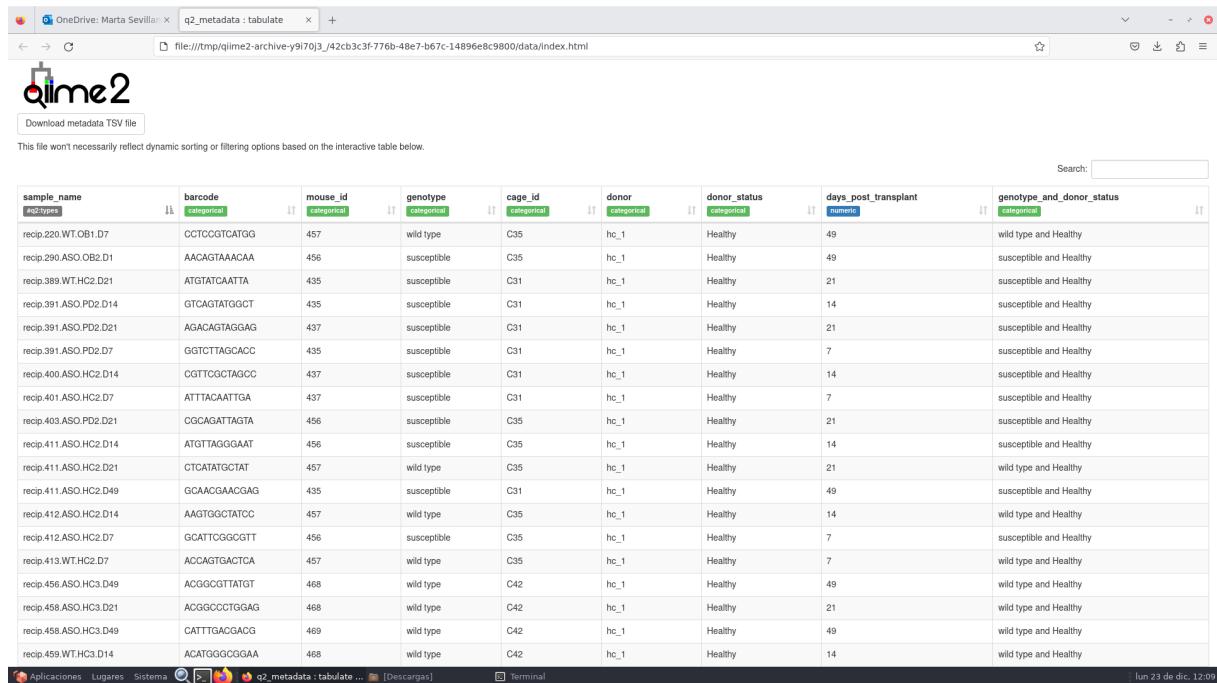
wild type and Healthy
susceptible and Healthy
wild type and PD
susceptible and PD

Por un lado, está el genotipo “wild type”, que es el tipo salvaje (BDF1), que no tiene ningún riesgo a nivel genético de sufrir la enfermedad. Y el tipo “susceptible” (Thy1-aSyn), que está genéticamente predisponente a la enfermedad.

Por otro lado, el estado del donante, “PD” (Parkinson’s disease), si el donante tiene enfermedad de Parkinson o “Healthy”, si el donante está neurológicamente sano y no tiene la enfermedad.



```
Archivo Editar Ver Buscar Terminal Ayuda
(qiime2-2021.4) [UNIVERSIDADVIU\msevillanogonzalez@a-1keohkce3uhb4 Actividad_1]$ qiime metadata tabulate --m-input-file metadata.tsv --o
-visualization metadata.qzv
Saved Visualization to: metadata.qzv
(qiime2-2021.4) [UNIVERSIDADVIU\msevillanogonzalez@a-1keohkce3uhb4 Actividad_1]$ ls
demultiplexed_seqs(2)(1).zip manifest(1).tsv metadata.tsv
gg-13-8-99-515-806-nb-classifier(2).qza metadata.qzv sepp-refs-gg-13-8(2).qza
(qiime2-2021.4) [UNIVERSIDADVIU\msevillanogonzalez@a-1keohkce3uhb4 Actividad_1]$ qiime tools view metadata.qzv
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be accessible or work correctly after quitting. This tool has
been deprecated, use 'gio open' instead.
See 'gio help open' for more info.
```



sample_name	barcode	mouse_id	genotype	cage_id	donor	donor_status	days_post_transplant	genotype_and_donor_status
seq-type	categorical	categorical	categorical	categorical	categorical	categorical	numeric	categorical
recip.220.WT.0B1.D7	CCTCCGTCATGG	457	wild type	C35	hc_1	Healthy	49	wild type and Healthy
recip.290.ASO.0B2.D1	AACAGTAAACAA	456	susceptible	C35	hc_1	Healthy	49	susceptible and Healthy
recip.389.WT.HC2.D21	ATGTATCAATT	435	susceptible	C31	hc_1	Healthy	21	susceptible and Healthy
recip.391.ASO.PD2.D14	GTCAGTATGGCT	435	susceptible	C31	hc_1	Healthy	14	susceptible and Healthy
recip.391.ASO.PD2.D21	AGACAGTAGAG	437	susceptible	C31	hc_1	Healthy	21	susceptible and Healthy
recip.391.ASO.PD2.D7	GGTCCTTAGCACC	435	susceptible	C31	hc_1	Healthy	7	susceptible and Healthy
recip.400.ASO.HC2.D14	CGTTGGCTAGCC	437	susceptible	C31	hc_1	Healthy	14	susceptible and Healthy
recip.401.ASO.HC2.D7	ATTTACAAATGA	437	susceptible	C31	hc_1	Healthy	7	susceptible and Healthy
recip.403.ASO.PD2.D21	CGCAGATTAGTA	456	susceptible	C35	hc_1	Healthy	21	susceptible and Healthy
recip.411.ASO.HC2.D14	ATGTTAGGGAT	456	susceptible	C35	hc_1	Healthy	14	susceptible and Healthy
recip.411.ASO.HC2.D21	CTCATATGCTAT	457	wild type	C35	hc_1	Healthy	21	wild type and Healthy
recip.411.ASO.HC2.D49	GCAACGAAAGAG	435	susceptible	C31	hc_1	Healthy	49	susceptible and Healthy
recip.412.ASO.HC2.D14	AAGTGGCTATCC	457	wild type	C35	hc_1	Healthy	14	wild type and Healthy
recip.412.ASO.HC2.D7	GCATTCCGGCGTT	456	susceptible	C35	hc_1	Healthy	7	susceptible and Healthy
recip.413.WT.HC2.D7	ACCAAGTGAATCA	457	wild type	C35	hc_1	Healthy	7	wild type and Healthy
recip.456.ASO.HC3.D49	ACGGCGTTATGT	468	wild type	C42	hc_1	Healthy	49	wild type and Healthy
recip.458.ASO.HC3.D21	ACGGGCCCTGGAG	468	wild type	C42	hc_1	Healthy	21	wild type and Healthy
recip.458.ASO.HC3.D49	CATTGACGACG	469	wild type	C42	hc_1	Healthy	49	wild type and Healthy
recip.459.WT.HC3.D14	ACATGGCGGAA	468	wild type	C42	hc_1	Healthy	14	wild type and Healthy

2. ¿Cuál es el comando de QIIME 2 que deberíamos haber utilizado con lecturas single end si quisiéramos hacer un demultiplexado? Asume que tenemos las secuencias en un archivo llamado emp-single-end-sequences.qza, que el fichero con los metadatos se llama metadata.tsv, que tiene una columna llamada BarcodeSequence y que queremos que el archivo de salida se llame demux_seqs.qza.

Comando a usar para hacer el demultiplexado con lecturas single end

```
qiime demux emp-single \
--i-sequences emp-single-end-sequences.qza \
--m-barcodes-file metadata.tsv \
--m-barcodes-column BarcodeSequence \
--o-per-sample-sequences demux_seqs.qza
```

3. Tras haber realizado el filtrado con DADA 2, ¿Cuántas lecturas quedan en la muestra recip.460.WT.HC3.D49? ¿Qué te parece este resultado? Incluye un pantallazo.

Tras utilizar el comando siguiente:

```
qiime dada2 denoise-single --i-demultiplexed-sequences ./demux_seqs.qza --p-trunc-len 150 --o-table ./dada2_table.qza --o-representative-sequences dada2_rep_set.qza --o-denoising-stats ./dada2_stats.qza
```

con el parámetro de longitud de truncamiento establecido en 150 pares de bases, que recortará las secuencias en ese punto y eliminará secuencias que tengan baja calidad al final, de 16.327 lecturas que había, tras el filtrado quedan 9.919 lecturas, y tras quitar el ruido quedan 347 lecturas.

El porcentaje de lecturas que han pasado el filtrado es de 60.75%, por lo que no es muy malo, aunque eliminar casi un 30% significa que esa parte de las lecturas tenían una calidad baja. Además, tras quitar el ruido quedan 347 lecturas, que sí es un número bajo para los análisis de diversidad y taxonomía de muestras microbianas y puede limitar la representatividad de la muestra.

```
Terminal
Archivo Editar Ver Buscar Terminal Ayuda
(qiime2-2021.4) [UNIVERSIDADADVIU\msevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime tools import --type "SampleData[SequencesWithQuality]" --input-format SingleEndFastqManifestPhred33V2 --input-path ./manifest.tsv --output-path ./demux_seqs.qza
Imported ./manifest.tsv as SingleEndFastqManifestPhred33V2 to ./demux_seqs.qza
(qiime2-2021.4) [UNIVERSIDADADVIU\msevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime demux summarize --i-data ./demux_seqs.qza --o-visualization ./demux_seqs.qzv
Saved Visualization to: ./demux_seqs.qzv
(qiime2-2021.4) [UNIVERSIDADADVIU\msevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime tools view demux_seqs.qz
```

Aplicaciones Lugares Sistema [q2_metadata : tabulat... Terminal mié 25 de dic, 11:32

```
Terminal
Archivo Editar Ver Buscar Terminal Ayuda
(qiime2-2021.4) [UNIVERSIDADADVIU\msevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime dada2 denoise-single --i-demultiplexed-seqs ./demux_seqs.qza --p-trunc-len 150 --o-table ./dada2_table.qza --o-representative-sequences dada2_rep_set.qza --o-denoising-stats ./dada2_stats.qza
Saved FeatureTable[Frequency] to: ./dada2_table.qza
Saved FeatureData[Sequence] to: dada2_rep_set.qza
Saved SampleData[DADA2Stats] to: ./dada2_stats.qza
(qiime2-2021.4) [UNIVERSIDADADVIU\msevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime metadata tabulate --m-input-file ./dada2_stats.qza --o-visualization ./dada2_stats.qzv
Saved Visualization to: ./dada2_stats.qzv
(qiime2-2021.4) [UNIVERSIDADADVIU\msevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime tools view dada2_stats.qzv
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be accessible or work correctly after quitting. This tool has been deprecated, use 'gio open' instead.
See 'gio help open' for more info.
```

Aplicaciones Lugares Sistema [q2_metadata : tabulat... Terminal mié 25 de dic, 18:55

sample-id	input	filtered	percentage of input passed filter	denoised	non-chimeric	percentage of input non-chimeric
recip-460.WT.HC3.D49	16327	9919	60.75	347	347	2.13

Showing 1 to 1 of 1 entries (filtered from 48 total entries)

Search: recip-460.WT.HC3.D49

Previous Next

mié 25 de dic, 18:56

4. ¿Cuáles son las tres categorías taxonómicas más abundantes en frecuencia relativa si usamos 5 niveles de clasificación taxonómica? Adjunta un pantallazo.

Las categorías taxonómicas más abundantes en frecuencia relativa usando 5 niveles de clasificación taxonómica son:

Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Bacteroidaceae
 Bacteria; Firmicutes; Clostridia; Clostridiales; Lachnospiraceae
 Bacteria; Bacteroidetes; Bacteroidia; Bacteroidales; Porphyromonadaceae

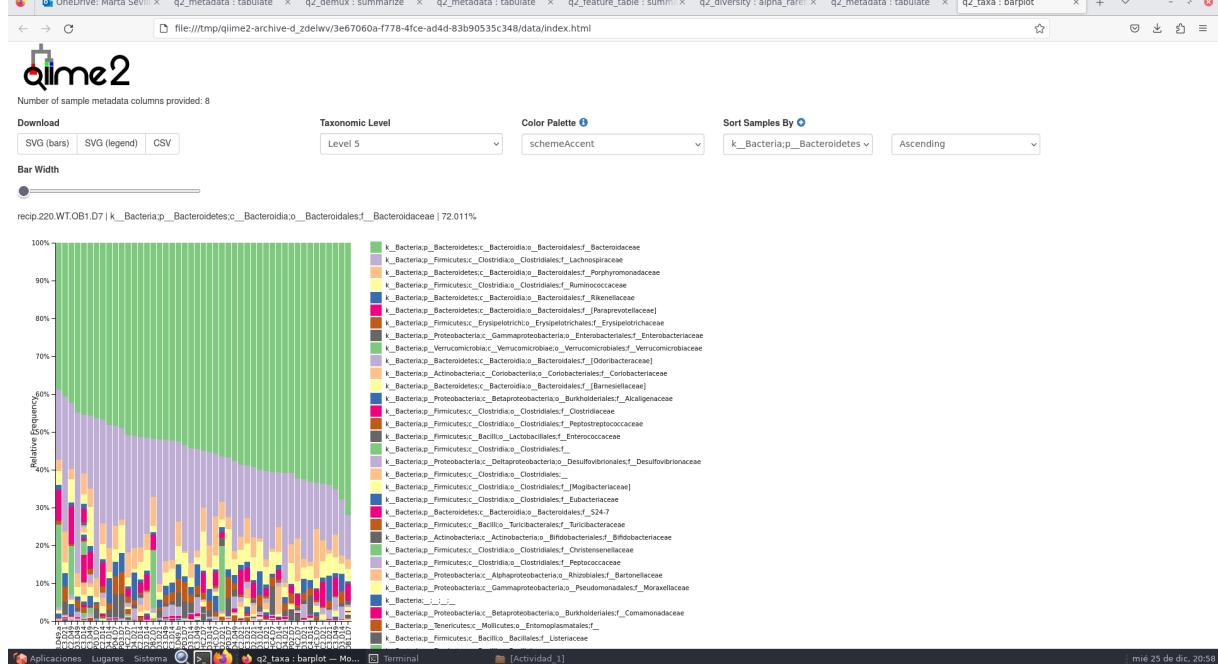
```

Archivo Editar Ver Buscar Terminal Ayuda
Terminal
(qiime2-2021.4) [UNIVERSIDADADIVIUmsevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime feature-classifier classify-sklearn --i-reads ./dada2_rep_set.qza --i-classifier ./gg-13-8-99-515-806-nb-classifier.qza --o-classification ./taxonomy.qza
Saved FeatureData[Taxonomy] to: ./taxonomy.qza
(qiime2-2021.4) [UNIVERSIDADADIVIUmsevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime metadata tabulate --m-input-file ./taxonomy.qza --o-visualization ./taxonomy.qzv
Saved Visualization to: ./taxonomy.qzv
(qiime2-2021.4) [UNIVERSIDADADIVIUmsevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime tools view taxonomy.qzv
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be accessible or work correctly after quitting. This tool has been deprecated, use 'gio open' instead.
See 'gio help open' for more info.

(qiime2-2021.4) [UNIVERSIDADADIVIUmsevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime feature-table filter-samples --i-table ./dada2_table.qza --p-min-frequency 2000 --o-filtered-table ./table_2k.qza
Saved FeatureTable[Frequency] to: ./table_2k.qza
(qiime2-2021.4) [UNIVERSIDADADIVIUmsevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime taxa barplot --i-table ./table_2k.qza --i-taxonomy ./taxonomy.qza --m-metadata-file ./metadata.tsv --o-visualization ./taxa_barplot.qzv
Saved Visualization to: ./taxa_barplot.qzv
(qiime2-2021.4) [UNIVERSIDADADIVIUmsevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime tools view taxa_barplot.qzv
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be accessible or work correctly after quitting. This tool has been deprecated, use 'gio open' instead.
See 'gio help open' for more info.

```

mié 25 de dic. 20:57



mié 25 de dic. 20:58

5. ¿En qué jaula hay más diversidad, según el análisis de rarefacción? Adjunta un pantallazo.

Los comandos usados para realizar el análisis de rarefacción y su visualización son:

```
qiime diversity alpha-rarefaction --i-table ./dada2_table.qza --m-metadata-file ./metadata.tsv --o-visualization ./alpha_rarefaction_curves.qzv --p-min-depth 10 --p-max-depth 4250
```

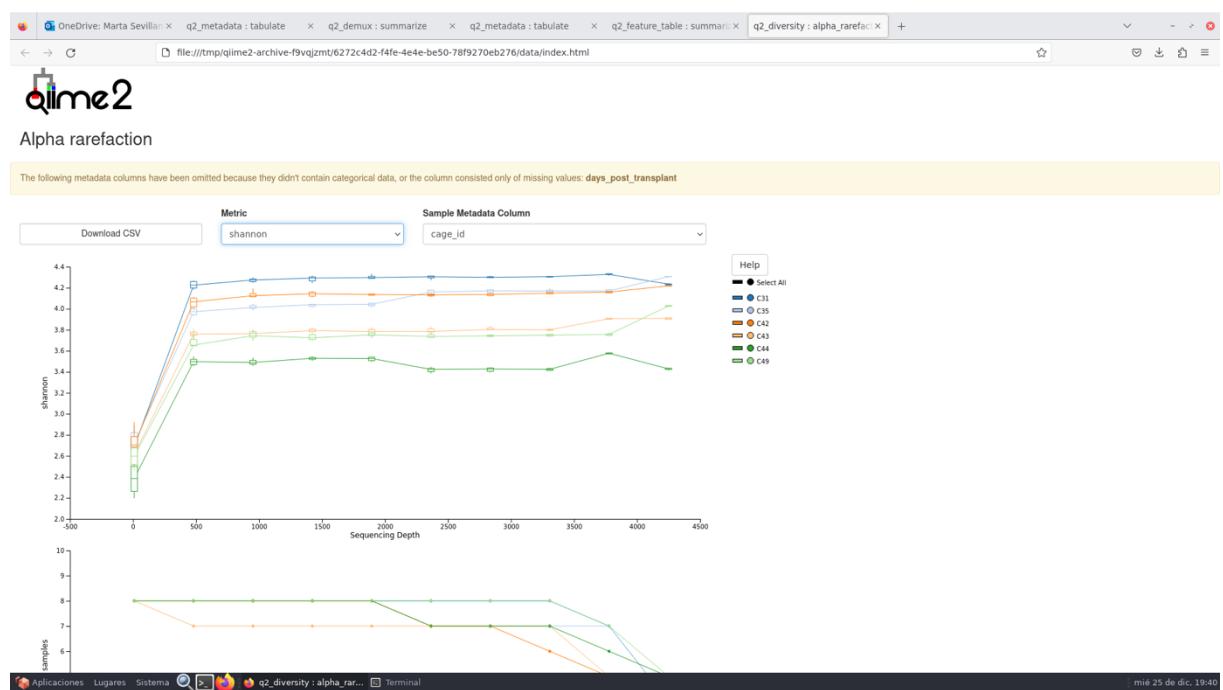
```
qiime tools view alpha_rarefaction_curves.qzv
```

Estableciendo unos parámetros de profundidad mínima de secuencias a considerar de 10 y de profundidad máxima de 4250.

```
Terminal
Archivo Editar Ver Buscar Terminal Ayuda
(qiime2-2021.4) [UNIVERSIDADADVIU\msevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime dada2 denoise-single --i-demultiplexed-seqs ./demux_seqs.qza --p-trunc-len 150 --o-table ./dada2_table.qza --o-representative-sequences dada2_rep_set.qza --o-denoising-stats ./dada2_stats.qza
Saved FeatureTable[Frequency] to: ./dada2_table.qza
Saved FeatureData[Sequence] to: dada2_rep_set.qza
Saved SampleData[DADA2Stats] to: ./dada2_stats.qza
(qiime2-2021.4) [UNIVERSIDADADVIU\msevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime metadata tabulate --m-input-file ./dada2_stats.qza --o-visualization ./dada2_stats.qzv
Saved Visualization to: ./dada2_stats.qzv
(qiime2-2021.4) [UNIVERSIDADADVIU\msevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime tools view dada2_stats.qzv
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be accessible or work correctly after quitting. This tool has been deprecated, use 'gio open' instead.
See 'gio help open' for more info.

(qiime2-2021.4) [UNIVERSIDADADVIU\msevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ 
(qiime2-2021.4) [UNIVERSIDADADVIU\msevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime feature-table summarize --i-table ./dada2_table.qza --m-sample-metadata-file ./metadata.tsv --o-visualization ./dada2_table.qzv
Saved Visualization to: ./dada2_table.qzv
(qiime2-2021.4) [UNIVERSIDADADVIU\msevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime tools view dada2_table.qzv
Press the 'q' key, Control-C, or Control-D to quit. This view may no longer be accessible or work correctly after quitting. This tool has been deprecated, use 'gio open' instead.
See 'gio help open' for more info.

(qiime2-2021.4) [UNIVERSIDADADVIU\msevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ 
(qiime2-2021.4) [UNIVERSIDADADVIU\msevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime diversity alpha-rarefaction --i-table ./dada2_table.qza --m-metadata-file ./metadata.tsv --o-visualization ./alpha_rarefaction_curves.qzv
Saved Visualization to: ./alpha_rarefaction_curves.qzv
(qiime2-2021.4) [UNIVERSIDADADVIU\msevillanogonzalez@-1keohkce3uhb4 Actividad_1]$ qiime tools view alpha_rarefaction_curves.qzv
```



Según el análisis de rarefacción, la jaula C31 es la que tiene mayor diversidad alfa, ya que tiene el índice de Shannon más alto y permanece estable incluso en los niveles más altos de profundidad de secuenciación. También se observa que casi todas las muestras alcanzan el máximo de diversidad con más o menos el mismo número de lecturas o profundidad de secuenciación y alcanzan la meseta rápidamente. La jaula C44 y la C35, dado que no es una meseta constante, puede ser que no se haya secuenciado con la profundidad suficiente para representar toda la diversidad existente en la muestra.

