



Caso de Estudio Complejivo

Fecha de inicio:	
Fecha de finalización	
Nombre del caso	Análisis metataxonómico de microbiota intestinal humana
Número Caso	3
Data para el caso	Secuencias.rar Examen Complejivo 2

Problemática:

El microbioma humano desempeña un papel fundamental en el mantenimiento de la homeostasis y la salud del hospedador. Alteraciones en la composición y estructura de estas comunidades microbianas, fenómeno conocido como disbiosis, han sido vinculadas con diversas patologías y estados fisiológicos alterados.

El presente caso de estudio se enmarca en una investigación clínica diseñada para evaluar el impacto de diferentes intervenciones o estados sobre la microbiota comensal. El objetivo principal es determinar si existen diferencias significativas en los perfiles taxonómicos de las comunidades bacterianas entre tres grupos de estudio: Condición 1, Condición 2 y un grupo Control.

A través de un análisis metataxonómico, se busca identificar molecularmente los microorganismos presentes y caracterizar los cambios en la abundancia y diversidad que definen a cada tratamiento. Para la obtención de los datos, se realizó la extracción de ADN total de muestras biológicas humanas, seguida de la amplificación y secuenciación de la región hipervariable del gen 16S ARNr utilizando la tecnología Illumina.

El set de datos provisto para este examen consiste en las secuencias crudas (archivos fastq) de tres réplicas por cada grupo:

Grupo Control: Muestras C1, C2, C3.

Tratamiento Condición 1: Muestras T1R1, T1R2, T1R3.

Tratamiento Condición 2: Muestras T2R1, T2R2, T2R3.

Se espera que el estudiante demuestre su capacidad para procesar estos datos masivos, interpretar las métricas ecológicas resultantes y determinar si existen taxones con abundancia diferencial que puedan actuar como biomarcadores para las condiciones estudiadas.

Objetivos:

- Generar variantes de secuencia de amplicón (ASVs) a partir de los datos crudos
- Determinar la abundancia y la diversidad alfa de los microorganismos en cada muestra
- Determinar si existen diferencias en diversidad beta entre tratamientos
- Comparar la abundancia de los principales géneros bacterianos presentes en los diferentes tratamientos
- Encontrar ASVs con abundancia diferencial

Proceso:

1. Utilizando el programa dada2 se procederá a limpiar las secuencias, generar los ASVs, eliminar quimeras, empalmar secuencias, crear tabla de secuencias, asignar taxonomía a las secuencias
2. Con Phyloseq se determinará la abundancia de géneros, la diversidad alfa y la diversidad beta
3. Con ampvis2 se creará un mapa de calor
4. Con deseq2 se determinarán los ASVs con abundancia diferencial

CONSIDERAR:

- El pipeline de dada2, complementado con el uso de phyloseq, ranacapa, microeco, microbiome, deseq2 y varias librerías para generación de gráficos y manipulación de datos que se revisaron en clase, pero también existen recursos en línea que pueden ser consultados, como por ejemplo: https://benjjneb.github.io/dada2/tutorial_1_8.html

- Se recomienda el uso de una computadora con por lo menos 8Gb en RAM para el análisis de datos
- Se entregará al estudiante un script como referencia. Este script contiene los códigos necesarios para cumplir con varias tareas, incluyendo las que se plantean en este caso. El estudiante deberá generar su propio script que contenga **únicamente** los códigos utilizados para este caso.

Retos:

- Procesamiento de datos de secuenciación masiva: La cantidad de secuencias a analizar es abundante y requiere de una capacidad de cómputo elevada
- Uso de lenguaje R: Los programas propuestos para el análisis requieren que el estudiante pueda ejecutarlos con comandos de R. Se recomienda el uso de RStudio

Entregables:

- Informe final que contenga al menos (las imágenes incluidas a continuación servirán de referencia para el estudiante):
 - Gráficos de calidad de las secuencias crudas (basta con mostrar 2 gráficos)



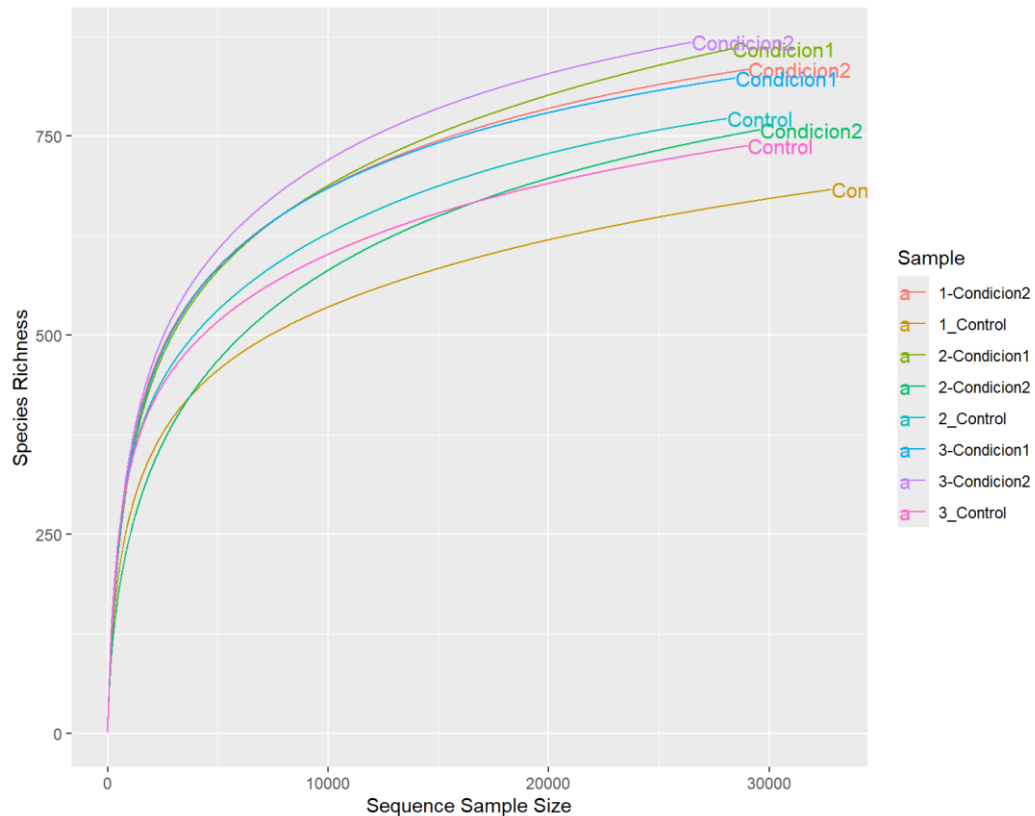
- Tabla que indique cuántas secuencias se fueron perdiendo durante el procesamiento en dada2 (ejemplo en la primera fila).

Muestra	Input	Filtered	Denoisedf	Denoisedr	Merged	Nonchim
1MR1	42630	13151	12766	12708	10025	6370

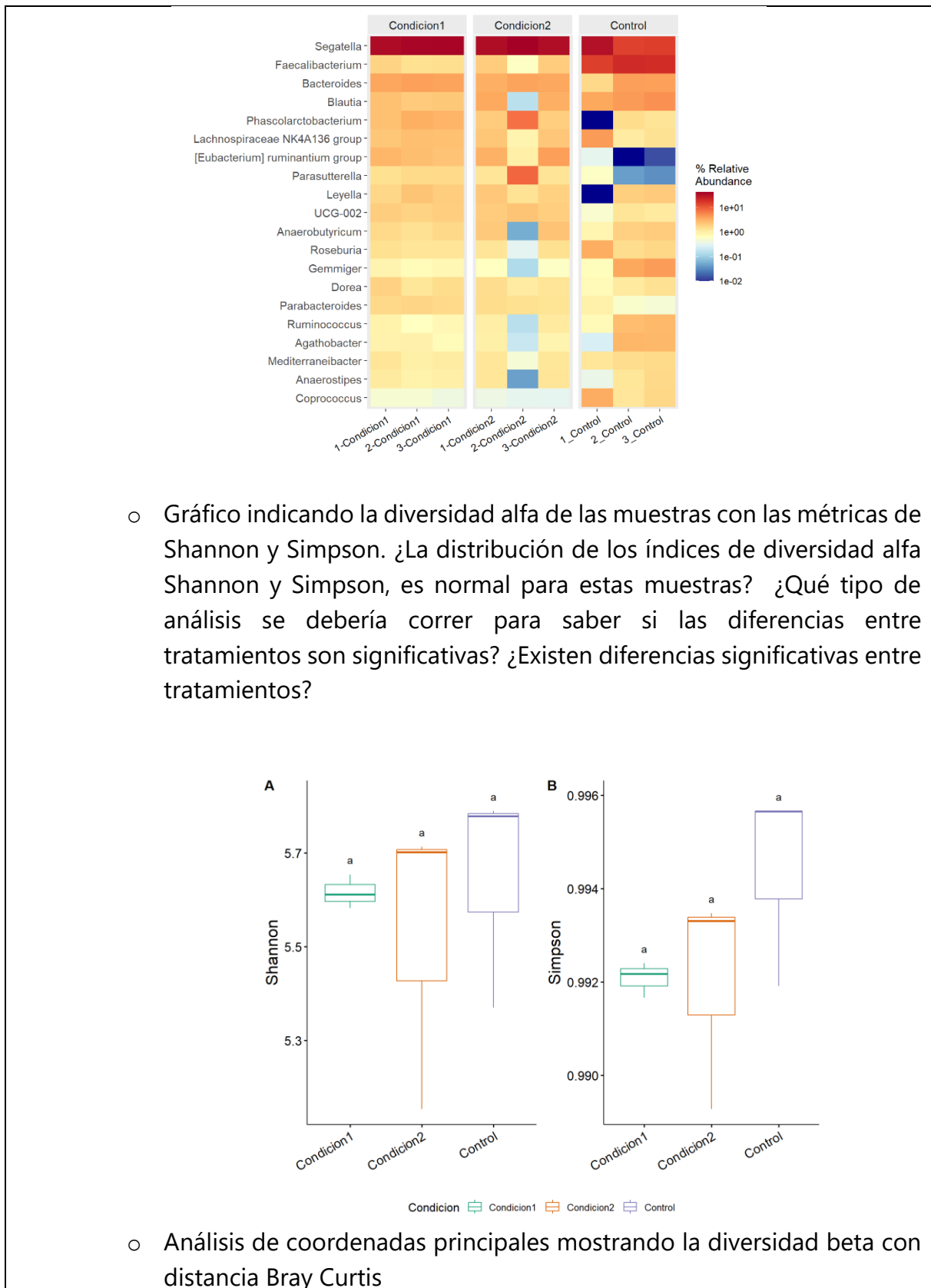


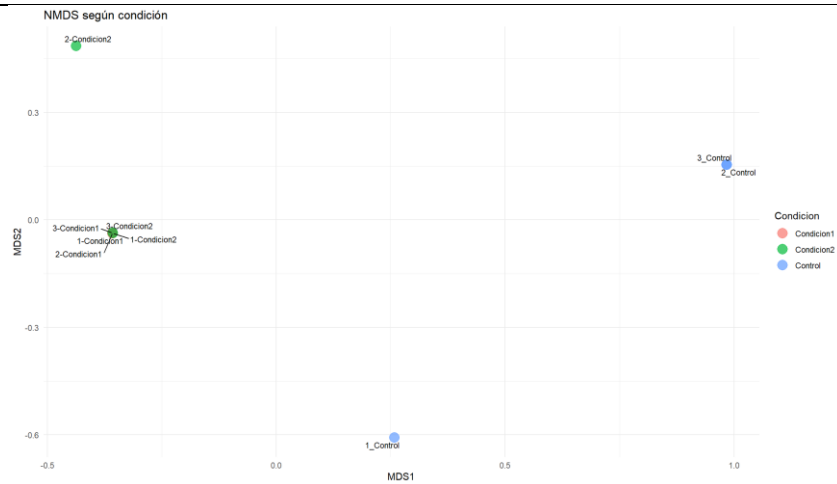
1MR2
1MR3
3MR1
3MR2
3MR3

- Curvas de rarefacción que muestren que el muestreo fue lo suficientemente profundo

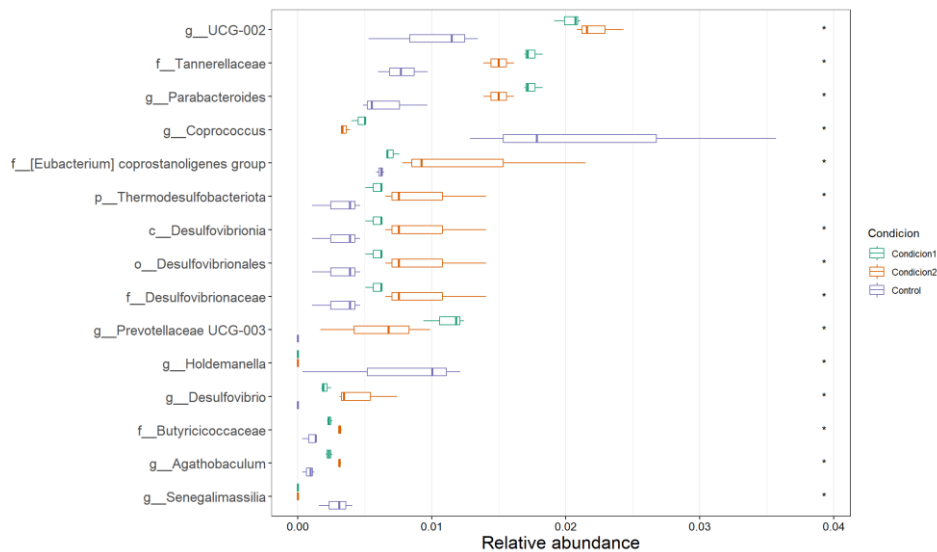


- Gráficos de barras que indiquen la abundancia de los principales géneros por tratamiento





- Gráficos que muestren cuales son los principales ASVs con abundancia diferencial

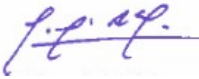



- Script con los comandos utilizados para procesar los datos, incluyendo comentarios básicos que indiquen para que sirve cada bloque de comandos (e.g. #empalme de secuencias, #eliminación de quimeras, #asignación taxonómica)



**Pontificia Universidad
Católica del Ecuador**

Seréis mis testigos

 		
Decano Facultad de Hábitat, Infraestructura y Creatividad	Coordinador Maestría en Biología Computacional	Tutor de titulación

