

# FICHA TÉCNICA DEL CURSO

## Análisis de Datos 16S rRNA con DADA2

TC-BIO-02

<b>Código:</b>	FOR-EDU-005
<b>Tipo de documento:</b>	FORMATO
<b>Fecha de emisión:</b>	Enero 2026
<b>Versión:</b>	1.0
<b>Elaboró:</b>	Dirección General
<b>Revisó:</b>	Dirección General
<b>Aprobó:</b>	Dirección General

## INFORMACIÓN DEL FORMATO

Este formato se utiliza para documentar la ficha técnica completa de cada curso o taller impartido por BioSeryl, especificando la estructura curricular, contenidos, metodología, evaluación y requisitos, conforme a NTC-ISO 21001:2025, Cláusula 8.1 (Planificación operacional) y Cláusula 8.5 (Prestación del servicio educativo). Una copia de esta ficha debe existir para cada curso activo.

*Retención: Durante vigencia del curso + 2 años después de obsolescencia.*

## IDENTIFICACIÓN DEL PROGRAMA

<b>Código del curso:</b>	TC-BIO-02
<b>Nombre completo:</b>	Análisis de Datos 16S rRNA con DADA2
<b>Nombre corto (para marketing):</b>	16S rRNA con DADA2
<b>Versión del currículo:</b>	1.0
<b>Fecha de creación:</b>	Enero 2026
<b>Última actualización:</b>	Enero 2026
<b>Próxima revisión:</b>	Enero 2027
<b>Estado:</b>	<input type="checkbox"/> En desarrollo <input checked="" type="checkbox"/> Activo

## CLASIFICACIÓN

<b>Línea de producto:</b>	<input checked="" type="checkbox"/> BIOINFORMATICA
<b>Tipo de programa:</b>	<input checked="" type="checkbox"/> Curso especializado (40–80h)
<b>Nivel:</b>	<input checked="" type="checkbox"/> Básico (sin prerrequisitos técnicos)
<b>Modalidad:</b>	<input checked="" type="checkbox"/> Sincrónico en vivo (videoconferencia)

## DURACIÓN Y ESTRUCTURA

<b>Duración total:</b>	40 horas
<b>Distribución:</b>	Teoría: 40 %      Práctica: 60 %
<b>Número de módulos:</b>	12
<b>Número de sesiones:</b>	12
<b>Duración por sesión:</b>	3 horas
<b>Frecuencia sugerida:</b>	2–3 veces/semana

## PÚBLICO OBJETIVO

<b>Perfil ideal:</b>	Biólogos, microbiólogos, investigadores en microbioma; estudiantes de posgrado con proyectos de diversidad microbiana
<b>Formación recomendada:</b>	Biología, microbiología, ciencias afines (mínimo pregrado)
<b>Prerrequisitos técnicos:</b>	Conocimientos básicos de biología molecular y comando de línea
<b>Prerrequisitos de software:</b>	R + RStudio
<b>Cupo mínimo:</b>	5 participantes



<b>Cupo máximo:</b>	25 participantes
---------------------	------------------

## DESCRIPCIÓN DEL PROGRAMA

### Descripción:

Este curso forma al participante en el análisis de datos de secuenciación 16S rRNA utilizando el pipeline DADA2, desde la inspección de calidad de las secuencias FASTQ hasta la generación de tablas de variantes de secuencia ampliada (ASV) y su asignación taxonómica. Se abordan conceptos fundamentales de diversidad microbiana, filtrado de secuencias, corrección de errores, chimera removal y análisis de diversidad alfa y beta. El participante aplicará estas técnicas sobre datasets reales de microbiota intestinal, oral y ambiental, desarrollando competencias para investigaciones de microbioma con estándares de calidad reproducibles.

## RESULTADOS DE APRENDIZAJE

Al finalizar este programa, el participante será capaz de:

No.	Resultado de aprendizaje
RA1	Ejecutar el pipeline DADA2 completo desde FASTQ hasta tabla ASV con parámetros optimizados.
RA2	Asignar taxonomía evaluando confidence scores y comprendiendo limitaciones por base de datos.
RA3	Analizar diversidad alfa y beta seleccionando métricas apropiadas según pregunta biológica.
RA4	Comparar paradigmas ASVs vs OTUs argumentando la selección según contexto del estudio.
RA5	Generar visualizaciones publication-ready que cumplan estándares editoriales.

**CONTENIDO MODULAR**

Módulo	Título	Horas	Temas principales	RA
1	Introducción a datos de amplicones 16S	3	Regiones variables, plataformas de secuenciación, formatos de datos, limpieza inicial	RA1
2	QIIME2: instalación y entorno	3	Instalación QIIME2, artefactos qza/qzv, primeros pasos en la interfaz	RA1
3	DADA2: denoising y ASVs	4	Algoritmo DADA2, calidad por posición, filtrado, inferencia de ASVs, fusión de lecturas	RA2
4	Asignación taxonómica	3	Bases de datos (SILVA, GTDB, RDP), clasificador naive Bayes, confidence thresholds	RA2
5	Árbol filogenético con SATé-enabled	3	Alineamiento MAFFT, FastTree, RAxML-NG, integración en QIIME2	RA2
6	Diversidad alfa	3	Índices Shannon, Chao1, Faith PD, curvas de rarefacción, comparación entre grupos	RA3
7	Diversidad beta	4	UniFrac (ponderado/no ponderado), PCoA, PERMANOVA, dispersión beta	RA3
8	Visualización de perfiles taxonómicos	3	Barras apiladas, heatmaps, gráficos de abundancia diferencial, phyloseq	RA4
9	Análisis de biomarcadores	4	LEfSe, ANCOM-BC, DESeq2 para 16S, random forest, curvas ROC	RA4
10	Co-ocurrencia y redes microbianas	3	SparCC, SPIEC-EASI, igraph, visualización de redes de co-ocurrencia	RA4
11	PICRUSt2: predicción funcional	4	PICRUSt2, inferencia de vías metabólicas, KEGG/EC, visualización	RA5
12	Publicación y reproducibilidad	3	Notebooks reproducibles, GitHub, reportes finales, repositorios de datos	RA5
<b>Total</b>		<b>40</b>		

## METODOLOGÍA

### Enfoque pedagógico

- |  |
|--|
| <input checked="" type="checkbox"/> Aprendizaje activo con ejercicios prácticos<br><input checked="" type="checkbox"/> Estudio de casos reales |
|--|

### Descripción metodológica

El curso utiliza un enfoque de aprendizaje activo donde el participante ejecuta un pipeline completo de análisis de 16S rRNA en cada sesión. Cada módulo incluye: (1) exposición teórica de conceptos clave (diversidad microbiana, ASVs vs OTUs, bases de datos taxonómicas), (2) demostración en vivo del pipeline DADA2, y (3) práctica supervisada con datasets reales. Se fomenta la discusión de casos y la resolución de problemas comunes en análisis de microbioma.

## EVALUACIÓN

### Instrumentos de evaluación

Instrumento	Peso ( %)	Descripción
Ejercicios durante sesiones	40 %	Ejercicios prácticos por módulo
Proyecto final/integrador	50 %	Pipeline/proyecto integrador con reporte
Participación activa	10 %	Asistencia y participación en sesiones
<b>Total</b>	<b>100 %</b>	

### Criterios de aprobación

<b>Modalidad de aprobación:</b>	<input checked="" type="checkbox"/> Ambos
<b>Criterio:</b>	Calificación mínima: 4.0/5.0 Asistencia mínima: 80 %
<b>Certificado otorgado:</b>	<input checked="" type="checkbox"/> Digital

## MATERIALES Y RECURSOS

Tipo	Descripción
Scripts de código	Scripts R completos comentados (pipeline reproducible)
Datasets de práctica	Datasets de ejemplo: 16S V4 (microbioma), ITS2 (micobioma)
Lecturas complementarias	Bases de datos taxonómicas: SILVA 138, UNITE 9.0
Lecturas complementarias	Guía de interpretación de resultados



Tipo	Descripción
Lecturas complementarias	Checklist de buenas prácticas bioinformáticas
Lecturas complementarias	Certificado digital de participación BioSeryl

## BIBLIOGRAFÍA DE REFERENCIA

No.	Referencia bibliográfica
1	Callahan, B.J. et al. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. <i>Nature Methods</i> , 13(7), 581–583.
2	Carding, S.R. (ed.) (2026). <i>Best Practice in Microbiome Research</i> . Springer.
3	Mitra, S. (ed.) (2023). <i>Metagenomic Data Analysis</i> . Springer (Methods in Molecular Biology, 2649).
4	Pevsner, J. (2015). <i>Bioinformatics and Functional Genomics</i> . 3ra ed. Wiley-Blackwell.

**INSTRUCTOR(ES) ASIGNADO(S)**

Nombre	FOR-EDU-004

**CONTENIDO PROGRAMÁTICO**
**Módulo 1 — Introducción a datos de amplicones 16S (3 horas)**

Unidad	Contenido
1.1	Regiones variables del gen 16S rRNA: V1-V9, regiones hipervariables y conservadas
1.2	Plataformas de secuenciación para amplicones 16S: Illumina, PacBio, ONT
1.3	Formatos de datos: FASTQ, calidad, metadatos de muestras
1.4	Limpieza inicial y control de calidad de amplicones

**Módulo 2 — QIIME2: instalación y entorno (3 horas)**

Unidad	Contenido
2.1	Instalación de QIIME2: entornos Conda y Docker
2.2	Artefactos qza/qzv: estructura y manipulación
2.3	Primeros pasos en la interfaz QIIME2
2.4	Importación de datos y validación de metadatos

**Módulo 3 — DADA2: denoising y ASVs (4 horas)**

Unidad	Contenido
3.1	Algoritmo DADA2: modelo de error y corrección
3.2	Calidad por posición y filtrado de secuencias
3.3	Inferencia de ASVs y fusión de lecturas
3.4	Eliminación de quimeras y tabla de ASVs

**Módulo 4 — Asignación taxonómica (3 horas)**

Unidad	Contenido
4.1	Bases de datos taxonómicas: SILVA, GTDB, RDP, UNITE
4.2	Clasificador naive Bayes: entrenamiento y parámetros
4.3	Confidence thresholds y evaluación de asignaciones
4.4	Limitaciones por base de datos y sesgos taxonómicos



### Módulo 5 — Árbol filogenético con SATé-enabled (3 horas)

Unidad	Contenido
5.1	Alineamiento MAFFT para secuencias de amplicones
5.2	FastTree: inferencia filogenética rápida
5.3	RAXML-NG: filogenia de máxima verosimilitud
5.4	Integración de árbol filogenético en QIIME2

### Módulo 6 — Diversidad alfa (3 horas)

Unidad	Contenido
6.1	Índices Shannon, Chao1, Faith PD: fundamentos y cálculo
6.2	Curvas de rarefacción: estandarización de profundidad
6.3	Comparación de diversidad alfa entre grupos
6.4	Visualización e interpretación de resultados

### Módulo 7 — Diversidad beta (4 horas)

Unidad	Contenido
7.1	UniFrac ponderado y no ponderado: cálculo e interpretación
7.2	PCoA y ordenamiento de comunidades
7.3	PERMANOVA: prueba de diferencias entre grupos
7.4	Dispersión beta y homogeneidad de varianzas

### Módulo 8 — Visualización de perfiles taxonómicos (3 horas)

Unidad	Contenido
8.1	Barras apiladas de abundancia taxonómica
8.2	Heatmaps de perfiles de abundancia
8.3	Gráficos de abundancia diferencial
8.4	Phyloseq: integración y visualización en R

### Módulo 9 — Análisis de biomarcadores (4 horas)

Unidad	Contenido
9.1	LEfSe: identificación de biomarcadores por LDA
9.2	ANCOM-BC: análisis de composición diferencial
9.3	DESeq2 para datos 16S: consideraciones y limitaciones



Unidad	Contenido
9.4	Random forest y curvas ROC para clasificación de muestras

### Módulo 10 — Co-ocurrencia y redes microbianas (3 horas)

Unidad	Contenido
10.1	SparCC: correlación para datos composicionales
10.2	SPIEC-EASI: inferencia de redes microbianas
10.3	igraph: construcción y análisis de redes
10.4	Visualización de redes de co-ocurrencia

### Módulo 11 — PICRUSt2: predicción funcional (4 horas)

Unidad	Contenido
11.1	PICRUSt2: inferencia de perfiles funcionales desde 16S
11.2	Predicción de vías metabólicas KEGG
11.3	Asignación EC y enriquecimiento de vías
11.4	Visualización e interpretación de resultados funcionales

### Módulo 12 — Publicación y reproducibilidad (3 horas)

Unidad	Contenido
12.1	Notebooks reproducibles con Jupyter/RMarkdown
12.2	GitHub para versionado y colaboración
12.3	Reportes finales y repositorios de datos
12.4	Estándares de publicación en revistas de microbioma

## CONTROL DE VERSIONES DEL CURRÍCULO

Versión	Fecha	Cambios realizados	Autor
1.0	Enero 2026	Versión inicial	Dirección General

## APROBACIÓN

<b>Diseño:</b>  Nombre: _____ Cargo: Director Departamento de Bioinformática Fecha: Enero 2026	<b>Revisó:</b>  Nombre: _____ Cargo: Dirección General Fecha: Enero 2026	<b>Aprobó:</b>  Nombre: _____ Cargo: Dirección General Fecha: Enero 2026
--	--	--



---

Documento controlado. FOR-EDU-005 v1.0 – TC-BIO-02 – Referenciado en PR-EDU-001, conforme a NTC-ISO 21001:2025 Cláusulas 8.1 y 8.5



## CONTROL DE CAMBIOS

<b>Versión</b>	<b>Fecha</b>	<b>Descripción del Cambio</b>	<b>Autor</b>
1.0	Enero 2026	Emisión inicial del formato	Dirección General