



UNIVERSIDAD AUTÓNOMA DE SINALOA

UNIDAD ACADÉMICA

FACULTAD DE BIOLOGÍA

MAESTRÍA EN CIENCIAS BIOLÓGICAS

“Evaluación de la variabilidad del gen ARN Ribosomal 16S en *Enterococcus* spp. mediante estimaciones filogenéticas y entropía de Shannon”

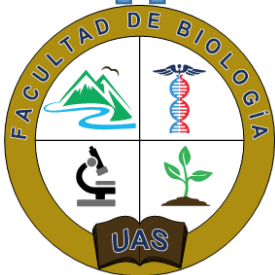
**Tesis que presenta:
Biom. Laura de Jesús Delgado Díaz**

Para obtener el grado de Maestra en Ciencias Biológicas

**Directores de tesis
Dr. Samuel Campista León
Dra. Luz Isela Peinado Guevara**

**Asesor externo
Dr. Gaspar Pérez Martínez**

**Asesores internos
Dr. Víctor Manuel Salomón Soto
M. en C. Yamel Guadalupe Rubio Rocha**



Culiacán, Sin. a diciembre de 2022.



Dirección General de Bibliotecas
Ciudad Universitaria
Av. de las Américas y Blvd. Universitarios
C. P. 80010 Culiacán, Sinaloa, México.
Tel. (667) 713 78 32 y 712 50 57
dgbuas@uas.edu.mx

UAS-Dirección General de Bibliotecas

Repositorio Institucional Buelna

Restricciones de uso

Todo el material contenido en la presente tesis está protegido por la Ley Federal de Derechos de Autor (LFDA) de los Estados Unidos Mexicanos (México).

Queda prohibido la reproducción parcial o total de esta tesis. El uso de imágenes, tablas, gráficas, texto y demás material que sea objeto de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente correctamente mencionando al o los autores del presente estudio empírico. Cualquier uso distinto, como el lucro, reproducción, edición o modificación sin autorización expresa de quienes gozan de la propiedad intelectual, será perseguido y sancionado por el Instituto Nacional de Derechos de Autor.

Esta obra está bajo una Licencia Creative Commons Atribución-No Comercial
Compartir Igual, 4.0 Internacional



DEDICATORIA

A mi familia

Que han estado conmigo en cada decisión que he tomado en la vida y en mi formación profesional. Brindándome su apoyo y confianza incondicionales. Gracias a esto, la culminación de mi proyecto significa la conclusión y obtención del grado de Maestría en Ciencias y el comienzo de un nuevo capítulo en mi carrera.

A mi esposo

Por sus palabras de aliento, su comprensión y confianza, por su amor y total soporte en mi realización profesional y por creer en mí, incondicionalmente.

Para Violeta

Espero que cuando leas estas palabras, sea un ejemplo para ti, que todo lo que desees lo puedes lograr con esfuerzo y dedicación, además recuerda que siempre contarás con mi apoyo y mi amor.

A mis maestros

Por su apoyo brindado en todos los aspectos relacionados a esta investigación, por su conocimiento, paciencia y su confianza que fueron esenciales para mi formación profesional.

A todos ustedes dedico este proyecto, por ser parte importante en mi vida y hacer que todo esto fuera posible.

AGRADECIMIENTOS

Para poder realizar esta investigación y concluirla de manera satisfactoria varias personas estuvieron involucradas durante todo el proceso de este trabajo a los que deseo agradecer. En primer lugar, quisiera agradecerles a mis directores de tesis, al Dr. Samuel Campista León y a la Dra. Luz Isela Peinado Guevara, en principio por permitirme formar parte de un gran grupo de trabajo y de este proyecto, por corregirme y enseñarme cómo abordar los problemas a lo largo de la investigación, gracias por todo su apoyo tanto formativo como económico. A mi asesor, Dr. Gaspar Pérez Martínez por sus enseñanzas y consejos. De igual manera a mis revisores la M. en C. Yamel Guadalupe Rubio Rocha y el Dr. Víctor Manuel Salomón por sus aportaciones para la mejora y el desarrollo de este proyecto.

Al Biom. Joel Tohevaris García Guerrero por su colaboración en la elaboración de este proyecto.

También quisiera agradecer a la UAS por su apoyo, por permitirme trabajar en sus instalaciones en el laboratorio de Microbiología y Biología Aplicada en la Facultad de Biología.

Finalmente quiero agradecer a mi familia y amigos por apoyarme en todo momento, por ayudarme en mi formación; a mis compañeros de laboratorio, gracias por sus consejos y su amistad.

ÍNDICE GENERAL

RESUMEN	xv
ABSTRACT	xvii
I. INTRODUCCIÓN	1
II. ANTECEDENTES	3
2.1. <i>Enterococcus</i>	3
2.1.1. Filogenia	3
2.1.2. Estructura filogenética dentro del género <i>Enterococcus</i> spp.....	5
2.1.3. Colonizadores del tracto gastrointestinal.....	8
2.1.4. Factores de virulencia.....	8
2.1.5. Resistencia a los antibióticos.....	9
2.1.5.1. Resistencia en <i>Enterococcus</i>	12
2.1.6. Importancia clínica	13
2.2. Identificación de <i>Enterococcus</i> spp.....	15
2.2.1. Métodos fenotípicos	15
2.2.2. Métodos genotípicos.....	16
2.2.2.1. Genes de interés en la identificación de <i>Enterococcus</i>	17
2.2.2.2. Gen ARNr 16S	20
2.2.2.3. Regiones hipervariables.....	22
2.3. Análisis filogenético.....	23
2.3.1. Métodos de construcción de arboles.....	24
2.3.1.1. Métodos basados en alineamientos	24
2.3.1.2. Métodos basados en distancias	25
Unión de vecinos	26
2.3.1.3. Métodos discretos o Estado de carácter.....	27
Máxima parsimonia	27
Máxima verosimilitud	28
Inferencia bayesiana	28
III. PLANTEAMIENTO DE PROBLEMA	30
IV. JUSTIFICACIÓN.....	32
VI. HIPÓTESIS	34
VI. OBJETIVOS.....	35
8.1. Objetivo general	35
8.2. Objetivos específicos.....	35

VII. MATERIALES Y MÉTODOS	36
7.1. LUGAR DE OBTENCIÓN DE LAS MUESTRAS.....	36
7.2. TIPO DE ESTUDIO.....	36
7.3. MUESTRA	36
7.3.1. Tamaño de muestra.....	36
7.3.2. Tipo de muestreo	36
7.3.3. Criterios de Inclusión	36
7.3.4. Criterios de Eliminación.....	37
7.4. METODOLOGÍA.....	37
7.4.1. Definición de las posiciones del gen ARNr 16S	37
7.4.2. Colecta de secuencias del gen ARNr 16S	38
7.4.3. Alineamiento de múltiples secuencias y mapeo posicional.....	39
7.4.4. Análisis filogenético basado en el gen ARNr 16S	39
7.4.4.1. Definición del perfil de alineamiento	40
7.4.4.2. Búsqueda de modelos de sustitución nucleotídica	40
7.4.4.3. Estimación filogenética mediante Unión de Vecinos.....	41
7.4.4.4. Estimación filogenética mediante Máxima Verosimilitud	41
7.4.4.5. Estimación filogenética mediante inferencia Bayesiana	41
7.4.4.6. Consenso de topologías	42
7.4.5. Estimación de entropía posicional del gen ARNr 16S	42
7.4.6. Análisis estadístico	43
7.4.6.1. Definición de parámetros estadísticos	43
7.4.6.2. Correlación de marcadores	43
7.4.6.3. Agrupamiento Jerárquico	44
7.4.6.4. Evaluación de marcadores	44
7.4.6.5. Oligotipificación del conjunto oligotipado.....	45
7.5. LUGAR DE REALIZACIÓN	46
7.6. FINANCIAMIENTO.....	46
VIII. RESULTADOS	47
8.1. Descripción de los conjuntos de datos.....	47
8.2. Estimación filogenética del género <i>Enterococcus</i>	47
8.2.1. Árbol filogenético.....	48
8.3 Entropía posicional del gen ARNr 16S	52
8.4. Clasificación de sitios mediante K-medias con base en la Entropía	57

8.5. Análisis de Componentes principales.....	61
8.6. Correlación y agrupamiento jerárquico de los marcadores conservados y variables .	66
8.7. Evaluación de marcadores mediante resolución filogenética.....	73
8.8 Oligotipificación.....	79
8.8.1. Estimación de la Entropía de Shannon del conjunto oligotipado.....	79
8.8.1.2. Oligotipos	82
8.8.2. Comparación de los resultados del valor de expectación y el porcentaje de identidad del Oligotipado con respecto al gen ARNr 16S en BLAST.....	84
IX. DISCUSIÓN.....	88
9.1. Analisis filogenético del genero <i>Enterococcus</i>	88
9.2. Entropía posicional del gen ARNr 16S de <i>Enterococcus</i>	94
9.3. Correlación y agrupamiento jerárquico de los marcadores conservados y variables .	94
9.4. Estimación filogenética de los marcadores similares y disimilares	95
9.5. Oligotipado y regiones discriminatorias.....	96
X. CONCLUSIONES.....	98
XI. PERSPECTIVAS	100
XII. REFERENCIAS BIBLIOGRÁFICAS.....	101
XIII. ANEXOS	119
13.1. Características del alineamiento de secuencias del conjunto taxonómico	119
13.2. Resultados de la construcción filogenética.....	119
13.3. Lista de modelos ordenados por puntajes CIB	120
13.4. Clasificación mediante K-medias conjunto SILVA	123
13.4. Clasificación mediante K-medias conjunto SILVA	123
13.5. Clasificación mediante K-medias conjunto Ribosomal Database Project.....	135
13.5. Clasificación mediante K-medias conjunto Ribosomal Database Project.....	135
13.6. Clasificación mediante K-medias conjunto Greengenes	147
13.6. Clasificación mediante K-medias conjunto Greengenes	147
13.7. Clasificación mediante K-medias conjunto Fusionado	159
13.7. Clasificación mediante K-medias conjunto Fusionado	159
13.8. Lista de modelos ordenados por puntajes CIB árbol ARNr 16S.....	171
13.9. Lista de modelos ordenados por puntajes CIB árbol de marcadores Similares.....	172
13.10. Lista de modelos ordenados por puntajes CIB árbol de marcadores Disimilares	174

Índice de Cuadros

Cuadro		Pág.
1	Lista de la OMS de patógenos prioritarios para la investigación y desarrollo de nuevos antibióticos.	11
2	Clasificación de los métodos de análisis filogenéticos y sus estrategias	26
3	Marcadores conservados a usar para la determinación de las regiones variables en el gen ARN ribosomal 16S	38
4	Descripción de los conjuntos de secuencias generadas en este estudio.	47
5	Oligotipado de especies y grupo de especies	83
6	Valor de expectación y porcentaje de identidad del Oligotipado y del ARNr 16S en BLAST	86
7	Valor de expectación y porcentaje de identidad del Oligotipado por grupos o parejas y del ARNr 16S en BLAST	87

Índice de Figuras

Figura		Pág.
1	Árbol de unión de vecinos que demuestra las relaciones filogenéticas dentro del género <i>Enterococcus</i> .	7
2	Regiones hipervariables del gen ARNr 16S.	20
3	Jerarquía de los métodos basados en alineamientos	25
4	Consenso filogenético de Inferencia Bayesiana, Máxima Verosimilitud y Unión de Vecinos de 1,252 posiciones compartidas del ARNr 16S del género <i>Enterococcus</i> .	50
5	Entropía del Conjunto SILVA	53
6	Entropía del Conjunto RDP	54
7	Entropía del Conjunto Greengenes	55
8	Entropía de los 3 conjuntos fusionados, SILVA, RDP y Greengenes	56
9	Clasificación de Sitios mediante K-medias con base en la Entropía conjunto SILVA	57
10	Clasificación de Sitios mediante K-medias con base en la Entropía conjunto RDP	58
11	Clasificación de Sitios mediante K-medias con base en la Entropía conjunto Greengenes	59
12	Clasificación de Sitios mediante K-medias con base en la Entropía conjunto Fusionado	60
13	Análisis de componentes principales del conjunto SILVA	61
14	Análisis de componentes principales del conjunto RDP	62
15	Análisis de componentes principales del conjunto Greengenes	63
16	Análisis de componentes principales del conjunto Fusionado	65
17	Análisis de Regresión de los dos componentes principales del conjunto Fusionado	65
18	Agrupamiento jerárquico mediante UPGMA y mapa de calor del conjunto SILVA	67
19	Agrupamiento jerárquico mediante UPGMA y mapa de calor del conjunto RDP	69
20	Agrupamiento jerárquico mediante UPGMA y mapa de calor del conjunto Greengenes	71
21	Agrupamiento jerárquico mediante UPGMA y mapa de calor del conjunto Fusionado	72
22	Árbol filogenético por ML ARNr 16S completo	74
23	Árbol filogenético por ML marcadores similares	76
24	Árbol filogenético por ML marcadores disimilares	78
25	Entropía de Shannon de conjunto oligotipado (panel a)	80
	Entropía de Shannon de conjunto oligotipado (panel b)	

Abreviaturas

Ace:	Del inglés, “ <i>collagen adhesin Enterococcus</i> ”, proteína de adhesión al colágeno presente en <i>Enterococcus</i>
Acm:	Del inglés, “ <i>collagen adhesin E. faecium</i> ” proteína de adhesión al colágeno presente en <i>E. faecium</i>
ADN:	Ácido desoxirribonucleico
AIC:	Del inglés, “ <i>Akaike information criterion</i> ”, Criterio de información Akaike
AICc:	Del inglés, “ <i>Corrected Akaike information criterion</i> ”, Criterio de Información Akaike Corregido
AtpA:	ATP sintasa
ARNr:	Ácido ribonucleico ribosomal
ATCC:	Del inglés, “ <i>American Type Culture Collection</i> ”
CIB:	Criterio de información Bayesiano
BLAST:	Del inglés, “ <i>Basic Local Alignment Search Tool</i> ”
CAUTI:	Del inglés, “ <i>Catheter-Associated Urinary Tract Infections</i> ”, Infecciones del tracto urinario asociadas con el catéter
CLABSI:	Del inglés, “ <i>Central Line-Associated BloodStream Infections</i> ”, Infecciones del torrente sanguíneo
ECDC	Del inglés, “ <i>European Centre for Disease Prevention and Control</i> ”, Centro de control y prevención de enfermedades Europeo
GTR:	Del inglés, “ <i>General Time Reverse</i> ”, <i>Tiempo general reversible</i>
IB:	Inferencia Bayesiana
iToL	Del inglés, “ <i>Interactive Tree of Life</i> ”, Árbol de la vida interactivo
ITU:	Infección del tracto urinario
LPSN:	Del inglés, “ <i>List of Prokaryotic names with Standing in Nomenclature</i> ”, listado de nombres procariotas con posición en nomenclatura
MALDI-TOF:	Del inglés, “ <i>Matrix-Assisted Laser Desorption/Ionization-Time of Flight</i> ”, espectrometría de masas -desorción/ionización por láser tiempo de vuelo
MC	Marcador Conservado del gen ARNr 16S
MCMC:	Del inglés, “ <i>Markov Chain Monte Carlo</i> ”, Cadena de Markov Monte Carlo
Mb:	Megabases
ML:	Del inglés, “ <i>Maximum Likelihood</i> ”, Máxima Verosimilitud
MP:	Del inglés, “ <i>Maximum Parsimony</i> ”, Máxima Parsimonia
MSMCRAMM:	Del inglés, “ <i>Microbial Surface Components Recognizing Adhesive Matrix Molecules</i> ”, Componentes de la superficie microbiana que reconocen las moléculas de la matriz adhesiva
NAAT:	Del inglés, “ <i>Nucleic Acid Amplification Tests</i> ”, pruebas de amplificación de ácido nucleicos
NCBI:	Del inglés, “ <i>National Center for Biotechnology Information</i> ”, Centro Nacional para la Información Biotecnológica
OMS:	Organización Mundial de la Salud
OTU’s:	Del inglés, “ <i>Operational Taxonomic Unit</i> ”, Unidades operacionales taxonómicas
pb:	Pares de bases
PCR:	Del inglés, “ <i>Polymerase Chain Reaction</i> ”, reacción en cadena de la polimerasa
PGC:	Genes de pilina
PheS:	Subunidad alfa de la fenilalanil-tRNA sintasa
PLS:	Del inglés, “ <i>Partial least squares regression</i> ”, Regresión de Mínimos Cuadrados Parciales
PNA-FISH:	Del inglés, “ <i>Fluorescence in situ hybridization using peptide nucleic acid probes</i> ”, hibridación <i>in situ</i> fluorescente de ácido nucleico peptídico

PYR:	Prueba de pirrolidonil arilamidasa/pirrolidonil aminopeptidasa
RDP:	Del inglés, “ <i>Ribosomal Database Project</i> ”
rpoA:	Subunidad alfa de la ARN polimerasa
sod	superóxido dismutasa
tuf	factor de elongación <i>Tu</i>
UPGMA:	Del inglés, “ <i>Unweighted pair group method with arithmetic mean</i> ”, Método de grupos de pares no ponderados con media aritmética
V	Región hipervariable del gen ARNr 16S

RESUMEN

Introducción: Las especies del género *Enterococcus* han evolucionado como miembros altamente adaptados de los consorcios gastrointestinales de una amplia variedad de hospederos: mamíferos, aves, reptiles e insectos, sin embargo, por razones que no están del todo claras, en la década de 1970 fueron causa de infecciones hospitalarias multirresistentes. *E. faecalis* y *E. faecium* son clínicamente prevalentes e importantes patógenos nosocomiales. Son especies vigorosas capaces de sobrevivir en nichos biológicos importantes, como el tracto gastrointestinal humano, y en condiciones ambientales rigurosas, facilitando su propagación entre el ambiente hospitalario y comunitario. El gen que codifica para ARNr 16S es utilizado en estudios filogenéticos, debido a su bajas tasa de mutación. Sin embargo, la diferenciación entre especies del género *Enterococcus* es difícil, al utilizar este gen. **Objetivo:** Evaluar la variabilidad del gen ARNr 16S en *Enterococcus* spp. mediante estimaciones filogenéticas y entropía de Shannon. **Metodología:** Los conjuntos de datos *in-silico* de secuencias fueron tomadas de bases de datos públicas, y se analizaron las subregiones ARNr 16S comúnmente objetivo, mediante estimaciones filogenéticas y el cálculo de la entropía de Shannon, se comparó la precisión taxonómica lograda del gen ARNr 16S completo de cepas de *Enterococcus* spp. con la utilización de marcadores denominados como similares al ARNr 16S. Finalmente se utilizó el método de aprendizaje de máquinas supervisado oligotipificación de las regiones V2-V8. **Resultados:** Se colectaron 10,551 secuencias de las bases SILVA, RDP y Greengenes. Se obtuvo un árbol filogenético con una topología conformada por 4 supergrupos y dos grupos menores: supergrupo *E. faecium*, supergrupo *E. faecalis*, supergrupo *E. casseliflavus* y supergrupo *E. dispar*, así como los grupos *E. cecorum* y *E. aquimarinus*. Las regiones hipervariables V2, V3 y V8, V9 resultaron ser similares al ARNr 16S, la comparación entre los

arboles filogenéticos de los respectivos marcadores similares contra el gen completo, ofreció una resolución similar entre ambas estimaciones. La oligotipificación, determino la ubicación de 13 posiciones dentro de la región V2, que poseen el poder resolutivo para 55% de las especies del género *Enterococcus*. Además de generar 43 oligotipos, que lograron identificar a 31 *Enterococcus* spp. en la base de datos GenBank. **Conclusiones:** La implementación de la oligotipificación logro la obtención de un fragmento de 105 nucleótidos como posible marcador para la identificación de *Enterococcus* spp.

Palabras clave: ARNr 16S, Estimaciones filogenéticas, entropía de Shannon, oligotipificación.

ABSTRACT

Background: Species of the genus *Enterococcus* had evolved as highly adapted members of the gastrointestinal consortia of a wide variety of hosts: mammals, birds, reptiles, and insects; However, for reasons that are not entirely clear, in the 1970's they were the cause of Multidrug-resistant hospital infections. *E. faecalis* and *E. faecium* are clinically prevalent and nosocomial-significant pathogens. Species are vigorous and capable of biologically surviving in important niches, such as the human gastrointestinal tract, and harsh environmental conditions, facilitating their spread between hospital and community environments. The 16S rRNA has been used in phylogenetic analyses due to its low mutation rate. However, differentiation between species of the genus *Enterococcus* is difficult when using this gene. **Objective:** To evaluate the variability of the 16S rRNA gene in *Enterococcus* spp. using phylogenetic estimates and Shannon entropy. **Methods:** *In-silico* sequence datasets were taken from public databases, and 16S rRNA subregions were analyzed using phylogenetic estimates and Shannon entropy, comparing the achieved taxonomic precision of 16S rRNA complete gene in *Enterococcus* spp. against the 16S rRNA-like markers. Finally, has done a supervised machine learning method; oligotyping analyses of the V2-V8 regions. **Results:** 10,551 sequences from SILVA, RDP, and Greengenes databases were collected. A phylogenetic tree was estimated and showed four supergroups and two smaller groups topology: *E. faecium* supergroup, *E. faecalis* supergroup, *E. casseliflavus* supergroup, and *E. dispar* supergroup, as well as the *E. cecorum* and *E. aquimarinus* groups. The hypervariable regions V2, V3, V8, and V9, turned out to be similar to the 16S rRNA, and a comparison between both phylogenetic trees, the similar marker ones against the complete gene tree proposed a similar resolution in both estimates. Oligotyping determined the location of 13 positions within the V2 region, which has the resolving power for 55% of the species of

the genus *Enterococcus*. In addition, it generated 43 oligotypes and identity managed to 31 *Enterococcus* spp. in the GenBank database. **Conclusions:** The implementation of oligotyping managed to obtain a fragment of 105 nucleotides as a possible molecular identification marker for *Enterococcus* spp.

Keywords: 16S rRNA, phylogenetic estimates, Shannon entropy and oligotyping

I. INTRODUCCIÓN

Enterococcus se considera como un género bacteriano antiguo que co-evolucionó en una amplia gama de huéspedes, lo que sugiere que eran miembros centrales del microbioma intestinal del último ancestro común de mamíferos, reptiles, aves e insectos, que existió en el período Devónico temprano, hace más de 400 millones de años (Gilmore *et al.*, 2013). Están altamente adaptados a vivir en ambientes complejos y a sobrevivir en condiciones adversas. Su morfología se describe como cocos Gram positivos, catalasa negativos, que se agrupan en parejas o en cadenas cortas. Los *Enterococcus* son auxótrofos para muchos aminoácidos, vitaminas y micronutrientes, los genomas de las cepas comensales tienen solo 2.75 Mb. También pueden sobrevivir y crecer en una amplia gama de temperaturas y pH así mismo en condiciones tanto hipertónicas como hipotónicas; sus hábitats incluyen huéspedes animales, así como plantas, suelo, agua, productos fabricados por el hombre, incluyendo alimentos fermentados y productos lácteos. En el hábitat acuático, la presencia de algunas especies de *Enterococcus* en alimentos de origen marino es indicativo de contaminación fecal (Larrea-Murrell *et al.*, 2013). Han surgido cepas de *Enterococcus* resistentes a los antibióticos, y con una regularidad alarmante se aíslan cepas con nuevos mecanismos de resistencia. Por ello, la relación entre el uso no terapéutico de antibióticos y la aparición de *Enterococcus* en diversos hábitats no humanos reviste un gran interés.

La taxonomía de los *Enterococcus* ha cambiado considerablemente en los últimos diez años, el género incluye ahora más de cincuenta especies distintas con diversos hábitats, tropismos, características metabólicas y fenotípicas con base en la taxonomía bacteriana *LPSN - List of Prokaryotic names with Standing in Nomenclature* (Parte *et al.*, 2020; Disponible en lpsn.dsmz.de/genus/enterococcus).

La importancia del género *Enterococcus* se debe a aquellas especies que continúan siendo patógenos nosocomiales, por ejemplo *E. faecalis* y *E. faecium* que son clínicamente prevalentes causantes de graves infecciones en el humano que van desde endocarditis hasta bacteriemias (Gilmore *et al.*, 2013; O'Driscoll y Crank, 2015). Debido a la propagación de estos patógenos oportunistas que albergan genes de resistencia a vancomicina de alto nivel, más allá de los hospitales, hacia la comunidad en general y a los hábitats asociados al ser humano es una amenaza potencial para la salud pública (Gilmore *et al.*, 2014).

En labores de identificación bacteriana, el gen ARNr 16S ha sido un pilar en el análisis basado en secuencias durante décadas (Johnson *et al.*, 2019). A pesar de la amplia utilización del gen, dentro de este género, existen conflictos para establecer la filogenia en especies estrechamente relacionadas y como es bien sabido que las especies bacterianas con genes ARNr 16S idénticos pueden representar distintos ecotipos con diferencias en las propiedades de virulencia y otros rasgos fenotípicos (Jaspers y Overmann, 2004). En este sentido, la sensibilidad del gen ARNr 16S es limitada, pero específica; se ha demostrado que incluso una diferencia de nucleótidos a nivel del gen ARNr 16S puede ser un indicador de una cepa ecológicamente distinta (Ward, 1998). Además, existe una correlación entre la divergencia del gen ARNr 16S y el contenido general de genes (Konstantinidis y Tiedje, 2007). Por lo que en el presente trabajo se propuso la evaluación de la variabilidad del gen ARNr 16S en *Enterococcus* spp. mediante estimaciones filogenéticas y entropía de Shannon como indicador para determinar la presencia de oligotipos dentro del ARNr 16S, en el cual se usaron conjuntos de datos *in-silico* de secuencias tomadas de bases de datos públicas.

II. ANTECEDENTES

2.1. *Enterococcus*

El término "enterococo" tiene su origen a finales del siglo XIX, cuando Thiercelin describió un coco saprófito, de origen intestinal, capaz de causar infecciones (Thiercelin y Jouhaud, 1899; Thiercelin, 1899). Ese mismo año, MacCallum y Hastings caracterizaron un organismo similar, que ahora se conoce como *Enterococcus faecalis*, a partir de un caso letal de endocarditis, proporcionando así una primera descripción detallada de sus capacidades patógenas (MacCallum y Hastings, 1899). Estas primeras descripciones constituyen la base del paradigma de la patogénesis enterocócica como un oportunista comensal. La aparición de *Enterococcus* multirresistentes en los últimos 50 años ha agravado considerablemente este problema, y los *Enterococcus* resistentes a los antibióticos son ahora las principales causas de infección nosocomial en todo el mundo (Prabaker y Weinstein, 2011; Werner, *et al.*, 2008). Igualmente se aíslan de la superficie de plantas, así como de suelo, arena y agua (Gilmore, 2003; Mundt, 1963). En el agua los *Enterococcus* se consideran como indicadores de contaminación fecal y de calidad de la misma (Byappanahalli *et al.*, 2012).

2.1.1. Filogenia

Los *Enterococcus* se encuentran enmarcados en el filo Firmicutes, clase Bacilli, orden Lactobacillales familia Enterococcaceae (Ludwig, Schleifer y Whitman, 2009), junto con los géneros *Bavariicoccus*, *Catelicoccus*, *Melissococcus*, *Pilibacter*, *Tetragenococcus* y *Vagococcus*.

En 1984 se estableció como tal el género *Enterococcus*, en ese mismo año Schleifer y Kilpper-Balz utilizaron la hibridación ADN-ADN y la secuenciación del ARNr 16S, para

analizar a los *Streptococcus* del grupo D de Lancefield y descubrieron que *Streptococcus faecium* y *Streptococcus faecalis* eran distintos de otros *Streptococcus* y se transfirió a estas dos especies al género *Enterococcus* (Schleifer y Kilpper-Balz, 1984). Posteriormente, muchas especies fueron reclasificadas como *Enterococcus*, entre las que se incluyen *E. avium*, *E. casseliflavus*, *E. durans*, *E. gallinarum* y *E. malodoratus* (Collins *et al.*, 1984). Filogenéticamente, el género *Enterococcus* pertenece a la rama de bajo GC de las bacterias Gram-positivas (es decir, 37.5% de GC en el cromosoma de *E. faecalis*) (Paulsen *et al.*, 2003). En su morfofisiología se describen como cocos que se agrupan en parejas o en cadenas cortas, catalasa negativos, anaerobios facultativos quimioorganotróficos con metabolismo homofermentativo y productores de ácido láctico como producto final predominante de la fermentación de carbohidratos (Švec y Franz, 2014). Pertenecen al grupo de las bacterias ácido lácticas y producen bacteriocinas (Fisher y Phillips, 2009).

El género *Enterococcus* consta de especies que se encuentran en el tracto gastrointestinal (GI) humano y animal (Mundt, 1963; 1986), así como en el tubo digestivo de los insectos (Martin y Mundt, 1972), en los alimentos fermentados tradicionales y en los productos lácteos (Mundt, 1986), y en diversos entornos, como las plantas (Mundt, 1963), el suelo (Mundt, 1961) y el agua (Ator y Starzyk, 1976). A medida que se encuentran y describen más y más especies nuevas, se ha hecho evidente que son de naturaleza ubicua (Lebreton *et al.*, 2014) y pueden aislarse de las plantas, el suelo y el agua (Sistek *et al.*, 2012). Sin duda, una clave importante de su amplia distribución en la naturaleza son microorganismos robustos capaces de crecer y mantenerse viables en condiciones extremas: su rango de temperatura de crecimiento es de 10 a 45 °C, crecen a concentraciones de cloruro de sodio de hasta el 6.5% y a un pH de 9.6, además de sobrevivir a un tratamiento térmico de 60 °C de 30 minutos de duración y resistir la presencia de sales biliares. En resumen, las

características fisiológicas de los *Enterococcus* permiten que estén presentes en una diversidad de hábitats, tanto ambientes bióticos y abióticos, teniendo alta persistencia en los mismos.

2.1.2. Estructura filogenética dentro del género *Enterococcus* spp.

Los análisis filogenéticos basados en las secuencias del gen ARNr 16S separa a ciertas especies en distintos linajes, pero la mayoría de los *Enterococcus* spp. están asignados en varios grupos filogenéticos (Figura 1). Similitudes de la secuencia del gen del ARNr 16S entre especies en estos grupos puede ser tan alto como 99.8% como se revela entre *E. casseliflavus* y *E. gallinarum* (Williams *et al.*, 1991) o incluso el 99.9% entre *E. devriesei* y *E. pseudoavium* (Švec *et al.*, 2005). Asignación de *Enterococcus* spp. en grupos filogenéticos revelados por la secuenciación del gen ARNr 16S se corresponde con *atpA*, *rpoA* (subunidad alfa de la ARN polimerasa) y, hasta cierto punto, con análisis de genes *pheS* (que codifica la subunidad alfa de la fenilalanil-tRNA sintasa) y ARNr 23S (Tsiodras *et al.*, 2000; Naser *et al.*, 2005a; 2005b), y las especies albergadas en estos grupos comparten ciertas características fenotípicas (Devriese *et al.*, 1993). El grupo *E. faecalis* contiene a *E. caccae*, *E. faecalis*, *E. haemoperoxidus*, *E. moraviensis*, *E. plantarum*, *E. quebecensis*, *E. rivorum*, *E. rotai*, *E. silesiacus*, *E. termitis*, *E. ureasiticus* y *E. ureilyticus* es el grupo más grande.

El grupo *E. faecium* engloba a especies como *E. canis*, *E. durans*, *E. faecium*, *E. hirae*, *E. lactis*, *E. mundtii*, *E. ratti*, *E. thailandicus* y *E. villorum*. El grupo de *E. avium* contiene a *E. avium*, *E. devriesei*, *E. gilvus*, *E. malodoratus*, *E. pseudoavium*, *E. raffinosus*, *E. viikkiensis*. Se reconocen otros cuatro grupos formados por pares de especies: el grupo *E. gallinarum* alberga a *E. gallinarum* y *E. casseliflavus*; el grupo *E. italicus* contiene a *E. italicus* y *E. camelliae*; el grupo *E. cecorum* agrupa las especies *E. cecorum* y *E. columbae*;

y el grupo *E. dispar* integrado por *E. dispar* con la reciente descripción de *E. canintestini*. Las especies del grupo *E. cecorum* revelan ciertos rasgos fenotípicos únicos en el género, están relacionados más lejanamente con los restantes *Enterococcus* spp. En la actualidad el género *Enterococcus* tiene hasta la fecha 58 especies descritas con publicaciones válidas, la mayoría en la era genómica (con base en la información recopilada de *LPSN - List of Prokaryotic names with Standing in Nomenclature* (Parte *et al.*, 2020; Disponible en lpsn.dsmz.de/genus/enterococcus).

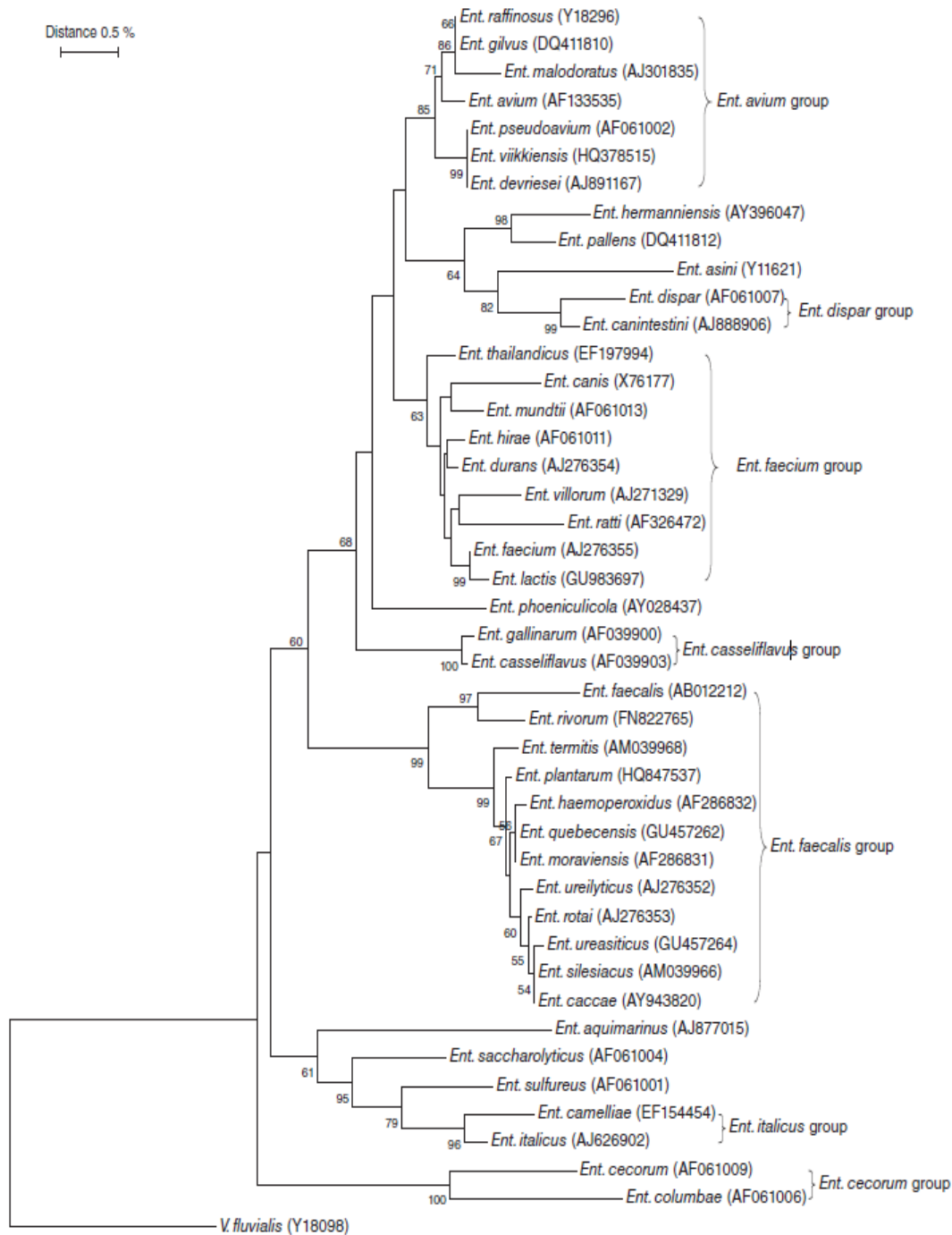


Figura 1. Relaciones filogenéticas dentro del género *Enterococcus* con base en el método de unión de vecinos. Porcentaje de arranque los valores (1000 réplicas de árboles) superiores al 50% se indican en los nodos de las ramas. La secuencia de *Vagococcus fluvialis* (Y18098) se utilizó como grupo externo. Los números de acceso de GenBank se dan entre paréntesis. La barra de distancia indica un 0.5% de divergencia de secuencia. El árbol se construyó utilizando el software MEGA5 (Tamura *et al.*, 2011) (Tomado de Bacterias del ácido láctico: biodiversidad y taxonomía Holzapfel y Wood, 2014).

2.1.3. Colonizadores del tracto gastrointestinal

Los *Enterococcus* son omnipresentes en el tracto gastrointestinal, aunque constituyen una pequeña proporción del intestino, menos del 1% de la microbiota de individuos en edad adulta (Finegold *et al.*, 1983; Sghir *et al.*, 2000). Se pueden aislar tanto del sistema digestivo de invertebrados como del tracto gastrointestinal de vertebrados, lo que implica que pertenecen al núcleo del microbioma intestinal desde el último antecesor común de mamíferos, reptiles, aves e insectos, que existió hace unos 400 millones de años, en el periodo devónico (Gilmore *et al.*, 2013).

2.1.4. Factores de virulencia

Los componentes de la superficie microbiana que reconocen las moléculas de la matriz adhesiva (MSCRAMM por sus siglas en inglés, *Microbial Surface Components Recognizing Adhesive Matrix Molecules*) son elementos de la superficie que ayudan a los *Enterococcus* a adherirse a los tejidos del hospedero, lo que ayuda a iniciar la infección (Patti y Hook, 1994).

En *E. faecalis*, los genes MSCRAMM se encuentran en la mayoría de las cepas y se expresan *in vivo* durante la infección humana (Sillanpaa *et al.*, 2004). Uno de los MSCRAMM mejor caracterizados es “Ace”, una proteína de unión al colágeno de *Enterococcus* (Rich *et al.*, 1999) que mejora la colonización temprana de las válvulas cardíacas, lo que sugiere un papel importante en el establecimiento temprano de la endocarditis (Singh *et al.*, 2010).

Los aislados clínicos de *E. faecium* son ricos en genes MSCRAMM, y los genes de esta familia presentes en el genoma fueron más abundantes en los aislados de endocarditis

(Sillanpaa *et al.*, 2009). En *E. faecium*, “Acm” (una proteína de unión al colágeno) es el MSCRAMM mejor caracterizado. El gen Acm está presente principalmente en los aislamientos asociados con la atención de la salud (presente en el 99% de los aislamientos analizados), aunque un estudio encontró que fue alterado por un transposón en los aislamientos comensales, volviéndose no funcional (Nallapareddy *et al.*, 2008).

Los grupos de genes de pilina (PGC) están presentes tanto en *E. faecalis* como en *E. faecium* y codifican proteínas de superficie con motivos similares a LPxTG que son responsables del ensamblaje a la pared celular bacteriana de estructuras filamentosas largas que se extienden desde la superficie, llamadas pili. Al igual que los MSCRAMM, los pili pueden funcionar como adhesinas (Telford *et al.*, 2006). En *E. faecalis*, el PGC “ebp” (pilus asociado a endocarditis y biopelículas) se asocia con la adherencia inicial y la formación de biopelículas y se ha implicado en la patogenia de la endocarditis y las infecciones urinarias (Nallapareddy *et al.*, 2006). En *E. faecium*, el papel del pili no está del todo claro; sin embargo, parece haber una regulación diferencial del PGC y el ensamblaje de las proteínas del pilus entre los aislados clínicos y las cepas comensales (Hendrickx *et al.*, 2008; Hendrickx *et al.*, 2010).

2.1.5. Resistencia a los antibióticos

Los antibióticos son los medicamentos utilizados para prevenir y tratar las infecciones bacterianas. Por lo tanto, la resistencia antibacteriana se produce cuando las bacterias mutan e inhiben la acción de estos fármacos y adquieren la capacidad para sobrevivir en presencia de dicho compuesto tóxico (antibiótico o antiséptico), permitiendo que las bacterias se multipliquen en presencia del fármaco (Troncoso *et al.*, 2017; OMS, 2018).

Existen diversos factores que derivan en la resistencia a los antibacterianos. En muchos entornos de atención médica existe una deficiencia en la identificación temprana del agente casual y sus patrones de susceptibilidad a antibacterianos, por lo que es común recurrir a antibióticos de amplio espectro y en su mayoría innecesariamente (Frieri *et al.*, 2017). Así también, el uso generalizado e inapropiado de los antibacterianos dentro de la población y su uso extensivo como potenciadores del crecimiento en alimentos para animales, que en su mayoría se administran sin supervisión de un profesional, son acciones que predisponen a la dispersión de resistencia (Mathur y Singh, 2013; OMS, 2018).

Como consecuencia, se obtiene un aumento en la resistencia emergente y, junto con ello, tratamientos más complejos y prolongados, que causan mayores costos económicos, estadías hospitalarias más largas, secuelas y una mortalidad más alta. Además, cuando se combinan con prácticas de control de infección deficientes, las bacterias resistentes tienden a infiltrarse hacia otros entornos (Alós, 2015; Frieri *et al.*, 2017). Hoy en día la resistencia a antibacterianos se convierte en una amenaza cada vez mayor para la salud pública mundial, ya que han logrado diseminarse hacia la comunidad, sea por la transmisión de un paciente infectado es dado de alta y que tuvo contacto con el exterior, o bien el desarrollo de una resistencia de *novo* por un tratamiento en pacientes de la comunidad (Mathur y Singh, 2013; OMS, 2018).

En 2017, la OMS publicó su primera lista de patógenos prioritarios resistentes a los antibióticos, en la que se incluyen las 12 especies de bacterias más peligrosas para la salud humana (Cuadro 1). Se destaca la amenaza que suponen las bacterias Gram negativas resistentes a múltiples antibióticos, consecuentes de la capacidad innata de estas bacterias de evadir diversos fármacos y de transmitir el material genético a otras bacterias. El grupo de prioridad crítica incluye las bacterias multirresistentes que se asocian principalmente a

infecciones nosocomiales. Los niveles segundo y tercero de la lista pertenecen a las categorías de prioridad alta y media, las cuales contienen otras bacterias con farmacorresistencia creciente y provocan enfermedades comunes como la gonorrea o intoxicaciones alimentarias.

Cuadro 1.- Lista OMS de patógenos prioritarios para la investigación y desarrollo de nuevos antibióticos

Estado	Patógeno	Resistencia
Crítico	<i>Acinetobacter baumannii</i>	Carbapenémicos
	<i>Pseudomonas aeruginosa</i>	Carbapenémicos
	<i>Enterobacteriaceae</i>	Carbapenémicos; BLEE
Elevada	<i>Enterococcus faecium</i>	Vancomicina
	<i>Staphylococcus aureus</i>	Meticilina; Vancomicina
	<i>Helicobacter pylori</i>	Claritromicina
	<i>Campylobacter</i> spp	Fluoroquinolonas
	<i>Salmonella</i>	Fluoroquinolonas
	<i>Neisseria gonorrhoeae</i>	Cefalosporinas; Fluoroquinolonas
Media	<i>Streptococcus pneumoniae</i>	Penicilina
	<i>Haemophilus influenzae</i>	Ampicilina
	<i>Shigella</i> spp	Fluoroquinolonas

2.1.5.1. Resistencia en *Enterococcus*

La resistencia de *Enterococcus* spp. a algunos antibacterianos es intrínseca, las bacterias del género pueden presentar baja susceptibilidad a antibióticos como penicilina y ampicilina y elevada resistencia a la mayoría de las cefalosporinas y todas las penicilinas semisintéticas, esta característica está relacionada con la expresión de proteínas de unión a la penicilina de baja afinidad (PBP) (Kristich *et al.*, 2014).

Es importante mencionar que *Enterococcus* spp. son tolerantes a la actividad bactericida de los agentes activos, como los antibióticos β -lactámicos y la vancomicina, esto implica que la bacteria puede ser inhibida por las concentraciones clínicamente alcanzables del antibiótico, pero solo será eliminada por concentraciones muy superiores a la concentración inhibitoria (Gagetti *et al.*, 2019).

La resistencia a los antibióticos β -lactámicos es un tema de gran relevancia, debido a que estos medicamentos se usan frecuentemente para el tratamiento de infecciones ocasionadas por *Enterococcus*. La diseminación de la resistencia marcada a ampicilina puede ser el resultado de la propagación clonal de cepas con genes mutados y la transferencia horizontal de genes. Los *Enterococcus* se caracterizan por tener una gran capacidad para adquirir nuevos mecanismos de resistencia (Gagetti *et al.*, 2019).

Por otra parte, entre las especies de *Enterococcus*, la transferencia de genes de resistencia a antibióticos mediada por fagos se ha probado con éxito (Devirgilis *et al.*, 2013).

2.1.6. Importancia clínica

Los *Enterococcus* se consideran organismos comensales del tracto gastrointestinal humano; sin embargo, también pueden ser patógenos, principalmente relacionados con las infecciones asociadas al hospital, que comúnmente causan infección del tracto urinario (ITU), bacteriemia, endocarditis, infecciones por quemaduras y heridas en el sitio quirúrgico, infecciones del abdomen y del tracto biliar e infección de catéteres y otros dispositivos médicos implantados. En la mayoría de las encuestas, los *Enterococcus* son la tercera causa más común de endocarditis de válvula nativa, después de *Staphylococcus aureus* y *Streptococcus* del grupo *viridans* (Murdoch *et al.*, 2009; Slipczuk *et al.*, 2013).

Parte de la importancia de *Enterococcus* radica en su implicación clínica, ya que las especies de este género causan frecuentemente infecciones en el humano que van desde endocarditis hasta bacteriemias concurrentes (Lee *et al.*, 2005; Jenkins, 2007; O'Driscoll y Crank, 2015).

Se sabe poco sobre los principales mecanismos utilizados por los *Enterococcus* para colonizar el tracto gastrointestinal de los individuos sanos o de los pacientes hospitalizados. Sin embargo, la exposición de los pacientes hospitalizados a antibióticos da lugar a importantes modificaciones del microbiota intestinal, que facilitan colonización del tracto gastrointestinal por *Enterococcus* resistentes a los antibióticos (Donskey, *et al.*, 2000; Ubeda, *et al.*, 2010).

La incidencia de infecciones por *Enterococcus* ha aumentado de manera constante desde fines de la década de 1970 (Arias y Murray, 2012; Teixeira *et al.*, 2016; Guzmán-Prieto *et al.*, 2016). Tanto en Europa como en los Estados Unidos, los *Enterococcus* a nivel de género son el segundo patógeno más común asociado con las infecciones asociadas al hospital, *E. faecalis* fue el quinto organismo aislado con mayor frecuencia en las infecciones

del tracto urinario asociadas con el catéter (CAUTI por sus siglas en inglés, Catheter-Associated Urinary Tract Infections) y el tercero en las infecciones del torrente sanguíneo (CLABSI por sus siglas en inglés, Central Line-Associated Blood Stream Infections) y para *E. faecium* se reportó como el 11^{vo} y el 5^o, en dichos tipos de infección respectivamente (European Centre for Disease Prevention and Control, 2013 ECDPC; Hidron *et al.*, 2008; Weiner *et al.*, 2016).

2.2. Identificación de *Enterococcus* spp.

2.2.1. Métodos fenotípicos

La diferenciación e identificación de *Enterococcus* spp. mediante las características bioquímicas y fisiológicas es difícil. Se han probado diferentes medios selectivos para el aislamiento e identificación de *Enterococcus*; sin embargo, no existen pruebas bioquímicas definitivas para diferenciar *Enterococcus* de otros cocos Gram positivos catalasa negativos. La mayoría de los *Enterococcus* son oxidasa y catalasa negativas, tolerantes a la sal (hasta 6.5%), resistentes al 40% de bilis, hidrolítica de esculina y capaces de crecer en presencia de azida sódica (hasta 0.4%). Además de las características descritas anteriormente, todas las especies descritas y probadas producen β -glucosidasa; leucina arilamidasa; ácido de los azúcares d -fructosa, galactosa, β -gentiobiosa, glucosa, lactosa, maltosa, d -manosa, ribosa, trehalosa, celobiosa y N acetil glucosamina; y los glucósidos salicina, metil β -D -glucósido, amigdalina y arbutina. En general, los *Enterococcus* son ureasa negativos y no producen ácido a partir de D- arabinosa, eritritol, D- y L- fucosa, metil α - D -xilosido y L -xilosa; estas características metabólicas se han utilizado en el desarrollo de kits de prueba comerciales. El crecimiento ocurre entre 10 °C y 45 °C, con un crecimiento óptimo para la mayoría de las especies entre 35 °C y 37 °C (Švec y Devriese, 2015). Los *Enterococcus* son notablemente resistentes a la desecación (Švec y Devriese, 2015). Aunque se ha tratado de implementar métodos bioquímicos convencionales o mediante kits comerciales, y aun así se presenta dificultades, ya que pueden revelarse diferentes rasgos fenotípicos entre cepas de la misma especie originarias de diferentes fuentes.

Si se cultivan en agar sangre de caballo, los *Enterococcus* pueden ser α , β o no hemolíticos y formar colonias de 1 a 2 mm con apariencia húmeda. Con base a sus capacidades metabólicas, se han desarrollado diferentes medios de cultivo selectivos para el aislamiento

de *Enterococcus*; estos medios selectivos contienen frecuentemente sales biliares, azida sódica, antibióticos y sales de esculina o tetrazolio. No todas las especies de *Enterococcus* pueden crecer en estos medios selectivos, pero las especies más relevantes clínicamente crecen bien (Švec y Devriese, 2015). La mayoría de las pruebas clínicas para la identificación de *Enterococcus* incluyen la prueba de catalasa, la prueba de pirrolidonil arilamidasa / pirrolidonil aminopeptidasa (PYR) y una prueba de hidrólisis de bilis esculina. Se han desarrollado kits comerciales para estandarizar y optimizar la detección de *Enterococcus* en el entorno clínico, y todos requieren aislamiento y cultivo previos de los organismos, lo que puede retrasar el diagnóstico (Švec y Devriese, 2015).

2.2.2. Métodos genotípicos

Entre los sistemas más nuevos para la clasificación e identificación de *Enterococcus* se encuentran espectrometría de masas -desorción/ionización por láser asistida por matriz (MALDI-TOF MS, por sus siglas en inglés Matrix-Assisted Laser Desorption/Ionization-Time of Flight Mass Spectrometry), pruebas de amplificación de ácido nucleicos (NAAT, por sus siglas en inglés Nucleic Acid Amplification Tests), hibridación *in situ* fluorescente de ácido nucleico peptídico (PNA-FISH, por sus siglas en inglés Fluorescence *in situ* hybridization using peptide nucleic acid probes). El ARNr específico de especie dirigido a PNA-FISH permite la detección rápida de la presencia de *Enterococcus* en los frascos de hemocultivo. Estas pruebas permiten diferenciar *E. faecalis*, *E. faecium* y otras especies de *Enterococcus* menos comunes (Deck *et al.*, 2014).

La identificación basada en MALDI-TOF MS es un método poderoso, rápido y confiable que está comenzando a ganar terreno de manera más amplia para la detección de rutina en los laboratorios de microbiología clínica para la identificación de especies (Bizzini

y Greub, 2010; Singhal *et al.*, 2015). El uso clínico de métodos basados en MALDI-TOF MS permite la identificación rápida de *Enterococcus* directamente de los frascos de hemocultivo, lo que potencialmente reduce el tiempo hasta el inicio del tratamiento antibacteriano (Lagace-Wiens *et al.*, 2012). MALDI-TOF MS tiene una alta sensibilidad, pudiendo identificar alrededor del 94% de los aislamientos a nivel de especie, incluida la diferenciación entre especies estrechamente relacionadas (Quintela-Baluja *et al.*, 2013; Stępień-Pyśniak *et al.*, 2017).

Los métodos NAAT se basan en la amplificación por PCR y la posterior secuenciación o matriz / hibridación o amplificación por PCR en tiempo real (Ljungstrom *et al.*, 2015) de uno o más genes que son útiles para la identificación de organismos a nivel de género o especie e igualmente importantes para detectar genes de resistencia a los antibacterianos.

2.2.2.1. Genes de interés en la identificación de *Enterococcus*

Se han utilizado diferentes genes con fines diagnósticos y filogenéticos. Se han propuesto varios genes para ayudar a diferenciar especies de *Enterococcus*. Un marcador propuesto fue el gen *tuf* (factor de elongación *Tu*) específico del género *Enterococcus* (Ke *et al.*, 1999). Sin embargo, los cebadores utilizados para este método también amplificaron los genes de seis especies bacterianas pertenecientes a otros dos géneros que se probaron como controles negativos en el estudio (Ke *et al.*, 2000).

En el caso de los genes de la ligasa *ddl* (d- alanina: d- alanina), se logró la caracterización de un fragmento interno, que permitió la identificación de 8 especies de *Enterococcus* (*E. columbae*, *E. durans*, *E. malodoratus*, *E. mundtii*, *E. raffinosus*, *E. seriolicida*, *E. solitarius*

y *E. sulfureus*). Cuyo análisis propuso un árbol evolutivo con una topología similar a la basada en secuencias de ARNr 16S (Ozawa *et al.*, 2000).

Además, la obtención de un fragmento de ADN de 438 pb de longitud interno (*sodA* (*int*)) al gen *sodA* que codifica la superóxido dismutasa dependiente de manganeso en 19 cepas de tipo enterococo (*E. avium*, *E. casseliflavus*, *E. cecorum*, *E. columbae*, *E. dispar*, *E. durans*, *E. faecalis*, *E. faecium*, *E. flavescens*, *E. gallinarum*, *E. hirae*, *E. malodoratus*, *E. mundtii*, *E. pseudoavium*, *E. raffinosus*, *E. saccharolyticus* *E. solitarius* y *E. sulfureus*). Logró la identificación correcta de 18 cepas, por lo que se propuso al gen *sodA* como una secuencia diana más discriminatoria que el gen ARNr 16S para diferenciar especies bacterianas estrechamente relacionadas (Poyart, Quesnes y Trieu-Cuot, 2000).

La chaperonina 60 (Cpn60), una región de aproximadamente 600 pb, amplificado por PCR con un solo par de cebadores degenerados, tuvo utilidad como un objetivo potencialmente universal para la identificación bacteriana. Este método de identificación del gen Cpn60 identificó correctamente entre 17 especies de *Enterococcus* (*E. asini*, *E. rattus*, *E. dispar*, *E. gallinarum*, *E. hirae*, *E. durans*, *E. cecorum*, *E. faecalis*, *E. mundtii*, *E. casseliflavus*, *E. faecium*, *E. malodoratus*, *E. raffinosus*, *E. avium*, *E. pseudoavium*, *E. nueva sp.* cepa Facklam y *E. saccharolyticus*) (Goh *et al*, 2000).

Estos son solo algunos marcadores y métodos moleculares basados en ADN que se han descrito para identificar *Enterococcus* a nivel de especie. La mayoría de estos estudios presentados anteriormente solo incluyeron el análisis de un número limitado de especies de *Enterococcus* y se limitaron a especies con importancia clínica (Goh *et al*, 2000). En una evaluación realizada por Kosecka-Strojek y cols. 2019, en dicho trabajo se utilizó una combinación de cuatro objetivos genéticos (ARNr 16S, *sodA*, *tuf* y *rpoB*) para confirmar sin

ambigüedades la identificación a nivel de especie para todas las cepas de *Streptococcus* y *Enterococcus*, encontrando que el gen ARNr 16S es el óptimo.

Más importante aún, la base de datos de secuencias para los genes de función específica de los *Enterococcus* ambientales y otros géneros filogenéticamente relacionados es mucho más limitante que la de los genes ARNr 16S. Además, la conservación de la secuencia en los genes funcionales es considerablemente menor que en los genes de ARNr, lo que explica por qué es difícil desarrollar ensayos específicos de género y grupo a menos que se desarrollen bases de datos de secuencias integrales (Ryu *et al.*, 2013).

2.2.2.2. Gen ARNr 16S

El gen ARNr 16S es considerado el más importante para la identificación bacteriana, ya que se trata de una molécula altamente conservada, presente en todos los organismos bacterianos, la información genética contenida en dicho gen es vital para las funciones básicas de supervivencia de la célula. Además, mantiene una tasa de mutación baja, por lo que se mantiene estable a través de las siguientes generaciones (Rodicio y Mendoza, 2004; Ajawatanawong, 2016).

El gen ARNr 16S está compuesto por aproximadamente 1,542 pb, organizado entre diez regiones conservadas que flanquean nueve regiones altamente variables (Figura 2). De manera que, la variabilidad de dichas regiones convierte a dicha región en una excelente herramienta para la estimación taxonómica y filogenética de diversas especies bacterianas (Grice *et al.*, 2008; Kong, 2011; Barb *et al.*, 2016; Yang *et al.*, 2016). Sin embargo, debido a las limitantes de secuenciación, la utilización de las regiones parciales (V1-V9) del gen ARNr 16S ha sido la metodología empleada para el análisis bioinformático. Por lo tanto, la selección de cebadores adecuados es fundamental para estudiar la filogenia bacteriana en diversos entornos (Yang *et al.*, 2016).

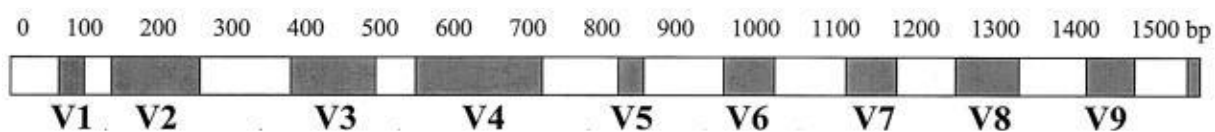


Figura 2.- Regiones hipervariables del gen ARNr 16S (Gillman *et al.*, 2001).

Desde hace décadas la amplia utilización del gen ARNr 16S ha sido de gran utilidad en la identificación bacteriana, por lo que Vasileiadis y cols., en 2012, se dieron a la tarea de evaluar los efectos de las propiedades de cuatro regiones variables consecutivas (V3-V6) sobre las metodologías analíticas comúnmente aplicadas en los estudios de ecología bacteriana. Considerando un enfoque *in silico*, se observó el rendimiento de cada región variable comparándolo con el gen completo de ARNr 16S mediante el cálculo la entropía para cada posición en la alineación, y las regiones conservadas / variables corresponden a la entropía baja / alta.

El enfoque *in silico* también fue utilizado en el trabajo de Vinje y cols., en 2014, en el cual se analizó si el poder discriminatorio del marcador ARNr16S reside solo en las regiones variables, mediante un método de aprendizaje supervisado para buscar sistemáticamente sitios que contribuyan a la clasificación correcta, ya sea a nivel de filo o de género. Se encontró que al clasificar los phyla, el algoritmo de selección de sitios localizó 50 sitios discriminativos. Estos estaban dispersos en la mayoría de los alineamientos y solo alrededor de la mitad de ellos estaban ubicados en las regiones variables. Sin embargo, los sitios seleccionados tenían una entropía significativamente mayor de lo esperado, lo que significa que son sitios de gran variación. Se descubrió que los sitios discriminativos suelen tener una gran entropía en comparación con sus vecinos más cercanos a lo largo de los alineamientos. Al clasificar los géneros, el algoritmo de selección de sitios necesitaba alrededor del 80% de los sitios en el gen ARNr 16S antes de que el error de clasificación alcanzara un mínimo. Esto significa que toda la variación, tanto en las regiones variables como en las conservadas, es necesaria para separar los géneros. Estos hallazgos no apoyan la afirmación de que el poder discriminatorio del gen 16S se encuentra solo en las regiones variables. Las regiones variables son importantes, pero también se encuentran muchos sitios discriminativos en las partes más

conservadas. El poder discriminativo se encuentra típicamente en sitios de gran variación ubicados dentro de regiones más cortas de mayor conservación.

La secuenciación del gen ARNr 16S se usa comúnmente para identificar especies bacterianas y permite la discriminación de *Enterococcus* a nivel de especie (Bosshard *et al.*, 2004, Angeletti *et al.*, 2001); sin embargo, diferenciar entre las especies estrechamente relacionadas dentro de un grupo, como el grupo de *E. faecium*, puede ser menos preciso (Devriese *et al.*, 2012; Quintela-Baluja *et al.*, 2013; Poyart, Quesnes y Trieu-Cuot, 2000).

2.2.2.3. Regiones hipervariables

En la actualidad, la gran mayoría de los estudios secuencian solo una parte del gen ARNr 16S, porque la plataforma de secuenciación Illumina ampliamente utilizada (mayor rendimiento, menor costo, menor esfuerzo) en comparación con Sanger produce secuencias cortas (≤ 300 bases). Por lo tanto, se apunta a diferentes subregiones del gen, que van desde regiones de una sola variable, como V4 o V6, a tres regiones variables, como V1-V3 o V3-V5 (utilizadas en el Proyecto de microbioma humano junto con la secuenciación 454 plataforma) (Johnson *et al.*, 2019).

Campista-León y cols. en 2021 caracterizaron filogenéticamente bacterias multirresistentes aisladas de cócteles de mariscos, en donde las regiones V4 y V6 fueron amplificadas para la identificación filogenética, encontrando *Enterococcus*, *Proteus*, *Vibrio*, *Staphylococcus*, *Lactococcus*, *Vagococcus*, *Micrococcus*, *Acinetobacter*, *Enterobacter* y *Brevibacterium*. Y reportaron que para el género *Enterococcus*, no se logró resolver su filogenia, ya que solo se presentaron como grupo de especies.

2.3. Análisis filogenético

El análisis filogenético es el estudio de las relaciones evolutivas entre diferentes entidades biológicas basándose en la evidencia presente en su material genético (Challa y Neelapu, 2019; Farrar, 2019), indica qué tan estrechamente relacionados están los organismos entre sí, organizando cada uno de los componentes en un esquema taxonómico jerárquico (Morrison, 2011). Un árbol filogenético se define como el diagrama utilizado para visualizar las relaciones evolutivas entre un conjunto de unidades taxonómicas operacionales (OTU's, por sus siglas en inglés Operational Taxonomic Units), las cuales generalmente representan las especies u organismos individuales dentro de una población (Ajawatanawong, 2016).

Los árboles filogenéticos han sido un artefacto clave para el avance en varios tipos de investigación; desde la clasificación taxonómica de organismos, hasta estudios de epidemiología en el área de la salud (Munjal *et al.*, 2019). De manera que, el desarrollo de árboles filogenéticos confiables es un paso importante para caracterizar nuevos patógenos y desarrollar nuevos tratamientos en biomedicina (Farrar, 2019).

En un contexto clínico, el análisis filogenético puede usarse para inferir si el trasfondo genético común de los patógenos significa que surgen de una fuente similar. Comprender la posible asociación de la variación del patógeno con el resultado de la infección es muy prometedor para el diagnóstico y el tratamiento (Hartfield *et al.*, 2014). Por lo tanto, la identificación correcta de las especies es crítica para la elección correcta de antibióticos o para emplearse novedosamente como probióticos (Farrar, 2019).

Debido a la importancia de *Enterococcus* como bacterias ácido-lácticas, Zhong y cols., en 2017 realizaron un análisis genómico comparativo del género *Enterococcus*, en el cual, se

secuenciaron los genomas de 29 cepas de especies de *Enterococcus*. En combinación con ocho secuencias genómicas completas de la base de datos Genbank, los genomas completos de 37 cepas de *Enterococcus* se analizaron comparativamente, y se describió que las relaciones genéticas fueron más estrechas en cepas que provienen de hábitats similares. De acuerdo con la topología del árbol del tiempo, se determinó que los humanos y los mamíferos pueden ser los huéspedes originales de *Enterococcus*. La longitud media de los genomas de *Enterococcus* fue de 3.20 Mb y el contenido medio de GC fue del 37.99%.

2.3.1. Métodos de construcción de arboles

2.3.1.1. Métodos basados en alineamientos

Dentro de los métodos basados en alineamientos se encuentran algunos de los primeros algoritmos planteados para la reconstrucción de filogenias, dado un conjunto de datos moleculares. Como el nombre lo indica, estos algoritmos reciben como entrada un alineamiento múltiple de secuencias (Figura 3). El alineamiento de secuencias consiste en organizar las secuencias, una encima de la otra, y resaltar las posiciones en las cuales las secuencias difieren, basándose en algunos parámetros de alineamiento u operaciones (como las inserciones, deleciones o sustituciones) (Munjal *et al.*, 2019).

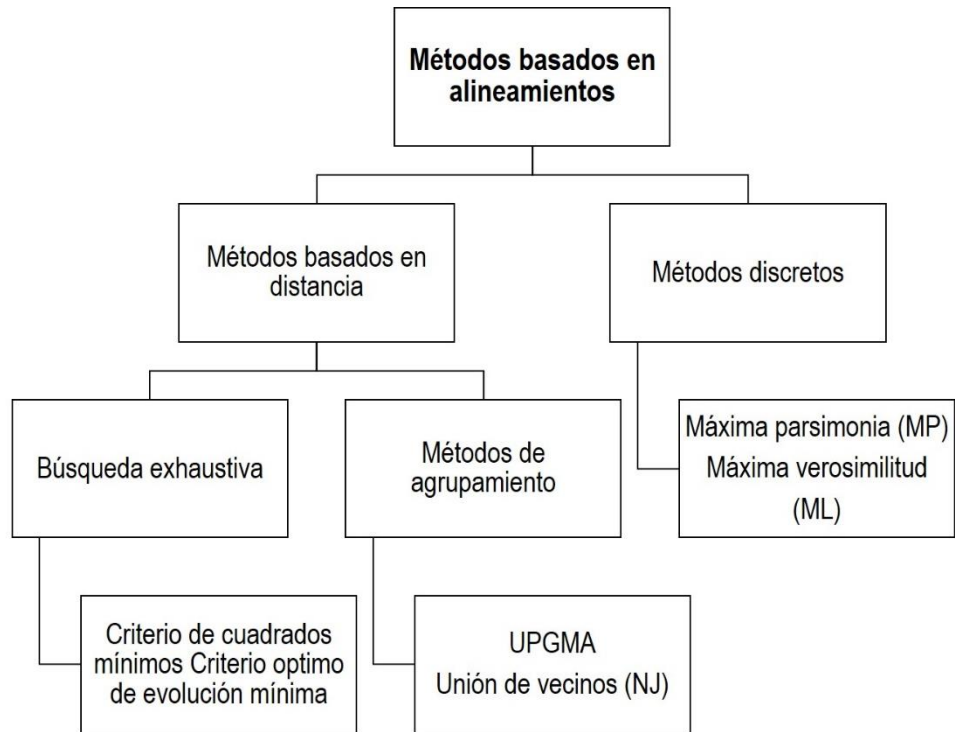


Figura 3.- Jerarquía de los métodos basados en alineamientos (adaptado de Munjal *et al.*, 2019).

2.3.1.2. Métodos basados en distancias

Los algoritmos de agrupamiento para construcción de árboles filogenéticos son métodos basados en distancias que realizan una construcción progresiva de subárboles (aglomeraciones), de acuerdo con qué tan “cerca” están unos de otros. De nuevo, el criterio de cercanía o similitud tomará la forma de una función que evalúe una matriz de distancias. (Hong *et al.*, 2020). Este tipo de algoritmos se pueden implementar con complejidad temporal polinomial, por lo cual, en general, son más rápidos que los algoritmos que realizan búsquedas sobre el espacio de árboles (Munjal *et al.*, 2019).

Unión de vecinos

El método de unión de vecinos o NJ (por sus siglas en inglés Neighbor Joining) es un algoritmo originalmente ideado por Saitou y Nei (1987), las descripciones, diagramas y notación están basados completamente en dicho trabajo, donde en cada iteración se busca encontrar las parejas de secuencias OTU's que minimicen la longitud total del árbol.

Cuadro 2.- Clasificación de los métodos de análisis filogenéticos y sus estrategias

	Optimización del criterio de búsqueda	Agrupamiento
Estado de carácter	Máxima parsimonia (MP) Máxima verosimilitud (ML) Inferencia Bayesiana	
Matriz de distancia	Fitch-Margoliash	UPGMA Unión de vecinos (NJ)

2.3.1.3. Métodos discretos o Estado de carácter

Entre los métodos cladistas comúnmente empleados para la inferencia filogenética destacan los métodos discretos o métodos de estado carácter que pueden usar cualquier conjunto de caracteres discretos como: caracteres morfológicos, propiedades fisiológicas, mapas de restricción o conjunto de secuencias. Cuando se comparan secuencias, cada posición de la secuencia en el alineamiento es un “carácter”, y los nucleótidos o aminoácido en la posición son los “estados”. Usualmente, cada posición/carácter es analizada independientemente, y cada columna en el alineamiento asume ser independiente en la realización del proceso evolutivo. Por lo cual, estos métodos retienen el estado del carácter original de los taxones por lo tanto se pueden utilizar para reconstruir el estado de carácter de los nodos ancestrales. Dentro de estos, los algoritmos más conocidos son el de Máxima parsimonia (MP por sus siglas en inglés Maximum Parsimony) y máxima verosimilitud (ML por sus siglas en inglés Maximum Likelihood). Tanto MP como ML realizan búsquedas en el espacio de árboles (tree space) para tratar de encontrar el mejor árbol de acuerdo con algún criterio de comparación. En la literatura son descritos como problemas de optimización difíciles y son clasificados como NP-Hard (Roshan *et al.*, 2004; Lemey *et al.*, 2009; Peña, 2011; Ajawatanawong, 2016).

Máxima parsimonia

La parsimonia es un concepto que se refiere a encontrar la solución o respuesta a un problema, que contenga la explicación más simple. En filogenética, existen muchas topologías de árbol que pueden describir un conjunto de secuencias; en este caso MP escogerá el árbol con menor cantidad de eventos con mutaciones (Richer *et al.*, 2013).

El método de MP asume que los caracteres compartidos por las diferentes entidades provienen de un descendiente en común. Los grupos son construidos en base a dicha compatibilidad, la longitud de caracteres es la mínima cantidad de cambios requeridos para dicho sitio, el árbol que requiere el menor número de cambios evolutivos para ser construido se considera el árbol con mayor parsimonia (Yang y Rannala, 2012).

Máxima verosimilitud

En el área de probabilidad, la estimación por Máxima Verosimilitud (Maximum Likelihood) es una técnica utilizada para aproximar los valores de los parámetros de una distribución de probabilidad determinada. La verosimilitud de un modelo también es vista como la probabilidad condicional de los datos recibidos, dado el modelo o hipótesis. El objetivo es encontrar los valores de los parámetros que maximizan una función de verosimilitud (Zwickl, 2006).

ML es un método estadístico que estima la longitud de las ramas y la topología del árbol basado en el modelo de sustitución y en el alineamiento de secuencias mediante el ajuste de un modelo y la estimación sus parámetros mediante la función de verosimilitud. El valor obtenido del análisis es la probabilidad de que la topología y el modelo empleado se ajusten a las secuencias tras la modificación de los parámetros. El árbol con mayor puntaje corresponde al árbol con la máxima verosimilitud (Yang y Rannala, 2012; Ajawatanawong, 2016).

Inferencia bayesiana

Por su parte, los métodos de inferencia bayesiano (IB) son métodos de estado de carácter que usan un criterio de optimización, pero son conceptualmente muy diferentes de MP y ML en que no intentan buscar solo el mejor árbol, donde también se evalúan los

posibles árboles al modificar los parámetros y calculando la verosimilitud de cada árbol. La diferencia se observa en que antes del análisis, a cada parámetro se le asigna una probabilidad *a priori*, es decir la probabilidad de cada hipótesis, la cual se combina con los datos de verosimilitud generando así una probabilidad *a posteriori*. Las probabilidades posteriores se obtienen explorando el espacio del árbol (tree space) usando una técnica de muestreo, llamado Cadena de Markov Monte Carlo (MCMC, por sus siglas en inglés Markov Chain Monte Carlo). Este método de muestreo comienza simulando un conjunto aleatorio de parámetros y propone un nuevo "estado", que es un nuevo conjunto de parámetros, cambiando los parámetros hasta cierto punto utilizando operadores aleatorios. En cada paso, se calcula la relación de verosimilitud y la probabilidad *a priori* para el nuevo estado en relación con el estado actual. Cuando el producto combinado es mejor, se aceptan los parámetros y se propone un siguiente paso; si el resultado es peor, la probabilidad de que el estado sea rechazado es inversamente proporcional a cuánto peor es el nuevo estado. Después de una convergencia inicial a un conjunto de soluciones probables de modelo/árbol ("burn-in", que debe descartarse), se espera que este algoritmo estocástico tome muestras de la distribución de probabilidad "posterior". Así el árbol resultante será el que maximice la probabilidad *a posteriori* (Lemey *et al.*, 2009; Yang y Rannala, 2012).

III. PLANTEAMIENTO DE PROBLEMA

Los *Enterococcus* spp. tienen un papel bien conocido como miembros en la microbiota intestinal y está documentada su resistencia intrínseca a vancomicina desde la década de los 80's. Las cepas multirresistentes como *E. faecalis* y *E. faecium*, ambas especies causante de ITU cuya mortalidad hospitalaria es de 16.5% (Álvarez-Artero *et al.*, 2021). Además, actúan como reservorio de genes de resistencia que se transfieren a otros patógenos importantes, como *Staphylococcus aureus*. Debido a esto han tomado cierta relevancia ya que son considerados como uno de los principales causantes de las infecciones nosocomiales, indicadores de contaminación por materia fecal y de calidad de las fuentes de agua.

Existen una variedad de genes que se utilizan en la identificación bacteriana como ARNr 16S, genes de la ligasa ddl (d- alanina: d- alanina) (Ozawa *et al.*, 2000), el gen tuf (factor de elongación *Tu*) (Ke *et al.*, 1999), la chaperonina 60 (Cpn60) (Goh *et al.*, 2000). Y el fragmento interno (sodA (int)) al gen sodA que codifica la superóxido dismutasa dependiente de manganeso (Poyart, Quesnes y Trieu-Cuot, 2000).

En una evaluación realizada por Kosecka-Strojek y cols. 2019, en dicho trabajo se utilizó una combinación de cuatro objetivos genéticos (ARNr 16S, sodA, tuf y rpoB) para confirmar sin ambigüedades la identificación a nivel de especie para todas las cepas de *Streptococcus* y *Enterococcus*, encontrando que el gen ARNr 16S es el óptimo.

Definida su importancia surge la problemática, ya que la taxonomía del género como *Enterococcus* basada en el ARNr 16S resulta conflictiva al presentar problemas de discriminación entre grupos de especies de alta similitud en este taxón (Zhong *et al.*, 2017; Church *et al.*, 2020; Campista-León *et al.*, 2021). Este marcador molecular es de primera elección para la identificación de bacterias y la descripción de comunidades microbianas

(Chakravorty *et al.*, 2007; Manaka *et al.*, 2017; Johnson *et al.*, 2019). Debido a su alta conveniencia en su uso, este marcador se seguirá empleando rutinariamente en identificación, descripción y clasificación taxonómica bacteriana. Por lo que es necesario definir un método de utilidad para identificar aquellas especies muy cercanas, de manera rápida y con una alta especificidad en el ámbito clínico y ecológico. Por lo que se plantea la siguiente pregunta de investigación *¿La información genética contenida en el gen ARNr 16S es suficiente para la discriminación entre especies del género Enterococcus?*

IV. JUSTIFICACIÓN

Entre las bacterias Gram positivas, *Enterococcus* se considera un género primitivo, por ser reconocido como miembro del núcleo de la microbiota intestinal, al colonizar primero a mamíferos y después a una amplia variedad de aves, reptiles y plantas, además de su presencia en recursos abióticos como el suelo y el agua e incluso en actividades económicas como la fermentación de alimentos. En la actualidad han tomado notoriedad debido al aumento en la prevalencia de especies como *E. faecalis* y *E. faecium*, y al surgimiento de cepas multirresistentes en el ambiente hospitalario y que llegan a propagarse hacia la comunidad en general. Un ejemplo de lo anterior es la aparición de *E. faecium* con prioridad elevada para investigación y desarrollo de nuevos tratamientos en el listado de patógenos publicado por la OMS, 2017. Debido a estas características se denota como un género único y de importancia tanto a nivel ecológico, económico y de salud pública.

En años recientes, se han descrito y clasificado un gran número de nuevas especies dentro del género *Enterococcus*, esto ha dificultado la comprensión de las relaciones filogenéticas entre especies que provienen de hábitats similares o que están estrechamente relacionadas y además podrían presentar multirresistentes. Por lo cual es importante y necesario poder diferenciar entre una especie patogénica y no patogénica, ya que no todas están reportadas como patogénicas.

Por lo que es necesario una identificación eficiente y confiable, ya que las deficiencias o dificultades en esta tarea, conlleva al diagnóstico incorrecto y a un inadecuado tratamiento de los pacientes o al mal manejo de las fuentes de agua o alimentos contaminados por *Enterococcus*. Debido a esto resulta relevante el estudio de dichas relaciones entre los integrantes de este género, con la implementación de análisis bioinformáticos exhaustivos

del gen ARNr 16S y sus regiones hipervariables, ya que es marcador molecular más utilizado en la identificación bacteriana y del cual existe un mayor número de bases de datos disponibles.

VI. HIPÓTESIS

Existen regiones patrón dentro del gen ARNr 16S que indican la variación a nivel de especie específicas para el género *Enterococcus*.

VI. OBJETIVOS

8.1. Objetivo general

Evaluar la variabilidad del gen ARNr 16S en *Enterococcus* spp. mediante estimaciones filogenéticas y entropía de Shannon.

8.2. Objetivos específicos

1. Establecer una asociación taxonómica de las especies del género *Enterococcus* utilizando el gen ARNr 16S.
2. Cuantificar y definir las posiciones variantes del ARNr 16S en las especies del género *Enterococcus* utilizando Entropía de Shannon como indicador.
3. Evaluar los marcadores similares y diferentes al ARNr 16S por Resolución Filogenética.
4. Determinar los oligotipos presentes en el gen ARNr 16S en las especies del género *Enterococcus*.

VII. MATERIALES Y MÉTODOS

7.1. LUGAR DE OBTENCIÓN DE LAS MUESTRAS

Las secuencias de ADN perteneciente a *Enterococcus* spp. fueron recuperadas de cuatro bases de datos de acceso libre: GenBank, Ribosomal Database Project (RDP, por sus siglas en inglés), la base de datos SILVA y Greengenes versión 13.5.

7.2. TIPO DE ESTUDIO

Estudio descriptivo, transversal, cuasi-experimental. y explicativo

7.3. MUESTRA

7.3.1. Tamaño de muestra

Se incluyeron las secuencias que cumplieron con los criterios de inclusión.

En la estimación filogenética del conjunto GenBank se seleccionaron 142 secuencias. En el caso de los análisis de entropía se utilizaron un total de 2, 800 no redundantes, distribuidas en tres conjuntos pertenecientes a tres bases de datos diferentes SILVA: 713, RDP: 1,819 y Greengenes: 425.

7.3.2. Tipo de muestreo

Por conveniencia

7.3.3. Criterios de Inclusión

- En la base de datos GenBank se colectaron secuencias pertenecientes a cepas tipo de especies taxonómicamente aceptadas de *Enterococcus* depositadas en NCBI GenBank y NCBI 16S RefSeq versión 204
- Para la base de datos Ribosomal Database Project las secuencias colectadas incluyeron cepas tipo y no tipo, $\geq 1,200$ nucleótidos y de buena calidad.

- En la base de datos de SILVA, los criterios de colecta incluyeron secuencias con anotación taxonómica SILVA, calidad de secuencia ≥ 90 , calidad de pintail ≥ 90 y longitud de secuencia $\geq 1,200$ nucleótidos.
- Otro conjunto de secuencias del género se obtendrá de la base de datos Greengenes versión 13.5

7.3.4. Criterios de Eliminación

Las secuencias sin nombre taxonómico binomial o nombre taxonómico inválido fueron removidas manualmente de dicho conjunto.

En todos los conjuntos, exceptuando el conjunto taxonómico, se removieron las secuencias con caracteres del alfabeto extendido ADN/ARN, así como las columnas de gaps $\geq 90\%$.

7.4. METODOLOGÍA

7.4.1. Definición de las posiciones del gen ARNr 16S

Se utilizó la secuencia referencia del 16S de *Escherichia coli* K-12 subcepa MG1655 (gen rrsB; Gene ID en NCBI: 948466) como plantilla para mapear las posiciones nucleotídicas en las especies de *Enterococcus*. La determinación de las regiones variables se logró a través de una detección inicial de regiones conservadas dentro de la secuencia referencia. Se buscaron los marcadores conservados de acuerdo con lo reportado en otros estudios (Baker *et al.*, 2003; Wang y Quian, 2009; Yang *et al.*, 2016). Las secuencias patrón se describen en el Cuadro 3. La búsqueda de concordancias de los marcadores conservados con la secuencia referencia se realizó en Unipro UGENE (Okonechnikov *et al.*, 2012); disponible en ugene.net). Las regiones flanqueadas por estos marcadores conservados fueron anotadas como regiones variables del gen. Para identificar con mayor precisión las posiciones

en las secuencias del género, se extrapolaron las anotaciones a la secuencia de *Enterococcus faecalis* ATCC 19433 (acceso Genbank: NR_115765).

Cuadro 3. Marcadores conservados utilizados para la determinación de las regiones variables en el gen ARNr 16S

Marcador conservado	Secuencia*	Posición de inicio esperada en el gen
1	AGAGTTTGATCATGGCTCA	8
2	GGCGVACGGGTGAGTAA	104
3	CAYTGGRACCTGAGACACGGYCC	314
4	GGCTAACTHCGTGCCAGCAGC	505
5	CGAAAGYGTGGGKAKCRCAGG	764
6	ACTCAAAGAATTGACGGGGRC	909
7	GTGSTGCATGGYTGTCGTCAGCT	1048
8	GGAAGGYGGGGAYGACGTCAA	1177
9	TGTACACACCGCCCGTCACACAC	1391
10/Fin del 16S	AAGTCGTAACAAGGTAACCGTA	1492

*V= A, G o C, Y= C o T, H= A, C o T, R= A o G, K= G o T, S= C o G.

7.4.2. Colecta de secuencias del gen ARNr 16S

Se colectaron cinco conjuntos de secuencias correspondientes al 16S: un *conjunto taxonómico* compuesto por secuencias pertenecientes a cepas tipo de especies taxonómicamente aceptadas de *Enterococcus* depositadas en NCBI GenBank y NCBI 16S RefSeq versión 204 (Clark *et al.*, 2016; O'Leary *et al.*, 2016). La validación de la nomenclatura taxonómica de dichas especies se realizó consultando la referencia en taxonomía bacteriana *LPSN* (Parte *et al.*, 2020; Disponible en lpsn.dsmz.de/genus/enterococcus). Un conjunto de secuencias correspondiente a *Enterococcus* fue recuperado de la base de datos correspondiente a la subunidad menor, conjunto de referencia no redundante (SSU Ref NR 99, por sus siglas en inglés), versión r138.1 SILVA (Quast *et al.*, 2013; Disponible en arb-silva.de). Los criterios de colecta

incluyeron secuencias con anotación taxonómica SILVA, calidad de secuencia ≥ 90 , calidad de pintail ≥ 90 y longitud de secuencia $\geq 1,200$ nucleótidos. Un conjunto de secuencias de *Enterococcus* fue recuperado del Ribosomal Database Project (RDP, por sus siglas en inglés; Cole *et al.*, 2014; Disponible en rdp.cme.msu.edu). Las secuencias colectadas incluyeron cepas tipo y no tipo, $\geq 1,200$ nucleótidos y de buena calidad. Finalmente, el último conjunto de secuencias del género se obtuvo de la base de datos Greengenes versión 13.5 (DeSantis *et al.*, 2006; Disponible en: Greengenes.lbl.gov). Las secuencias sin nombre taxonómico binomial o nombre taxonómico inválido fueron removidas manualmente de dicho conjunto. En todos los conjuntos, exceptuando el conjunto taxonómico, se removieron las secuencias con caracteres del alfabeto extendido ADN/ARN, así como las columnas de gaps $\geq 90\%$.

7.4.3. Alineamiento de múltiples secuencias y mapeo posicional

Todos los conjuntos de secuencias colectados fueron alineados en Unipro UGENE usando MAFFT versión 7.475 (Kato y Standley, 2013.) bajo parámetros predeterminados. Las columnas de gaps $\geq 90\%$ fueron removidas. Las posiciones del 16S comprendidas en los alineamientos se determinaron usando las secuencias anotadas con las regiones variables y conservadas de *E. coli* K-12 subcepa MG1655 y *E. faecalis* ATCC 19433.

7.4.4. Análisis filogenético basado en el gen ARNr 16S

Para estimar las relaciones taxonómicas de las especies del género *Enterococcus* se realizó un análisis filogenético con métodos de árbol basados en distancias y en criterios de optimización bajo los siguientes pasos: (1) Definición del perfil de alineamiento, (2) Búsqueda de modelos de sustitución nucleotídica, (3) Estimación filogenética mediante Unión de Vecinos, (4) Máxima Verosimilitud, (5) Inferencia Bayesiana y (6) Consenso de topologías.

7.4.4.1. Definición del perfil de alineamiento

Partiendo del conjunto taxonómico alineado, refinado y mapeado, se extrajo la secuencia consenso (modo estricto, límite mínimo 50%) para la búsqueda de grupos externos usando NCBI BLAST con el programa blastn, bajo algoritmo MegaBLAST (Morgulis *et al.*, 2008; Disponible en: blast.ncbi.nlm.nih.gov/Blast.cgi) contra la base 16S Ribosomal RNA Sequences (16S RefSeq). Las búsquedas iniciaron con la exclusión de secuencias del género *Enterococcus* (NCBI taxid: 1350), limitando a miembros de la familia Enterococcaceae (NCBI taxid: 81852). Las secuencias se recuperaron bajo criterio de porcentaje de identidad <95%. Las secuencias obtenidas de la búsqueda fueron añadidas y alineadas con los parámetros de alineamiento usados previamente.

7.4.4.2. Búsqueda de modelos de sustitución nucleotídica

En la determinación de los modelos se usó el complemento ModelFinder (Kalyaanamoorthy *et al.*, 2017) de IQ-TREE versión 1.6.12 (Nguyen *et al.*, 2015) considerando 22 esquemas de sustitución (Núm. de modelos=154), frecuencias nucleotídicas desiguales ($A \neq T \neq C \neq G$), tasas de variación Gamma (+ Γ), proporciones de sitios invariantes (+I). Obtenidos los modelos, el mejor modelo ajustado se seleccionó bajo los criterios de Información de Akaike corregido (AICc; Akaike, 1974), y Criterio de Información Bayesiano (CIB; Schwarz, 1978) siendo el modelo con las ponderaciones más bajas por ambos criterios el aceptado para simular los eventos de sustitución nucleotídica.

7.4.4.3. Estimación filogenética mediante Unión de Vecinos

La construcción del árbol filogenético mediante Unión de Vecinos (Neighbor Joining; Saitou y Nei, 1987) se utilizó el programa PAUP versión 4.0 (Swofford, 2002) de construcción de árboles filogenéticos bajo el modelo seleccionado en ModelFinder, enraizando el árbol en un grupo parafilético compuesto por los grupos de referencia colectados. La confianza estadística de nodo se estimó mediante una búsqueda heurística con 1,000 réplicas de bootstrap (Felsenstein, 1985). La topología obtenida fue condensada a un corte del 70% de soporte de nodo (≥ 700 réplicas).

7.4.4.4. Estimación filogenética mediante Máxima Verosimilitud

Se utilizó IQ-TREE para la construcción del árbol filogenético basado en Máxima Verosimilitud usando el modelo especificado. La búsqueda de topologías se realizó a partir de Intercambios de Vecinos Cercanos (NNI). 1,000 réplicas de bootstrap fueron usadas para estimar la confianza de nodos. La topología obtenida fue condensada a un corte del 70% de soporte de nodo (≥ 700 réplicas).

7.4.4.5. Estimación filogenética mediante inferencia Bayesiana

El árbol filogenético mediante IB fue generado usando BEAST versión 1.8.4 (Drummond *et al.*, 2012). En el análisis se utilizó una Cadena de Markov Monte Carlo (MCMC) diagnóstica de 10^7 generaciones bajo el modelo seleccionado por ModelFinder, descartando el 10% de las generaciones como burn-in. La estimación de los tamaños de muestra efectivos se realizó en Tracer, versión 1.6 (Rambaut *et al.*, 2018). Se consideraron tamaños de muestra efectivos $\geq 2,000$ ESS. Las generaciones en la cadena fueron incrementadas hasta alcanzar dicho criterio, aplicando el mismo porcentaje de burn-in.

7.4.4.6. Consenso de topologías

Los árboles filogenéticos obtenidos se fusionaron mediante TreeGraph2 (Stöver y Müller, 2010) utilizando la topología obtenida por IB como base. Para generar un dendrograma consenso, se removieron las topologías conflictivas de cada árbol. El orden jerárquico de soporte se especificó en el siguiente orden: IB>MV>NJ. El árbol obtenido fue exportado a IToL versión 4 (Letunic y Bork, 2019) para su edición final.

7.4.5. Estimación de entropía posicional del gen ARNr 16S

Para determinar la variación posicional a través del gen ARNr 16S se utilizaron las calculadoras de Resumen Numérico Posicional de Nucleótidos (Positional Nucleotide Numerical Summary en inglés), Entropía del Instituto Nacional de Salud Pública, República Checa (Disponible en <http://www1.szu.cz:8080/EntropyCalcWeb/>; Nagy *et al*, 2015) y el programa Bioedit versión 7.2 (Hall, 1999) partiendo de los alineamientos colectados de SILVA, RDP y Greengenes. La fórmula de cálculo de entropía se describe a continuación:

$$H(i) = - \sum f(x, i) \ln f(x, i)$$

Donde $H(i)$ corresponde a la entropía en la posición i , $f(x, i)$ es la frecuencia de cada base x en la posición i en un alineamiento de múltiples secuencias. Aunque las posiciones con valores de entropía cercanos o iguales a 0 se consideraron de alta o total conservación (Nagy *et al*, 2015), la clasificación de los sitios conservados/variantes se realizó mediante un enfoque de K-medias usando el software de aprendizaje de máquinas WEKA versión 3.8.5. (Frank *et al.*, 2016) bajo parámetros predeterminados. Los sitios clasificados como variantes fueron seleccionados de los alineamientos y filtrados para la generación de alineamientos de sitios variantes.

El sesgo por la presencia de gaps en los extremos del alineamiento por variaciones de longitud se ajustará a partir de la entropía inicial obtenida mediante la fórmula:

$$H(i)_a = \frac{H(i)_p [C_i]}{100\%}$$

Donde $H(i)_a$ corresponde a la entropía ajustada en la posición i , $H(i)_p$ es la entropía inicial en la posición i y C_i es la cobertura (porcentaje de nucleótidos) en la posición i .

7.4.6. Análisis estadístico

7.4.6.1. Definición de parámetros estadísticos

Se utilizó la entropía media, proporción de sitios variantes, proporción de sitios conservados, número de sitios variantes y número de sitios conservados de cada región conservada y variable del ARNr 16S como parámetros para la discriminación de la variabilidad génica de cada región. Los parámetros de cada región conservada/variable fueron contrastados contra los obtenidos de la totalidad del gen ARNr 16S. La estimación de los parámetros fue obtenida de cada alineamiento (SILVA, RDP, Greengenes y el conjunto fusionado). Además, se realizó un análisis de componentes principales con la finalidad de reducir la dimensión de la cantidad de variables para explicar la varianza de los datos en el programa IBM SPSS Statistics versión 22 (IBM Corp, 2015).

7.4.6.2. Correlación de marcadores

Se estimó la correlación de los marcadores conservados/variables y gen ARNr 16S completo a partir de los parámetros evaluados mediante el índice de correlación de Spearman. Se consideró una significancia estadística $p < 0.05$. Para dicho análisis se utilizó el programa IBM SPSS Statistics versión 22 (IBM Corp, 2015).

7.4.6.3. Agrupamiento Jerárquico

Los parámetros de las regiones conservadas/variables y gen ARNr 16S fueron agrupados mediante el método de grupos de pares no ponderados con media aritmética (UPGMA, por sus siglas en inglés) en PAST 4 (Hammer *et al.*, 2001). Para estimar la relación entre los marcadores y el gen ARNr 16S se utilizó similitud por distancia euclidiana. La confianza estadística de los clados obtenidos se estimó mediante 10,000 réplicas de bootstrap, con un corte de nodo $\geq 7,000$ réplicas. La confianza estadística del agrupamiento se evaluó bajo el índice de correlación cofenética ≥ 0.95 . Aquellos marcadores agrupados junto al gen ARNr 16S se consideraron representativos de dicho gen. El árbol obtenido fue exportado a IToL (Interactive Tree Of Life) v6 (Letunic y Bork, 2021) para su edición final y la creación de un mapa de calor para visualizar los cambios en los valores de correlación con base en un gradiente de colores.

7.4.6.4. Evaluación de marcadores

Aquellos marcadores similares y disimilares al gen ARNr 16S por correlación y agrupamiento jerárquico fueron sometidos a evaluación de marcadores mediante resolución filogenética. Se utilizó alineamiento del conjunto fusionado integrado por secuencias de las tres bases de datos (SILVA, RDP y Greengenes) fue fragmentado en sus respectivos marcadores. Se generaron 2 conjuntos adicionales de secuencias: un conjunto concatenado con los marcadores similares al gen ARNr 16S y un segundo concatenado con los marcadores disimilares al gen ARNr 16S. Estos alineamientos, junto con el alineamiento inicial del gen ARNr 16S fusionado, fueron simulados en IQ-TREE versión 2.0 (Nguyen *et al.*, 2015) utilizando el complemento Alisim (Ly-Trong *et al.*, 2022) bajo el comando `-alisim alignment_mimic -s example.phy --no-copy-gaps`, los conjunto simulados se usaron para la

generación de árboles filogenéticos mediante IQ-TREE bajo el método de Máxima verosimilitud (ML) considerando 1,000 réplicas basadas en prueba de razón de verosimilitud aproximada similar a SH (Shimodaira y Hasegawa, 1999; Guidon *et al.*, 2010) como prueba de confianza estadística. El árbol obtenido fue exportado a TreeGraph2 (Stöver y Müller, 2010) para la revisión del soporte de los nodos (>70%), después fue exportado a IToL (Interactive Tree Of Life) v6 (Letunic y Bork, 2021) para su edición final. Obtenidos los árboles filogenéticos, se estimó la sensibilidad de agrupamiento en los grupos monofiléticos encontrados utilizando la siguiente formula:

$$S = \frac{VP}{VP + FN}$$

Donde *S* corresponde a la sensibilidad; VP verdaderos positivos y FN: falsos negativos.

7.4.6.5. Oligotipificación del conjunto oligotipado

Se construyó un alineamiento múltiple del conjunto oligotipado con solo 60 secuencias pertenecientes a 60 especies diferentes de *Enterococcus* válidamente aceptadas a la fecha de acuerdo con la consulta en LPSN. La longitud común de estas secuencias fue reducida a 1,252 sitios, debido a longitud mínima (1,260 pb) de las secuencias de *E. alcedinis* CCM 8433^T (MT760203), que comprende las regiones V2-V8. Dicho alineamiento se utilizó para el análisis de entropía de Shannon.

Con base en dicho análisis, se implementó la tubería de oligotipificación (Eren *et al.*, 2011) con el fin de identificar aquellos componentes/posiciones que aportan el mayor grado de información para la identificación de especies estrechamente relacionadas, resultando en oligotipos específicos por especie, es decir, una secuencia particular para cada especie.

7.5. LUGAR DE REALIZACIÓN

Este trabajo de investigación se realizó en el Laboratorio de Microbiología y Biología Aplicada de la Facultad de Biología de la Universidad Autónoma de Sinaloa. Durante los meses de septiembre de 2021 a septiembre de 2022.

7.6. FINANCIAMIENTO

Número de beca 777739 otorgada por CONACYT.

VIII. RESULTADOS

8.1. Descripción de los conjuntos de datos

Se colectaron 3,659 de *Enterococcus* spp. de las bases de datos SILVA, RDP y Greengenes sin caracteres del alfabeto DNA/RNA extendido, agrupadas en cinco conjuntos de datos. La composición de los conjuntos de datos se detalla en el Cuadro 4. Para la estimación filogenética, las secuencias obtenidas a partir de la base de datos NCBI Genbank (16S RefSeq + nt), se les denominó *Conjunto Taxonómico*. El alineamiento final estuvo constituido por 142 secuencias con una longitud de 1,253 sitios (Anexo 10.1), el cual se utilizó para la construcción del árbol filogenético (Cuadro 4).

Cuadro 4. Descripción de los conjuntos de secuencias generados en este estudio.

Análisis	Organismo	Conjuntos de Datos	Origen	Composición					
				Inicial		Filtrado		Final	
				Núm. de secuencias	Pos	Núm. de secuencias	Pos	Núm. de secuencias	Pos
Estimación Filogenética	<i>Enterococcus</i> spp.	ENT-CT	NCBI Genbank (16S RefSeq + nt)	181	1,631	142	1,631	142	1,253
		ENT-SIL	SILVA	1,332	3,278	1,129	2,690	713	1,571
Variabilidad por Entropía	<i>Enterococcus</i> spp.	ENT-RDP	RDP	2,943	3,709	2,090	3,081	1,819	1,534
		ENT-GG	Greengenes	622	2,045	440	1,828	425	1,526
		ENT-FUS						2,800	1,526

Pos=posiciones

8.2. Estimación filogenética del género *Enterococcus*

La estimación filogenética comprendió a 60 especies de *Enterococcus* válidamente aceptadas a la fecha de acuerdo con la consulta en LPSN. La longitud común de estas secuencias fue reducida a 1,253 sitios, debido a longitud mínima (1,260 pb) de las secuencias de *E. alcedinis* CCM 8433^T (MT760203) y L34^T (NR_109727).

Los árboles filogenéticos fueron inferidos bajo el modelo TPM2+F+I+Γ4, siendo seleccionado de acuerdo con el puntaje del CIB (Anexo 10.2 y 10.3) en Máxima Verosimilitud y equiparado con GTR+F+I+Γ4 (tasas AC=AT, AG=CT, CG=GT, frecuencias de bases equivalentes) en Inferencia Bayesiana y Unión de vecinos.

8.2.1. Árbol filogenético

El árbol filogenético consenso se muestra en la Figura 4 y se compone de 4 supergrupos y dos grupos menores: Supergrupo *E. faecium*, *E. faecalis*, *E. casseliflavus* y *E. dispar*, así como los grupos *E. cecorum* y *E. aquimarinus*. Las especies no agrupadas incluyen a *E. phoeniculicola* (parafilética al Supergrupo *E. faecium*), *E. xinjiangensis*, *E. nangangensis* y *E. songbeiensis*.

Supergrupo E. faecium

Representa la agrupación más grande del género. Se divide en los grupos *E. avium-pseudoavium* (*E. avium*, *E. hulanensis*, *E. malodoratus*, *E. gilvus*, *E. raffinosus*, *E. pseudoavium*, *E. devriesei*, *E. pingfangensis*, *E. viikkiensis*, *E. xiangfangensis* y *E. dongliensis*) y el grupo base *E. faecium* (*E. faecium*, *E. lactis*, *E. hirae*, *E. thailandicus*, *E. durans*, *E. ratti*, *E. mundtii*, *E. villorum* y *E. canis*).

Supergrupo E. faecalis

Se compone de los grupos *E. faecalis-E. rivorum* y *E. haemoperoxidus* (*E. haemoperoxidus*, *E. crotali*, *E. rotai*, *E. caccae*, *E. silesiacus*, *E. ureilyticus*, *E. ureasiticus*, *E. termitis*, *E. quebecensis*, *E. moraviensis*, *E. plantarum* y *E. wangshanyuanii*).

Supergrupo E. casseliflavus

Está constituido por las agrupaciones *E. gallinarum*-*E. casseliflavus*, *E. saccharolyticus* (*E. saccharolyticus saccharolyticus*, *E. saccharolyticus taiwanensis* y *E. olivae*), *E. sulfureus*-*E. bulliens* y *E. italicus*-*E. camelliae*.

Supergrupo E. dispar

Se integra de los grupos *E. dispar* (*E. dispar*, *E. canintestini* y *E. saigonensis*), *E. pallens* (*E. pallens*, *E. hermanniensis* y *E. florum*) y *E. asini* (*E. asini* y *E. diestrammenae*).

Grupos E. cecorum y E. aquimarinus

El grupo *E. cecorum* lo constituyen las especies *E. cecorum* y *E. columbae*. El grupo *E. aquimarinus* incluye a *E. aquimarinus*, *E. lemanii*, *E. alcedinis* y *E. eurekaensis*.

Soporte de nodo:
PPB/BS(MV)/BS(NJ)

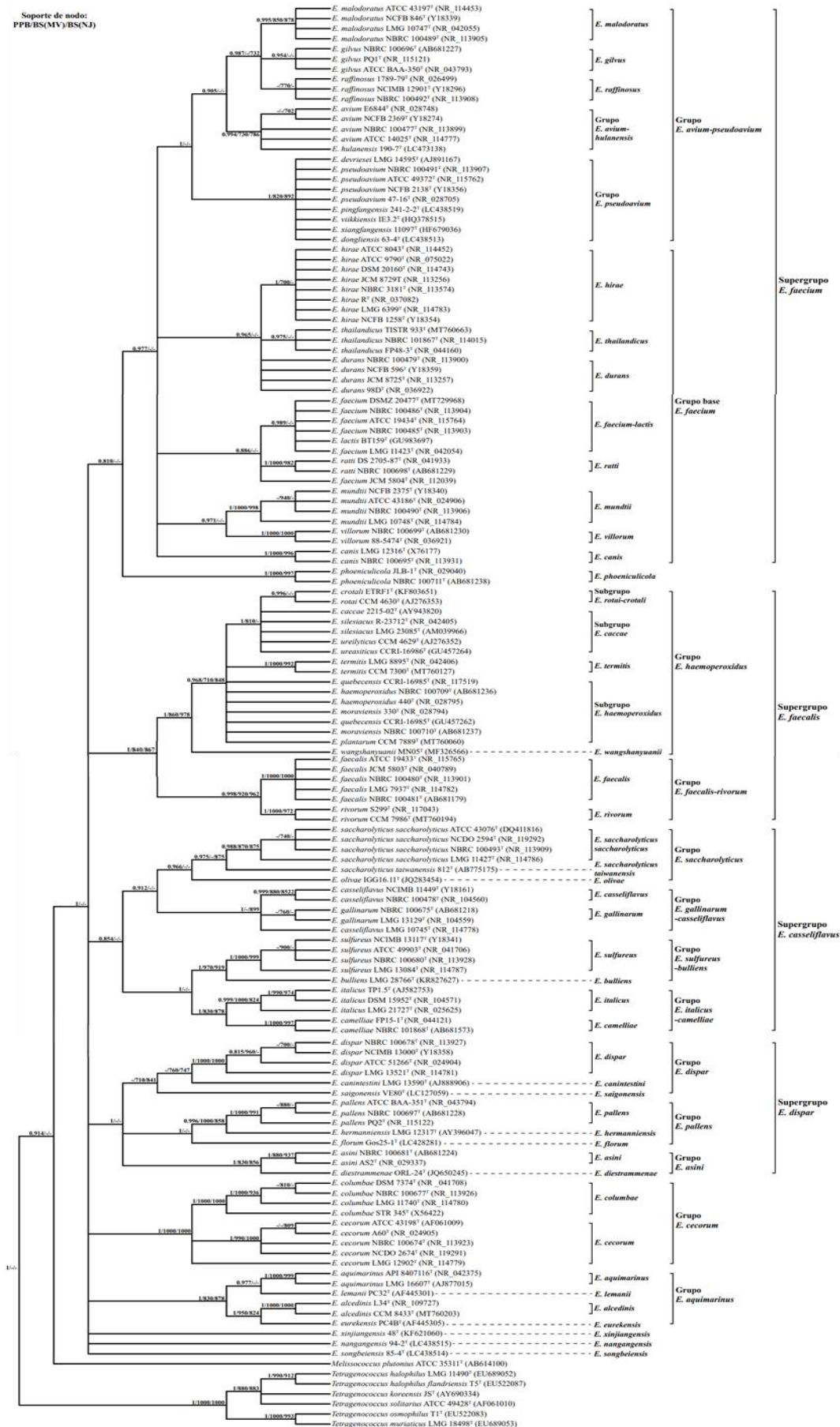


Figura 4. Consenso filogenético de Inferencia Bayesiana, Máxima Verosimilitud y Unión de Vecinos de 1,253 posiciones compartidas del ARNr 16S del género *Enterococcus*. El dendrograma de consenso muestra el soporte de nodo en orden de probabilidad posterior (IB) y bootstraps (ML y NJ; 1,000 réplicas). La topología obtenida mediante inferencia Bayesiana fue utilizada como base para la generación del consenso. Las secuencias de *Melissococcus plutonius* ATCC 35311^T (AB614100), *Tetragenococcus halophilus* LMG 11490^T (EU689052), *Tetragenococcus halophilus flandriensis* T5^T (EU522087), *Tetragenococcus koreensis* JS^T (AY690334), *Tetragenococcus solitarius* ATCC 49428^T (AF061010), *Tetragenococcus osmophilus* T1^T (EU522083) y *Tetragenococcus muriaticus* LMG 18498^T (EU689053) fueron escogidas como grupos de referencia.

8.3 Entropía posicional del gen ARNr 16S

El análisis de entropía para las secuencias alineadas dio como resultado un contenido de información específico de cada posición usando la expresión de entropía para tres conjuntos de secuencias provenientes de las bases de datos SILVA, RDP y Greengenes. Además de un cuarto conjunto que es la concatenación de los conjuntos antes mencionados.

SILVA

El número total de secuencias provenientes de SILVA fueron 713 secuencias del gen ARNr 16S pertenecientes al género *Enterococcus*, el alineamiento final tuvo una longitud de 1,571 posiciones. Los resultados de la determinación de la entropía se muestran en la Figura 5.

De acuerdo con el mapeo, la longitud del gen ARNr 16S fue de 1,521 sitios, el valor de entropía más alta registrada fue 1.03 en la región V2. Un total de 12 posiciones dentro de las regiones V1, V2 y V8, es decir, el 0.78% resultaron con una entropía >0.8 . El 2.038% de las posiciones (treinta y uno) tuvieron una entropía >0.6 ; que se ubicaron en V1 (once posiciones), V2 (siete posiciones), V3 (cuatro posiciones), V5 (una posición), V6 (una posición), V7 (dos posiciones), V8 (cuatro posiciones) y MC7 (una posición). El 97.1% restantes de las posiciones tuvieron una entropía <0.4 .

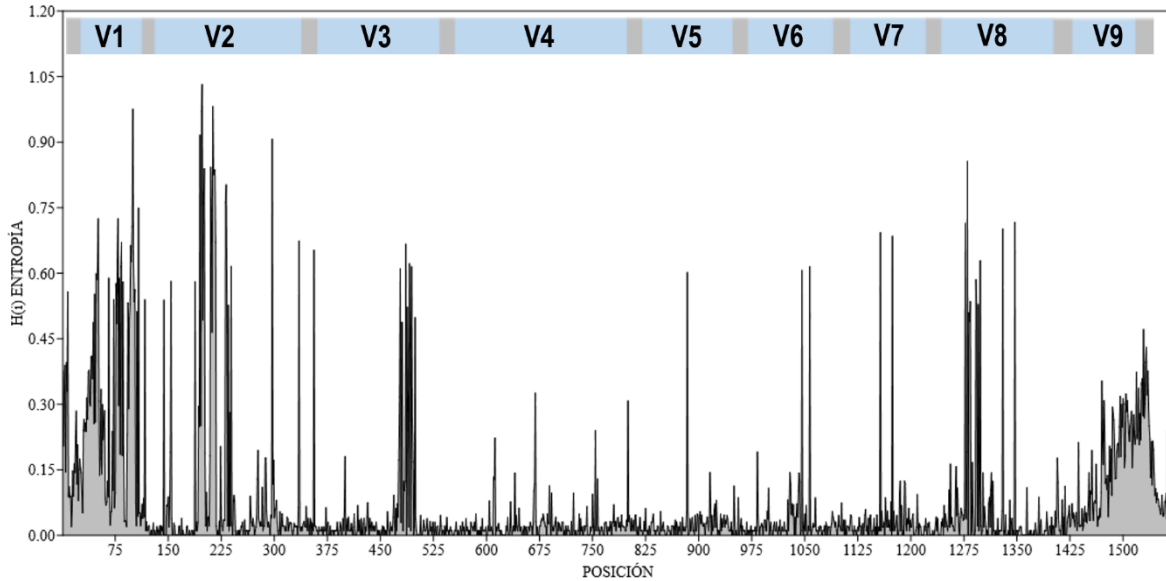


Figura 5. Entropía del conjunto SILVA. La longitud del alineamiento es de 1,571 posiciones; la barra superior enmarca las 9 regiones hipervariables flanqueadas por los marcadores conservados.

Ribosomal Database Project

El número total de secuencias provenientes de RDP fueron 1,819 secuencias del gen ARNr 16S, el alineamiento final tuvo una longitud de 1,534 posiciones. Los resultados de la determinación de la entropía se muestran en la Figura 6.

Con base en el mapeo, la longitud del gen ARNr 16S fue de 1,525 sitios. La región V2 en una posición presentó la entropía más alta registrada con un valor de 1.00. El 0.59% de las posiciones (nueve) tuvieron un valor de entropía >0.8 , V2 (seis posiciones), V1 (una posición), V8 (una posición) y MC10 (una posición). El 3.016% de las posiciones (cuarenta y seis) tuvieron una entropía >0.6 , V1 (diecinueve posiciones), V2 (nueve posiciones), V3 (cuatro posiciones), V5 (una posición), V6 (dos posiciones), V7 (dos posiciones), V8 (cinco posiciones), MC2 (una posición), MC3 (una posición) y MC10 (dos posiciones). El resto de las posiciones que son el 96.39% tuvo una entropía <0.4 .

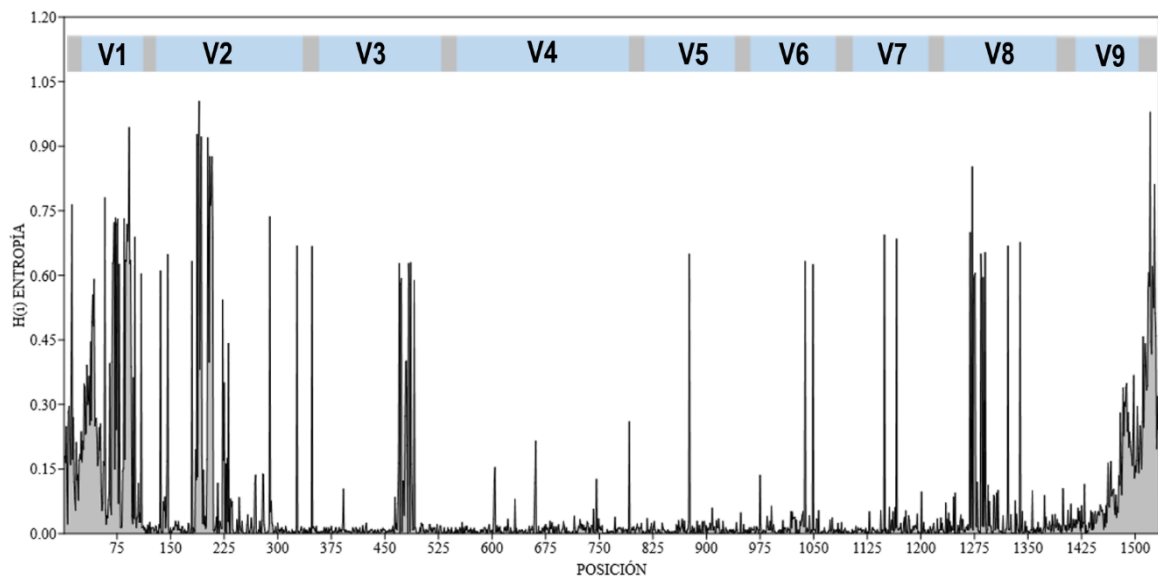


Figura 6. Entropía del conjunto *Ribosomal Database Project* (RDP). La longitud del alineamiento es de 1,534 posiciones, la barra superior enmarca las 9 regiones hipervariables flanqueadas por los marcadores conservados.

Greengenes

El número total de secuencias provenientes de Greengenes fueron 425 secuencias del gen ARNr 16S pertenecientes al género *Enterococcus*, el alineamiento final tuvo una longitud de 1,526 posiciones. Los resultados de la determinación de la entropía se muestran en la Figura 7.

Con base al mapeo, la longitud del gen ARNr 16S fue de 1,526 sitios. El 0.15% de las posiciones tuvieron una entropía >1 , V1 y V2 una posición respectivamente. El 0.65% de las posiciones (diez) en V2 tuvieron una entropía >0.8 . El 2.42% de las posiciones (treinta y siete) tuvieron una entropía >0.6 , V1 (diez posiciones), V2 (seis posiciones), V3 (seis posiciones), V5 (una posición), V6 (dos posiciones), V7 (dos posiciones), V8 (siete posiciones), V9 (una posición) y MC3 (una posición). El 96.78% restantes de las posiciones tuvieron una entropía <0.4 .

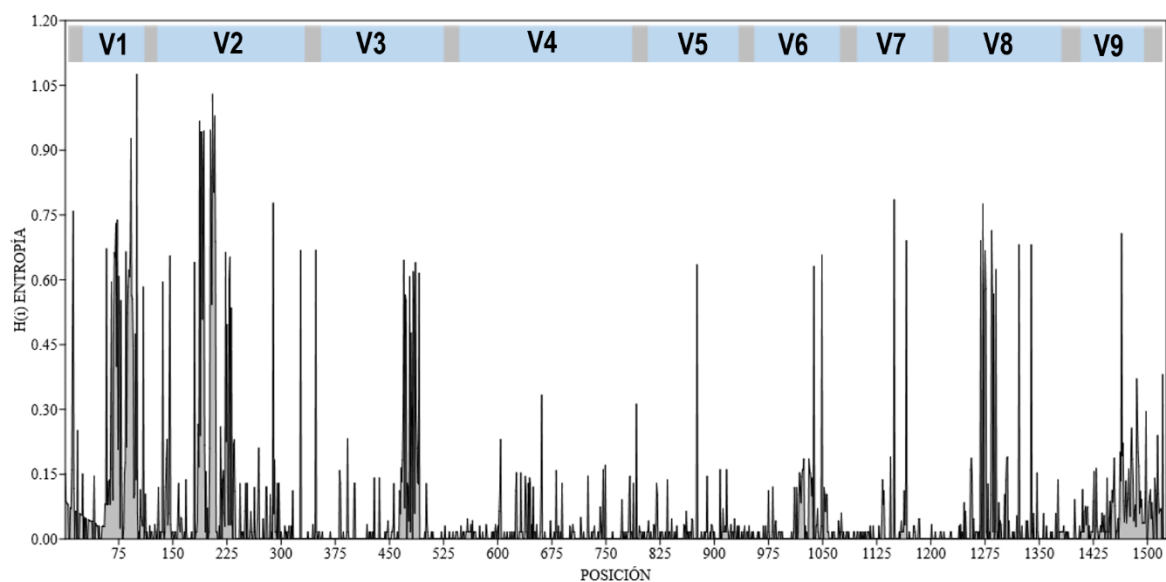


Figura 7. Entropía del conjunto Greengenes. La longitud del alineamiento es de 1,526 posiciones, la barra superior enmarca las 9 regiones hipervariables flanqueadas por los marcadores conservados.

Fusionado

A partir de una fusión de los tres conjuntos de datos SILVA, RDP y Greengenes, con un total de 2,800 secuencias del gen ARNr 16S pertenecientes al género *Enterococcus* y una longitud de alineamiento de 1,526 sitios. Los resultados de la determinación de la entropía se muestran en la Figura 8.

Con base al mapeo, la longitud del gen ARNr 16S fue de 1,513 sitios. Una posición en V2 presento el valor más alto de entropía 1.029. El 0.92% de las posiciones (doce) tuvieron una entropía >0.8 ; V1 (una posición), V2 (once posiciones). El 1.69% de las posiciones (veintidós) tuvieron una entropía >0.6 , V2 (cinco posiciones), V3 (cuatro posiciones), V5 (una posición), V6 (dos posiciones), V7 (dos posiciones), V8 (siete posiciones) y MC3 (una posición). El 98.07% del resto de las posiciones tuvieron una entropía <0.4 .

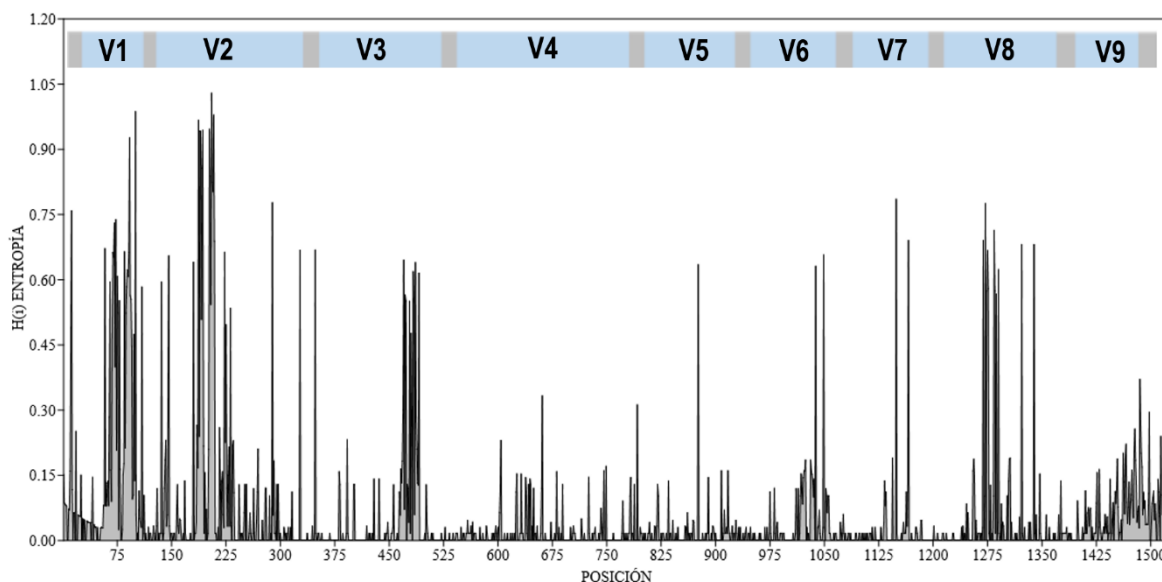


Figura 8. Entropía de los 3 conjuntos fusionados, SILVA, RDP y Greengenes. La longitud del alineamiento es de 1,526 posiciones, la barra superior enmarca las 9 regiones hipervariables flanqueadas por los marcadores conservados.

8.4. Clasificación de sitios mediante K-medias con base en la Entropía

Los valores de entropía por sitio fueron utilizados para el agrupamiento mediante K-medias, dando como resultado tres distintos grupos o categorías: conservado, variante y altamente variante.

SILVA

Los valores de entropía para cada uno de los 1,521 sitios del conjunto SILVA fueron analizados, con base en el algoritmo de K-medias resultando como centroides para cada grupo los siguientes valores de entropía: 0.0232, 0.2308 y 0.6439, que corresponden al grupo conservado, variante y altamente variante, respectivamente (Figura 9). Un total de 1,304 sitios (86%) fueron clasificados como conservados, 137 (9%) como sitios variantes y 80 (5%) como sitios altamente variantes. En el Anexo 13.4 se enlista cada posición, valor de entropía y su clasificación correspondiente.

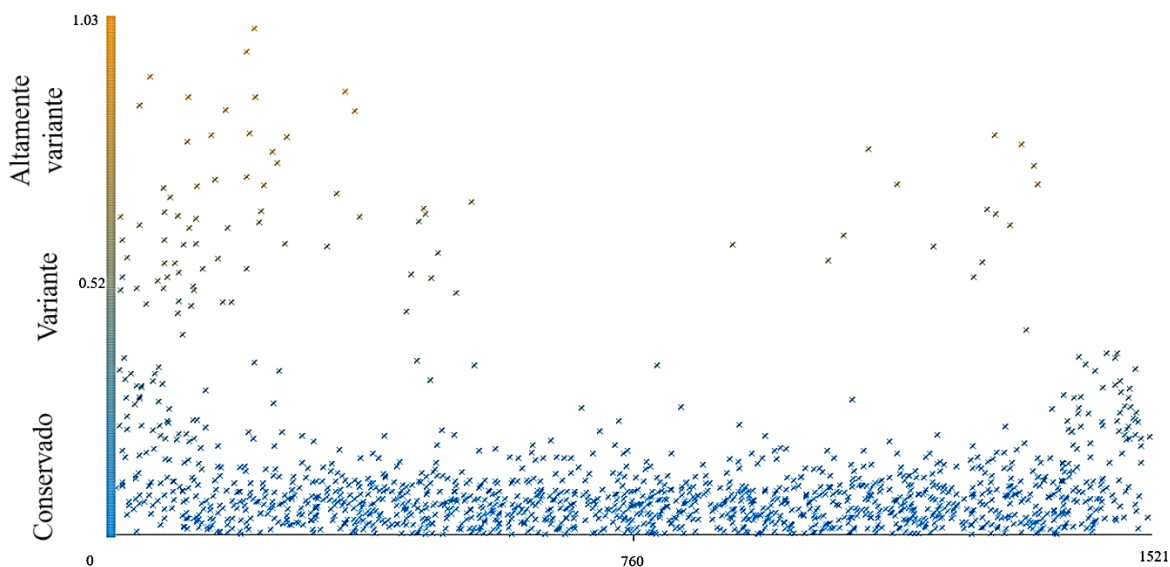


Figura 9. Clasificación de sitios mediante K-medias con base en la Entropía conjunto SILVA. El gráfico muestra el agrupamiento de los sitios: Conservado, Variante y Altamente variante, representado de diferente color, azul, verde y anaranjado, respectivamente.

Ribosomal Database Project

Los valores de entropía para cada uno de los 1,525 sitios del conjunto RDP fueron analizados, se clasificaron con base en el algoritmo K-medias resultando como centroides los valores de entropía: 0.0147, 0.2429 y 0.6847, que corresponden al grupo conservado, variante y altamente variante respectivamente (Figura 10). Un total de 1,331 sitios (87%) fueron clasificados como conservados, 122 (8%) como sitios variantes y 72 (5%) como sitios altamente variantes. En el Anexo 13.5 se enlista cada posición, valor de entropía y su clasificación correspondiente.

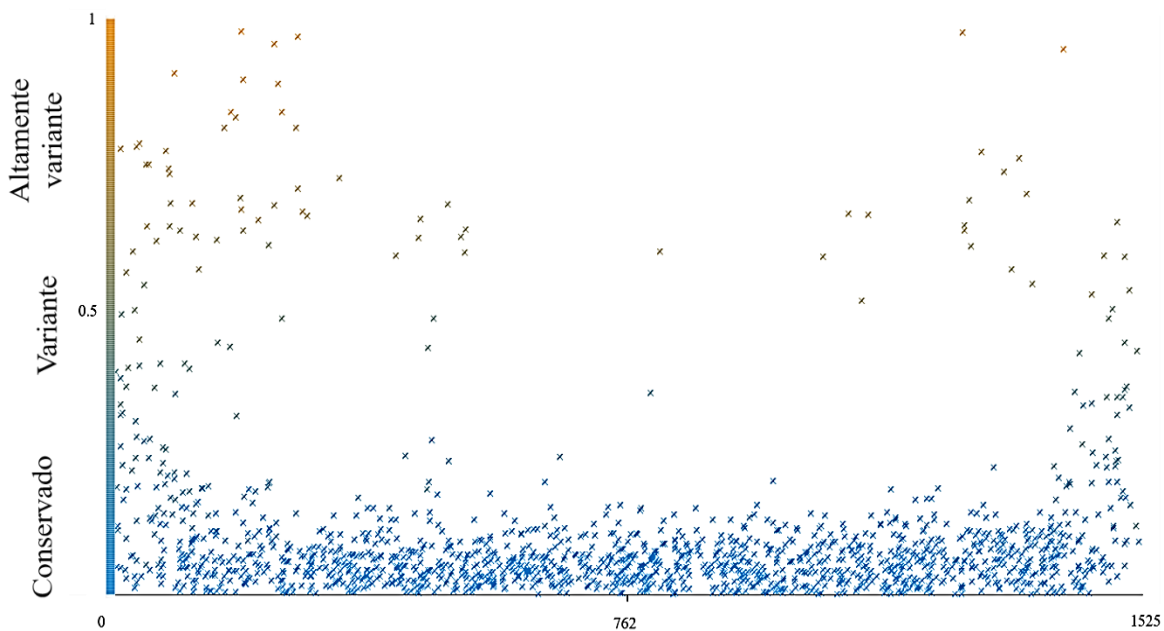


Figura 10. Clasificación de sitios mediante K-medias con base en la Entropía conjunto RDP. El gráfico muestra el agrupamiento de los sitios: Conservado, Variante y Altamente variante, representado de diferente color, azul, verde y anaranjado, respectivamente.

Greengenes

Los valores de entropía para cada uno de los 1,525 sitios del conjunto Greengenes, fueron analizados. Resultaron como centroides para cada grupo los valores: 0.0107, 0.1538 y 0.6713 que corresponden al grupo conservado, variante y altamente variante respectivamente (Figura 11). Un total de 1,259 sitios (82.5%) fueron clasificados como conservados, 191 (12.5%) como sitios variantes y 76 (5%) como sitios altamente variantes. En el Anexo 13.6 se enlista cada posición, valor de entropía y su clasificación correspondiente.

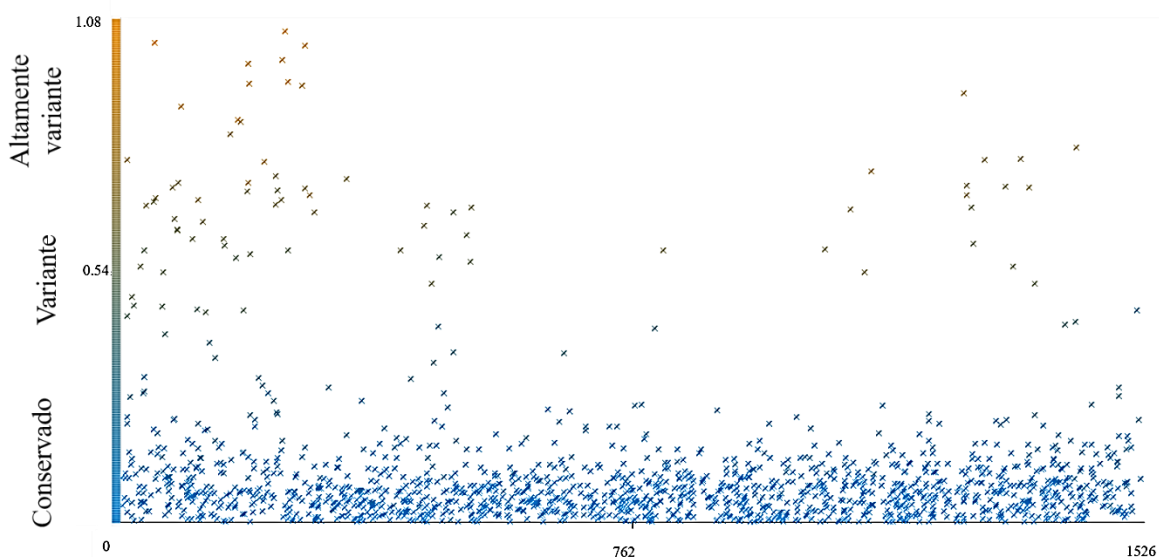


Figura 11. Clasificación de sitios mediante K-medias con base en la Entropía conjunto Greengenes. El gráfico muestra el agrupamiento de los sitios: Conservado, Variante y Altamente variante, representado de diferente color, azul, verde y anaranjado, respectivamente.

Fusionado

Los valores de entropía para cada sitio del conjunto fusionado se clasificaron en tres categorías: conservado, variante y altamente variante (Figura 12). Los valores de entropía fueron analizados en 1,513 sitios con base en el algoritmo de K-medias resultando como centroides de cada grupo: 0.012, 0.1501 y 0.6701, respectivamente. Un total de 1,243 sitios (82%) fueron clasificados como conservados, 197 (13%) como sitios variantes y 73 (5%) como sitios altamente variantes. En el Anexo 13.7 se enlista cada posición, valor de entropía y su clasificación correspondiente.

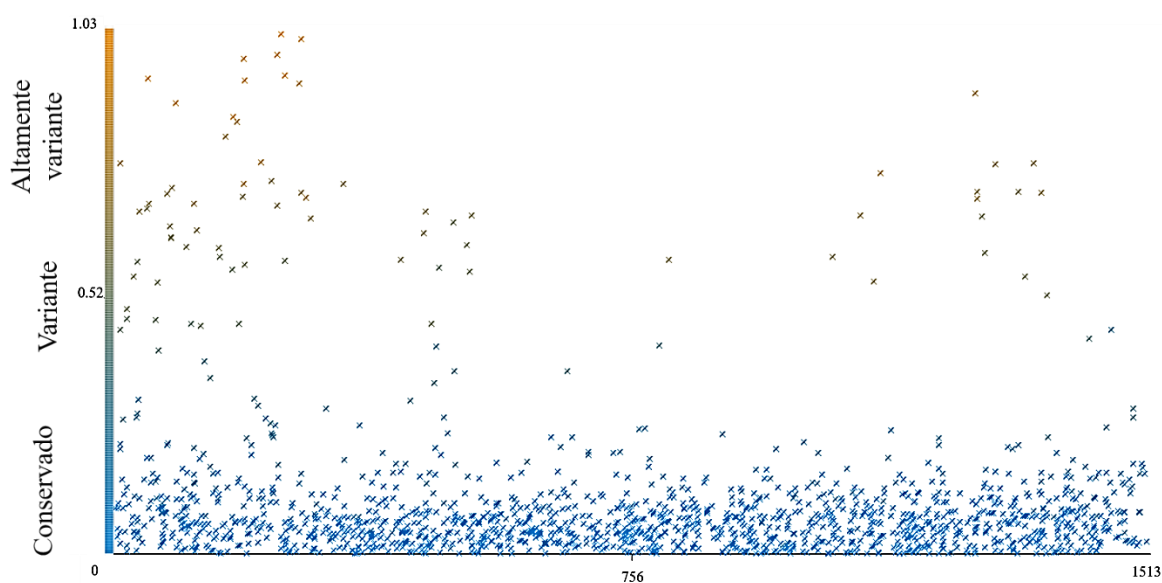


Figura 12. Clasificación de sitios mediante K-medias con base en la Entropía conjunto fusionado. El gráfico muestra el agrupamiento de los sitios: Conservado, Variante y Altamente variante, representado de diferente color, azul, verde y anaranjado, respectivamente.

8.5. Análisis de Componentes principales

SILVA

Dentro de los análisis de componentes principales (ACP), el componente 1 (MC1) explica el 84.92% de la variabilidad de los datos y el componente 2 (V1) explica el 15.04% de los datos, en conjunto explican el 99.96%. Es decir, la variación de estos dos marcadores, explica la varianza total del conjunto de 20 marcadores evaluados (MC1, V1, MC2, V2, MC3, V3, MC4, V4, MC5, V5, MC6, V6, MC7, V7, MC8, V8, MC9, V9, MC10 y gen ARNr 16S) (Figura 13).

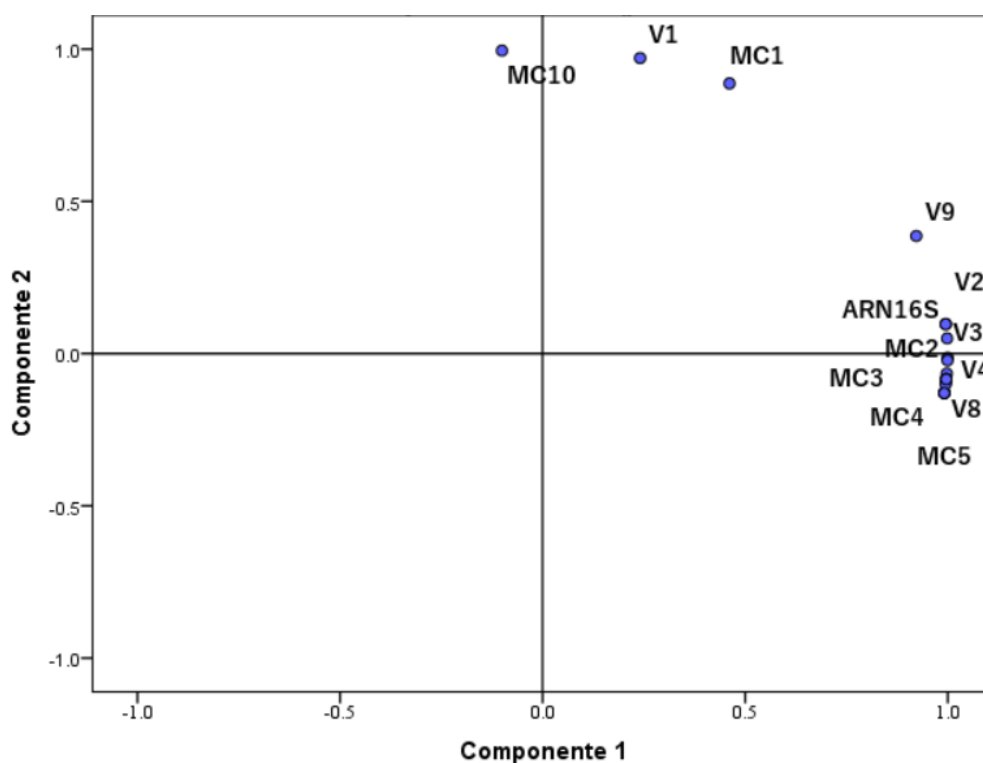


Figura 13. Análisis de componentes principales del conjunto SILVA. El gráfico muestra el resultado del ACP para los 20 marcadores con respecto a los 2 componentes extraídos del conjunto SILVA.

RDP

Con base en el análisis de componentes principales, se describe que el componente 1 (MC1) explica el 84.65% de la variabilidad de los datos, el componente 2 (V1) explica el 15.31%, en conjunto explican el 99.96%. Es decir, la variación de estos dos marcadores, explica la varianza total del conjunto de 20 marcadores evaluados (Figura 14).

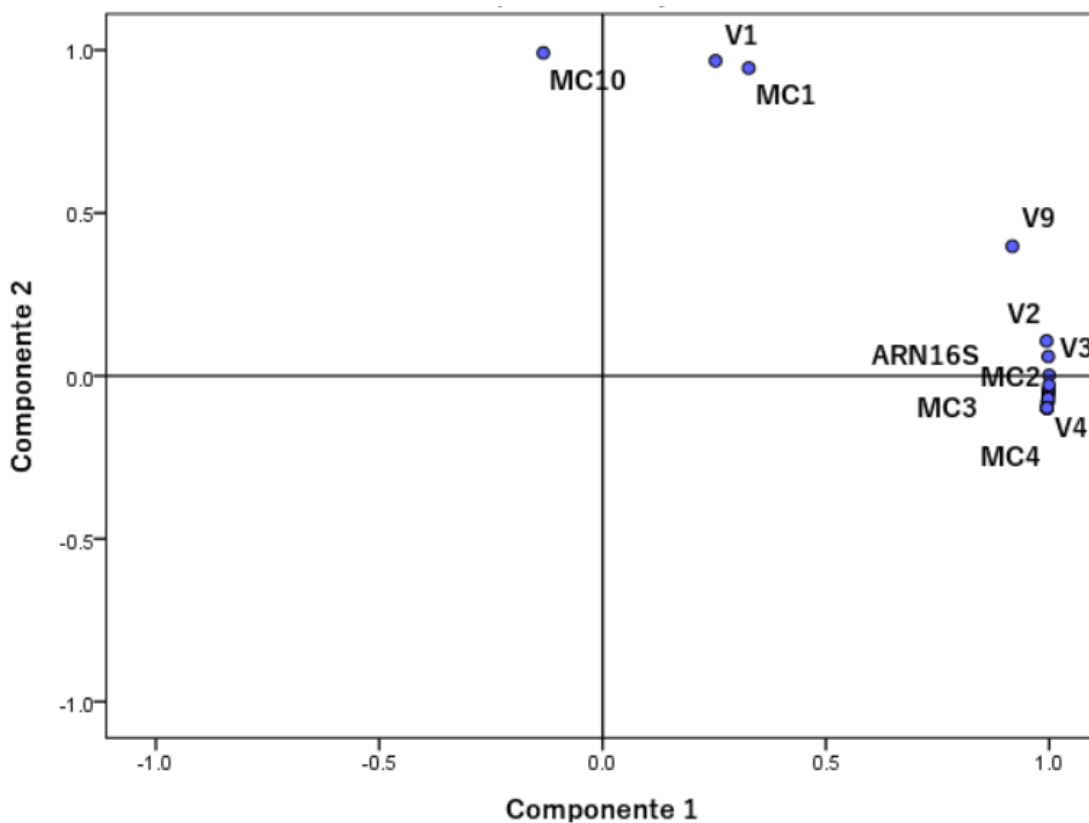


Figura 14. Análisis de componentes principales del conjunto RDP. El gráfico muestra el resultado del ACP para los 20 marcadores con respecto a los 2 componentes extraídos del conjunto RDP.

Fusionado

Los resultados del análisis de componentes principales, demuestran que el componente 1 (MC1) explica el 93.32% de la variabilidad de los datos, mientras que el componente 2 (V1) explica el 6.63%, en total explican el 99.95% de la variabilidad de los datos. Es decir, la variación de estos dos marcadores, explica la varianza total del conjunto de 20 marcadores evaluados (Figura 16).

Además, en la Figura 17, se muestra gráficamente la ubicación de los parámetros explicado por los dos componentes principales extraídos del análisis anterior. El parámetro denominado como número de sitios variantes es cercano al componente 2, con respecto al gráfico de componentes principales (Figura 16) los marcadores V1, MC1, V9, MC10, V2, MC5 y ARN 16S son los más cercanos a este parámetro.

En contraposición está el parámetro denominado como número de sitios conservados con respecto al componente 1, que es donde se acerca al resto de marcadores.

Y en el cuadrante inferior izquierdo se encuentran el resto de los parámetros como media, mediana, proporción de sitios conservados y proporción de sitios variantes.

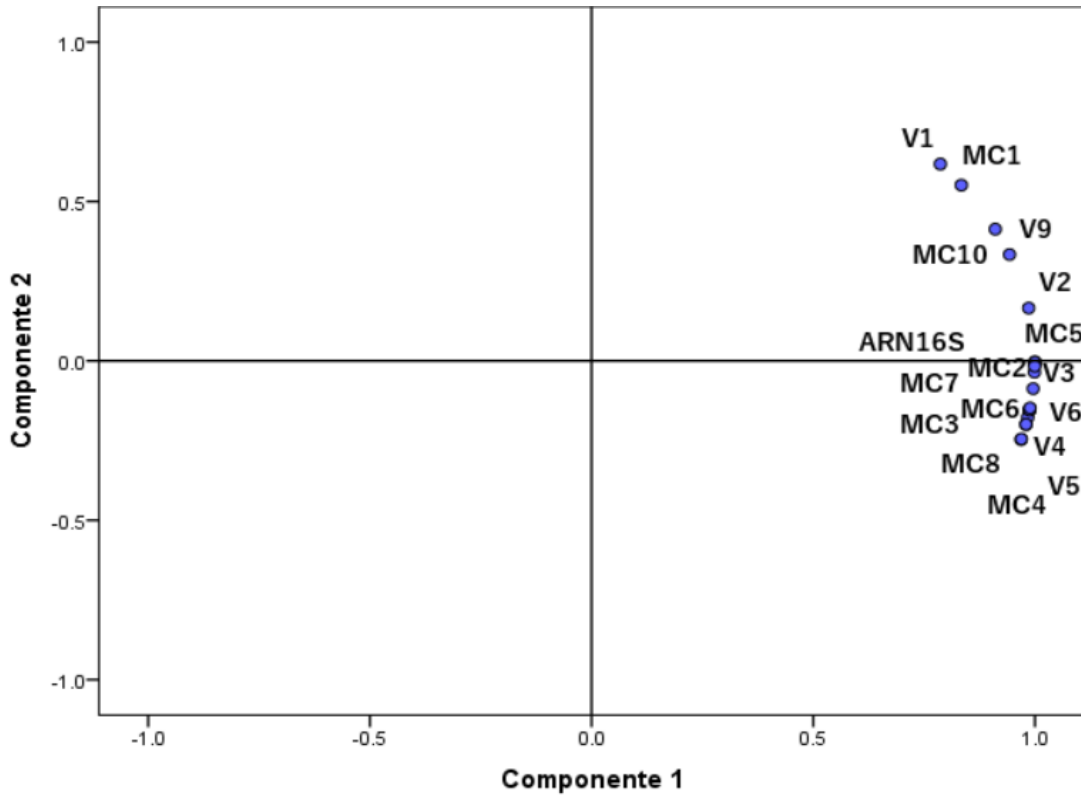


Figura 16. Análisis de componentes principales del conjunto Fusionado. El gráfico muestra el resultado del ACP para los 20 marcadores con respecto a los 2 componentes extraídos del conjunto Fusionado.

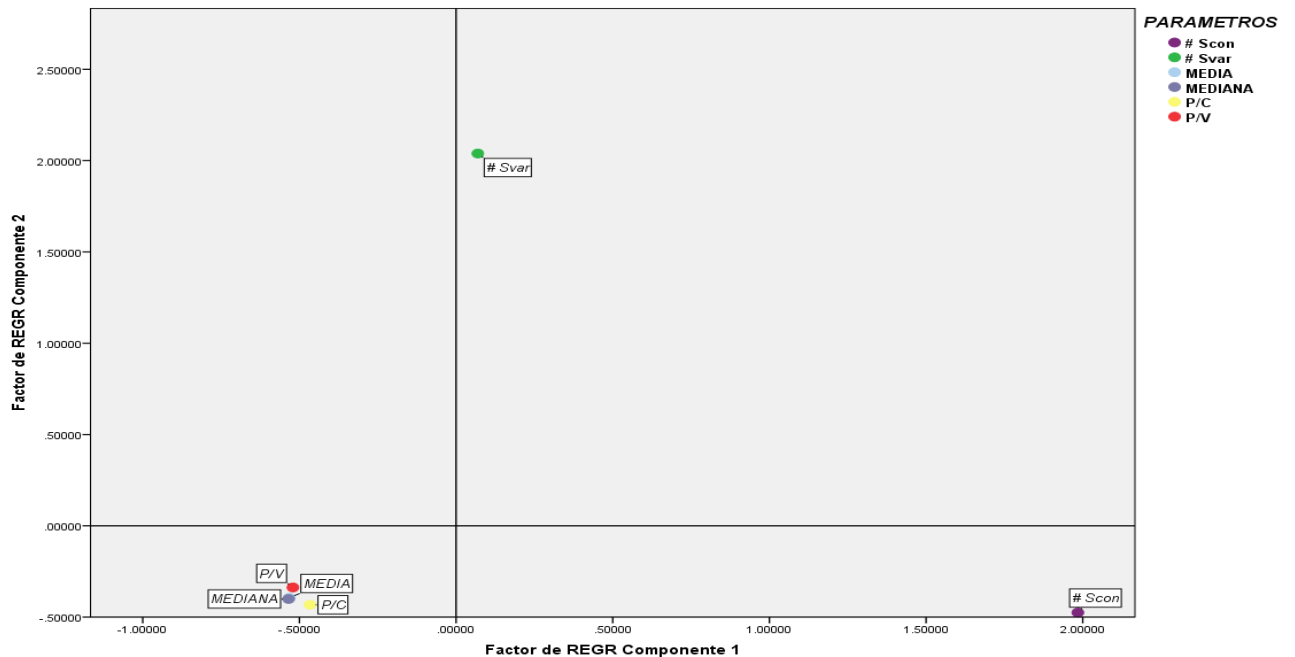


Figura 17. Análisis de Regresión de los dos componentes principales del conjunto Fusionado. El gráfico muestra el resultado del ACP de los parámetros respecto a los 2 componentes extraídos del conjunto Fusionado. #Svar= número de sitios variantes, #Scon= número de sitios conservados, P/V= proporción de sitios variantes, P/C= proporción de sitios conservados, media y mediana.

8.6. Correlación y agrupamiento jerárquico de los marcadores conservados y variables

El cálculo de los parámetros de variación (entropía media y mediana, la proporción de sitios conservados, la proporción de sitios variantes, el número de sitios variantes y el número de sitios conservados) para un total de 20 marcadores, resultó en una matriz de datos, a partir de la cual se le realizó la prueba de correlación de Spearman que llevó a la construcción de un dendrograma y mapa de calor para cada conjunto respectivamente.

El grado de correlación entre los marcadores conservados/variables y el gen 16S completo para los cuatro conjuntos (SILVA, RDP, Greengenes y Fusionado) fue obtenida utilizando la correlación de Spearman, estos resultados se muestran en un mapa de calor y debido al agrupamiento jerárquico se logró identificar a los marcadores similares y distantes al ARN 16S

SILVA

En la Figura 18 se muestra el dendrograma, el cual presentó una longitud de rama de 1.68 constituido por 2 ramas principales donde se agrupan a los marcadores diferentes y los similares y en la parte superior se encontró al MC10 como parafilético que obtuvo la correlación más baja respecto al 16S (-0.14) indicando una correlación negativa débil. Seguido por un clado integrado por los 4 marcadores conservados (20%) (MC8, MC6, MC5 y MC4), con una correlación de 0.32, indicando una correlación positiva débil.

Dentro de la rama de los marcadores similares al gen ARNr 16S, aparecen agrupados 2 marcadores (10%) MC1 y V1 con una correlación de 0.89 y 0.6, respectivamente. Otros 2 marcadores V7 y V5 están en el clado más cercano a los similares al gen ARNr 16S, con una correlación de 0.83 y 0.85, respectivamente.

En el clado de los similares al gen ARNr 16S, se integran 9 marcadores (45%) V9, MC9, V8, MC7, V4, V3, V2, MC2 con una correlación de 1 la cual se denomina altamente positiva. MC3 resultó parafilético a dicho clado con una correlación de 0.94.

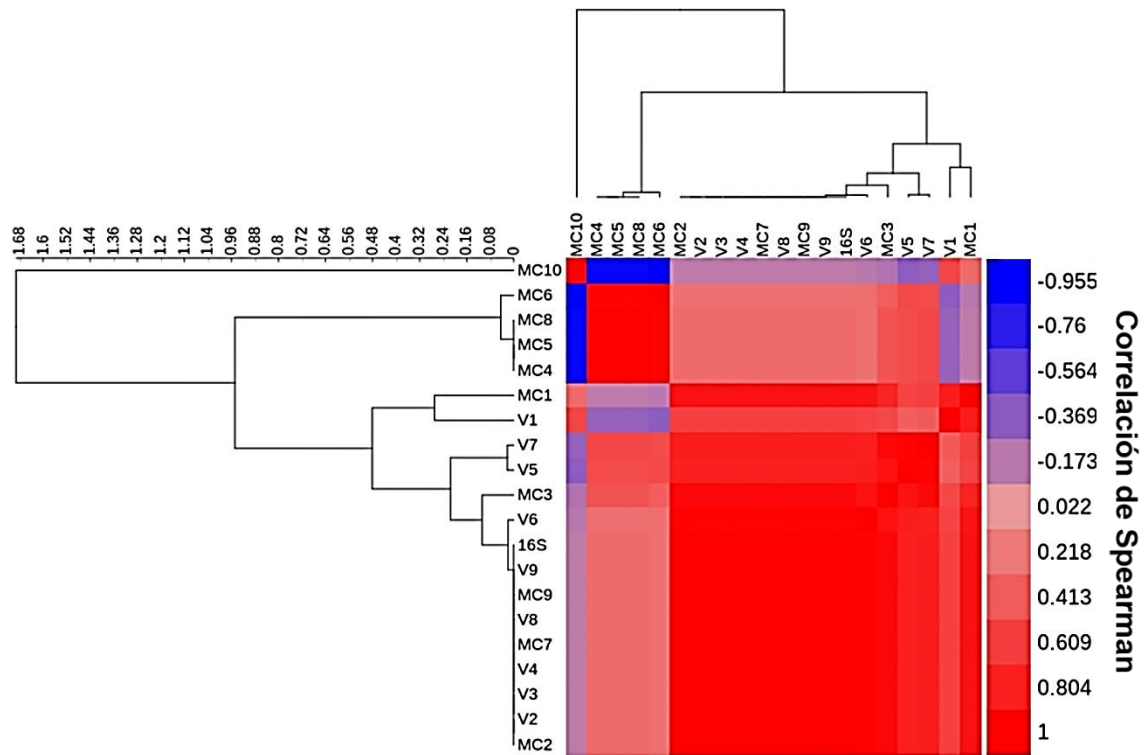


Figura 18. Agrupamiento jerárquico mediante UPGMA y mapa de calor del Conjunto SILVA. El gráfico muestra el agrupamiento jerárquico de los marcadores conservados/variables y el gen ARNr 16S completo por el método UPGMA y mapa de calor que muestra el gradiente de variación de la correlación de Spearman, del lado derecho la barra indica el grado de variación de color azul a rojo (-0.955 a 1).

RDP

El dendrograma resultante para el conjunto RDP presento una longitud de rama de 4.8, está constituido por 2 ramas principales (Figura 19). En la rama superior se ubican a los marcadores que resultaron ser diferentes al gen ARNr 16S completo, se encontró al MC10 que resultó parafilético que obtuvo la correlación más baja respecto al gen ARNr 16S (-0.14) indicando una correlación negativa débil, seguido por un clado integrado de 5 (25%) marcadores conservados (MC9, MC8, MC7, MC6 y MC4), los cuales obtuvieron una correlación de 0.32, lo que indica una correlación positiva débil.

En la rama de los marcadores similares, aparecen 2 marcadores (10%), MC1 y V1 agrupados dentro del mismo clado con una correlación 0.89 y 0.77, respectivamente.

El marcador V5 resultó parafilético al clado de los similares al gen ARNr 16S, con una correlación altamente positiva de 0.83. Seguido por los marcadores V7 y V4, en donde ambos obtuvieron una correlación altamente positiva de 0.94, pero no fue suficiente para agruparse dentro del clado de los marcadores similares al gen ARNr 16S, V9, V8, V6, MC5, V3, MC3, V2, MC2; un total de 10 marcadores (50%) con una correlación altamente positiva de 1 de los marcadores similares al gen ARNr 16S.

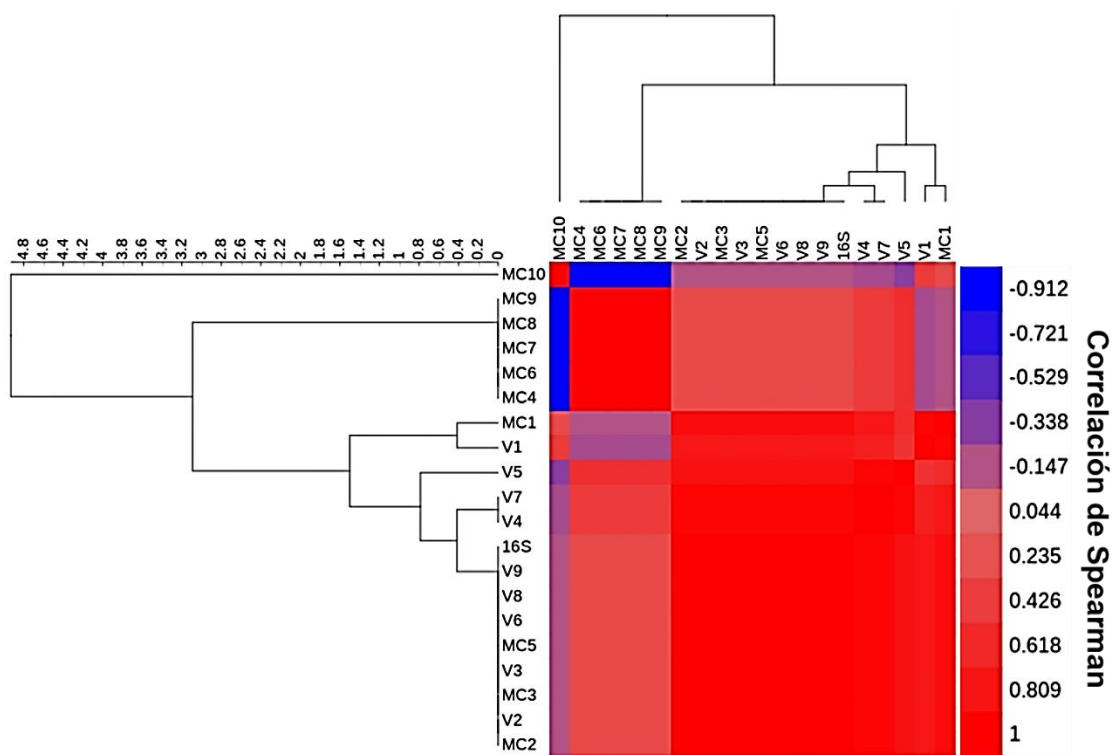


Figura 19. Agrupamiento jerárquico mediante UPGMA y mapa de calor del Conjunto RDP. El gráfico muestra el agrupamiento jerárquico de los marcadores conservados/variables y el gen ARNr 16S completo por el método UPGMA y mapa de calor que muestra el gradiente de variación de la correlación de Spearman, del lado derecho la barra indica el grado de variación de color azul a rojo (-0.912 a 1).

Greengenes

El dendrograma del conjunto Greengenes presentó una longitud de rama de 2.4, y está constituido por 2 ramas principales (Figura 20).

En la rama superior se ubican a los marcadores que fueron diferentes al gen ARNr 16S completo, se encontró al MC8 que resultó parafilético que obtuvo la correlación más baja respecto al gen ARNr 16S (0.29), seguido por un clado integrado 3 marcadores conservados (MC7, MC6 y MC4), los cuales obtuvieron una correlación de 0.52.

Dentro de la rama de los similares al gen ARNr 16S, en el primer clado se ubicaron los marcadores MC9 y MC2 con una correlación de 0.99 (10%) y en el clado más grande se encontraron el 70% (14/20) de los marcadores con una correlación de 1, denominándola como altamente positiva con respecto al gen ARNr 16S, incluyendo al MC10, V9, V8, V7, V6, V5, MC5, V4, V3, MC3, V2, MC1 y V1.

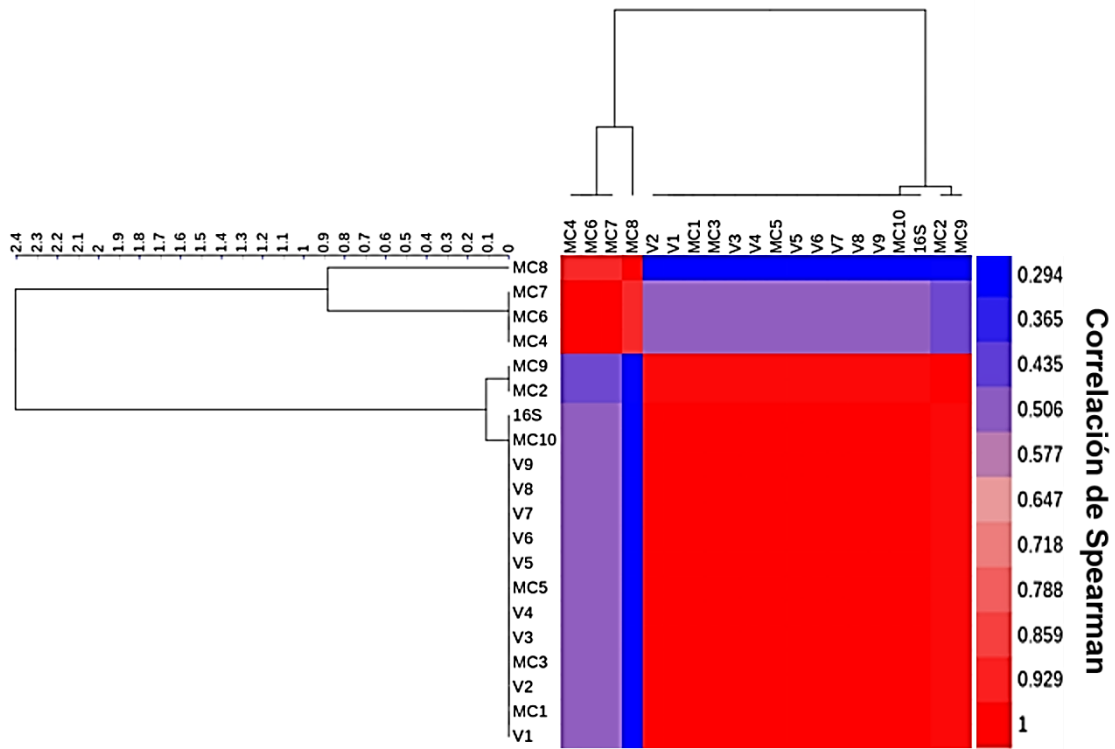


Figura 20. Agrupamiento jerárquico mediante UPGMA y mapa de calor del Conjunto Greengenes. El gráfico muestra el agrupamiento jerárquico de los marcadores conservados/variables y el gen ARNr 16S completo por el método UPGMA y mapa de calor que muestra el gradiente de variación de la correlación de Spearman, del lado derecho la barra indica el grado de variación de color azul a rojo (0.294 a 1).

Fusionado

El dendrograma resultante para el conjunto fusionado una longitud de rama de 2.2 y está constituido por 2 ramas principales que separan a los marcadores diferentes y similares al gen ARNr 16S (Figura 21). En la rama superior se encontró al MC8 y MC4 (2/20), con una correlación de 0.52 respecto al gen ARNr 16S.

Después aparece V1 como parafilético al clado de los marcadores similares, ya que obtuvo una correlación de 0.89, que se considera como positiva muy fuerte pero que no fue suficiente para entrar en el clado de los similares con una correlación de 1, que engloban al 85% (17/20) de los marcadores del gen ARNr 16S, MC10, V9, MC9, V8, V7, MC7, V6, MC6, V5, MC5, V4, V3, MC3, V2, MC1 y MC2.

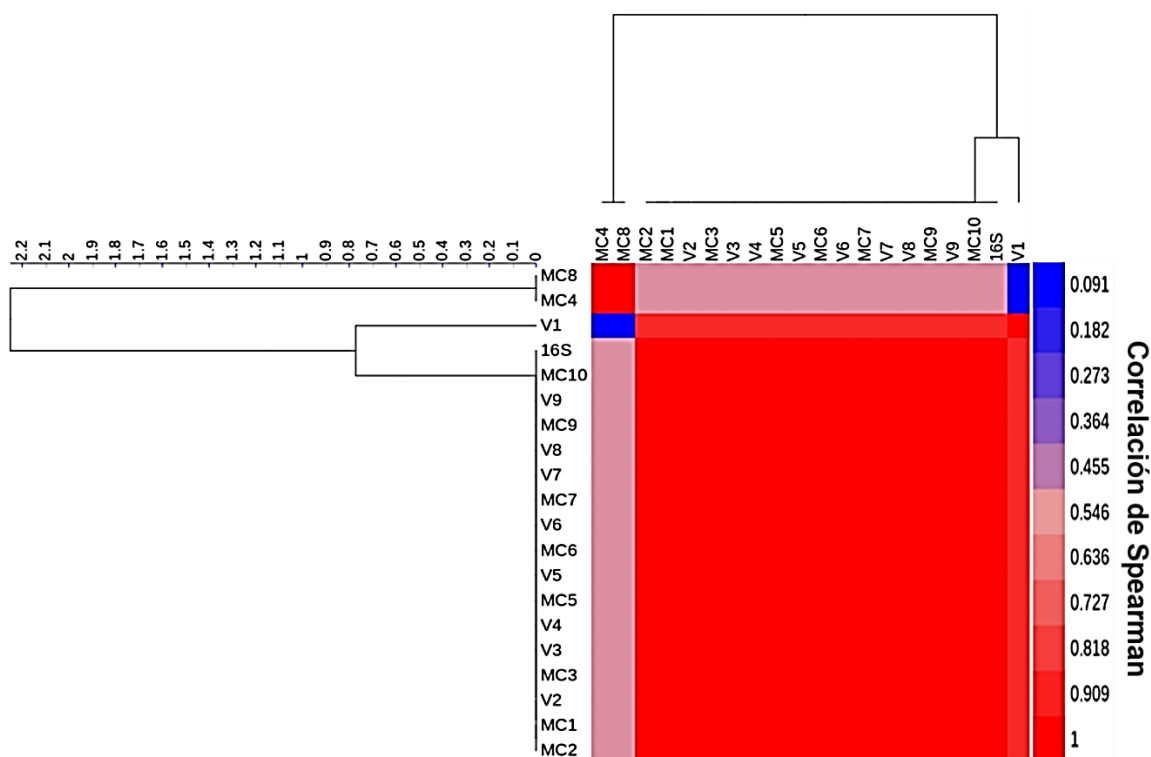


Figura 21. Agrupamiento jerárquico mediante UPGMA y mapa de calor del Conjunto Fusionado. El gráfico muestra el agrupamiento jerárquico de los marcadores conservados/variables y el gen ARNr 16S completo por el método UPGMA y mapa de calor que muestra el gradiente de variación de la correlación de Spearman, del lado derecho la barra indica el grado de variación de color azul a rojo (0.091 a 1).

8.7. Evaluación de marcadores mediante resolución filogenética

Gen ARNr 16S

Un total de 2,800 secuencias y una longitud 1,528 posiciones comprendieron el conjunto del gen ARNr 16S, el cual fue simulado bajo el modelo TNe+I+R7 (Del inglés Tamura-Nei Equal base Frequencies con sitios invariantes) más 7 tasas de heterogeneidad a través de los sitios y a partir de este conjunto se obtuvo el árbol filogenético del gen ARNr 16S completo realizado mediante el método de Máxima verosimilitud (ML), fue inferido bajo el modelo TIM3e+I+G4 (Transition Model Equal base frequencies) Modelo de Transición con frecuencia de tasas iguales más proporción de sitios invariantes más 4 tasas Gamma, siendo seleccionado de acuerdo con el puntaje del Criterio de Información Bayesiano (CIB) (Anexo 10.8). Dentro del cual se obtuvieron 31 clados con un soporte de nodo >70% bootstraps (Figura 22).

El 80.64% de clados fueron monofiléticos *E. wangshanyuanii*, *E. cecorum*, *E. columbae*, *E. bulliens*, *E. lemanii*, *E. hiraе*, *E. pallens*, *E. rivorum*, *E. hermanniensis*, *E. canis* y *E. rattii*, el 44% de estos agrupamientos obtuvieron una sensibilidad del 100%. El 66% de los clados formados por especies obtuvieron los siguientes valores de sensibilidad *E. avium* (97.56%), *E. faecalis* (93.7%), *E. sulfureus* (92.86%) *E. mundtii* (92.31%) *E. eurekaensis* (71.43%) *E. dispar* (88.89%), *E. phoeniculicola* (83.33%), *E. thailandicus* (55%), *E. malodoratus* (30%), *E. gilvus* (20%) y *E. faecium* (15.9%). Y debido a la similitud entre parejas de especies se formaron 4 clados *E. gallinarum-casseliflavus*, *E. italicus-camelliae*, *E. saccharolyticus-olivae* y *E. asini-diestrammenae*, cuyo valor de sensibilidad fue de 100% para cada una de las parejas.

Se formaron 2 clados que agruparon a dos grandes grupos de especies; el grupo *haemoperoxidus* (*E. haemoperoxidus*, *E. crotali*, *E. rotai*, *E. caccae*, *E. silesiacus*, *E.*

ureilyticus, *E. ureasiticus*, *E. quebecensis*, *E. moraviensis*, *E. plantarum*) y el grupo *pseudoavium* (*E. pseudoavium*, *E. devriesei*, *E. pingfangensis*, *E. viikkiensis*, *E. xiangfangensis* y *E. dongliensis*).

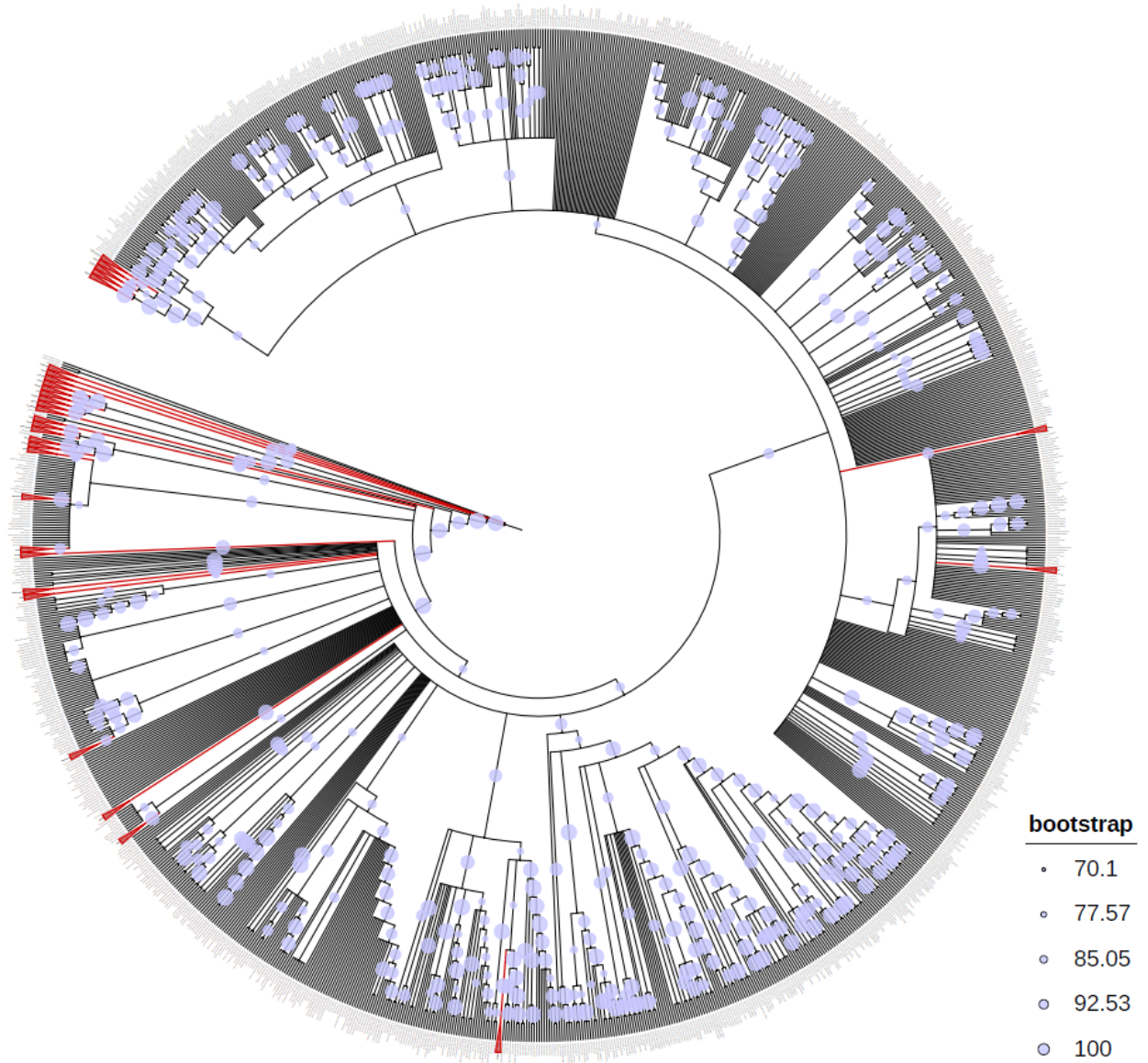


Figura 22. Árbol filogenético por ML ARNr 16S completo. Se muestra una representación circular del árbol filogenética del ARNr 16S por ML con valores de Bootstrap >70%. De color rojo están marcados los clados descritos.

Marcadores similares

Un total de 2,800 secuencias y una longitud 648 posiciones fueron analizadas en el conjunto de marcadores similares, simulado bajo el modelo TPM3+R5 (frecuencias de tasas iguales) más 5 tasas libres y a partir de este conjunto se obtuvo el árbol filogenético de los marcadores similares (V2, V3, V8 y V9) realizado mediante el método de Máxima verosimilitud (ML), fue inferido bajo el modelo TIM3e+G4 (Modelo de Transición con tasas iguales más 4 tasas Gamma), siendo seleccionado de acuerdo con el puntaje del CIB (Anexo 10.9). Dentro del cual se obtuvieron 30 clados con un soporte de nodo >70% bootstraps (Figura 23).

El 96.66% de clados agruparon a una sola especie; *E. wangshanyuanii*, *E. cecorum*, *E. camelliae*, *E. lemanii*, *E. mundtii*, *E. rivorum*, *E. eurekensis*, *E. pallens*, *E. hermanniensis*, *E. asini*, *E. diestrammenae*, *E. phoeniculicola*, *E. ratti*, *E. olivae*, *E. villorum*, *E. termitis*, *E. plantarum*, *E. columbae* y *E. canis* estos agrupamientos obtuvieron una sensibilidad del 100% (65.51%).

El 34.49% de los clados formados por especies obtuvieron los siguientes valores de sensibilidad *E. hirae* (99.55%), *E. avium* (97.56%), *E. faecalis* (95.89%), *E. sulfureus* (92.86%), *E. italicus* (85.71%), *E. gallinarum* (83.5%), *E. devriesei* (66.67%), *E. caccae* (50%), *E. durans* (43.49%) y *E. faecium* (39.51%). Y debido a la similitud entre dos especies se creó el clado *E. dispar-canintestini*, cuyo valor de sensibilidad fue de 90.91%.

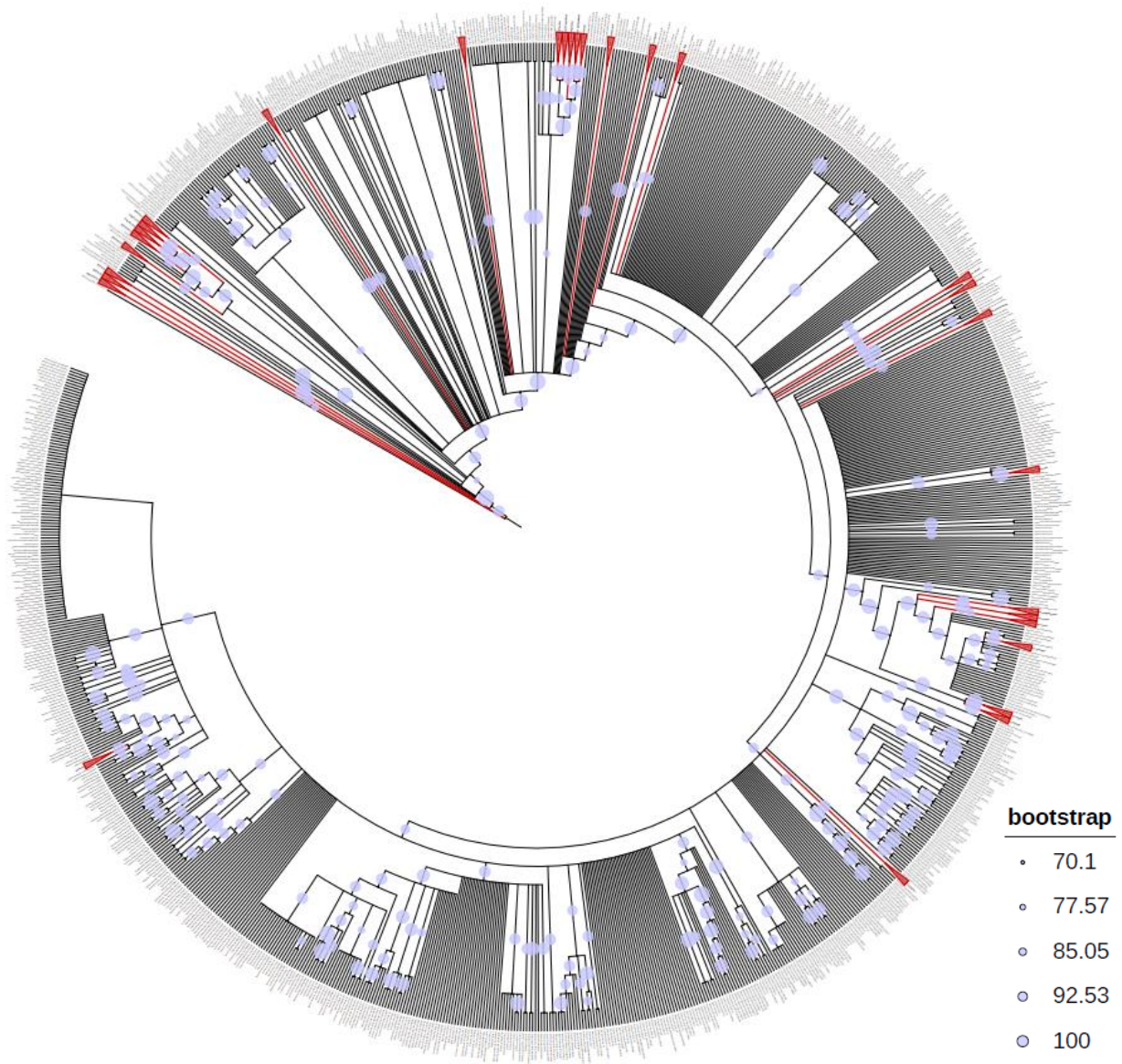


Figura 23. Árbol filogenético por ML marcadores similares. Se muestra una representación circular del árbol filogenética de los marcadores similares al ARNr 16S por ML con valores de Bootstrap >70%. De color rojo están marcados los clados descritos.

Marcadores disimilares

Un total de 2,800 secuencias y una longitud 881 posiciones fueron analizadas en el conjunto de simulado bajo el modelo K2P+I+R6 (Kimura 2 parámetros) más 6 tasas de heterogeneidad a través de los sitios y a partir de cual se construyó el árbol filogenético de los marcadores disimilares (MC1, V1, MC2, MC3, MC4, V4, MC5, V5, MC6, V6, MC7, V7, MC8, MC9 y MC10), realizado mediante Máxima verosimilitud (ML), inferido bajo el modelo TNe+I+G4 (Tamura y Nei) más proporción de sitios invariantes más 4 tasas Gamma, siendo seleccionado de acuerdo con el puntaje del CIB (Anexo 10.10). Dentro del cual se obtuvieron 17 clados con un soporte de nodo >70% bootstraps (Figura 24).

Los clados que resultaron ser monofiléticos se denominaron como el nombre de la especie por el que está integrado, se describen los clados *E. phoeniculicola*, *E. diestrammenae*, *E. lemanii*, *E. eurekaensis*, *E. wangshanyuanii*, *E. hermanniensis* y los clados que están formado por parejas de especies que debido a la similitud entre ellas no se pudieron separar tenemos a *E. cecorum-columbae*, *E. saccharolyticus-olivae*, *E. italicus-camelliae* y *E. gallinarum-casseliflavus*. El 58.8% de los clados se agruparon con una sensibilidad del 100%.

El 41.17% de los clados obtuvieron los siguientes porcentajes de sensibilidad *E. faecalis* (95.47%), *E. villorum* (71.43%), *E. mundtii* (37.84%), *E. asini* (37.5%), *E. durans* (8.89%) y *E. faecium* (3.7%). Además de la pareja de *E. sulfureus-bulliensi* (93.75%).

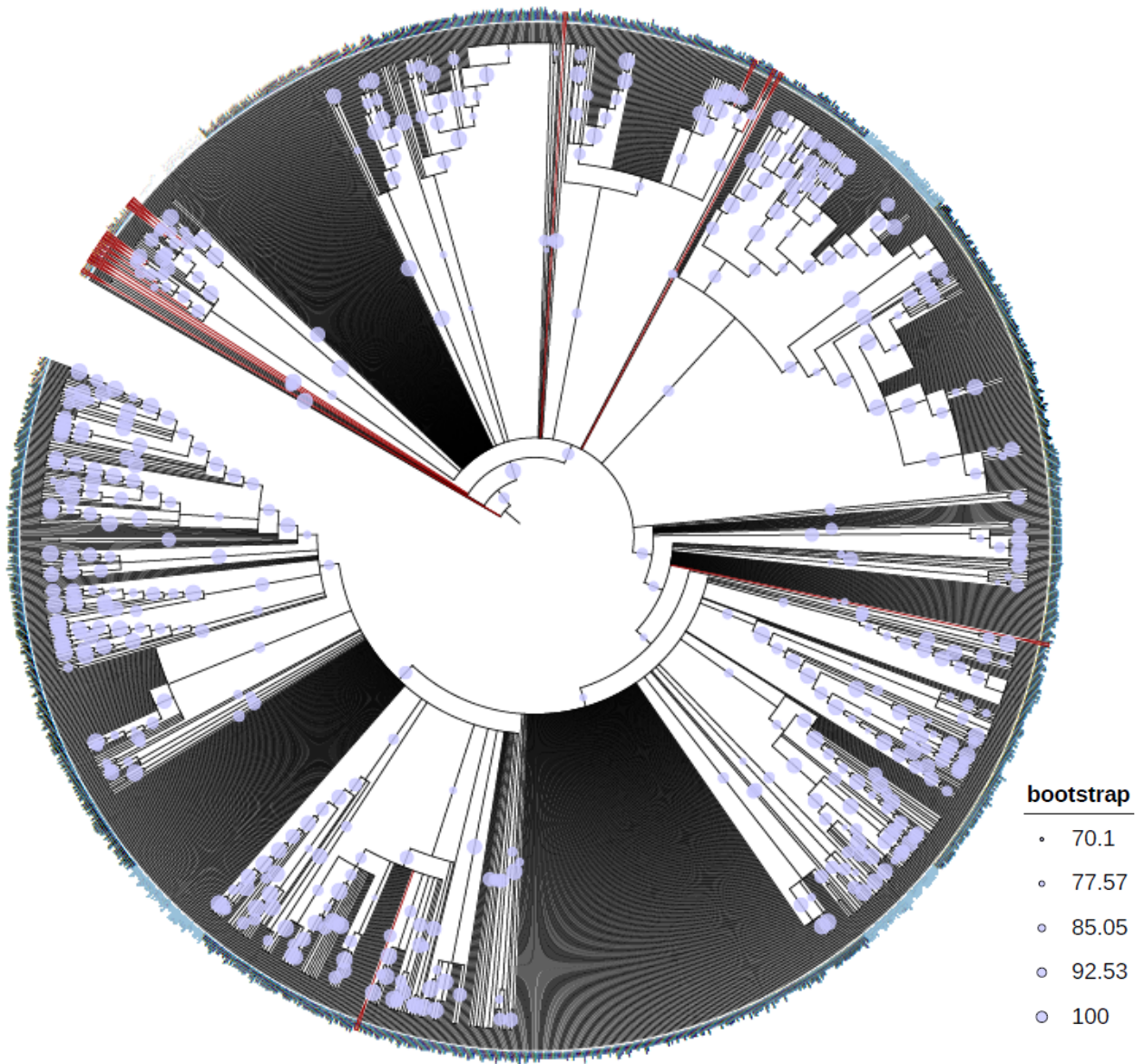


Figura 24. Árbol filogenético por ML marcadores disimilares. Se muestra una representación circular del árbol filogenética de los marcadores disimilares al gen ARNr 16S por ML con valores de Bootstrap >70%. De color rojo están marcados los clados descritos.

8.8 Oligotipificación

8.8.1. Estimación de la Entropía de Shannon del conjunto oligotipado

Con base en el análisis de entropía de Shannon, del conjunto oligotipado (60 especies) 60 secuencias de una longitud de 1,249 que comprende desde V2 a V8 se determinó que la región con mayor entropía correspondía a la región V2 (Figura 25-panel a), seguida de la región V8 (Figura 25-panel b).

La oligotipificación mostró un total de 13 componentes/posiciones que debido al valor de entropía aportaban el mayor grado de información discriminatoria entre especies, las cuales están dentro de la región hipervariable V2.

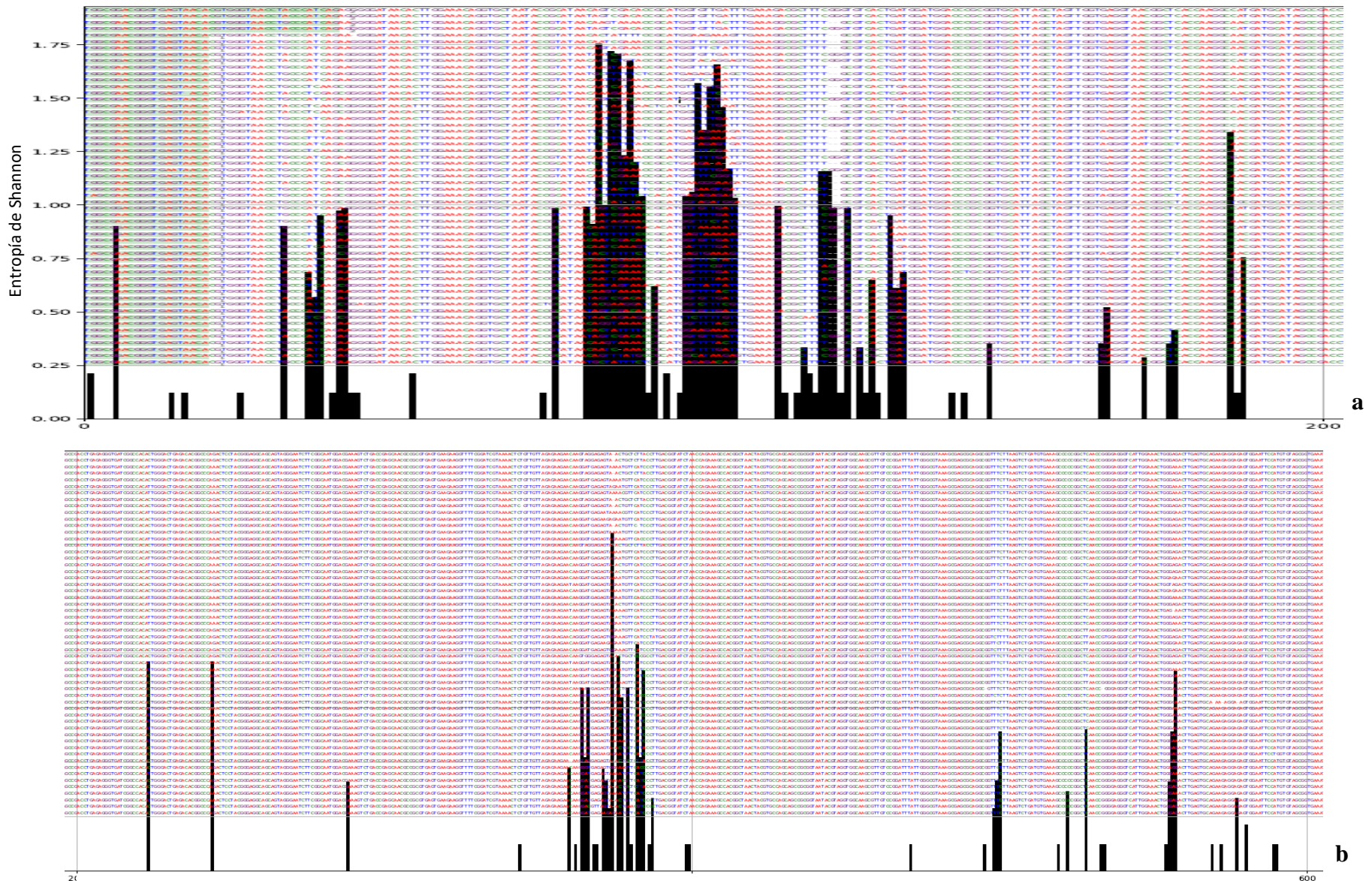


Figura 25- panel a y b. Entropía de Shannon de conjunto oligotipado. El gráfico representa el análisis de entropía de Shannon del alineamiento múltiple de secuencias del conjunto taxonómico de 60 especies con una longitud de 1,252 posiciones.

8.8.1.2. Oligotipos

Se obtuvieron un total de 42 oligotipos con una longitud de 13 nucleótidos, 33 fueron específicos por especie con una pureza de 1 que se enlistan en el cuadro 5. Además 3 oligotipos fueron específicos para 3 grupos de especies. El grupo uno denominado como Haemoperoxidus que incluye a 6 especies (*E. haemoperoxidus*, *E. moraviensis*, *E. caccae*, *E. quebecensis*, *E. silesiacus* y *E. ureasiticus*). Seguido por el grupo Durans conformado por 5 especies (*E. durans*, *E. hiraе*, *E. lactis*, *E. thailandicus* y *E. xinjiangensis*), y finalmente el grupo Pseudoavium que está integrado por 4 especies (*E. devriesei*, *E. pseudoavium*, *E. viikkiensis* y *E. xiangfangensis*) (Cuadro 5).

También resultaron 6 parejas de especies que comparten un oligotipo, entre ellas están las especies *E. avium* y *E. hulanensis*, *E. gilvus* y *E. raffinosus*, *E. crotali* y *E. rotai*, *E. alcedinis* y *E. eurekaensis*, *E. casseliflavus* y *E. gallinarum* y finalmente *E. asini* y *E. saigonensis* (Cuadro 5).

Cuadro 5. Oligotipado de especies y grupo de especies

Especie	Oligotipos	Especie	Oligotipos
<i>E. aquimarinus</i>	ACGCTTTTCA-GGC	<i>E. haemoperoxidus</i>	
<i>E. bulliens</i>	GCGCTCTTTAAAGC	<i>E. moraviensis</i>	
<i>E. camelliae</i>	GCGCTCTTTGACGC	<i>E. caccae</i>	ACGCTGCGCGATA
<i>E. canintestini</i>	GAGCCAAGATATGC	<i>E. quebecensis</i>	
<i>E. canis</i>	GCATCACGCGATGC	<i>E. silesiacus</i>	
<i>E. cecorum</i>	GCGCTCATTTGAGC	<i>E. ureasiticus</i>	
<i>E. columbae</i>	GCGCTTACTTGTGC	<i>E. durans</i>	
<i>E. diestrammenae</i>	AGGCTAAGATGTTC	<i>E. hirae</i>	
<i>E. dispar</i>	AAGCTTAAGTATGC	<i>E. lactis</i>	GCATCACGAGATG
<i>E. dongliensis</i>	GCATTATAAACTGC	<i>E. thailandicus</i>	
<i>E. faecalis</i>	ACGCCGTTTAGGGC	<i>E. xinjiangensis</i>	
<i>E. faecium</i>	GCATCAC-AGATG	<i>E. devriesei</i>	
<i>E. florum</i>	GAGCTACGAGATGC	<i>E. pseudoavium</i>	GCATCATAAACTG
<i>E. hermanniensis</i>	AAGCCAGAACTTGC	<i>E. viikkiensis</i>	
<i>E. italicus</i>	GCGCTCTTTAATGC	<i>E. xiangfangensis</i>	
<i>E. lemanii</i>	ACGCTCTTCAAAC	<i>E. casseliflavus</i>	GCATCCATTAATG
<i>E. malodoratus</i>	GCATCGAGATATGC	<i>E. gallinarum</i>	
<i>E. mundtii</i>	GCATCACGAGTTGC	<i>E. avium</i>	GCATCACGAGGT
<i>E. nangangensis</i>	GTGCCATTTAATAC	<i>E. hulanensis</i>	
<i>E. olivae</i>	GCATTCATTCACGC	<i>E. gilvus</i>	GCATCAAGATATG
<i>E. pallens</i>	AGGCCAGAACTTGC	<i>E. raffinosus</i>	
<i>E. phoeniculicola</i>	GCATCATGAAATAC	<i>E. crotali</i>	ACGCCGCGCGATA
<i>E. pingfangensis</i>	GCATCGTAAACTGC	<i>E. rotai</i>	
<i>E. plantarum</i>	ACGCTACGCGATAC	<i>E. alcedinis</i>	GTGCTTTTTAAGAC
<i>E. ratti</i>	ACATCACAAGATAC	<i>E. eurekaensis</i>	
<i>E. rivorum</i>	ACGCCGCTTGGGAC	<i>E. asini</i>	GAGCCACAAGATGC
<i>E. saccharolyticus</i>	GCGCTCTTTAACGC	<i>E. saigonensis</i>	
<i>E. songbeiensis</i>	GCGCTCAAGTCTGC		
<i>E. sulfureus</i>	GCGCTCGTTGTAGC		
<i>E. termitis</i>	GCGCTGCGCGATAC		
<i>E. ureilyticus</i>	ACGCTGCGCGATGC		
<i>E. villorum</i>	GCATTATAAAATGC		
<i>E. wangshanyuanii</i>	GCATCGCGCGATAC		

8.8.2. Comparación de los resultados del valor de expectación y el porcentaje de identidad del Oligotipado con respecto al gen ARNr 16S en BLAST

Una vez determinadas las 13 posiciones discriminatorias, se optó por una longitud de secuencia de 105 nucleótidos. Las secuencias (105 nt) fueron analizadas en BLAST al igual que las secuencias completas del gen ARNr 16S. En el caso de las 60 secuencias que contiene a los oligotipos (105 nt) el 60% se logró identificar con una única secuencia con un porcentaje de identidad del 100% con distintos niveles valores de expectación, el 36.6% de las secuencias obtuvieron un valor de expectación de $7.00E-49$, seguido del 18.33% ($5.00E-50$) y finalmente 2 secuencias se identificaron como *E. cecorum* ($1.00E-46$) y *E. florum* ($1.00E-48$) obtuvieron los valores de expectación más bajos (Cuadro 6).

Dentro de las cuales también se incluyen a parejas que comparten oligotipo como *E. alcedinis* y *E. eurekaensis*, además de *E. asini* y *E. saigonensis* que mediante la búsqueda en BLAST, lograron identificarse plenamente. Por su parte utilizando el gen ARNr 16S completo se obtuvieron porcentajes de identidad $>99\%$ y con un valor de expectación de 0.0, con la excepción de la especie *E. florum* que obtuvo una identidad de 98.53% (Cuadro 7). En cuanto a los agrupamientos los valores de expectación y porcentajes de identidad para el grupo Pseudoavium (*E. devriesei*, *E. pseudoavium*, *E. viikkiensis* y *E. xiangfangensis*) obtuvo ($7.00E-49$; 100.00%). En el caso del grupo Haemoperoxidus (*E. haemoperoxidus*, *E. moraviensis*, *E. caccae*, *E. quebecensis*, *E. silesiacus* y *E. ureasiticus*) y el grupo durans (*E. durans*, *E. hirae*, *E. lactis*, *E. thailandicus* y *E. xinjiangensis*) los valores fueron de $5.00E-50$; 100.00% para ambos grupos, por lo que no se puede discernir entre las especies de estos grupos. Además, la secuencia de *E. faecium* que tiene su propio oligotipo (GCATCAC-AGATG), en la búsqueda en BLAST resultó con un valor de expectación de $9.00E-48$ y 99.05% de identidad para las especies del grupo Durans. Los valores de identidad para el gen

completo fueron >99% para todas las especies de cada grupo. Además, para la especie *E. xiangfangensis* obtuvo una 95.98% con el grupo Durans (Cuadro 7).

De las 6 parejas que se formaron por el oligotipado, 3 de ellas no pudieron identificarse adecuadamente en la búsqueda en BLAST, en primera instancia las parejas de *E. avium* y *E. hulanensis* y la pareja de *E. gilvus* y *E. raffinosus* obtuvieron los mismos valores de expectación e identidad (7.00E-49; 100.00%), por lo que no fue posible separar a estas parejas. De igual manera en el caso de la pareja integrada por *E. crotali* y *E. rotai* (5.00E-50; 100.00%) y que se consideran como la misma secuencia (Cuadro 7).

Finalmente, los valores de expectación e identidad fueron de 2.00E-48; 100.00% y para el gen completo fue de 98.91% con 0.0 de valor de expectación *E. casseliflavus* y 7.00E-49; 100.00% *E. gallinarum* en donde esta última tuvo los mismos valores con la especie *E. casseliflavus*. Para el gen completo obtuvo una identidad de 98.91% (0.0)

Cuadro 6. Valor de expectación y porcentaje de identidad del Oligotipado y del ARNr 16S en BLAST

Especie	Longitud 105 nt			ARNr 16S >1,200	
	Oligotipo	Valor-E	ID %	Valor-E	ID %
<i>E. nangangensis</i>	GTGCCATTTAATAC	7.00E-49	100.00%	0.0	99.50%
<i>E. columbae</i>	GCGCTTACTTGTGC	7.00E-49	100.00%	0.0	99.55%
<i>E. termitis</i>	GCGCTGCGGATAAC	5.00E-50	100.00%	0.0	99.47%
<i>E. camelliae</i>	GCGCTCTTTGACGC	7.00E-49	100.00%	0.0	99.66%
<i>E. italicus</i>	GCGCTCTTTAATGC	7.00E-49	100.00%	0.0	99.60%
<i>E. saccharolyticus</i>	GCGCTCTTTAACGC	7.00E-49	100.00%	0.0	99.16%
<i>E. bulliens</i>	GCGCTCTTTAAAGC	7.00E-49	100.00%	0.0	99.66%
<i>E. sulfureus</i>	GCGCTCGTTGTAGC	7.00E-49	100.00%	0.0	99.16%
<i>E. cecorum</i>	GCGCTCATTGAGC	1.00E-46	99.03%	0.0	99.42%
<i>E. songbeiensis</i>	GCGCTCAAGTCTGC	7.00E-49	100.00%	0.0	99.93%
<i>E. olivae</i>	GCATTCATTCACGC	7.00E-49	100.00%	0.0	99.93%
<i>E. dongliensis</i>	GCATTATAAACTGC	7.00E-49	100.00%	0.0	99.44%
<i>E. villorum</i>	GCATTATAAAAATGC	5.00E-50	100.00%	0.0	99.40%
<i>E. pingfangensis</i>	GCATCGTAAACTGC	7.00E-49	100.00%	0.0	99.44%
<i>E. wangshanyuanii</i>	GCATCGCGGATAAC	5.00E-50	100.00%	0.0	99.47%
<i>E. malodoratus</i>	GCATCGAGATATGC	7.00E-49	100.00%	0.0	99.29%
<i>E. phoeniculicola</i>	GCATCATGAAATAC	5.00E-50	100.00%	0.0	99.39%
<i>E. canis</i>	GCATCACGCGATGC	5.00E-50	100.00%	0.0	99.39%
<i>E. mundtii</i>	GCATCACGAGTTGC	5.00E-50	100.00%	0.0	99.04%
<i>E. florum</i>	GAGCTACGAGATGC	1.00E-48	100.00%	0.0	98.53%
<i>E. canintestini</i>	GAGCCAAGATATGC	7.00E-49	100.00%	0.0	99.54%
<i>E. diestrammenae</i>	AGGCTAAGATGTTT	7.00E-49	100.00%	0.0	99.65%
<i>E. pallens</i>	AGGCCAGAACTTGC	7.00E-49	100.00%	0.0	99.46%
<i>E. aquimarinus</i>	ACGCTTTTCA-GGC	5.00E-45	100.00%	0.0	99.27%
<i>E. lemanii</i>	ACGCTCTTCAAAC	7.00E-49	100.00%	0.0	99.68%
<i>E. plantarum</i>	ACGCTACGCGATAAC	5.00E-50	100.00%	0.0	99.37%
<i>E. faecalis</i>	ACGCCGTTTAGGGC	5.00E-50	100.00%	0.0	99.29%
<i>E. rivorum</i>	ACGCCGCTTGGGAC	5.00E-50	100.00%	0.0	99.32%
<i>E. ratti</i>	ACATCACAAAGATAAC	5.00E-50	100.00%	0.0	99.47%
<i>E. dispar</i>	AAGCTTAAGTATGC	7.00E-49	100.00%	0.0	99.36%
<i>E. hermanniensis</i>	AAGCCAGAACTTGC	7.00E-49	100.00%	0.0	99.30%
<i>E. ureilyticus</i>	ACGCTGCGGATGC	5.00E-50	100.00%	0.0	99.41%
<i>E. alcedinis</i>	GTGCTTTTTAAGAC	7.00E-49	100.00%	0.0	99.92%
<i>E. eurekaensis</i>		7.00E-49	100.00%	0.0	99.68%
<i>E. asini</i>		7.00E-49	100.00%	0.0	99.55%
<i>E. saigonensis</i>	GAGCCACAAGATGC	7.00E-49	100.00%	0.0	99.78%

Cuadro 7. Valor de expectación y porcentaje de identidad del Oligotipado por grupos o parejas y del ARNr 16S en BLAST

Especie	Longitud 105 nt			ARNr 16S >1,200	
	Oligotipo	Valor-E	ID %	Valor-E	ID %
<i>E. haemoperoxidus</i>				0.0	99.01%
<i>E. moraviensis</i>				0.0	99.07%
<i>E. caccae</i>	ACGCTGCGCGATA	5.00E-50	100.00%	0.0	99.39%
<i>E. quebecensis</i>				0.0	99.77%
<i>E. silesiacus</i>				0.0	99.41%
<i>E. ureasiticus</i>				0.0	99.69%
<i>E. durans</i>				0.0	99.48%
<i>E. hirae</i>				0.0	99.29%
<i>E. lactis</i>	GCATCACGAGATG	5.00E-50	100.00%	0.0	99.44%
<i>E. thailandicus</i>				0.0	99.39%
<i>E. xinjiangensis</i>				0.0	95.98%
<i>E. faecium</i>				0.0	99.36%
<i>E. devriesei</i>	GCATCAC-AGATG	9.00E-48	99.05%	0.0	99.36%
<i>E. pseudoavium</i>				0.0	99.47%
<i>E. viikkiensis</i>				0.0	99.35%
<i>E. viikkiensis</i>				0.0	99.53%
<i>E. xiangfangensis</i>				0.0	99.44%
<i>E. casseliflavus</i>	GCATCATAAACTG	7.00E-49	100.00%	0.0	98.91%
<i>E. gallinarum</i>				0.0	99.36%
<i>E. avium</i>	GCATCCATTAATG	7.00E-49	100.00%	0.0	99.42%
<i>E. hulanensis</i>				0.0	99.57%
<i>E. gilvus</i>	GCATCACGAGGT	7.00E-49	100.00%	0.0	99.46%
<i>E. raffinosus</i>				0.0	99.33%
<i>E. crotali</i>	GCATCAAGATATG	7.00E-49	100.00%	0.0	99.39%
<i>E. rotai</i>				0.0	99.41%
	ACGCCGCGCGATA	5.00E-50	100.00%	0.0	99.39%
				0.0	99.41%

IX. DISCUSIÓN

9.1. Análisis filogenético del genero *Enterococcus*

Los resultados del consenso filogenético de Inferencia Bayesiana, Máxima Verosimilitud y Unión de Vecinos de 1,253 posiciones compartidas del gen ARNr 16S del género *Enterococcus* (Figura 4) se compararon con la filogenia de *Enterococcus* spp. reportadas por Zhong y cols. en 2017. Comparando dichas estimaciones, en el presente trabajo se analizaron un total de 60 spp., 11 más que Zhong *et al.* (2017), las cuales son actualizaciones filogenéticas, es decir, nuevas especies que fueron reportadas en años recientes como *E. hulanensis*, *E. pingfangensis*, *E. dongliensis*, *E. crotali*, *E. wangshanyuanii*, *E. bulliens*, *E. saigonensis*, *E. florum*, *E. nangangensis*, *E. xinjiangensis* y *E. songbeiensis* (Figura 4).

En el presente trabajo se describieron 4 supergrupos y dos grupos menores: Supergrupo *E. faecium*, *E. faecalis*, *E. casseliflavus* y *E. dispar*, así como los grupos *E. cecorum* y *E. aquimarinus*, la cual es similar a la estimación realizada por Zhong *et al.* (2017) y Gooch *et al.* (2021) ambos utilizando el gen ARNr 16S mediante el método de máxima verosimilitud con un corte de nodo del 70% (1,000 réplicas) y enraizado en el punto medio en el trabajo de Gooch *et al.* (2021).

Zhong y cols. en 2017 reportaron una segunda estimación filogenética con los genes concatenados del core y pan-genoma. Cuya topología de 6 ramas conformada por las especies *E. faecium*, *E. faecalis*, *E. dispar*, *E. casseliflavus*, *E. pallens* y *E. canis* es similar a la presentada en este estudio (4 supergrupos), debido a que la rama de *E. pallens* y *E. canis*

están dentro del Supergrupo *E. casseliflavus*, y se discuten a continuación con base a la topología obtenida en el presente estudio.

Supergrupo *E. faecium*.

En este supergrupo se incluye al grupo *E. avium-pseudoavium*, el cual contiene a las especies *E. malodoratus*, *E. gilvus* y *E. raffinosus* que aparecen con un ancestro en común con un soporte de (0.987/-/732) como grupos independientes. Mientras que Kosecka-Strojek y cols. (2020) no pudieron separar a *E. raffinosus* / *E. gilvus* ya que eran idénticas o casi idénticas (identidad $\geq 99.8\%$). Sin embargo, en el presente estudio la topología del árbol filogenético contiene a *E. gilvus* y *E. raffinosus* en dos grupos bien definidos con un soporte de nodo de -/0.954 y -/770/-, respectivamente, lo cual coincide con Li y Gu (2019). Dicho agrupamiento es el que tiene más definición respecto a las estimaciones filogenéticas de Zhong *et al.* (2017) y Gooch *et al.* (2021). Seguido por el grupo *E. avium-hulanensis* conformado por dichas especies con un soporte de nodo (0.994/730/786), similar a lo reportado por Li y Gu en 2019.

Uno de los grupos más grandes y con menor definición es el grupo *E. pseudoavium* que contiene a las especies *E. xiangfangensis*, *E. devriesei*, *E. pingfangensis*, *E. dongliensis*, *E. pseudoavium* y *E. viikkiensis* (1/820/892). Esto es similar a lo reportado anteriormente por Gooch *et al.* (2021) y Li y Gu (2019) solo que difiere en que la especie *E. dongliensis* se logró separar de este grupo en ambas estimaciones. Pero aun así dicho grupo no tuvo buena resolución debido a la similitud entre las secuencias, por lo que concuerda con la hipótesis que menciona las deficiencias discriminatorias del gen ARNr 16S para especies que provienen del mismo hábitat (Gilmore *et al.*, 2013).

El grupo base *E. faecium* está conformado por las especies *E. hiraе*, *E. thailandicus* y *E. durans* que tiene el mismo ancestro en común (965/-/-). Seguido por el grupo *E. faecium-lactis* que incluye a dichas especies, dicha topología concuerda con Gilmore y cols. (2013) y Gooch y cols. (2021). En la rama contigua se encuentran las especies *E. ratti*, *E. mundtii*, *E. villorum* y *E. canis*. Zhong y cols. (2017) concuerda con el agrupamiento descrito en el presente estudio solo en 5 especies (*E. faecium*, *E. thailandicus*, *E. durans*, *E. mundtii* y *E. villorum*) dicho agrupamiento se atribuye a que son aislados principalmente de sangre y tracto gastrointestinal.

De acuerdo con Li y Gu. (2019), *E. villorum* y *E. canis* tienen una relación más estrecha, mientras que Gooch *et al.* (2021) reporta que *E. canis* y *E. hiraе* son quienes tienen dicha relación. Por su parte, en este estudio se reporta que *E. canis* tiene su propia rama con un soporte de (1/1000/996), contrastando con la de Zhong *et al.* (2017) en cuya estimación dicha especie resultó parafilética.

Además, se reporta a la especie *E. phoeniculicola* como parafilética al Supergrupo *E. faecium* con un soporte (1/1000/997), esto concuerda con Gooch *et al.* (2021) y Zhong *et al.* (2017), en cambio Li y Gu (2019) reporta que fue divergente al grupo *E. faecalis*.

Supergrupo *E. faecalis*

El primer grupo que se describe es *E. haemoperoxidus*, está integrado por las especies *E. ureilyticus*, *E. rotai*, *E. ureasiticus*, *E. caccae*, *E. silesiacus* y *E. crotali* (1/810/-) similar a lo reportado por Zhong *et al.* (2017). Dicho grupo está dividido en 3 subgrupos, el primer subgrupo surge a partir de la especie *E. rotai* que se unió a la nueva actualización taxonómica

E. crotali en una rama independiente (0.996/-/-) concordando con Li y Gu (2019) y Gooch *et al.* (2021) con un soporte >70% formando el subgrupo *E. rotai-crotali*.

El subgrupo *E. caccae* está conformado por las especies *E. ureilyticus*, *E. ureasiticus*, *E. caccae*, *E. silesiacus*, se muestra una politomía que concuerda con Gooch *et al.* (2021) al integrar las mismas especies; sin embargo, Li y Gu (2019) reportaron a *E. ureilyticus* fuera del subgrupo. El subgrupo *E. haemoperoxidus* está integrado por las especies *E. termitis*, *E. plantarum*, *E. haemoperoxidus*, *E. quebecensis*, *E. moraviensis* de acuerdo con lo reportado por Zhong y cols. (2017); sin embargo, en el presente estudio *E. termitis* formó su propio grupo (1/1000/992), lo cual concuerda con Li y Gu (2019); Gooch *et al.* (2021) y la especie de *E. wangshanyuanii* una nueva actualización filogenética se reporta como una rama divergente al subgrupo *E. haemoperoxidus* reportado previamente por Gooch *et al.* (2021). El grupo *E. faecalis-rivorum* (1/840/867) integrado por dichas especies tienen un ancestro común. Esto ya había sido reportado por Li y Gu (2019) y Gooch *et al.* (2021).

En el estudio de Gooch *et al.* (2021) la especie *E. nangangensis* se mostró divergente a la rama de *E. faecalis*, contrastando con lo reportado en este estudio, en el que dicha especie se muestra parafilética a todo el árbol.

En general, la comparación entre el presente estudio y lo reportado por Zhong y cols. (2017), en el supergrupo *E. faecalis* coinciden en 7 especies (*E. faecalis*, *E. termitis*, *E. quebecensis*, *E. moraviensis*, *E. caccae*, *E. haemoperoxidus* y *E. silesiacus*), y ellos atribuyen dicho agrupamiento a que estas especies son aisladas principalmente de agua y tracto gastrointestinal.

Supergrupo *E. casseliflavus*

Está integrado por el grupo *E. saccharolyticus* (0.966/-/-) que incluye a las especies *E. saccharolyticus saccharolyticus*, *E. saccharolyticus taiwanensis* y *E. olivae*, el agrupamiento entre *E. saccharolyticus* y *E. olivae* concuerda con Zhong y cols. (2017), aunque no se pueden hacer más comparaciones ya que el presente estudio, es el único en el que se incluyeron subespecies de *E. saccharolyticus*, además se reporta que *E. saccharolyticus taiwanensis* diverge del grupo de las subespecies; es importante resaltar que de acuerdo con Chen y cols. (2013) entre *E. saccharolyticus taiwanensis* y *E. saccharolyticus*, se observaron al menos cuatro diferencias de nucleótidos entre las secuencias del gen ARNr 16S.

Además de 3 agrupamientos más, el grupo *E. italicus-camelliae* está conformado por dichas especies (1/830/878), en la estimación realizada por Zhong *et al.* (2017) se reporta el mismo agrupamiento, con la diferencia que la especie *E. sulfureus* diverge, mientras que en este estudio dicha especie se agrupa con la nueva actualización taxonómica *E. bulliens* y forman el grupo *E. sulfureus-bulliens* (1/970/919); otro grupo importante es el *E. gallinarum-casseliflavus* (1/-/899), ambos agrupamiento concuerdan con Li y Gu (2019) y Gooch *et al.* (2021).

En dicho agrupamiento se concuerda con lo reportado por Zhong y cols. (2017) en 5 especies (*E. casseliflavus*, *E. gallinarum*, *E. saccharolyticus*, *E. sulfureus* y *E. italicus*) y se atribuye a la fuente de aislamiento como plantas y tracto gastrointestinal de aves.

Supergrupo *E. dispar*

Dicho supergrupo (1/-/) está dividido en 3 grupos: el primero denominado grupo *E. dispar*, conformado por *E. dispar*, *E. canintestini* y *E. saigonensis* (-/710/841), el segundo grupo *E. asini* (1/830/856) integrado por *E. asini* y *E. diestrammenae*, y el tercer grupo *E. pallens* conformado por *E. pallens* y *E. hermanniensis* (1/-/). Dichos agrupamientos concuerdan con los reportes de Li y Gu (2019) y Gooch *et al.* (2021), además en base a lo reportado por Zhong y cols. (2017) sobre la explicación de dichas especies, se debe a que han sido aisladas del tracto gastrointestinal de humanos y mamíferos.

E. cecorum

El grupo *E. cecorum* conformado por las especies *E. columbae* y *E. cecorum* con soportes de nodo 1/1000/1000, dicho agrupamiento concuerda con los reportes de Gilmore *et al.* (2013), Zhong *et al.* (2017) y Gooch *et al.* (2021).

E. aquimarinus

En el grupo denominado *E. aquimarinus*, se agruparon las especies *E. aquimarinus*, *E. lemanii*, *E. eurekensis* y *E. alcedinis* con un soporte nodo (1/830/878), dicho agrupamiento concuerda las estimaciones realizadas por Li y Gu (2019) y Gooch *et al.* (2021).

9.2. Entropía posicional del gen ARNr 16S de *Enterococcus*

En el presente estudio los resultados del cálculo de la entropía para cuatro conjuntos (SILVA, RDP, Greengenes y Fusionado) de secuencias pertenecientes al género *Enterococcus*, lograron identificar posiciones altamente variantes, dentro de los 4 conjuntos (100, 190, 198, 205) con valor de entropía de 1. De igual manera Eren y cols. (2011) utilizaron la entropía de Shannon como indicador de variabilidad dentro del gen ARNr 16S a partir de la cual identificaron posiciones con alta variación, que sirvió de base para generar oligotipos que se implementaron para la identificación de *Gardnerella vaginalis*.

9.3. Correlación y agrupamiento jerárquico de los marcadores conservados y variables

De manera global, comparando los marcadores similares al gen ARNr 16S completo en los cuatro conjuntos, los marcadores que obtuvieron un coeficiente de correlación de 1, fueron V2, V3, V8 y V9. El marcador MC2, mostró correlación en 3 de los 4 conjuntos (SILVA, RDP y Fusionado).

Chaudhary *et al.* (2015) a partir del modelado de las regiones hipervariables (RHV) respecto al gen ARNr 16S completo, mediante bosques aleatorios (RF, Random Forest por sus siglas en inglés), reportaron los valores óptimos para una clasificación taxonómica más rápida y precisa RHV cortas. Entre los modelos de RHV individuales, los modelos para RHV V2, V4 y V8 mostraron los valores de precisión más altos de 0.85, 0.87 y 0.85, respectivamente. Esto concuerda con lo reportado en este estudio mediante correlación y agrupamiento, así como estimación de resolución filogenética en que las RHV V2, V4 y V8 tienen una correlación de 1 con respecto a la similaridad con el ARNr 16S. Estos RHV también eran más largos (>200 pb) en comparación con los otros RHV individuales. Los modelos para las regiones V6 y V7 mostraron los valores de precisión más bajos (0.63 y 0.65,

respectivamente) y también tuvieron la longitud más pequeña (86 y 107 pb, respectivamente) en comparación con otros RHV individuales. Y se reporta que los resultados indican que el valor de la precisión es directamente proporcional ($R = 0.85, p \leq 0$) a la longitud de la RHV. Esto último no concuerda con lo reportado en este trabajo ya que los marcadores V6 y V7, también se agruparon dentro de los marcadores similares al del gen ARNr 16S completo en los conjuntos de Greengenes y Fusionado. Además, el modelo de RF del gen ARNr 16S completo también mostró el valor de precisión más alto de 0.91.

9.4. Estimación filogenética de los marcadores similares y disimilares

Con la obtención de tres árboles filogenéticos (Figura 22, 23 y 24), el primero denominado como el árbol del gen ARNr 16S completo en la Figura 22 (1,528 posiciones), el segundo árbol de los marcadores similares en la Figura 23 (V2, V3, V8 y V9) y el tercero de los marcadores disimilares en la Figura 24 (MC1, V1, MC2, MC3, MC4, V4, MC5, V5, MC6, V6, MC7, V7, MC8, MC9 y MC10) con una longitud de 861 posiciones, se demostró la similitud entre el árbol del gen ARNr 16S y el de los marcadores similares sin importar la diferencia de 880 posiciones entre los alineamientos, esto remarca la importancia discriminatoria de los marcadores similares que a pesar de su corta longitud (648 posiciones) aportó una resolución altamente similar al del gen completo, ya que el árbol de los marcadores similares solo se creó una pareja de especies *E. dispar-canintestini*, cuyo valor de sensibilidad fue de 90.91%. y el 96.66% de los clados fueron agrupamiento por especie con un soporte de nodo >70% y el 65.51% de los agrupamientos resultó con una sensibilidad del 100%.

9.5. Oligotipado y regiones discriminatorias

La oligotipificación mostró un total de 13 componentes/posiciones que debido al valor de entropía aportaban el mayor grado de información discriminatoria entre especies, las cuales resultan estar dentro de la región hipervariable V2, y con base a la búsqueda en BLAST, los 33 oligotipos correspondían con 33 especies diferentes con una identidad del 100%. En este trabajo se refleja la variabilidad entre *Enterococcus* spp. por su parte Srinivasan y cols., 2019 en su estudio demuestran y validan el uso de un enfoque de alto rendimiento de análisis de oligotipado de secuencias del gen ARNr 16S para dilucidar la diversidad a nivel de clado para *Candidatus Accumulibacter phosphatis* (CAP).

De manera similar Berry y cols., 2017 realizaron un análisis exhaustivo de la significancia ecológica y filogenética, examinaron cómo se interpretan los datos de oligotipado en la literatura reciente, un estudio de caso de la dañina cianobacteria *Microcystis* que forma floraciones. Identificaron tres oligotipos de *Microcystis* a través del gen ARNr 16S (región V4) de la comunidad bacteriana occidental del lago Erie que se había agrupado previamente en una OTU (Unidad Taxonómica Operativa). Encontraron los mismos tres oligotipos y dos variantes de secuencia adicionales en 46 cultivos de *Microcystis* aislados de lagos interiores de Michigan que abarcan un gradiente trófico. Dicho estudio respalda la utilidad de la oligotipificación para distinguir tipos de secuencias a lo largo de ciertas características ecológicas, mientras que enfatiza que los tipos de secuencias del gen ARNr 16S pueden no reflejar poblaciones cohesivas ecológica o filogenéticamente.

Un caso especial es *E. faecium* que a pesar de tener su oligotipo, el resultado de la búsqueda en BLAST con un valor de expectación de 9.00E-48 y 99.05% de identidad para las especies del grupo durans (*E. durans*, *E. hirae*, *E. lactis*, *E. thailandicus* y *E.*

xinjiangensis). En contraste con otras especies que comparte oligotipo y en BLAST resultan identificarse correctamente.

De acuerdo con Eren *et al.*, 2013, en cuyo trabajo desarrollaron el oligotipado, un nuevo método computacional supervisado que permite investigar la diversidad de organismos bacterianos estrechamente relacionados pero distintos en unidades taxonómicas operativas finales identificadas en conjuntos de datos ambientales a través de datos de genes de ARNr 16S mediante enfoques canónicos. Y concluyen que la oligotipificación puede resolver la distribución de organismos estrechamente relacionados entre entornos y revelar patrones ecológicos previamente pasados por alto para las comunidades microbianas. Como se observó en el presente estudio se logró resolver o llegar a la comprensión de las pequeñas diferencias entre las secuencias del ARNr 16S entre *Enterococcus* spp., en particular de la región V2, y los oligotipos encontrados en esta región 13 posiciones en las que albergar la mayor cantidad de información discriminatoria a nivel de género.

X. CONCLUSIONES

1. Se clasificó a las 60 spp del género *Enterococcus* a diferentes niveles de agrupamiento intragénero, identificándose los supergrupos *E. faecium*, *E. faecalis*, *E. casseliflavus* y *E. dispar*.
2. *E. phoeniculicola*, *E. xinjiangensis*, *E. nangangensis* y *E. songbeiensis* son especies parafiléticas y representan linajes evolutivos divergentes en el género *Enterococcus*.
3. Con base en el cálculo de la entropía y la clasificación de los sitios mediante K-medias, indica que la región con mayor variación es V1. Además, los análisis de componentes principales, demuestra que los marcadores MC1 y V1, son los que acumulan mayor variación, por lo que explican el 99.96% de la varianza total de los 20 marcadores.
4. El número de sitios conservados y número de sitios variables son los parámetros determinantes en el análisis de correlación, son los que tiene mayor poder discriminatorio entre los marcadores.
5. El porcentaje aproximado de sitios conservados es >82%, seguido un porcentaje del >8% de sitios variantes y los sitios altamente variantes corresponden al 5% en los 4 conjuntos del total de sitios evaluados.
6. Los marcadores que resultaron ser similares al ARNr 16S con base en el coeficiente de correlación de Spearman fueron V2, V3, V8 y V9 dentro de los 4 conjuntos.
7. La evaluación de los marcadores mediante estimación filogenética demostró que los marcadores similares al ARNr 16S tienen una resolución adecuada para algunos clados, con respecto a la estimación del gen ARNr 16S completo.

8. La oligotipificación demostró que las 13 posiciones dentro de la región V2, poseen el poder resolutivo para 55% de las especies del género *Enterococcus*.

XI. PERSPECTIVAS

1. Analizar el oligotipado de la región V8, como otro posible marcador informativo.
2. Determinar oligos específicos para esta región discriminadora y hacer el análisis *in vitro* con cepas identificadas como *Enterococcus* spp.

XII. REFERENCIAS BIBLIOGRÁFICAS

1. Ajawatanawong P. 2016. **Molecular Phylogenetics: Concepts for a Newcomer.** *Advances in Biochemical Engineering/Biotechnology.* 160:185-196. DOI: 10.1007/10_2016_49.
2. Akaike, H. 1974. **A new look at the statistical model identification.** *IEEE transactions on automatic control.* 19(6):716-723.
3. Alós J. I. 2015. **Resistencia bacteriana a los antibióticos: una crisis global.** *Enfermedades Infecciosas y Microbiología Clínica.* 33(10): 692-699.9. DOI: 10.1016/j.eimc.2014.10.004
4. Álvarez-Artero, E., Campo-Núñez, A., García-García, I., García-Bravo, M., *et al.* 2021. **Infección urinaria por enterococos: Factores de riesgo y mortalidad. Estudio observacional.** *Revista Clínica Española.* 375-383, ISSN 0014-2565. <https://doi.org/10.1016/j.rce.2020.09.005>
5. Angeletti, S., Lorino, G., Gherardi, G., Battistoni, F., De Cesaris, M. and Dicuonzo, G. 2001. **Routine molecular identification of enterococci by gene specific PCR and 16S ribosomal DNA sequencing.** *J Clin Microbiol.* 39:794–797. <https://doi.org/10.1128/JCM.39.2.794-797.2001>.
6. Arias, C.A. and Murray, B.E. 2012. **The rise of the *Enterococcus*: beyond vancomycin resistance.** *Nat Rev Microbiol.* 10:266 –278. <https://doi.org/10.1038/nrmicro2761>.
7. Ator, L. L. and Starzyk, M. J. 1976. **Distribution of group D streptococci in rivers and streams.** *Microbios.* 16(64):91–104. PubMed PMID: 829267.
8. Barb J. J., Oler A. J., Kim H-S., Chalmers N., Wallen GR., Cashion A, Munson P. J. and Ames N. J. 2016. **Development of an Analysis Pipeline Characterizing Multiple Hypervariable Regions of 16S rRNA Using Mock Samples.** *PLoS one.* 11(2): e0148047. DOI: 10.1371/journal.pone.0148047
9. Baker, G. C., Smith, J. J., and Cowan, D. A. 2003. **Review and re-analysis of domain-specific 16S primers.** *Journal of microbiological methods,* 55(3), 541-555.
10. Berry M. A., White J. D., Davis T. W., Jain S., Johengen T. H., Dick G. J., Sarnelle O., Denev V. J. 2017. **Are Oligotypes Meaningful Ecological and Phylogenetic Units? A**

- Case Study of Microcystis in Freshwater Lakes.** *Frontiers in Microbiology*. 8. 1664-302X. doi:10.3389/fmicb.2017.00365
11. Bizzini, A., and Greub, G. 2010. **Matrix-assisted laser desorption ionization time-of-flight mass spectrometry, a revolution in clinical microbial identification.** *Clinical microbiology and infection: the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 16(11), 1614–1619. <https://doi.org/10.1111/j.1469-0691.2010.03311.x>
 12. Bosshard, P.P., Abels, S., Altwegg, M., Böttger, E.C. and Zbinden, R. 2004. **Comparison of conventional and molecular methods for identification of aerobic catalase-negative gram-positive cocci in the clinical laboratory.** *J Clin Microbiol*. 42:2065–2073. <https://doi.org/10.1128/JCM.42.5.2065-2073.2004>.
 13. Byappanahalli, M. N., M. B. Nevers, A. Korajkic, Z. R. Staley, and V. J. Harwood, 2012. **Enterococci in the environment:** *Microbiol Mol Biol Rev*. 76:685-706.
 14. Campista-León, S., Rivera-Serrano, B. V., García-Guerrero, J. T., and Peinado-Guevara, L. I. 2021. **Phylogenetic characterization and multidrug resistance of bacteria isolated from seafood cocktails.** *Archives of microbiology*. 203(6):3317-3330. <https://doi.org/10.1007/s00203-021-02319-1>
 15. Chakravorty, S., Helb, D., Burday, M., Connell, N., and Alland, D. 2007. **A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria.** *Journal of Microbiological Methods*. 69(2):330–339.
 16. Challa, S., and Neelapu, N. R. R. 2019. **Phylogenetic Trees: Applications, Construction, and Assessment.** In *Essentials of Bioinformatics, Volume III* (pp. 167–192). *Springer International Publishing*. https://doi.org/10.1007/978-3-030-19318-8_10
 17. Chaudhary, N., Sharma, A, K., Agarwal, P., Gupta, A. and Sharma, V.K. 2015. **16S Classifier: A Tool for Fast and Accurate Taxonomic Classification of 16S rRNA Hypervariable Regions in Metagenomic Datasets.** *PLoS ONE*. 10(2): e0116106. <https://doi.org/10.1371/journal.pone.0116106>
 18. Chen, Y. S., Lin, Y. H., Pan, S. F., Ji, S. H., Chang, Y. C., Yu, C. R., Liou, M. S., Wu, H. C., Otaguro, M., Yanagida, F., Liao, C. C., Chiu, C. M., and Huang, B. Q. 2013. ***Enterococcus saccharolyticus* subsp. taiwanensis subsp. nov., isolated from**

- broccoli.** *International journal of systematic and evolutionary microbiology*. 63(Pt 12), 4691–4697. <https://doi.org/10.1099/ijs.0.052761-0>
19. Church, D. L., Cerutti, L., Gürtler, A., Griener, T., Zelazny, A., and Emler, S. 2020. **Performance and Application of 16S rRNA Gene Cycle Sequencing for Routine Identification of Bacteria in the Clinical Microbiology Laboratory.** *Clinical microbiology reviews*. 33(4): e00053-19. <https://doi.org/10.1128/CMR.00053-19>
20. Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Sayers, E. W. 2016. **GenBank.** *Nucleic acids research*. 44(D1): D67–D72. <https://doi.org/10.1093/nar/gkv1276>.
21. Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., Brown, C. T., Porras-Alfaro, A., Kuske, C. R., and Tiedje, J. M. 2014. **Ribosomal Database Project: data and tools for high throughput rRNA analysis.** *Nucleic acids research*. 42(Database issue): D633–D642. <https://doi.org/10.1093/nar/gkt1244>.
22. Collins, M.D., Jones, D., Farrow, J.A.E., Kilpper-Balz, R. and Schleifer, K.H. 1984. ***Enterococcus avium* nom. rev., comb. nov.; *E. casseliflavus* norn. rev., comb. nov.; *E. durans* norn. rev., comb. nov.; *E. gallinarum* comb. nov.; and *E. malodoratus* sp. nova.** *Int. J. Syst. Bacteriol.* 34: 220–223 F.
23. Deck, M.K., Anderson, E.S., Buckner, R.J., Colasante, G., Davis, T.E., Coull, J.M., Crystal, B., Latta, P.D., Fuchs, M., Fuller, D., Harris, W., Hazen, K., Klimas, L.L., Lindao, D., Meltzer, M.C., Morgan, M., Shepard, J., Stevens, S., Wu, F. and Fiandaca, M.J. 2014. **Rapid detection of *Enterococcus* spp. direct from blood culture bottles using *Enterococcus* QuickFISH method: a multicenter investigation.** *Diagn Microbiol Infect Dis.* 78:338 –342. <https://doi.org/10.1016/j.diagmicrobio.2013.12.004>
24. DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., Huber, T., Dalevi, D., Hu, P., and Andersen, G. L. 2006. **Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB.** *Applied and environmental microbiology*. 72(7): 5069–5072. <https://doi.org/10.1128/AEM.03006-05>.
25. Devriese, L.A., Pot, B. and Collins, M.D. 1993. **Phenotypic identification of the genus *Enterococcus* and differentiation of phylogenetically distinct enterococcal species and species groups.** *J. Appl. Microbiol.* 75: 399–408.

26. Devriese, L.A., Vancanneyt, M., Descheemaeker, P., Baele, M., Van Landuyt, H.W., Gordts, B., Butaye, P., Swings, J. and Haesebrouck, F. 2002. **Differentiation and identification of *Enterococcus durans*, *E. hirae* and *E. villorum*.** *J Appl Microbiol.* 92:821–827. <https://doi.org/10.1046/j.1365-2672.2002.01586.x>.
27. Donskey, C. J., Chowdhry, T. K., Hecker, M. T., Hoyen, C. K., Hanrahan, J. A., Hujer, A. M., *et al.* 2000. **Effect of antibiotic therapy on the density of vancomycin-resistant enterococci in the stool of colonized patients.** *The New England Journal of Medicine.* 343(26):1925–1932. PubMed PMID: 11136263.
28. Drummond, A. J., Suchard, M. A., Xie, D., and Rambaut, A. 2012. **Bayesian phylogenetics with BEAUti and the BEAST 1.7.** *Molecular biology and evolution.* 29(8): 1969–1973. <https://doi.org/10.1093/molbev/mss075>
29. Eren, A. M., Zozaya, M., Taylor, C. M., Dowd, S. E., Martin, D. H., and Ferris, M. J. 2011. **Exploring the diversity of *Gardnerella vaginalis* in the genitourinary tract microbiota of monogamous couples through subtle nucleotide variation.** *PloS one.* 6(10): e26732. <https://doi.org/10.1371/journal.pone.0026732>
30. Eren, A.M., Maignien, L., Sul, W.J., Murphy, L.G., Grim, S.L., Morrison, H.G. and Sogin, M.L. 2013. **Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data.** *Methods Ecol. Evol.* 4: 1111-1119. <https://doi.org/10.1111/2041-210X.12114>
31. European Centre for Disease Prevention and Control. 2013. **Point prevalence survey of healthcare-associated infections and antimicrobial use in European acute care hospitals 2011–2012.** European Centre for Disease Prevention and Control, Solna, Sweden.
32. Farrar Stoakes S. 2019. **Importance of Phylogeny in Microbiology.** *News Medical Life Sciences.* Disponible en: <https://www.news-medical.net/life-sciences/Importance-of-Phylogeny-in-Microbiology.aspx> (Accesado en Junio 2021)
33. Felsenstein, J. 1985. **Confidence limits on phylogenies: an approach using the bootstrap.** *Evolution.* 39(4): 783-791.

34. Finegold, S. M., Sutter, V. L., and Mathisen, G. E. 1983. **Normal indigenous intestinal flora.** In D. J. Hentges, *Human intestinal microflora in health and disease* (pp. 3-29). Waltham: Academic Press.
35. Fisher, K., and Phillips, C. 2009. **In vitro inhibition of vancomycin-susceptible and vancomycin resistant *Enterococcus faecium* and *E. faecalis* in the presence of citrus essential oils:** *Br J Biomed Sci.* 66:180-185.
36. Frank, E., Hall, M. A. and Witten I. H. 2016. **The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques",** Morgan Kaufmann, Fourth Edition, 2016.
37. Frieri M., Kumar K. and Boutin A. 2017. Antibiotic resistance. *Journal of Infection and Public Health.* 10(4):369–378. DOI: 10.1016/j.jiph.2016.08.007
38. Ggetti, P., Bonofiglio, L., García Gabarrot, G., Kaufman, S., Mollerach, M., Vigliarolo, L., von Specht, M., Toresani, I. and Lopardo, H.A. 2019. **Resistance to β -lactams in enterococci.** *Rev Argent Microbiol.* 51(2):179-183. doi: 10.1016/j.ram.2018.01.007. Epub 2018 Sep 20. PMID: 30243525.
39. Gilmore, M. S., Lebreton, F., and van Schaik, W. 2013. **Genomic transition of enterococci from gut commensals to leading causes of multidrug-resistant hospital infection in the antibiotic era.** *Current opinion in microbiology.* 16(1):10–16. <https://doi.org/10.1016/j.mib.2013.01.006>
40. Gilmore, M.S., Clewell, D.B., Ike, Y., *et al.* editors. 2014. **Enterococci: From Commensals to Leading Causes of Drug Resistant Infection** [Internet]. Boston: Massachusetts Eye and Ear Infirmary.
41. Grice E. A., Kong H. H., Renaud G., Young A. C., NISC Comparative Sequencing Program, Bouffard, G. G. and Segre, J. A. 2008. **A diversity profile of the human skin microbiota.** *Genome research.* 18(7):1043–1050. DOI:10.1101/gr.075549.107
42. Goh, S. H., Facklam, R. R., Chang, M., Hill, J. E., Tyrrell, G. J., Burns, E. C., Chan, D., He, C., Rahim, T., Shaw, C., and Hemmingsen, S. M. 2000. **Identification of *Enterococcus* species and phenotypically similar *Lactococcus* and *Vagococcus* species by reverse checkerboard hybridization to chaperonin 60 gene sequences.** *Journal of clinical microbiology.* 3953–3959. <https://doi.org/10.1128/JCM.38.11.3953-3959.2000>

43. Gooch, H., Kiu, R., Rudder, S., Baker, D. J., Hall, L. J., and Maxwell, A. 2021. ***Enterococcus innesii* sp. nov., isolated from the wax moth *Galleria mellonella*.** *International journal of systematic and evolutionary microbiology*. 71(12):005168. <https://doi.org/10.1099/ijsem.0.005168>
44. Guindon, S., Dufayard, J., Lefort, V., Anisimova, M., Hordijk, W., y Gascuel, O. 2010. **New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of phylml 3.0.** *Systematic Biology*, 59(3), 307-321. doi:10.1093/sysbio/syq010
45. Guzmán Prieto, A.M., van Schaik. W., Rogers, M.R.C., Coque, T.M., Baquero, F., Corander, J. and Willems, R.J.L. 2016. **Global emergence and dissemination of enterococci as nosocomial pathogens: attack of the clones?** *Front Microbiol.* 7:788. <https://doi.org/10.3389/fmicb.2016.00788>
46. Hall, T.A. 1999. **BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT.** *Nucleic Acids Symposium Series*. 41:95-98.
47. Hammer, O., Harper, D. and Ryan, P. 2001. **PAST: paleontological statistics software package for education and data analysis.** *Palaeontologia Electronica*. 4:1-9.
48. Hartfield, M., Murall, C. L. and Alizon, S. 2014. **Clinical applications of pathogen phylogenies.** *Trends in Molecular Medicine*. 20(7):394-404. DOI: 10.1016/j.molmed.2014.04.002
49. Hendrickx, A.P., Bonten, M.J., van Luit-Asbroek, M., Schapendonk, C.M., Kragten, A.H. and Willems, R.J. 2008. **Expression of two distinct types of pili by a hospital-acquired *Enterococcus faecium* isolate.** *Microbiology*. 154: 3212–3223. <https://doi.org/10.1099/mic.0.2008/020891-0>.
50. Hendrickx, A.P., Schapendonk, C.M., van Luit-Asbroek, M., Bonten, M.J., van Schaik, W. and Willems, R.J. 2010. **Differential Pila pilus assembly by a hospital acquired and a community-derived *Enterococcus faecium* isolate.** *Microbiology*. 156:2649 – 2659. <https://doi.org/10.1099/mic.0.041392-0>.
51. Hidron, A.I., Edwards, J.R., Patel, J., Horan, T.C., Sievert, D.M., Pollock, D.A. and Fridkin, S.K. 2008. **Antimicrobial-resistant pathogens associated with healthcare-associated infections: annual summary of data reported to the National Healthcare**

- Safety Network at the Centers for Disease Control and Prevention, 2006–2007.** *Infect Control Hosp Epidemiol.* 29:996–1011. <https://doi.org/10.1086/591861>.
52. Holzapfel, W. H. and Wood, B.J.B. (ed). 2014. **The genus *Enterococcus***, p 180–190. In *Lactic acid bacteria: biodiversity and taxonomy*. John Wiley & Sons, Ltd, Chichester, England. <https://doi.org/10.1002/9781118655252.ch15>.
53. Hong, Y., Guo, M., and Wang, J. 2020. **ENJ algorithm can construct triple phylogenetic trees.** *Molecular Therapy - Nucleic Acids.* 23:286–293. <https://doi.org/10.1016/j.omtn.2020.11.004>
54. IBM Corp. 2015. **Version 22 of IBM SPSS Statistics for Windows.** IBM Corp. is headquartered in Armonk, New York.
55. Jaspers, E., Overmann, J. 2004. **Ecological significance of microdiversity: identical 16S rRNA gene sequences can be found in bacteria with highly divergent genomes and ecophysologies.** *Appl Environ Microbiol.* 70:4831–4839.
56. Jenkins, I. 2007. **Linezolid-and vancomycin-resistant *Enterococcus faecium* endocarditis: successful treatment with tigecycline and daptomycin.** *Journal of hospital medicine.* 2(5): 343. DOI: 10.1002/jhm.236.
57. Johnson, J. S., Spakowicz, D. J., Hong, B. Y., Petersen, L. M., Demkowicz, P., Chen, L. and Weinstock, G. M. 2019. **Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis.** *Nature communications.* 10(1):1-11.
58. Kalyaanamoorthy, S., Minh, B. Q., Wong, T., von Haeseler, A. and Jeremiin, L. S. 2017. **ModelFinder: fast model selection for accurate phylogenetic estimates.** *Nature methods.* 14(6): 587–589. <https://doi.org/10.1038/nmeth.4285>
59. Katoh, K., and Standley, D. M. 2013. **MAFFT multiple sequence alignment software version 7: improvements in performance and usability.** *Molecular biology and evolution.* 30(4): 772–780. <https://doi.org/10.1093/molbev/mst010>.
60. Ke, D., Picard, F. J., Martineau, F., Ménard, C., Roy, P. H., Ouellette, M. and Bergeron, M. G. 1999. **Development of a PCR assay for rapid detection of enterococci.** *Journal of clinical microbiology,* 37(11), 3497–3503. <https://doi.org/10.1128/JCM.37.11.3497-3503.1999>
61. Ke, D., Boissinot, M., Huletsky, A., Picard, F. J., Frenette, J., Ouellette, M., Roy, P. H. and Bergeron, M. G. 2000. **Evidence for horizontal gene transfer in evolution of**

- elongation factor Tu in enterococci.** *Journal of bacteriology*, 6913–6920. <https://doi.org/10.1128/JB.182.24.6913-6920.2000>
62. Kong H. H. 2011. **Skin microbiome: genomics-based insights into diversity and role of skin microbes.** *Trends in molecular medicine*. 17(6): 320–328. doi: 10.1016/j.molmed.2011.01.013
63. Konstantinidis, K.T. and Tiedje, J.M. 2007. **Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead.** *Curr Opin Microbiol*. 10:504–509.
64. Kosecka-Strojek, M., Sabat, A. J., Akkerboom, V., Kooistra-Smid, A., Miedzobrodzki, J., and Friedrich, A. W. 2019. **Development of a reference data set for assigning *Streptococcus* and *Enterococcus* species based on next generation sequencing of the 16S-23S rRNA region.** *Antimicrobial resistance and infection control*. 8:178. <https://doi.org/10.1186/s13756-019-0622-3>.
65. Kosecka-Strojek, M., Wolska, M., Żabicka, D., Sadowy, E., and Międzobrodzki, J. 2020. **Identification of Clinically Relevant *Streptococcus* and *Enterococcus* Species Based on Biochemical Methods and 16S rRNA, *sodA*, *tuf*, *rpoB*, and *recA* Gene Sequencing.** *Pathogens (Basel, Switzerland)*. 9(11): 939. <https://doi.org/10.3390/pathogens9110939>
66. Kristich, C.J., Rice, L.B. and Arias, C.A. 2014. **Enterococcal Infection—Treatment and Antibiotic Resistance.** In: Gilmore MS, Clewell DB, Ike Y, Shankar N, editors. *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection* [Internet]. Boston: Massachusetts Eye and Ear Infirmary. PMID: 24649502.
67. Lagace-Wiens, P.R., Adam, H.J., Karlowsky, J.A., Nichol, K.A., Pang, P.F., Guenther, J., Webb, A.A., Miller, C. and Alfa, M.J. 2012. **Identification of blood culture isolates directly from positive blood cultures by use of matrix-assisted laser desorption ionization–time of flight mass spectrometry and a commercial extraction system: analysis of performance, cost, and turnaround time.** *J Clin Microbiol*. 50:3324–3328. <https://doi.org/10.1128/JCM.01479-12>.
68. Larrea-Murrell, J. A., Rojas-Badía, M. M., Romeu-Álvarez, B., Rojas-Hernández, N. M. and Heydrich-Pérez, M. 2013. **Bacterias indicadoras de contaminación fecal en la evaluación de la calidad de las aguas: revisión de la literatura.** *Revista CENIC*.

- Ciencias Biológicas*. 44(3): 24-34. [Fecha de Consulta 19 de junio de 2021]. ISSN: 0253-5688. Disponible en: <https://www.redalyc.org/articulo.oa?id=181229302004>
69. Lebreton, F., Willems, R.J.L. and Gilmore, M.S., 2014. **Enterococcus Diversity, Origins in Nature, and Gut Colonization**. In: Gilmore, M.S., Clewell, D.B., Ike, Y., Shankar, N. (Eds.), *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection*. Massachusetts Eye and Ear Infirmary, Boston, pp. 5–63.
70. Lee, I. K., Liu, J. W. and Yang, K. D. 2005. **Clinical characteristics and risk factors for concurrent bacteremia in adults with dengue hemorrhagic fever**. *The American journal of tropical medicine and hygiene*. 72(2):221-226. DOI: <https://doi.org/10.4269/ajtmh.2005.72.221>
71. Lemey, P., Salemi, M. and Vandamme, AM. 2009. **Basic concepts of molecular evolution. The Phylogenetic Handbook; A Practical Approach to Phylogenetic Analysis and Hypothesis Testing** (20-25). United States of America: Cambridge University Press, New York.
72. Lemoine, F., Correia, D., Lefort, V., Doppelt-Azeroual, O., Mareuil, F., Cohen-Boulakia, S., and Gascuel, O. 2019. **NGPhylogeny.fr: new generation phylogenetic services for non-specialists**. *Nucleic acids research*, 47(W1), W260–W265. <https://doi.org/10.1093/nar/gkz303>.
73. Letunic, I., and Bork, P. 2021. **Interactive Tree Of Life (iTOL) v6: online tool for phylogenetic tree display and annotation**. *Nucleic acids research*. 47(W1), doi: **10.1093/nar/gkab301**.
74. Li, Y. Q., and Gu, C. T. 2019. **Enterococcus pingfangensis sp. nov., Enterococcus dongliensis sp. nov., Enterococcus hulanensis sp. nov., Enterococcus nangangensis sp. nov. and Enterococcus songbeiensis sp. nov., isolated from Chinese traditional pickle juice**. *International journal of systematic and evolutionary microbiology*. 69(10): 3191–3201. <https://doi.org/10.1099/ijsem.0.003608>
75. Ljungstrom, L., Enroth, H., Claesson, B.E., Ovemyr, I., Karlsson, J., Froberg, B., Brodin, A.K., Pernestig, A.K., Jacobsson, G., Andersson, R. and Karlsson, D. 2015. **Clinical evaluation of commercial nucleic acid amplification tests in patients with suspected sepsis**. *BMC Infect Dis*. 15:199. <https://doi.org/10.1186/s12879-015-0938-4>.

76. Ludwig, W., Schleifer, K., y Whitman, W. B. 2009. **Family IV. Enterococcaceae fam. nov.** In P. Vos, G. Garrity, D. Jones, N. R. Krieg, W. Ludwig, F. A. Rainey, *et al.*, *Bergey's Manual of Systematic Bacteriology.*, 2nd edn, vol. 3 (The Firmicutes) (p. 594). New York: Springer.
77. Ly-Trong, N., Naser-Khdour, S., Lanfear, R., y Minh, B. Q. 2022. **AliSim: A fast and versatile phylogenetic sequence simulator for the genomic era.** *Molecular Biology and Evolution*, 39(5). doi:10.1093/molbev/msac092
78. Maccallum, W. G., and T. W. Hastings, 1899. **A case of acute endocarditis caused by *Micrococcus zymogenes* (nov. spec.), with a description of the microorganism:** *J Exp Med.* 4:521-34.
79. Manaka, A., Tokue, Y., and Murakami, M. 2017. **Comparison of 16S ribosomal RNA gene sequence analysis and conventional culture in the environmental survey of a hospital.** *Journal of Pharmaceutical Health Care and Sciences.* 3(1): 8.
80. Martin, J. D. and Mundt, J. O. 1972. **Enterococci in insects.** *Journal of Applied Microbiology.* 24(4):575–580. PubMed PMID: 4628796.
81. Mathur P. and Singh S. 2013. **Multidrug resistance in bacteria: a serious patient safety challenge for India.** *Journal of laboratory physicians.* 5(1):5–10. doi:10.4103/0974-2727.115898
82. Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T. L., Agarwala, R., and Schäffer, A. A. 2008. **Database indexing for production MegaBLAST searches.** *Bioinformatics (Oxford, England).* 24(16): 1757–1764. <https://doi.org/10.1093/bioinformatics/btn322>.
83. Morrison D. A. 2011. **Phylogenetic Analysis of Pathogens.** *Genetics and Evolution of Infectious Disease.* 203–231. DOI: 10.1016/b978-0-12-384890-1.00008-x
84. Mundt J. O. 1961. **Occurrence of Enterococci:** Bud., Blossom., and Soil Studies. *Journal of Applied Microbiology.* 9(6):541–544. PubMed PMID: 16349612.
85. Mundt J. O., Coggin J. H. Jr., and Johnson L. F. 1962. **Growth of *Streptococcus faecalis* var. liquefaciens on plants.** *Journal of Applied Microbiology.* 10(6):552–555. PubMed PMID: 13936612.
86. Mundt, J. O. 1963a. **Occurrence of enterococci in animals in a wild environment.** *Journal of Applied Microbiology.* 11(2):136–140. PubMed PMID: 13936610.

87. Mundt, J. O. 1963b. **Occurrence of enterococci on plants in a wild environment.** *Journal of Applied Microbiology*. 11(2):141–144. PubMed PMID: 13936611.
88. Mundt, J.O. and Graham, W.F. 1968. ***Streptococcus faecium* var. casseliflavus nov. var.** *J. Bacteriol.* 95: 2005–9.
89. Mundt, J.O. 1986. **Enterococci.** In: Sneath, P.H.A., Mair, N.S., Sharpe, M.E. y Holt, J.G. (eds), *Bergey's Manual of Systematic Bacteriology*, Vol. 2. Baltimore: Williams & Wilkins, pp. 1063–5.
90. Munjal, G., Hanmandlu, M., and Srivastava, S. 2019. **Phylogenetics Algorithms and Applications. In Advances in Intelligent Systems and Computing** (pp. 187–194). Springer, Singapore. https://doi.org/10.1007/978-981-13-5934-7_17
91. Murdoch, D. R., Corey, G. R., Hoen, B., Miró, J. M., Fowler, V. G., Jr, Bayer, A. S., Karchmer, A. W., Olaison, L., Pappas, P. A., Moreillon, P., Chambers, S. T., Chu, V. H., Falcó, V., Holland, D. J., Jones, P., Klein, J. L., Raymond, N. J., Read, K. M., Tripodi, M. F., Utili, R.. 2009. **International Collaboration on Endocarditis-Prospective Cohort Study (ICE-PCS) Investigators. Clinical presentation, etiology, and outcome of infective endocarditis in the 21st century: the International Collaboration on Endocarditis-Prospective Cohort Study.** *Archives of internal medicine*. 463–473. <https://doi.org/10.1001/archinternmed.2008.603>
92. Nagy A., Jiřinec T., Černíková L., Jiřincová H. and Havlíčková M. 2015. **Large-Scale Nucleotide Sequence Alignment and Sequence Variability Assessment to Identify the Evolutionarily Highly Conserved Regions for Universal Screening PCR Assay Design: An Example of Influenza A Virus.** En: Basu C. (eds) PCR Primer Design. Methods in Molecular Biology. vol 1275. Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-2365-6_4
93. Nallapareddy, S.R., Singh, K.V., Sillanpää, J., Garsin, D.A., Höök, M., Erlandsen, S.L. and Murray, B.E. 2006. **Endocarditis and biofilm-associated pili of *Enterococcus faecalis*.** *J Clin Invest*. 116:2799 –2807. <https://doi.org/10.1172/JCI29021>.
94. Nallapareddy, S.R., Singh, K.V., Okhuysen, P.C. and Murray, B.E. 2008. **A functional collagen adhesin gene, acm, in clinical isolates of *Enterococcus faecium* correlates with the recent success of this emerging nosocomial pathogen.** *Infect Immun*. 76:4110 – 4119. <https://doi.org/10.1128/IAI.00375-08>.

95. Naser, S., Thompson, F.L., Hoste, B. *et al.* 2005a. **Phylogeny and identification of enterococci by *atpA* gene sequence analysis.** *J. Clin. Microbiol.* 43: 2224–2230.
96. Naser, S.M., Thompson, F.L., Hoste, B. *et al.* 2005b. **Application of multilocus sequence analysis (MLSA) for rapid identification of *Enterococcus* species based on *rpoA* and *pheS* genes.** *Microbiol.* 151: 2141–50.
97. National Institute of Public Health (Czech Republic). **Alignment Explorer.** Recuperado 8 Mayo 2021, de <http://www1.szu.cz:8080/EntropyCalcWeb/>
98. Nguyen, L. T., Schmidt, H. A., von Haeseler, A., and Minh, B. Q. 2015. **IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies.** *Molecular biology and evolution.* 32(1):268–274. <https://doi.org/10.1093/molbev/msu300>
99. Okonechnikov, K., Golosova, O., Fursov, M., and UGENE team. 2012. **Unipro UGENE: a unified bioinformatics toolkit.** *Bioinformatics (Oxford, England).* 28(8): 1166–1167. <https://doi.org/10.1093/bioinformatics/bts091>.
100. OMS. 2017. **La OMS publica la lista de las bacterias para las que se necesitan urgentemente nuevos antibióticos.** Comunicado de prensa, Ginebra.
101. OMS. 2018. **Resistencia a los antibióticos.** Recuperado de <http://www.who.int/es/news-room/fact-sheets/detail/resistencia-a-los-antibi%C3%B3ticos>
102. O'Driscoll, T. and Crank, CW. 2015. **Infecciones por *Enterococcus* resistentes a la vancomicina: epidemiología, manifestaciones clínicas y manejo óptimo.** *Infección y resistencia a los medicamentos.* 8. 217. doi: 10.1016/j.idc.2016.07.009.
103. O'leary, N. A., Wright, M. W., Brister, J. R., Ciufu, S., Haddad, D., Mcveigh, R., Rajput, B., Robbertse, B., Smith-White, *et al.* 2016. **Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation.** *Nucleic acids research.* 44(D1): D733–D745. <https://doi.org/10.1093/nar/gkv1189>.
104. Ozawa, Y., Courvalin, P. and Gaiimand, M. 2000. **Identification of enterococci at the species level by sequencing of the genes for D-alanine: D-alanine ligases.** *Systematic and applied microbiology*, 23(2), 230–237. [https://doi.org/10.1016/s0723-2020\(00\)80009-0](https://doi.org/10.1016/s0723-2020(00)80009-0)

105. Parte, A. C., Sardà Carbasse, J., Meier-Kolthoff, J. P., Reimer, L. C., and Göker, M. 2020. **List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ.** *International Journal of Systematic and Evolutionary Microbiology*. ijsem004332.
106. Patti, J.M. and Hook, M. 1994. **Microbial adhesins recognizing extracellular matrix macromolecules.** *Curr Opin Cell Biol.* 6:752–758. [https://doi.org/10.1016/0955-0674\(94\)90104-X](https://doi.org/10.1016/0955-0674(94)90104-X).
107. Paulsen, I. T. L., Banerjee, G. S., Myers, K. E., Nelson, R., Seshadri, T. D., Read, D. E., Fouts, J. A., *et al.* 2003. **Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*:** *Science.* 299:2071-2074.
108. Peña, C. 2011. **Métodos de inferencia filogenética.** *Revista Peruana de Biología.* 18(2): 265 – 267.
109. Poyart, C., Quesnes, G. and Trieu-Cuot P. 2000. **Sequencing the gene encoding manganese-dependent superoxide dismutase for rapid species identification of enterococci.** *J Clin Microbiol.* 38:415– 418.
110. Price, M. N., Dehal, P. S., and Arkin, A. P. 2010. **FastTree 2—approximately maximum-likelihood trees for large alignments.** *PloS one.* 5(3): e9490.
111. Prabaker, K. and Weinstein, R. A. 2011. **Trends in antimicrobial resistance in intensive care units in the United States.** *Current Opinion in Critical Care.* 17(5):472–479. PubMed PMID: 21900766.
112. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F. O. 2013. **The SILVA ribosomal RNA gene database project: improved data processing and web-based tools.** *Nucleic acids research.* 41(Database issue): D590–D596. <https://doi.org/10.1093/nar/gks1219>.
113. Quintela-Baluja, M., Böhme, K., Fernández-No, I.C., Morandi, S., Alnakip, M.E., Caamaño-Antelo, S., Barros-Velázquez, J. and Calo-Mata, P. 2013. **Characterization of different food-isolated *Enterococcus* strains by MALDI-TOF mass fingerprinting.** *Electrophoresis.* 34:2240–2250. <https://doi.org/10.1002/elps.201200699>.
114. Rambaut, A., Drummond, A. J., Xie, D., Baele, G., and Suchard, M. A. 2018. **Posterior summarization in Bayesian phylogenetics using Tracer 1.7.** *Systematic biology.* 67(5): 901.

115. Rich, R.L., Kreikemeyer, B., Owens, R.T., LaBrenz, S., Narayana, S.V., Weinstock, G.M., Murray, B.E. and Hook, M. 1999. **Ace is a collagen-binding MSCRAMM from *Enterococcus faecalis*.** *J Biol Chem.* 274:26939 –26945. <https://doi.org/10.1074/jbc.274.38.26939>.
116. Richer, J.M., Rodriguez-Tello, E. and Vazquez-Ortiz, K.E. 2013. **Maximum Parsimony Phylogenetic Inference Using Simulated Annealing.** In: Schütze O. *et al.* (eds) *EVOLVE - A Bridge between Probability, Set Oriented Numerics, and Evolutionary Computation II. Advances in Intelligent Systems and Computing*, vol 175. Springer, Berlin, Heidelberg. https://doi-org.ezproxy.uniandes.edu.co:8443/10.1007/978-3-642-31519-0_12
117. Rodicio M.R. y Mendoza M.C. 2004. **Identificación bacteriana mediante secuenciación del ARNr 16S: fundamento, metodología y aplicaciones en microbiología clínica.** *Enfermedades Infecciosas y Microbiología Clínica.* 22 (4): 238-245.
118. Roshan, U. W., Moret, B. M. E., Warnow, T., and Williams, T. L. 2004. **Rec-I-DCM3: a fast algorithmic technique for reconstructing large phylogenetic trees.** *Proceedings IEEE Computational Systems Bioinformatics Conference.* 98-109. <https://doi.org/10.1109/csb.2004.1332422>
119. Ryu, H., Henson, M., Elk, M., Toledo-Hernandez, C., Griffith, J., Blackwood, D., Noble, R., Gourmelon, M., Glassmeyer, S., and Santo Domingo, J. W. 2013. **Development of quantitative PCR assays targeting the 16S rRNA genes of *Enterococcus* spp. and their application to the identification of enterococcus species in environmental samples.** *Applied and environmental microbiology*, 79(1), 196–204. <https://doi.org/10.1128/AEM.02802-12>
120. Saitou, N. and Nei, M. 1987. **The neighbor-joining method: a new method for reconstructing phylogenetic trees.** *Mol. Biol. Evol.* 4:406–425. <https://doi.org/10.1093/oxfordjournals.molbev.a040454>
121. Schleifer, K. H., and R. Kilpper-Bälz, 1984, **Transfer of *Streptococcus faecalis* and *Streptococcus faecium* to the Genus *Enterococcus* nom. rev. as *Enterococcus faecalis* comb. nov. and *Enterococcus faecium* comb. nov.** *International Journal of Systematic and Evolutionary Microbiology.* 34: 31-34.

122. Schwarz, G. 1978. **Estimating the dimension of a model.** *Annals of statistics.* 6(2):461-464.
123. Sghir, A., Gramet, G., Suaud, A., Violaine, R., Pochart, P. and Dore, J. 2000. **Quantification of bacterial groups within human fecal flora by oligonucleotide probe hybridization.** *Applied and Environmental Microbiology.* 66(5):2263–2266. PubMed PMID: 10788414.
124. Shimodaira, H. y Hasegawa, M. (1999) **Multiple comparisons of log-likelihoods with applications to phylogenetic inference.** *Molecular Biology and Evolution,* 16, 1114–1116.
125. Sillanpaa, J., Xu, Y., Nallapareddy, S.R., Murray, B.E. and Hook, M. 2004. **A family of putative MSCRAMMs from *Enterococcus faecalis*.** *Microbiology.* 150: 2069 –2078. <https://doi.org/10.1099/mic.0.27074-0>.
126. Sillanpaa, J., Prakash, V.P., Nallapareddy, S.R. and Murray, B.E. 2009. **Distribution of genes encoding MSCRAMMs and pili in clinical and natural populations of *Enterococcus faecium*.** *J Clin Microbiol.* 47:896–901. <https://doi.org/10.1128/JCM.02283-08>.
127. Singh, K.V., Nallapareddy, S.R., Sillanpaa, J. and Murray, B.E. 2010. **Importance of the collagen adhesin Ace in pathogenesis and protection against *Enterococcus faecalis* experimental endocarditis.** *PLoS Pathog.* 6: e1000716. <https://doi.org/10.1371/journal.ppat.1000716>.
128. Singhal, N., Kumar, M., Kanaujia, P.K. and Viridi, J.S. 2015. **MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis.** *Front Microbiol.* 6:791. <https://doi.org/10.3389/fmicb.2015.00791>.
129. Sistek, V., Maheux, A.F., Boissinot, M., Bernard, K.A., Cantin, P., Cleenwerck, I., *et al.*, 2012. ***Enterococcus ureasiticus* sp. nov. and *Enterococcus quebecensis* sp. nov., isolated from water.** *Int. J. Syst. Evol. Microbiol.* 62:1314–1320.
130. Slipczuk, L., Codolosa, J. N., Davila, C. D., Romero-Corral, A., Yun, J., Pressman, G. S., and Figueredo, V. M. 2013. **Infective endocarditis epidemiology over five decades: a systematic review.** *PloS one,* 8(12), e82665. <https://doi.org/10.1371/journal.pone.0082665>

131. Srinivasan, V. N., Li, G., Wang, D., Tooker, N. B., Dai, Z., Onnis-Hayden, A., Pinto, A. Gu, A. Z., 2019. **Oligotyping and Genome-Resolved Metagenomics Reveal Distinct Candidatus Accumulibacter Communities in Full-Scale Side-Stream versus Conventional Enhanced Biological Phosphorus Removal (EBPR) Configurations** *bioRxiv* 596692. <https://doi.org/10.1101/596692>
132. Stępień-Pyśniak, D., Hauschild, T., Rózański, P., and Marek, A. 2017. **MALDI-TOF Mass Spectrometry as a Useful Tool for Identification of *Enterococcus* spp. from Wild Birds and Differentiation of Closely Related Species.** *Journal of microbiology and biotechnology*, 27(6):1128–1137. <https://doi.org/10.4014/jmb.1612.12036>
133. Stöver, B. C., and Müller, K. F. 2010. **TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses.** *BMC bioinformatics.*, 11(1): 1-9.
134. Švec, P., Vancanneyt, M., Koort, J. *et al.* 2005. ***Enterococcus devriesei* sp. nov., associated with animal sources.** *Int. J. Syst. Evol. Microbiol.* 55: 2479–84.
135. Švec, P. and Franz, CMAP. 2014. **The genus *Enterococcus***, p 175–211. In Holzapfel WH, Wood BJB (ed), *Lactic acid bacteria: biodiversity and taxonomy*. John Wiley & Sons, Ltd, Chichester, England. <https://doi.org/10.1002/9781118655252.ch15>.
136. Švec, P. and Devriese, L.A. 2015. ***Enterococcus***, p 1–25. In Whitman WB, Rainey F, Kämpfer P, KeTrujillo M, Chun J, DeVos P, Hedlund B, Dedysh S (ed). *Bergey's manual of systematics of archaea and bacteria*. John Wiley & Sons, Ltd, Chichester, England. <https://doi.org/10.1002/9781118960608.gbm00600>
137. Swofford, D. L. 2002. PAUP* Ver 4.0. B10. **Phylogenetic analysis using parsimony and other methods.** *Sinauer and Associates, Sunderland, Massachusetts.*
138. Teixeira, L.M., Carvalho, M.D.G.S., Facklam, R.R. and Shewmaker, P.L. 2015. ***Enterococcus***, p 403– 421. In Jorgensen JH, Pfaller MA, Carroll KC, Funke G, Landry ML, Richter SS, Warnock DW (ed), *Manual of clinical microbiology*, 11th ed, vol 1. ASM Press, Washington, DC.
139. Telford, J.L., Barocchi, M.A., Margarit, I., Rappuoli, R. and Grandi, G. 2006. **Pili in gram-positive pathogens.** *Nat Rev Microbiol.* 4:509–519. <https://doi.org/10.1038/nrmicro1443>.
140. Thiercelin, M. E. 1899. **Morphologie et modes de reproduction de l'enterocoque.** *Comptes Rendus des Seances de la Societe de Biologie et des ses Filiales.* 11:551–553.

141. Thiercelin, M. E. and Jouhaud, L. 1899. **Sur un diplococque saprophyte de l'intestin susceptible de devenir pathogene.** *CR Soc. Biol.* 50:269–271.
142. Thiercelin, M.E. and Jouhaud, L. 1903. **Reproduction de l'enterocoque: taches centrales; granulations peripheriques et microblastes.** *C. R. S Θ ances Soc. Biol.* **55**: 686–8.
143. Troncoso, C., Pavez, M., Santos, A., Salazar, R. y Barrientos, L. 2017. **Implicancias Estructurales y Fisiológicas de la Célula Bacteriana en los Mecanismos de Resistencia Antibiótica.** *International Journal of Morphology.* 35(4):1214-1223. DOI: 10.4067/S0717-95022017000401214.
144. Tsiodras, S., Gold, H.S., Coakley, E.P.G., Wennersten, C., Moellering, R.C. Jr., and Eliopoulos, G.M. 2000. **Diversity of domain V of 23S rRNA gene sequence in different *Enterococcus* species.** *J. Clin. Microbiol.* 38: 3991–3.
145. Ubeda, C., Taur, Y., Jeng, R. R., Equinda, M. J., Son, T., Samstein, M., *et al.* 2010. ***Enterococcus* domination of intestinal microbiota is enabled by antibiotic treatment in mice and precedes bloodstream invasion in humans.** *The Journal of Clinical Investigation.* 120(12):4332–4341. PubMed PMID: 21099116.
146. Vasileiadis, S., Puglisi, E., Arena, M., Cappa, F., Cocconcelli, P.S. and Trevisan, M. 2012. **Soil bacterial diversity screening using single 16S rRNA gene V regions coupled with multi-million read generating sequencing technologies.** *PLoS One.* 7 (8): e42671. doi: 10.1371/journal.pone.0042671.
147. Vinje, H., Almøy, T., Liland, K.H., *et al.* 2014. **A systematic search for discriminating sites in the 16S ribosomal RNA gene.** *Microb Informatics Exp*4, 2. <https://doi.org/10.1186/2042-5783-4-2>
148. Wang, Y., and Qian, P. Y. 2009. **Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies.** *PloS one*, 4(10), e7401.
149. Ward, D.M. 1998. **A natural species concept for prokaryotes.** *Curr Opin Microbiol.* 1:271–277.
150. Weiner, L.M., Webb, A.K., Limbago, B., Dudeck, M.A, Patel, J., Kallen, A.J., Edwards, J.R. and Sievert, D.M. 2016. **Antimicrobial-resistant pathogens associated with healthcare-associated infections: summary of data reported to the National**

- Healthcare Safety Network at the Centers for Disease Control and Prevention, 2011-2014.** *Infect Control Hosp Epidemiol.* 37:1288–1301. <https://doi.org/10.1017/ice.2016.174>.
151. Werner, G., Coque, T. M., Hammerum, A. M., Hope, R., Hryniewicz, W., Johnson, A., *et al.* 2008. **Emergence and spread of vancomycin resistance among enterococci in Europe.** *Eurosurveillance.* 13 (47) PubMed PMID: 19021959.
152. Williams, A.M., Rodrigues, U.M. and Collins, M.D. 1991. **Intragenetic relationships of enterococci as determined by reverse transcriptase sequencing of small-subunit rRNA.** *Res. Microbiol.* 142: 67–74.
153. Yang, B., Wang, Y., and Qian, P. Y. 2016. **Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis.** *BMC bioinformatics.* 17:135. <https://doi.org/10.1186/s12859-016-0992-y>.
154. Yang, Z. and Rannala, B. 2012. **Molecular phylogenetics: principles and practice.** *Nature Review Genetics* 13: 303–314. DOI: 10.1038/nrg3186
155. Zhong, Z., Zhang, W., Song, Y., Liu, W., Xu, H., Xi, X., and Sun, Z. 2017. **Comparative genomic analysis of the genus *Enterococcus*.** *Microbiological research,* 196, 95-105.
156. Zwickl, D.J. 2006. **Genetic Algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion** (Publication No. 85839175) [Thesis, University of Texas at Austin]. UT Electronic Theses and Dissertations.

XIII. ANEXOS

13.1. Características del alineamiento de secuencias del conjunto taxonómico

Parámetros	Resultados
Datos de entrada	142 secuencias con 1253 sitios
Número de sitios constantes	957 (= 76.3767% de todos los sitios)
Número de sitios invariantes (constante o constante ambiguo)	957 (= 76.3767% de todos los sitios)
Número de sitios informativos parsimoniosos	185
Número de sitios de patrones distintos	305

13.2. Resultados de la construcción filogenética

Árbol utilizado para ModelFinder	Resultado
Log de verosimilitud del árbol	-6569.0501 (s.e. 304.9008)
Log de verosimilitud sin restricciones (sin árbol)	-3974.6649
Número de parámetros libres (#ramas + #parámetros del modelo)	264
Criterio de información Akaike	13666.1003
Criterio de Información Akaike Corregido	13807.7197
Criterio de información Bayesiano	15021.2904
Longitud total del árbol (suma de las longitudes de las ramas)	1.0124
Suma de la longitud de las ramas internas	0.4863 (48.0317% de longitud del árbol)

13.3. Lista de modelos de sustitución de nucleótidos ordenados por puntajes CIB

Modelo	LogL	AIC	w-AIC	AICc	w-AICc	CIB	w- CIB
TPM2+F+I+G4	-6569.2354	13666.4708	-0.028	13808.0903	0.1019	15021.661	0.4541
TPM2u+F+I+G4	-6569.2471	13666.4942	-0.0277	13808.1137	0.1007	15021.6844	0.4488
TIM2+F+I+G4	-6567.9544	13665.9088	-0.0371	13808.7456	0.0734	15026.2322	-0.0462
TVM+F+I+G4	-6564.5829	13661.1658	0.3974	13805.2267	0.4267	15026.6226	-0.038
GTR+F+I+G4	-6563.3349	13660.6698	0.5093	13805.9611	0.2955	15031.2598	-0.0037
TPM3u+F+I+G4	-6574.4226	13676.8452	-0.0002	13818.4646	-0.0006	15032.0353	-0.0025
TPM3+F+I+G4	-6574.433	13676.866	-0.0002	13818.4855	-0.0006	15032.0562	-0.0025
HKY+F+I+G4	-6578.1198	13682.2396	0	13822.6481	-0.0001	15032.2965	-0.0022
K2P+I+G4	-6589.8083	13699.6166	0	13836.4311	0	15034.2735	-0.0008
TIM3+F+I+G4	-6573.109	13676.2181	-0.0002	13819.0549	-0.0004	15036.5415	-0.0003
TN+F+I+G4	-6576.8249	13681.6498	0	13823.2692	-0.0001	15036.8399	-0.0002
TIM2e+I+G4	-6584.109	13692.2179	0	13831.4219	0	15037.1414	-0.0002
TVMe+I+G4	-6581.0557	13688.1114	0	13828.5199	0	15038.1682	-0.0001
K3Pu+F+I+G4	-6577.9919	13683.9838	0	13825.6032	0	15039.1739	-0.0001
TNe+I+G4	-6588.7911	13699.5822	0	13837.5882	0	15039.3724	-0.0001
TIM3e+I+G4	-6585.6416	13695.2832	0	13834.4872	0	15040.2067	0
K3P+I+G4	-6589.5696	13701.1391	0	13839.1452	0	15040.9294	0
SYM+I+G4	-6580.0003	13688.0007	0	13829.6201	0	15043.1908	0
TIM+F+I+G4	-6576.6389	13683.2778	0	13826.1147	0	15043.6012	0
TIMe+I+G4	-6588.5364	13701.0728	0	13840.2768	0	15045.9963	0
TPM2u+F+G4	-6642.3037	13810.6074	0	13951.0159	0	15160.6642	0
TPM2+F+G4	-6642.3041	13810.6082	0	13951.0167	0	15160.665	0
TVM+F+G4	-6636.7383	13803.4765	0	13946.3134	0	15163.7999	0
TIM2+F+G4	-6641.7667	13811.5335	0	13953.1529	0	15166.7236	0
GTR+F+G4	-6636.2454	13804.4909	0	13948.5517	0	15169.9476	0
TPM3+F+G4	-6647.2076	13820.4153	0	13960.8238	0	15170.4721	0
TPM3u+F+G4	-6647.208	13820.416	0	13960.8245	0	15170.4729	0
HKY+F+G4	-6651.8774	13827.7547	0	13966.9588	0	15172.6783	0
TIM2e+G4	-6657.5021	13837.0042	0	13975.0103	0	15176.7945	0
TIM3+F+G4	-6646.8164	13821.6328	0	13963.2522	0	15176.8229	0
K2P+G4	-6664.8429	13847.6858	0	13983.3152	0	15177.2094	0
TVMe+G4	-6654.9366	13833.8732	0	13973.0772	0	15178.7967	0
TIM3e+G4	-6658.5611	13839.1222	0	13977.1282	0	15178.9124	0
TN+F+G4	-6651.4289	13828.8577	0	13969.2662	0	15178.9146	0
K3Pu+F+G4	-6651.6463	13829.2926	0	13969.7011	0	15179.3494	0
TNe+G4	-6662.4873	13844.9746	0	13981.7891	0	15179.6316	0
SYM+G4	-6652.6803	13831.3606	0	13971.7691	0	15181.4174	0
K3P+G4	-6664.5424	13849.0848	0	13985.8993	0	15183.7417	0
TIM+F+G4	-6651.1994	13830.3988	0	13972.0182	0	15185.5889	0
TIMe+G4	-6662.1978	13846.3957	0	13984.4017	0	15186.1859	0
TPM2u+F+I	-6729.4392	13984.8784	0	14125.2869	0	15334.9352	0
TPM2+F+I	-6729.441	13984.882	0	14125.2905	0	15334.9388	0

TVM+F+I	-6722.6522	13975.3045	0	14118.1414	0	15335.6279	0
TIM2+F+I	-6728.482	13984.964	0	14126.5835	0	15340.1542	0
GTR+F+I	-6721.7317	13975.4635	0	14119.5243	0	15340.9202	0
TPM3+F+I	-6734.2912	13994.5823	0	14134.9908	0	15344.6392	0
TPM3u+F+I	-6734.2937	13994.5874	0	14134.9959	0	15344.6442	0
HKY+F+I	-6740.1744	14004.3487	0	14143.5528	0	15349.2723	0
TIM3+F+I	-6733.4172	13994.8344	0	14136.4539	0	15350.0246	0
TVMe+I	-6741.4654	14006.9308	0	14146.1348	0	15351.8543	0
F81+F+I+G4	-6742.1414	14008.2828	0	14147.4868	0	15353.2063	0
TN+F+I	-6739.2509	14004.5018	0	14144.9103	0	15354.5586	0
TIM2e+I	-6746.4565	14014.9131	0	14152.9192	0	15354.7033	0
K3Pu+F+I	-6739.9658	14005.9316	0	14146.3401	0	15355.9885	0
SYM+I	-6740.3262	14006.6525	0	14147.0609	0	15356.7093	0
K2P+I	-6754.6687	14027.3375	0	14162.9669	0	15356.8611	0
TIM3e+I	-6748.3023	14018.6046	0	14156.6107	0	15358.3949	0
JC+I+G4	-6756.2683	14030.5366	0	14166.166	0	15360.0602	0
TIM+F+I	-6739.046	14006.0921	0	14147.7115	0	15361.2822	0
TNe+I	-6753.5134	14027.0267	0	14163.8412	0	15361.6837	0
K3P+I	-6754.429	14028.8579	0	14165.6724	0	15363.5149	0
Time+I	-6753.2818	14028.5637	0	14166.5697	0	15368.3539	0
F81+F+G4	-6814.9762	14151.9523	0	14289.9584	0	15491.7426	0
JC+G4	-6828.567	14173.134	0	14307.5847	0	15497.5243	0
F81+F+I	-6898.7555	14319.511	0	14457.517	0	15659.3012	0
JC+I	-6913.6296	14343.2593	0	14477.71	0	15667.6496	0
GTR+F	-7317.1024	15164.2047	0	15307.0416	0	16524.5281	0
TIM2+F	-7327.2406	15180.4813	0	15320.8897	0	16530.5381	0
TVM+F	-7325.852	15179.704	0	15321.3234	0	16534.8941	0
TIM3+F	-7329.8583	15185.7166	0	15326.1251	0	16535.7734	0
TPM2u+F	-7335.9524	15195.9047	0	15335.1088	0	16540.8283	0
TPM2+F	-7335.9535	15195.907	0	15335.1111	0	16540.8306	0
TPM3+F	-7338.6644	15201.3288	0	15340.5328	0	16546.2523	0
TPM3u+F	-7338.6661	15201.3323	0	15340.5363	0	16546.2558	0
TN+F	-7338.9199	15201.8397	0	15341.0437	0	16546.7632	0
TVMe	-7343.235	15208.4699	0	15346.476	0	16548.2602	0
TIM+F	-7338.6061	15203.2122	0	15343.6207	0	16553.2691	0
SYM	-7342.3409	15208.6818	0	15347.8859	0	16553.6054	0
TIM2e	-7351.2216	15222.4433	0	15359.2578	0	16557.1002	0
HKY+F	-7347.6911	15217.3822	0	15355.3882	0	16557.1724	0
TIM3e	-7352.146	15224.2919	0	15361.1065	0	16558.9489	0
K2P	-7360.8656	15237.7312	0	15372.1819	0	16562.1216	0
K3Pu+F	-7347.3707	15218.7413	0	15357.9454	0	16563.6649	0
TNe	-7359.9612	15237.9224	0	15373.5518	0	16567.446	0
K3P	-7360.572	15239.1441	0	15374.7735	0	16568.6677	0
Time	-7359.677	15239.354	0	15376.1686	0	16574.011	0
F81+F	-7493.8148	15507.6296	0	15644.4441	0	16842.2866	0

JC	-7514.1081	15542.2161	0	15675.4945	0	16861.4732	0
----	------------	------------	---	------------	---	------------	---

En el cuadro se enlistan todos los modelos probados para la construcción del árbol por Máxima Verosimilitud, ordenado con base en el mejor puntaje CIB. **AIC**, **w-AIC**: Puntaje de Criterio de información de Akaike y ponderado **AICc**, **w-AICc**: Puntaje de AIC corregido puntajes y ponderado. CIB, **w-** CIB: Puntaje de Criterio de información Bayesiano y ponderado.

13.4. Clasificación mediante K-medias conjunto SILVA

Pos	H(i)	Grupo						
1	0.376	V	40	0.296	V	80	0.297	V
2	0.205	V	41	0.409	V	81	0.588	AV
3	0.311	V	42	0.41	V	82	0.186	V
4	0.389	V	43	0.358	V	83	0.615	AV
5	0.145	V	44	0.487	AV	84	0.67	AV
6	0.395	V	45	0.256	V	85	0.183	V
7	0.33	V	46	0.552	AV	86	0.58	AV
8	0.557	AV	47	0.259	V	87	0.203	V
9	0.089	C	48	0.598	AV	88	0.032	C
10	0.111	C	49	0.586	AV	89	0.034	C
11	0.087	C	50	0.613	AV	90	0.032	C
12	0.09	C	51	0.724	AV	91	0.021	C
13	0.02	C	52	0.307	V	92	0.231	V
14	0.057	C	53	0.138	V	93	0.532	AV
15	0.147	V	54	0.176	V	94	0.29	V
16	0.09	C	55	0.334	V	95	0.497	AV
17	0.163	V	56	0.141	V	96	0.529	AV
18	0.144	V	57	0.299	V	97	0.663	AV
19	0.225	V	58	0.219	V	98	0.627	AV
20	0.284	V	59	0.252	V	99	0.707	AV
21	0.117	C	60	0.285	V	100	0.975	AV
22	0.207	V	61	0.084	C	101	0.619	AV
23	0.08	C	62	0.069	C	102	0.535	AV
24	0.15	V	63	0.077	C	103	0.562	AV
25	0.174	V	64	0.124	C	104	0.074	C
26	0.154	V	65	0.088	C	105	0.022	C
27	0.143	V	66	0.589	AV	106	0.512	AV
28	0.1	C	67	0.089	C	107	0.044	C
29	0.052	C	68	0.024	C	108	0.749	AV
30	0.24	V	69	0.054	C	109	0.051	C
31	0.265	V	70	0.054	C	110	0.051	C
32	0.238	V	71	0.238	V	111	0.022	C
33	0.257	V	72	0.103	C	112	0.043	C
34	0.238	V	73	0.539	AV	113	0.071	C
35	0.315	V	74	0.063	C	114	0.039	C
36	0.266	V	75	0.097	C	115	0.085	C
37	0.373	V	76	0.576	AV	116	0.042	C
38	0.377	V	77	0.485	AV	117	0.539	AV
39	0.304	V	78	0.626	AV	118	0.011	C
			79	0.724	AV	119	0.031	C

120	0.011	C	165	0.02	C	210	0.842	AV
121	0.011	C	166	0.021	C	211	0.743	AV
122	0.02	C	167	0	C	212	0.465	AV
123	0.022	C	168	0	C	213	0.981	AV
124	0	C	169	0.038	C	214	0.837	AV
125	0	C	170	0	C	215	0.826	AV
126	0.011	C	171	0.011	C	216	0.837	AV
127	0	C	172	0	C	217	0.671	AV
128	0.011	C	173	0	C	218	0.021	C
129	0.028	C	174	0	C	219	0	C
130	0	C	175	0	C	220	0.011	C
131	0.011	C	176	0.019	C	221	0	C
132	0	C	177	0.021	C	222	0.032	C
133	0	C	178	0	C	223	0	C
134	0.02	C	179	0	C	224	0.203	V
135	0.011	C	180	0.011	C	225	0.041	C
136	0	C	181	0	C	226	0.021	C
137	0.011	C	182	0	C	227	0.011	C
138	0.022	C	183	0.011	C	228	0.035	C
139	0	C	184	0.011	C	229	0.035	C
140	0.022	C	185	0	C	230	0.035	C
141	0.011	C	186	0.011	C	231	0.762	AV
142	0.02	C	187	0	C	232	0.802	AV
143	0	C	188	0.58	AV	233	0.559	AV
144	0.538	AV	189	0.011	C	234	0.021	C
145	0	C	190	0.03	C	235	0.526	AV
146	0	C	191	0.011	C	236	0.011	C
147	0	C	192	0.03	C	237	0.281	V
148	0.069	C	193	0.295	V	238	0.021	C
149	0.067	C	194	0.249	V	239	0.615	AV
150	0.088	C	195	0.916	AV	240	0.027	C
151	0.02	C	196	0.251	V	241	0	C
152	0.063	C	197	0.97	AV	242	0.079	C
153	0.19	V	198	1.031	AV	243	0.091	C
154	0.581	AV	199	0.493	AV	244	0.079	C
155	0.039	C	200	0.76	AV	245	0.011	C
156	0	C	201	0.839	AV	246	0	C
157	0	C	202	0.188	V	247	0	C
158	0.028	C	203	0.021	C	248	0.011	C
159	0.022	C	204	0.098	C	249	0	C
160	0	C	205	0.011	C	250	0.011	C
161	0	C	206	0.038	C	251	0	C
162	0	C	207	0.042	C	252	0.011	C
163	0	C	208	0.021	C	253	0.021	C
164	0	C	209	0.182	V	254	0.019	C

255	0.03	C	300	0.041	C	345	0.066	C
256	0	C	301	0.059	C	346	0.019	C
257	0.011	C	302	0.061	C	347	0.011	C
258	0.035	C	303	0.08	C	348	0.039	C
259	0	C	304	0.03	C	349	0.021	C
260	0.011	C	305	0.011	C	350	0	C
261	0.011	C	306	0	C	351	0.038	C
262	0.011	C	307	0.011	C	352	0.03	C
263	0	C	308	0.054	C	353	0	C
264	0.011	C	309	0	C	354	0.027	C
265	0.011	C	310	0.021	C	355	0.011	C
266	0.089	C	311	0.049	C	356	0.653	AV
267	0.049	C	312	0.021	C	357	0.021	C
268	0.03	C	313	0.021	C	358	0.021	C
269	0.021	C	314	0.021	C	359	0.019	C
270	0.038	C	315	0.011	C	360	0.011	C
271	0.021	C	316	0.019	C	361	0.011	C
272	0	C	317	0.011	C	362	0	C
273	0.011	C	318	0.042	C	363	0.019	C
274	0.019	C	319	0.038	C	364	0.011	C
275	0.041	C	320	0.011	C	365	0.011	C
276	0.129	V	321	0.021	C	366	0.011	C
277	0.194	V	322	0.03	C	367	0.019	C
278	0.038	C	323	0.027	C	368	0.011	C
279	0.03	C	324	0.027	C	369	0	C
280	0.021	C	325	0.011	C	370	0.019	C
281	0.03	C	326	0	C	371	0.011	C
282	0.021	C	327	0.011	C	372	0.011	C
283	0.11	C	328	0.027	C	373	0.063	C
284	0.011	C	329	0.021	C	374	0	C
285	0.021	C	330	0.019	C	375	0.019	C
286	0.021	C	331	0.021	C	376	0.03	C
287	0.176	V	332	0.021	C	377	0	C
288	0.177	V	333	0	C	378	0	C
289	0.055	C	334	0.065	C	379	0.021	C
290	0.027	C	335	0.673	AV	380	0	C
291	0.019	C	336	0.019	C	381	0.011	C
292	0.03	C	337	0.011	C	382	0.011	C
293	0.021	C	338	0.019	C	383	0.011	C
294	0	C	339	0.011	C	384	0	C
295	0.047	C	340	0.019	C	385	0.011	C
296	0.041	C	341	0.027	C	386	0.011	C
297	0.906	AV	342	0.032	C	387	0.011	C
298	0.118	C	343	0.021	C	388	0.011	C
299	0.172	V	344	0.019	C	389	0.03	C

390	0	C	435	0.039	C	480	0.487	AV
391	0	C	436	0.021	C	481	0.487	AV
392	0.011	C	437	0.011	C	482	0.011	C
393	0.03	C	438	0.03	C	483	0.123	C
394	0.021	C	439	0.019	C	484	0.102	C
395	0	C	440	0.011	C	485	0.124	C
396	0	C	441	0.021	C	486	0.666	AV
397	0.038	C	442	0.011	C	487	0	C
398	0.011	C	443	0.019	C	488	0.522	AV
399	0.063	C	444	0.03	C	489	0.284	V
400	0.18	V	445	0.011	C	490	0.019	C
401	0	C	446	0	C	491	0.621	AV
402	0.019	C	447	0.011	C	492	0.466	AV
403	0.038	C	448	0	C	493	0.011	C
404	0.011	C	449	0.011	C	494	0.615	AV
405	0.042	C	450	0	C	495	0.268	V
406	0.011	C	451	0	C	496	0.168	V
407	0.011	C	452	0.011	C	497	0	C
408	0.011	C	453	0.021	C	498	0.019	C
409	0.027	C	454	0	C	499	0.498	AV
410	0.011	C	455	0.019	C	500	0.038	C
411	0	C	456	0.011	C	501	0.011	C
412	0.019	C	457	0	C	502	0	C
413	0.049	C	458	0.011	C	503	0	C
414	0	C	459	0	C	504	0	C
415	0	C	460	0.054	C	505	0	C
416	0	C	461	0.021	C	506	0.011	C
417	0.011	C	462	0.011	C	507	0.045	C
418	0.068	C	463	0	C	508	0	C
419	0.011	C	464	0.03	C	509	0	C
420	0.035	C	465	0.011	C	510	0.011	C
421	0.011	C	466	0.019	C	511	0.035	C
422	0	C	467	0.011	C	512	0.021	C
423	0.038	C	468	0.042	C	513	0.011	C
424	0	C	469	0.091	C	514	0.011	C
425	0	C	470	0.03	C	515	0.019	C
426	0.027	C	471	0	C	516	0.038	C
427	0.039	C	472	0.062	C	517	0	C
428	0	C	473	0.011	C	518	0	C
429	0.019	C	474	0.072	C	519	0.011	C
430	0.011	C	475	0.042	C	520	0.032	C
431	0.042	C	476	0.172	V	521	0.011	C
432	0.075	C	477	0.316	V	522	0.019	C
433	0.019	C	478	0.61	AV	523	0	C
434	0	C	479	0.03	C	524	0.03	C

525	0.011	C	570	0.011	C	615	0.011	C
526	0	C	571	0.038	C	616	0.021	C
527	0.03	C	572	0.011	C	617	0.027	C
528	0.011	C	573	0.019	C	618	0.011	C
529	0	C	574	0.011	C	619	0.011	C
530	0.019	C	575	0	C	620	0	C
531	0.021	C	576	0.011	C	621	0.011	C
532	0	C	577	0	C	622	0	C
533	0	C	578	0.03	C	623	0	C
534	0.011	C	579	0.021	C	624	0	C
535	0.045	C	580	0.019	C	625	0	C
536	0.021	C	581	0.011	C	626	0	C
537	0	C	582	0.021	C	627	0	C
538	0.011	C	583	0.03	C	628	0.021	C
539	0.011	C	584	0.011	C	629	0.011	C
540	0	C	585	0.049	C	630	0.041	C
541	0.021	C	586	0.011	C	631	0	C
542	0	C	587	0	C	632	0	C
543	0.011	C	588	0.011	C	633	0.011	C
544	0.041	C	589	0	C	634	0.076	C
545	0.011	C	590	0.011	C	635	0.019	C
546	0.011	C	591	0	C	636	0	C
547	0	C	592	0.019	C	637	0.011	C
548	0.011	C	593	0	C	638	0.021	C
549	0.019	C	594	0.038	C	639	0.011	C
550	0.011	C	595	0.011	C	640	0.143	V
551	0.011	C	596	0.011	C	641	0.011	C
552	0	C	597	0.011	C	642	0.045	C
553	0	C	598	0.011	C	643	0.019	C
554	0	C	599	0.019	C	644	0.011	C
555	0.011	C	600	0	C	645	0.021	C
556	0.011	C	601	0	C	646	0.062	C
557	0.03	C	602	0.021	C	647	0.011	C
558	0.011	C	603	0.011	C	648	0.011	C
559	0	C	604	0.078	C	649	0.019	C
560	0.011	C	605	0.011	C	650	0	C
561	0	C	606	0.011	C	651	0	C
562	0	C	607	0.019	C	652	0.011	C
563	0.021	C	608	0.011	C	653	0.021	C
564	0.011	C	609	0	C	654	0	C
565	0	C	610	0.133	V	655	0.019	C
566	0.03	C	611	0.128	V	656	0.019	C
567	0.011	C	612	0.223	V	657	0	C
568	0.019	C	613	0	C	658	0	C
569	0	C	614	0.019	C	659	0.019	C

660	0.019	C	705	0.011	C	750	0.094	C
661	0.011	C	706	0.011	C	751	0.021	C
662	0.03	C	707	0.03	C	752	0	C
663	0.021	C	708	0.03	C	753	0.019	C
664	0.019	C	709	0	C	754	0.239	V
665	0	C	710	0.011	C	755	0.011	C
666	0.027	C	711	0	C	756	0.041	C
667	0.102	C	712	0.021	C	757	0.129	V
668	0.118	C	713	0	C	758	0.011	C
669	0.325	V	714	0.011	C	759	0.019	C
670	0.032	C	715	0.021	C	760	0	C
671	0	C	716	0.011	C	761	0.011	C
672	0.019	C	717	0.011	C	762	0.019	C
673	0	C	718	0	C	763	0.019	C
674	0	C	719	0	C	764	0.011	C
675	0.021	C	720	0.032	C	765	0	C
676	0.021	C	721	0	C	766	0	C
677	0.032	C	722	0	C	767	0.027	C
678	0.038	C	723	0.096	C	768	0.03	C
679	0.035	C	724	0.011	C	769	0.03	C
680	0.048	C	725	0	C	770	0.011	C
681	0.041	C	726	0.011	C	771	0	C
682	0.03	C	727	0	C	772	0.019	C
683	0.021	C	728	0	C	773	0.011	C
684	0.019	C	729	0.011	C	774	0.011	C
685	0	C	730	0.011	C	775	0.032	C
686	0.049	C	731	0.048	C	776	0.019	C
687	0.027	C	732	0	C	777	0	C
688	0.027	C	733	0.038	C	778	0.03	C
689	0.113	C	734	0.011	C	779	0.03	C
690	0.03	C	735	0.019	C	780	0.07	C
691	0	C	736	0	C	781	0.021	C
692	0.097	C	737	0	C	782	0.032	C
693	0.011	C	738	0.021	C	783	0.011	C
694	0.019	C	739	0.011	C	784	0.03	C
695	0.011	C	740	0.021	C	785	0.039	C
696	0.021	C	741	0.021	C	786	0.041	C
697	0.03	C	742	0.066	C	787	0.019	C
698	0.042	C	743	0	C	788	0.019	C
699	0	C	744	0	C	789	0.03	C
700	0	C	745	0	C	790	0.011	C
701	0.03	C	746	0	C	791	0.046	C
702	0.011	C	747	0.019	C	792	0.021	C
703	0.027	C	748	0.011	C	793	0.021	C
704	0.019	C	749	0.019	C	794	0.019	C

795	0.019	C	840	0.011	C	885	0	C
796	0.011	C	841	0.011	C	886	0	C
797	0.011	C	842	0.021	C	887	0.011	C
798	0.035	C	843	0.019	C	888	0.038	C
799	0.021	C	844	0	C	889	0.011	C
800	0.307	V	845	0	C	890	0.03	C
801	0.032	C	846	0.054	C	891	0.041	C
802	0.045	C	847	0.011	C	892	0.011	C
803	0.03	C	848	0.011	C	893	0.011	C
804	0.011	C	849	0.021	C	894	0	C
805	0.039	C	850	0.03	C	895	0.045	C
806	0.011	C	851	0.011	C	896	0.011	C
807	0.021	C	852	0.011	C	897	0.011	C
808	0.035	C	853	0	C	898	0.049	C
809	0.011	C	854	0	C	899	0.045	C
810	0.045	C	855	0	C	900	0	C
811	0.046	C	856	0.021	C	901	0.039	C
812	0.011	C	857	0.019	C	902	0.054	C
813	0	C	858	0.011	C	903	0.027	C
814	0.019	C	859	0.011	C	904	0.046	C
815	0	C	860	0	C	905	0.03	C
816	0.021	C	861	0	C	906	0.021	C
817	0.041	C	862	0.021	C	907	0.027	C
818	0.011	C	863	0.021	C	908	0.021	C
819	0.011	C	864	0.021	C	909	0	C
820	0.011	C	865	0	C	910	0.027	C
821	0	C	866	0.019	C	911	0.027	C
822	0.011	C	867	0.035	C	912	0.042	C
823	0.032	C	868	0	C	913	0.032	C
824	0.011	C	869	0.03	C	914	0.011	C
825	0.061	C	870	0.019	C	915	0.041	C
826	0.011	C	871	0	C	916	0.143	V
827	0.011	C	872	0.019	C	917	0.066	C
828	0	C	873	0.011	C	918	0	C
829	0.011	C	874	0.03	C	919	0.045	C
830	0.021	C	875	0	C	920	0.011	C
831	0	C	876	0	C	921	0.019	C
832	0.011	C	877	0.041	C	922	0.057	C
833	0.03	C	878	0.011	C	923	0.072	C
834	0.03	C	879	0.019	C	924	0.019	C
835	0.038	C	880	0.011	C	925	0.08	C
836	0.048	C	881	0.011	C	926	0.019	C
837	0.011	C	882	0.019	C	927	0	C
838	0.011	C	883	0.054	C	928	0.011	C
839	0	C	884	0.601	AV	929	0.011	C

930	0	C	975	0	C	1020	0.011	C
931	0.049	C	976	0.011	C	1021	0.03	C
932	0.011	C	977	0.011	C	1022	0.011	C
933	0.038	C	978	0.011	C	1023	0.011	C
934	0.019	C	979	0.011	C	1024	0.011	C
935	0.03	C	980	0.019	C	1025	0	C
936	0.047	C	981	0	C	1026	0.081	C
937	0.019	C	982	0.011	C	1027	0.011	C
938	0.035	C	983	0.191	V	1028	0.072	C
939	0.027	C	984	0.045	C	1029	0.143	V
940	0.045	C	985	0.011	C	1030	0.118	C
941	0.038	C	986	0	C	1031	0.049	C
942	0.019	C	987	0.019	C	1032	0.072	C
943	0.011	C	988	0.021	C	1033	0.045	C
944	0.011	C	989	0.027	C	1034	0.07	C
945	0.011	C	990	0.021	C	1035	0	C
946	0.03	C	991	0	C	1036	0.042	C
947	0	C	992	0.019	C	1037	0	C
948	0.011	C	993	0.035	C	1038	0.045	C
949	0	C	994	0.035	C	1039	0.072	C
950	0.113	C	995	0.011	C	1040	0.042	C
951	0.038	C	996	0.011	C	1041	0.102	C
952	0	C	997	0.055	C	1042	0.143	V
953	0.011	C	998	0.038	C	1043	0.123	C
954	0	C	999	0.108	C	1044	0	C
955	0	C	1000	0	C	1045	0.011	C
956	0.086	C	1001	0.019	C	1046	0.606	AV
957	0	C	1002	0.03	C	1047	0.03	C
958	0.011	C	1003	0.011	C	1048	0	C
959	0.021	C	1004	0	C	1049	0.019	C
960	0.038	C	1005	0	C	1050	0.049	C
961	0	C	1006	0.021	C	1051	0.011	C
962	0.021	C	1007	0.03	C	1052	0.068	C
963	0.011	C	1008	0.011	C	1053	0	C
964	0.011	C	1009	0.011	C	1054	0	C
965	0.011	C	1010	0.03	C	1055	0.021	C
966	0	C	1011	0.011	C	1056	0.011	C
967	0	C	1012	0	C	1057	0.615	AV
968	0.03	C	1013	0.019	C	1058	0.03	C
969	0	C	1014	0.011	C	1059	0	C
970	0.011	C	1015	0.035	C	1060	0.011	C
971	0.011	C	1016	0	C	1061	0.019	C
972	0	C	1017	0.011	C	1062	0.027	C
973	0.011	C	1018	0.019	C	1063	0.019	C
974	0.021	C	1019	0.011	C	1064	0.021	C

1065	0.085	C	1110	0.021	C	1155	0.011	C
1066	0.011	C	1111	0.019	C	1156	0.011	C
1067	0.011	C	1112	0.019	C	1157	0.693	AV
1068	0.011	C	1113	0.011	C	1158	0.011	C
1069	0	C	1114	0.03	C	1159	0.055	C
1070	0.019	C	1115	0.047	C	1160	0.019	C
1071	0.011	C	1116	0.042	C	1161	0.019	C
1072	0.021	C	1117	0.011	C	1162	0.03	C
1073	0.019	C	1118	0.011	C	1163	0	C
1074	0	C	1119	0	C	1164	0.085	C
1075	0	C	1120	0.021	C	1165	0.038	C
1076	0.011	C	1121	0.011	C	1166	0.019	C
1077	0.011	C	1122	0.011	C	1167	0.055	C
1078	0.022	C	1123	0	C	1168	0.019	C
1079	0	C	1124	0.011	C	1169	0.035	C
1080	0.03	C	1125	0.019	C	1170	0.019	C
1081	0.021	C	1126	0.011	C	1171	0.011	C
1082	0	C	1127	0.041	C	1172	0.076	C
1083	0.019	C	1128	0.021	C	1173	0	C
1084	0.019	C	1129	0.021	C	1174	0.684	AV
1085	0.019	C	1130	0	C	1175	0.011	C
1086	0	C	1131	0.019	C	1176	0.055	C
1087	0.019	C	1132	0.027	C	1177	0.011	C
1088	0.027	C	1133	0	C	1178	0.019	C
1089	0.054	C	1134	0.019	C	1179	0.027	C
1090	0.042	C	1135	0.046	C	1180	0.045	C
1091	0.035	C	1136	0.059	C	1181	0	C
1092	0.03	C	1137	0.011	C	1182	0	C
1093	0.019	C	1138	0.021	C	1183	0.011	C
1094	0.03	C	1139	0	C	1184	0.097	C
1095	0.011	C	1140	0.039	C	1185	0.124	C
1096	0	C	1141	0	C	1186	0.011	C
1097	0.021	C	1142	0.041	C	1187	0.068	C
1098	0.045	C	1143	0.047	C	1188	0	C
1099	0.021	C	1144	0.011	C	1189	0.011	C
1100	0.019	C	1145	0	C	1190	0	C
1101	0.011	C	1146	0.011	C	1191	0.124	C
1102	0.074	C	1147	0.011	C	1192	0.118	C
1103	0.011	C	1148	0.039	C	1193	0.027	C
1104	0.011	C	1149	0	C	1194	0.011	C
1105	0.035	C	1150	0.019	C	1195	0.038	C
1106	0	C	1151	0	C	1196	0.027	C
1107	0.035	C	1152	0.046	C	1197	0.062	C
1108	0	C	1153	0	C	1198	0.011	C
1109	0	C	1154	0.019	C	1199	0.035	C

1200	0.019	C	1245	0.038	C	1290	0	C
1201	0.019	C	1246	0	C	1291	0.011	C
1202	0.038	C	1247	0.068	C	1292	0.585	AV
1203	0.049	C	1248	0.019	C	1293	0.52	AV
1204	0	C	1249	0.048	C	1294	0.021	C
1205	0	C	1250	0.03	C	1295	0.528	AV
1206	0	C	1251	0.021	C	1296	0.011	C
1207	0.011	C	1252	0.011	C	1297	0.105	C
1208	0.019	C	1253	0.021	C	1298	0.628	AV
1209	0.093	C	1254	0.103	C	1299	0.021	C
1210	0.019	C	1255	0.019	C	1300	0	C
1211	0	C	1256	0.163	V	1301	0.011	C
1212	0.027	C	1257	0.027	C	1302	0.143	V
1213	0.021	C	1258	0.041	C	1303	0.019	C
1214	0.019	C	1259	0.021	C	1304	0.066	C
1215	0	C	1260	0.021	C	1305	0.041	C
1216	0	C	1261	0.011	C	1306	0.021	C
1217	0	C	1262	0.011	C	1307	0	C
1218	0	C	1263	0.113	C	1308	0.011	C
1219	0.021	C	1264	0.157	V	1309	0	C
1220	0.011	C	1265	0.045	C	1310	0.072	C
1221	0.038	C	1266	0.108	C	1311	0.087	C
1222	0.019	C	1267	0.049	C	1312	0.019	C
1223	0	C	1268	0.011	C	1313	0.102	C
1224	0.027	C	1269	0	C	1314	0.143	V
1225	0.019	C	1270	0.042	C	1315	0	C
1226	0.03	C	1271	0.062	C	1316	0.123	C
1227	0	C	1272	0.052	C	1317	0.011	C
1228	0	C	1273	0.047	C	1318	0.011	C
1229	0	C	1274	0.035	C	1319	0	C
1230	0	C	1275	0.038	C	1320	0.011	C
1231	0.011	C	1276	0.057	C	1321	0	C
1232	0	C	1277	0.714	AV	1322	0	C
1233	0	C	1278	0.038	C	1323	0.038	C
1234	0.011	C	1279	0.192	V	1324	0	C
1235	0.027	C	1280	0.856	AV	1325	0.011	C
1236	0.042	C	1281	0	C	1326	0.011	C
1237	0.011	C	1282	0.51	AV	1327	0.011	C
1238	0.038	C	1283	0.489	AV	1328	0.011	C
1239	0	C	1284	0.535	AV	1329	0.056	C
1240	0	C	1285	0	C	1330	0.701	AV
1241	0	C	1286	0.011	C	1331	0.011	C
1242	0	C	1287	0.167	V	1332	0	C
1243	0.03	C	1288	0.055	C	1333	0.011	C
1244	0.021	C	1289	0	C	1334	0.047	C

1335	0	C	1380	0	C	1425	0.02	C
1336	0	C	1381	0.087	C	1426	0.022	C
1337	0	C	1382	0	C	1427	0.047	C
1338	0.011	C	1383	0.022	C	1428	0.072	C
1339	0	C	1384	0.035	C	1429	0.04	C
1340	0.08	C	1385	0	C	1430	0.051	C
1341	0	C	1386	0	C	1431	0.033	C
1342	0.027	C	1387	0	C	1432	0.02	C
1343	0.039	C	1388	0	C	1433	0.044	C
1344	0	C	1389	0.011	C	1434	0.052	C
1345	0.011	C	1390	0.011	C	1435	0.011	C
1346	0	C	1391	0.011	C	1436	0.029	C
1347	0.716	AV	1392	0.053	C	1437	0.212	V
1348	0	C	1393	0.028	C	1438	0.022	C
1349	0.011	C	1394	0	C	1439	0.02	C
1350	0.011	C	1395	0	C	1440	0.022	C
1351	0	C	1396	0.02	C	1441	0.022	C
1352	0.011	C	1397	0.022	C	1442	0.061	C
1353	0.011	C	1398	0	C	1443	0.011	C
1354	0.019	C	1399	0.041	C	1444	0.022	C
1355	0.019	C	1400	0	C	1445	0.048	C
1356	0.019	C	1401	0.022	C	1446	0.066	C
1357	0	C	1402	0.02	C	1447	0.032	C
1358	0	C	1403	0	C	1448	0.051	C
1359	0	C	1404	0.055	C	1449	0.029	C
1360	0	C	1405	0.02	C	1450	0.048	C
1361	0	C	1406	0.04	C	1451	0.077	C
1362	0	C	1407	0.176	V	1452	0.143	V
1363	0	C	1408	0.121	C	1453	0.034	C
1364	0.109	C	1409	0.031	C	1454	0.101	C
1365	0	C	1410	0.011	C	1455	0.095	C
1366	0	C	1411	0.011	C	1456	0.194	V
1367	0	C	1412	0	C	1457	0.06	C
1368	0.011	C	1413	0.022	C	1458	0.034	C
1369	0	C	1414	0.081	C	1459	0.073	C
1370	0.011	C	1415	0	C	1460	0.079	C
1371	0	C	1416	0.042	C	1461	0.074	C
1372	0	C	1417	0	C	1462	0.163	V
1373	0	C	1418	0.112	C	1463	0.057	C
1374	0.011	C	1419	0	C	1464	0.021	C
1375	0	C	1420	0	C	1465	0.043	C
1376	0.028	C	1421	0.028	C	1466	0.038	C
1377	0.011	C	1422	0.042	C	1467	0.061	C
1378	0	C	1423	0.048	C	1468	0.042	C
1379	0	C	1424	0.068	C	1469	0.098	C

1470	0.353	V	1515	0.276	V
1471	0.208	V	1516	0.225	V
1472	0.162	V	1517	0.209	V
1473	0.308	V	1518	0.228	V
1474	0.248	V	1519	0.373	V
1475	0.119	C	1520	0.222	V
1476	0.13	V	1521	0.24	V
1477	0.105	C			
1478	0.137	V			
1479	0.043	C			
1480	0.097	C			
1481	0.204	V			
1482	0.124	C			
1483	0.197	V			
1484	0.092	C			
1485	0.293	V			
1486	0.279	V			
1487	0.264	V			
1488	0.198	V			
1489	0.129	V			
1490	0.131	V			
1491	0.192	V			
1492	0.208	V			
1493	0.211	V			
1494	0.179	V			
1495	0.248	V			
1496	0.318	V			
1497	0.18	V			
1498	0.3	V			
1499	0.248	V			
1500	0.312	V			
1501	0.3	V			
1502	0.156	V			
1503	0.182	V			
1504	0.324	V			
1505	0.283	V			
1506	0.308	V			
1507	0.273	V			
1508	0.236	V			
1509	0.216	V			
1510	0.25	V			
1511	0.264	V			
1512	0.284	V			
1513	0.155	V			
1514	0.216	V			

En el cuadro se enlistan las 1,521 posiciones del conjunto de SILVA con su respectiva entropía y la clasificación a la que corresponde cada posición; C: Conservado, V: Variante y AV: Altamente Variante; H(i): entropía, Pos: posición.

13.5. Clasificación mediante K-medias conjunto Ribosomal Database Project

Pos	H(i)	Grupo							
1	0.177	V	39	0.263	V	78	0.625	AV	
2	0.179	V	40	0.491	AV	79	0.165	V	
3	0.165	V	41	0.554	AV	80	0.015	C	
4	0.248	V	42	0.483	AV	81	0.014	C	
5	0.1	C	43	0.59	AV	82	0.018	C	
6	0.023	C	44	0.25	V	83	0.145	V	
7	0.282	V	45	0.261	V	84	0.151	V	
8	0.296	V	46	0.268	V	85	0.731	AV	
9	0.218	V	47	0.229	V	86	0.172	V	
10	0.16	V	48	0.158	V	87	0.635	AV	
11	0.39	V	49	0.176	V	88	0.633	AV	
12	0.764	AV	50	0.246	V	89	0.718	AV	
13	0.173	V	51	0.194	V	90	0.68	AV	
14	0.269	V	52	0.254	V	91	0.723	AV	
15	0.12	C	53	0.099	C	92	0.943	AV	
16	0.082	C	54	0.054	C	93	0.627	AV	
17	0.054	C	55	0.071	C	94	0.634	AV	
18	0.211	V	56	0.167	V	95	0.387	V	
19	0.125	C	57	0.159	V	96	0.041	C	
20	0.136	V	58	0.78	AV	97	0.023	C	
21	0.078	C	59	0.055	C	98	0.362	V	
22	0.17	V	60	0.022	C	99	0.036	C	
23	0.184	V	61	0.042	C	100	0.688	AV	
24	0.173	V	62	0.037	C	101	0.034	C	
25	0.236	V	63	0.078	C	102	0.014	C	
26	0.167	V	64	0.096	C	103	0.036	C	
27	0.215	V	65	0.395	V	104	0.04	C	
28	0.196	V	66	0.068	C	105	0.116	C	
29	0.347	V	67	0.044	C	106	0.027	C	
30	0.34	V	68	0.399	V	107	0.047	C	
31	0.231	V	69	0.628	AV	108	0.027	C	
32	0.236	V	70	0.635	AV	109	0.603	AV	
33	0.391	V	71	0.722	AV	110	0.014	C	
34	0.3	V	72	0.15	V	111	0.014	C	
35	0.312	V	73	0.733	AV	112	0.017	C	
36	0.365	V	74	0.242	V	113	0.009	C	
37	0.253	V	75	0.718	AV	114	0.014	C	
38	0.445	V	76	0.731	AV	115	0	C	
			77	0.07	C	116	0.01	C	

117	0	C	162	0.009	C	207	0.784	AV
118	0.02	C	163	0.009	C	208	0.875	AV
119	0	C	164	0.009	C	209	0.687	AV
120	0.014	C	165	0	C	210	0.009	C
121	0.026	C	166	0.016	C	211	0.005	C
122	0	C	167	0.016	C	212	0.009	C
123	0.013	C	168	0	C	213	0.016	C
124	0.014	C	169	0.005	C	214	0.038	C
125	0.014	C	170	0.009	C	215	0	C
126	0.013	C	171	0	C	216	0.117	C
127	0.005	C	172	0.005	C	217	0.016	C
128	0.009	C	173	0	C	218	0	C
129	0.017	C	174	0.005	C	219	0.028	C
130	0.016	C	175	0.024	C	220	0.021	C
131	0	C	176	0.019	C	221	0.012	C
132	0.005	C	177	0	C	222	0.012	C
133	0	C	178	0	C	223	0.542	AV
134	0.021	C	179	0.012	C	224	0.17	V
135	0.014	C	180	0.632	AV	225	0.35	V
136	0.61	AV	181	0	C	226	0.005	C
137	0.005	C	182	0.013	C	227	0.162	V
138	0	C	183	0.009	C	228	0	C
139	0.005	C	184	0.005	C	229	0.175	V
140	0.074	C	185	0.195	V	230	0.028	C
141	0.054	C	186	0.125	C	231	0.442	V
142	0.085	C	187	0.927	AV	232	0.025	C
143	0.013	C	188	0.133	V	233	0	C
144	0.042	C	189	0.904	AV	234	0.08	C
145	0.168	V	190	1.004	AV	235	0.063	C
146	0.648	AV	191	0.382	V	236	0.074	C
147	0.017	C	192	0.745	AV	237	0.009	C
148	0	C	193	0.921	AV	238	0.005	C
149	0.005	C	194	0.18	V	239	0	C
150	0.013	C	195	0.024	C	240	0	C
151	0	C	196	0.147	V	241	0	C
152	0.005	C	197	0.02	C	242	0.005	C
153	0.009	C	198	0.04	C	243	0.032	C
154	0.014	C	199	0.012	C	244	0.012	C
155	0.009	C	200	0.018	C	245	0	C
156	0.017	C	201	0.173	V	246	0.083	C
157	0.029	C	202	0.919	AV	247	0.017	C
158	0.019	C	203	0.716	AV	248	0.009	C
159	0.019	C	204	0.398	V	249	0.005	C
160	0.009	C	205	0.875	AV	250	0.028	C
161	0.028	C	206	0.763	AV	251	0.005	C

252	0	C	297	0.016	C	342	0	C
253	0	C	298	0.005	C	343	0.005	C
254	0	C	299	0	C	344	0.013	C
255	0	C	300	0.024	C	345	0	C
256	0	C	301	0.009	C	346	0.012	C
257	0.02	C	302	0.005	C	347	0	C
258	0.043	C	303	0.009	C	348	0.667	AV
259	0.005	C	304	0.005	C	349	0.013	C
260	0	C	305	0.012	C	350	0.016	C
261	0.005	C	306	0.013	C	351	0.012	C
262	0.014	C	307	0	C	352	0.009	C
263	0.036	C	308	0.009	C	353	0.005	C
264	0.016	C	309	0.005	C	354	0.017	C
265	0	C	310	0.005	C	355	0	C
266	0	C	311	0.017	C	356	0	C
267	0.014	C	312	0.009	C	357	0.005	C
268	0.116	C	313	0	C	358	0	C
269	0.136	V	314	0	C	359	0	C
270	0.005	C	315	0.005	C	360	0.005	C
271	0	C	316	0.005	C	361	0.005	C
272	0	C	317	0	C	362	0	C
273	0.009	C	318	0	C	363	0.005	C
274	0.005	C	319	0	C	364	0.005	C
275	0.028	C	320	0.009	C	365	0.014	C
276	0	C	321	0	C	366	0	C
277	0	C	322	0.005	C	367	0.009	C
278	0.03	C	323	0	C	368	0.013	C
279	0.138	V	324	0.005	C	369	0	C
280	0.136	V	325	0	C	370	0.005	C
281	0.021	C	326	0.016	C	371	0.013	C
282	0.009	C	327	0.668	AV	372	0	C
283	0	C	328	0.005	C	373	0.009	C
284	0	C	329	0.005	C	374	0.009	C
285	0	C	330	0	C	375	0.017	C
286	0.005	C	331	0.005	C	376	0.005	C
287	0.005	C	332	0.009	C	377	0	C
288	0.022	C	333	0.016	C	378	0.005	C
289	0.736	AV	334	0.009	C	379	0.005	C
290	0.051	C	335	0	C	380	0.005	C
291	0.075	C	336	0	C	381	0.013	C
292	0.025	C	337	0.012	C	382	0.013	C
293	0.019	C	338	0.005	C	383	0	C
294	0	C	339	0	C	384	0.005	C
295	0.005	C	340	0.009	C	385	0.005	C
296	0.005	C	341	0.005	C	386	0.014	C

387	0.005	C	432	0.009	C	477	0.043	C
388	0.009	C	433	0.005	C	478	0.273	V
389	0.005	C	434	0	C	479	0.399	V
390	0.009	C	435	0.005	C	480	0.4	V
391	0.028	C	436	0.009	C	481	0.205	V
392	0.103	C	437	0.005	C	482	0	C
393	0.009	C	438	0	C	483	0.627	AV
394	0.005	C	439	0.005	C	484	0.579	AV
395	0.009	C	440	0.009	C	485	0.005	C
396	0.005	C	441	0.005	C	486	0.629	AV
397	0.005	C	442	0	C	487	0.161	V
398	0.005	C	443	0	C	488	0.07	C
399	0.005	C	444	0.009	C	489	0.009	C
400	0.005	C	445	0.009	C	490	0.009	C
401	0.005	C	446	0.005	C	491	0.588	AV
402	0	C	447	0.012	C	492	0.013	C
403	0	C	448	0.009	C	493	0.013	C
404	0.005	C	449	0.009	C	494	0.005	C
405	0.013	C	450	0.009	C	495	0	C
406	0.009	C	451	0	C	496	0	C
407	0.005	C	452	0.005	C	497	0	C
408	0.005	C	453	0.005	C	498	0.005	C
409	0.014	C	454	0.005	C	499	0.013	C
410	0.012	C	455	0.005	C	500	0.005	C
411	0	C	456	0.009	C	501	0.024	C
412	0.009	C	457	0.005	C	502	0.019	C
413	0.005	C	458	0.005	C	503	0.021	C
414	0.005	C	459	0.009	C	504	0.005	C
415	0.013	C	460	0.005	C	505	0.009	C
416	0	C	461	0.017	C	506	0	C
417	0	C	462	0.009	C	507	0.005	C
418	0.005	C	463	0	C	508	0.009	C
419	0.018	C	464	0.084	C	509	0.005	C
420	0.005	C	465	0.012	C	510	0.005	C
421	0	C	466	0.024	C	511	0.005	C
422	0.005	C	467	0.009	C	512	0.02	C
423	0.016	C	468	0.07	C	513	0.02	C
424	0.024	C	469	0.163	V	514	0.009	C
425	0.005	C	470	0.627	AV	515	0.005	C
426	0	C	471	0	C	516	0.013	C
427	0.005	C	472	0.58	AV	517	0.005	C
428	0.005	C	473	0.593	AV	518	0.012	C
429	0.005	C	474	0.005	C	519	0.012	C
430	0.009	C	475	0.122	C	520	0.013	C
431	0.009	C	476	0.122	C	521	0	C

522	0.005	C	567	0.005	C	612	0	C
523	0.021	C	568	0.009	C	613	0.005	C
524	0	C	569	0.005	C	614	0	C
525	0	C	570	0.012	C	615	0	C
526	0.005	C	571	0	C	616	0	C
527	0.017	C	572	0.009	C	617	0.013	C
528	0.017	C	573	0	C	618	0.009	C
529	0.005	C	574	0.005	C	619	0.016	C
530	0	C	575	0.012	C	620	0.005	C
531	0	C	576	0	C	621	0.009	C
532	0.005	C	577	0.005	C	622	0.033	C
533	0.013	C	578	0.005	C	623	0.012	C
534	0	C	579	0	C	624	0	C
535	0	C	580	0.005	C	625	0	C
536	0.005	C	581	0	C	626	0.013	C
537	0	C	582	0	C	627	0.005	C
538	0.013	C	583	0	C	628	0	C
539	0.005	C	584	0.009	C	629	0	C
540	0.009	C	585	0.005	C	630	0.005	C
541	0	C	586	0.009	C	631	0.005	C
542	0	C	587	0.005	C	632	0.08	C
543	0.013	C	588	0.005	C	633	0	C
544	0.005	C	589	0.014	C	634	0.012	C
545	0	C	590	0.013	C	635	0	C
546	0	C	591	0.009	C	636	0	C
547	0.005	C	592	0	C	637	0.005	C
548	0	C	593	0.009	C	638	0.009	C
549	0.013	C	594	0	C	639	0.005	C
550	0.005	C	595	0.013	C	640	0.005	C
551	0	C	596	0.02	C	641	0.009	C
552	0.009	C	597	0.005	C	642	0.009	C
553	0	C	598	0.005	C	643	0.005	C
554	0.005	C	599	0.005	C	644	0	C
555	0.014	C	600	0.009	C	645	0.005	C
556	0.009	C	601	0	C	646	0.005	C
557	0	C	602	0.034	C	647	0.009	C
558	0.026	C	603	0.122	C	648	0	C
559	0	C	604	0.153	V	649	0	C
560	0.013	C	605	0.005	C	650	0	C
561	0	C	606	0.005	C	651	0.009	C
562	0.005	C	607	0.018	C	652	0.017	C
563	0.017	C	608	0.005	C	653	0	C
564	0.009	C	609	0.013	C	654	0.009	C
565	0.017	C	610	0.005	C	655	0.009	C
566	0.005	C	611	0.016	C	656	0.017	C

657	0.009	C	702	0.005	C	747	0.005	C
658	0	C	703	0	C	748	0.005	C
659	0.039	C	704	0.013	C	749	0.034	C
660	0.118	C	705	0	C	750	0.012	C
661	0.214	V	706	0	C	751	0.016	C
662	0	C	707	0.005	C	752	0.009	C
663	0	C	708	0.005	C	753	0.009	C
664	0	C	709	0	C	754	0.018	C
665	0.017	C	710	0	C	755	0	C
666	0	C	711	0	C	756	0	C
667	0.005	C	712	0.005	C	757	0	C
668	0.005	C	713	0	C	758	0	C
669	0	C	714	0	C	759	0	C
670	0.012	C	715	0.04	C	760	0	C
671	0.012	C	716	0.013	C	761	0.009	C
672	0.005	C	717	0	C	762	0	C
673	0.017	C	718	0.005	C	763	0.005	C
674	0.005	C	719	0	C	764	0	C
675	0.021	C	720	0.009	C	765	0.005	C
676	0.012	C	721	0.017	C	766	0.005	C
677	0.009	C	722	0.013	C	767	0.005	C
678	0.013	C	723	0.032	C	768	0.005	C
679	0	C	724	0.022	C	769	0.005	C
680	0	C	725	0.012	C	770	0	C
681	0.028	C	726	0.022	C	771	0.012	C
682	0.012	C	727	0.009	C	772	0.038	C
683	0.012	C	728	0.019	C	773	0	C
684	0.025	C	729	0.009	C	774	0.009	C
685	0.005	C	730	0.013	C	775	0.005	C
686	0	C	731	0.005	C	776	0.012	C
687	0.016	C	732	0.013	C	777	0.012	C
688	0	C	733	0.028	C	778	0.013	C
689	0.009	C	734	0.024	C	779	0.013	C
690	0.016	C	735	0.009	C	780	0.012	C
691	0	C	736	0	C	781	0.009	C
692	0.012	C	737	0.013	C	782	0	C
693	0.02	C	738	0.005	C	783	0.013	C
694	0.005	C	739	0.019	C	784	0.005	C
695	0	C	740	0.019	C	785	0.005	C
696	0	C	741	0.005	C	786	0.009	C
697	0.005	C	742	0.057	C	787	0.009	C
698	0.009	C	743	0.005	C	788	0.005	C
699	0.012	C	744	0	C	789	0.009	C
700	0.028	C	745	0.009	C	790	0	C
701	0.017	C	746	0.126	C	791	0.013	C

792	0.259	V	837	0	C	882	0.009	C
793	0	C	838	0.024	C	883	0.013	C
794	0	C	839	0.013	C	884	0	C
795	0.009	C	840	0.009	C	885	0	C
796	0.013	C	841	0.009	C	886	0.005	C
797	0.009	C	842	0.005	C	887	0.028	C
798	0.016	C	843	0.005	C	888	0	C
799	0.009	C	844	0	C	889	0.005	C
800	0	C	845	0.005	C	890	0.019	C
801	0.009	C	846	0.009	C	891	0.005	C
802	0.013	C	847	0	C	892	0	C
803	0.022	C	848	0.005	C	893	0.009	C
804	0.013	C	849	0	C	894	0.028	C
805	0.013	C	850	0	C	895	0	C
806	0.005	C	851	0.005	C	896	0.027	C
807	0.013	C	852	0	C	897	0.028	C
808	0.013	C	853	0	C	898	0.018	C
809	0.024	C	854	0.005	C	899	0.005	C
810	0	C	855	0.005	C	900	0.017	C
811	0.005	C	856	0.005	C	901	0	C
812	0.005	C	857	0	C	902	0.012	C
813	0	C	858	0.02	C	903	0.009	C
814	0	C	859	0.016	C	904	0.013	C
815	0.005	C	860	0	C	905	0.024	C
816	0	C	861	0.03	C	906	0.009	C
817	0.035	C	862	0.012	C	907	0.021	C
818	0.005	C	863	0.005	C	908	0.059	C
819	0.009	C	864	0.012	C	909	0.009	C
820	0	C	865	0.017	C	910	0.009	C
821	0.005	C	866	0.033	C	911	0.016	C
822	0.014	C	867	0.013	C	912	0	C
823	0.024	C	868	0.005	C	913	0.005	C
824	0	C	869	0.029	C	914	0.03	C
825	0.009	C	870	0.005	C	915	0.014	C
826	0.017	C	871	0.005	C	916	0.014	C
827	0.028	C	872	0	C	917	0.034	C
828	0.009	C	873	0.009	C	918	0.013	C
829	0.005	C	874	0.012	C	919	0	C
830	0	C	875	0.028	C	920	0	C
831	0	C	876	0.649	AV	921	0.009	C
832	0.009	C	877	0	C	922	0	C
833	0.005	C	878	0.005	C	923	0.028	C
834	0.009	C	879	0.005	C	924	0.009	C
835	0	C	880	0.017	C	925	0	C
836	0	C	881	0	C	926	0	C

927	0.013	C	972	0.009	C	1017	0	C
928	0.009	C	973	0	C	1018	0.05	C
929	0.005	C	974	0.009	C	1019	0.013	C
930	0	C	975	0.135	V	1020	0.05	C
931	0.005	C	976	0.021	C	1021	0.034	C
932	0.009	C	977	0	C	1022	0.037	C
933	0.005	C	978	0	C	1023	0.037	C
934	0.009	C	979	0.005	C	1024	0.038	C
935	0.005	C	980	0.009	C	1025	0	C
936	0.009	C	981	0.005	C	1026	0.031	C
937	0	C	982	0.005	C	1027	0	C
938	0.009	C	983	0	C	1028	0.009	C
939	0	C	984	0.016	C	1029	0	C
940	0.005	C	985	0.04	C	1030	0	C
941	0.005	C	986	0.017	C	1031	0.028	C
942	0.025	C	987	0.009	C	1032	0.019	C
943	0.009	C	988	0.009	C	1033	0.031	C
944	0	C	989	0.028	C	1034	0.043	C
945	0	C	990	0.009	C	1035	0.051	C
946	0	C	991	0.063	C	1036	0.017	C
947	0	C	992	0.005	C	1037	0.017	C
948	0.048	C	993	0.012	C	1038	0.632	AV
949	0	C	994	0.02	C	1039	0.005	C
950	0	C	995	0.005	C	1040	0.005	C
951	0	C	996	0.005	C	1041	0.012	C
952	0.021	C	997	0	C	1042	0.005	C
953	0	C	998	0	C	1043	0.009	C
954	0.005	C	999	0	C	1044	0.042	C
955	0.009	C	1000	0.005	C	1045	0.005	C
956	0.005	C	1001	0.005	C	1046	0	C
957	0	C	1002	0.009	C	1047	0.009	C
958	0.009	C	1003	0	C	1048	0.005	C
959	0	C	1004	0.009	C	1049	0.624	AV
960	0.013	C	1005	0	C	1050	0.009	C
961	0	C	1006	0.005	C	1051	0.009	C
962	0.005	C	1007	0.03	C	1052	0.009	C
963	0.013	C	1008	0	C	1053	0	C
964	0.009	C	1009	0.013	C	1054	0.034	C
965	0.009	C	1010	0.013	C	1055	0.031	C
966	0.009	C	1011	0.012	C	1056	0	C
967	0.016	C	1012	0.005	C	1057	0.053	C
968	0	C	1013	0.005	C	1058	0	C
969	0.005	C	1014	0.013	C	1059	0	C
970	0.016	C	1015	0.009	C	1060	0	C
971	0.013	C	1016	0.009	C	1061	0.005	C

1062	0	C	1107	0.005	C	1152	0	C
1063	0.005	C	1108	0.005	C	1153	0	C
1064	0.005	C	1109	0.017	C	1154	0.009	C
1065	0.013	C	1110	0	C	1155	0	C
1066	0.005	C	1111	0.005	C	1156	0.061	C
1067	0	C	1112	0.013	C	1157	0	C
1068	0	C	1113	0.009	C	1158	0.017	C
1069	0	C	1114	0.009	C	1159	0.031	C
1070	0.005	C	1115	0.009	C	1160	0.009	C
1071	0.005	C	1116	0.012	C	1161	0.05	C
1072	0.009	C	1117	0.005	C	1162	0.005	C
1073	0.032	C	1118	0.005	C	1163	0.009	C
1074	0.005	C	1119	0.009	C	1164	0.06	C
1075	0.009	C	1120	0.009	C	1165	0.013	C
1076	0.036	C	1121	0.005	C	1166	0.684	AV
1077	0	C	1122	0	C	1167	0.017	C
1078	0	C	1123	0.012	C	1168	0.009	C
1079	0.009	C	1124	0	C	1169	0.009	C
1080	0.005	C	1125	0.005	C	1170	0.012	C
1081	0.005	C	1126	0	C	1171	0.012	C
1082	0.009	C	1127	0	C	1172	0.024	C
1083	0.005	C	1128	0.05	C	1173	0	C
1084	0.024	C	1129	0.009	C	1174	0	C
1085	0	C	1130	0	C	1175	0.005	C
1086	0.009	C	1131	0	C	1176	0.025	C
1087	0.005	C	1132	0.013	C	1177	0.037	C
1088	0	C	1133	0.013	C	1178	0.005	C
1089	0.025	C	1134	0.009	C	1179	0.052	C
1090	0	C	1135	0.009	C	1180	0	C
1091	0.005	C	1136	0.005	C	1181	0.005	C
1092	0	C	1137	0.005	C	1182	0	C
1093	0.009	C	1138	0	C	1183	0.041	C
1094	0.012	C	1139	0.009	C	1184	0.03	C
1095	0.005	C	1140	0.005	C	1185	0.005	C
1096	0.005	C	1141	0	C	1186	0	C
1097	0	C	1142	0	C	1187	0.009	C
1098	0.005	C	1143	0	C	1188	0.005	C
1099	0.005	C	1144	0.053	C	1189	0.017	C
1100	0	C	1145	0	C	1190	0.009	C
1101	0	C	1146	0.009	C	1191	0.016	C
1102	0	C	1147	0.005	C	1192	0.019	C
1103	0	C	1148	0.009	C	1193	0.016	C
1104	0.009	C	1149	0.693	AV	1194	0.012	C
1105	0	C	1150	0.005	C	1195	0.045	C
1106	0.005	C	1151	0.012	C	1196	0	C

1197	0.005	C	1242	0.005	C	1287	0.595	AV
1198	0	C	1243	0.017	C	1288	0.017	C
1199	0	C	1244	0.009	C	1289	0.086	C
1200	0.005	C	1245	0.009	C	1290	0.653	AV
1201	0.096	C	1246	0.085	C	1291	0.02	C
1202	0	C	1247	0	C	1292	0	C
1203	0.009	C	1248	0.094	C	1293	0.009	C
1204	0.025	C	1249	0	C	1294	0.112	C
1205	0.005	C	1250	0	C	1295	0.019	C
1206	0.009	C	1251	0	C	1296	0.074	C
1207	0	C	1252	0.019	C	1297	0.018	C
1208	0	C	1253	0.009	C	1298	0.017	C
1209	0.005	C	1254	0	C	1299	0.005	C
1210	0.009	C	1255	0.028	C	1300	0	C
1211	0	C	1256	0.043	C	1301	0.005	C
1212	0	C	1257	0.009	C	1302	0.089	C
1213	0.016	C	1258	0.033	C	1303	0.084	C
1214	0.009	C	1259	0.031	C	1304	0.009	C
1215	0	C	1260	0.005	C	1305	0.034	C
1216	0.009	C	1261	0.009	C	1306	0.094	C
1217	0.014	C	1262	0.005	C	1307	0.012	C
1218	0.024	C	1263	0.013	C	1308	0.1	C
1219	0.005	C	1264	0.016	C	1309	0.012	C
1220	0	C	1265	0.009	C	1310	0.005	C
1221	0	C	1266	0.02	C	1311	0.005	C
1222	0.005	C	1267	0.009	C	1312	0.005	C
1223	0.034	C	1268	0.024	C	1313	0.005	C
1224	0	C	1269	0.699	AV	1314	0	C
1225	0	C	1270	0.013	C	1315	0.041	C
1226	0.009	C	1271	0.119	C	1316	0	C
1227	0.009	C	1272	0.852	AV	1317	0	C
1228	0.036	C	1273	0.009	C	1318	0	C
1229	0.009	C	1274	0.592	AV	1319	0.009	C
1230	0.024	C	1275	0.599	AV	1320	0.012	C
1231	0.009	C	1276	0.605	AV	1321	0.051	C
1232	0.005	C	1277	0.02	C	1322	0.668	AV
1233	0	C	1278	0.016	C	1323	0.009	C
1234	0	C	1279	0.119	C	1324	0.009	C
1235	0.071	C	1280	0.009	C	1325	0.012	C
1236	0	C	1281	0.005	C	1326	0.044	C
1237	0.033	C	1282	0.005	C	1327	0	C
1238	0	C	1283	0	C	1328	0	C
1239	0.047	C	1284	0.649	AV	1329	0	C
1240	0.005	C	1285	0.59	AV	1330	0	C
1241	0.028	C	1286	0.018	C	1331	0	C

1332	0.076	C	1377	0.013	C	1422	0.052	C
1333	0	C	1378	0.014	C	1423	0.013	C
1334	0.045	C	1379	0.01	C	1424	0.005	C
1335	0.021	C	1380	0.028	C	1425	0.065	C
1336	0.009	C	1381	0.018	C	1426	0.015	C
1337	0.005	C	1382	0.005	C	1427	0.026	C
1338	0	C	1383	0.01	C	1428	0.024	C
1339	0.676	AV	1384	0.043	C	1429	0.114	C
1340	0.005	C	1385	0.021	C	1430	0	C
1341	0.017	C	1386	0.005	C	1431	0.019	C
1342	0.051	C	1387	0.019	C	1432	0	C
1343	0	C	1388	0.044	C	1433	0.005	C
1344	0	C	1389	0.015	C	1434	0.005	C
1345	0.005	C	1390	0.018	C	1435	0.025	C
1346	0.028	C	1391	0.033	C	1436	0.015	C
1347	0.005	C	1392	0.009	C	1437	0.023	C
1348	0.016	C	1393	0.023	C	1438	0.053	C
1349	0.005	C	1394	0.02	C	1439	0.01	C
1350	0.005	C	1395	0.005	C	1440	0.049	C
1351	0.005	C	1396	0.018	C	1441	0.01	C
1352	0	C	1397	0	C	1442	0.014	C
1353	0	C	1398	0.01	C	1443	0.043	C
1354	0.009	C	1399	0.104	C	1444	0.051	C
1355	0.009	C	1400	0.023	C	1445	0.024	C
1356	0.099	C	1401	0.018	C	1446	0.03	C
1357	0	C	1402	0.009	C	1447	0.044	C
1358	0.012	C	1403	0.018	C	1448	0.05	C
1359	0.005	C	1404	0.005	C	1449	0.045	C
1360	0.009	C	1405	0.018	C	1450	0.043	C
1361	0	C	1406	0.054	C	1451	0.066	C
1362	0.012	C	1407	0.009	C	1452	0.061	C
1363	0.005	C	1408	0.022	C	1453	0.052	C
1364	0.026	C	1409	0.014	C	1454	0.055	C
1365	0.005	C	1410	0.068	C	1455	0.044	C
1366	0.012	C	1411	0.005	C	1456	0.01	C
1367	0.005	C	1412	0.01	C	1457	0.046	C
1368	0.013	C	1413	0.033	C	1458	0.016	C
1369	0	C	1414	0.014	C	1459	0.065	C
1370	0	C	1415	0.019	C	1460	0.026	C
1371	0.01	C	1416	0.032	C	1461	0.069	C
1372	0.014	C	1417	0.014	C	1462	0.163	V
1373	0.088	C	1418	0.031	C	1463	0.065	C
1374	0.005	C	1419	0.024	C	1464	0.062	C
1375	0.038	C	1420	0.059	C	1465	0.138	V
1376	0.022	C	1421	0.03	C	1466	0.167	V

1467	0.088	C	1512	0.247	V
1468	0.095	C	1513	0.39	V
1469	0.101	C	1514	0.441	V
1470	0.103	C	1515	0.388	V
1471	0.058	C	1516	0.334	V
1472	0.044	C	1517	0.343	V
1473	0.09	C	1518	0.559	AV
1474	0.038	C	1519	0.606	AV
1475	0.061	C	1520	0.575	AV
1476	0.05	C	1521	0.978	AV
1477	0.134	V	1522	0.446	V
1478	0.113	C	1523	0.418	V
1479	0.28	V	1524	0.62	AV
1480	0.157	V	<u>1525</u>	<u>0.581</u>	<u>AV</u>
1481	0.091	C			
1482	0.177	V			
1483	0.338	V			
1484	0.227	V			
1485	0.304	V			
1486	0.296	V			
1487	0.339	V			
1488	0.349	V			
1489	0.237	V			
1490	0.281	V			
1491	0.201	V			
1492	0.236	V			
1493	0.213	V			
1494	0.193	V			
1495	0.183	V			
1496	0.159	V			
1497	0.24	V			
1498	0.367	V			
1499	0.129	V			
1500	0.157	V			
1501	0.143	V			
1502	0.17	V			
1503	0.295	V			
1504	0.198	V			
1505	0.141	V			
1506	0.172	V			
1507	0.25	V			
1508	0.21	V			
1509	0.151	V			
1510	0.259	V			
1511	0.457	V			

En el cuadro se enlistan las 1,525 posiciones del conjunto de RDP con su respectiva entropía y la clasificación a la que corresponde cada posición; C: Conservado, V: Variante y AV: Altamente Variante; H (i): entropía, Pos: posición.

13.6. Clasificación mediante K-medias conjunto Greengenes

Pos	H(i)	Grupo						
1	0.152	V	41	0.1454	V	83	0.1513	V
2	0.0871	V	42	0	C	84	0.1766	V
3	0.0837	V	43	0.0366	C	85	0.6642	AV
4	0.0815	C	44	0.031	C	86	0.2133	V
5	0.0785	C	45	0.031	C	87	0.5232	AV
6	0	C	46	0.0306	C	88	0.5836	AV
7	0	C	47	0	C	89	0.6227	AV
8	0.0724	C	48	0.0297	C	90	0.6073	AV
9	0.0708	C	49	0	C	91	0.6627	AV
10	0.1599	V	50	0	C	92	0.9265	AV
11	0.4905	AV	51	0	C	93	0.5634	AV
12	0.7585	AV	52	0.0289	C	94	0.5554	AV
13	0.1169	V	53	0.0302	C	95	0.4472	AV
14	0	C	54	0.0289	C	96	0.0958	V
15	0.064	C	55	0.0289	C	97	0.105	V
16	0.064	C	56	0.0805	C	98	0.4743	AV
17	0	C	57	0.0805	C	99	0.1378	V
18	0.2509	V	58	0.6715	AV	100	1.0755	AV
19	0	C	59	0.0802	C	101	0.0602	C
20	0	C	60	0.0802	C	102	0	C
21	0.0584	C	61	0.132	V	103	0.017	C
22	0.0574	C	62	0.1369	V	104	0	C
23	0.0565	C	63	0.0805	C	105	0.114	V
24	0.0555	C	64	0.2842	V	106	0.0508	C
25	0.1502	V	65	0.5947	AV	107	0.0428	C
26	0	C	66	0.1621	V	108	0.0304	C
27	0.0503	C	67	0.0905	V	109	0.583	AV
28	0	C	68	0.5007	AV	110	0	C
29	0.0482	C	69	0.6628	AV	111	0.0168	C
30	0.0469	C	70	0.6519	AV	112	0.1034	V
31	0	C	71	0.7297	AV	113	0	C
32	0	C	72	0.3997	V	114	0.0167	C
33	0.0448	C	73	0.7379	AV	115	0	C
34	0.0434	C	74	0.2076	V	116	0	C
35	0.0422	C	75	0.6077	AV	117	0	C
36	0.0417	C	76	0.5236	AV	118	0.0302	C
37	0.0414	C	77	0.0254	C	119	0	C
38	0.0407	C	78	0.5516	AV	120	0.0167	C
39	0	C	79	0.164	V	121	0.0167	C
40	0.0417	C	80	0	C	122	0	C
			81	0	C	123	0	C
			82	0.0254	C	124	0	C

125	0.0333	C	170	0	C	215	0	C
126	0.0167	C	171	0	C	216	0.2593	V
127	0	C	172	0	C	217	0.1286	V
128	0	C	173	0	C	218	0	C
129	0.0333	C	174	0	C	219	0.1368	V
130	0.1196	V	175	0	C	220	0.1583	V
131	0	C	176	0.0166	C	221	0.0598	C
132	0.0167	C	177	0	C	222	0.0299	C
133	0	C	178	0	C	223	0.6633	AV
134	0.0167	C	179	0	C	224	0.2261	V
135	0.0167	C	180	0.6403	AV	225	0.4963	AV
136	0.5942	AV	181	0	C	226	0.1203	V
137	0.0167	C	182	0.0166	C	227	0.0332	C
138	0	C	183	0.0166	C	228	0.5994	AV
139	0.0167	C	184	0.042	C	229	0.6524	AV
140	0.1206	V	185	0.2653	V	230	0.0332	C
141	0.186	V	186	0.1669	V	231	0.5343	AV
142	0.2306	V	187	0.9667	AV	232	0.042	C
143	0.0332	C	188	0.2395	V	233	0	C
144	0.0587	C	189	0.9421	AV	234	0.2186	V
145	0.3261	V	190	0.9412	AV	235	0.2295	V
146	0.6549	AV	191	0.5088	AV	236	0.11	V
147	0.0332	C	192	0.8568	AV	237	0	C
148	0	C	193	0.9436	AV	238	0	C
149	0.0166	C	194	0.217	V	239	0	C
150	0.0166	C	195	0.0166	C	240	0	C
151	0	C	196	0.1557	V	241	0	C
152	0	C	197	0	C	242	0	C
153	0.0166	C	198	0.0719	C	243	0.128	V
154	0.0166	C	199	0	C	244	0.0166	C
155	0	C	200	0.0166	C	245	0	C
156	0	C	201	0.2097	V	246	0.0166	C
157	0.0743	C	202	0.9459	AV	247	0.0166	C
158	0.1282	V	203	0.8383	AV	248	0	C
159	0.0332	C	204	0.543	AV	249	0	C
160	0.0166	C	205	1.0291	AV	250	0.064	C
161	0.0499	C	206	0.8243	AV	251	0.1288	V
162	0.0466	C	207	0.8029	AV	252	0	C
163	0.0166	C	208	0.9788	AV	253	0.1286	V
164	0	C	209	0.6887	AV	254	0	C
165	0	C	210	0.0166	C	255	0	C
166	0.0166	C	211	0.0166	C	256	0	C
167	0	C	212	0	C	257	0.0166	C
168	0.137	V	213	0	C	258	0.0631	C
169	0.0166	C	214	0.0299	C	259	0.0166	C

260	0	C	305	0.0299	C	350	0	C
261	0	C	306	0.0166	C	351	0	C
262	0.0166	C	307	0	C	352	0.0166	C
263	0.1191	V	308	0.0166	C	353	0.0167	C
264	0	C	309	0	C	354	0	C
265	0	C	310	0.0166	C	355	0	C
266	0	C	311	0.0299	C	356	0	C
267	0.0465	C	312	0	C	357	0.0166	C
268	0.1115	V	313	0	C	358	0	C
269	0.2105	V	314	0.0299	C	359	0	C
270	0	C	315	0	C	360	0	C
271	0	C	316	0.1115	V	361	0	C
272	0	C	317	0	C	362	0	C
273	0	C	318	0	C	363	0	C
274	0	C	319	0	C	364	0	C
275	0.0465	C	320	0	C	365	0	C
276	0	C	321	0	C	366	0	C
277	0	C	322	0	C	367	0	C
278	0	C	323	0	C	368	0.0166	C
279	0.1201	V	324	0	C	369	0	C
280	0.1201	V	325	0	C	370	0	C
281	0	C	326	0	C	371	0	C
282	0	C	327	0.668	AV	372	0	C
283	0	C	328	0	C	373	0	C
284	0	C	329	0	C	374	0	C
285	0.1026	V	330	0	C	375	0	C
286	0	C	331	0	C	376	0	C
287	0.0166	C	332	0	C	377	0	C
288	0	C	333	0	C	378	0	C
289	0.777	AV	334	0	C	379	0	C
290	0.0586	C	335	0	C	380	0	C
291	0.1826	V	336	0	C	381	0.1583	V
292	0.0166	C	337	0.0166	C	382	0.1286	V
293	0.0166	C	338	0	C	383	0	C
294	0	C	339	0	C	384	0	C
295	0.1286	V	340	0	C	385	0	C
296	0	C	341	0	C	386	0.0166	C
297	0.1286	V	342	0	C	387	0	C
298	0.0166	C	343	0	C	388	0	C
299	0	C	344	0.0332	C	389	0	C
300	0.0166	C	345	0	C	390	0	C
301	0	C	346	0	C	391	0.0166	C
302	0	C	347	0	C	392	0.2314	V
303	0	C	348	0.6684	AV	393	0.0166	C
304	0	C	349	0.0166	C	394	0	C

395	0	C	440	0	C	485	0.0598	C
396	0	C	441	0	C	486	0.6393	AV
397	0	C	442	0	C	487	0.2842	V
398	0	C	443	0	C	488	0.2104	V
399	0	C	444	0.0166	C	489	0.1324	V
400	0	C	445	0	C	490	0.128	V
401	0.1286	V	446	0.0166	C	491	0.615	AV
402	0.1286	V	447	0.0166	C	492	0	C
403	0	C	448	0.042	C	493	0	C
404	0	C	449	0	C	494	0	C
405	0	C	450	0	C	495	0	C
406	0	C	451	0	C	496	0	C
407	0	C	452	0	C	497	0	C
408	0	C	453	0.0166	C	498	0	C
409	0	C	454	0.0166	C	499	0.0166	C
410	0	C	455	0.0299	C	500	0	C
411	0	C	456	0.128	V	501	0.128	V
412	0	C	457	0	C	502	0.042	C
413	0	C	458	0	C	503	0.0166	C
414	0	C	459	0	C	504	0	C
415	0	C	460	0	C	505	0	C
416	0	C	461	0.0166	C	506	0	C
417	0	C	462	0	C	507	0	C
418	0	C	463	0	C	508	0.0166	C
419	0.0332	C	464	0.1115	V	509	0	C
420	0	C	465	0.0586	C	510	0.0166	C
421	0	C	466	0.1645	V	511	0	C
422	0.0166	C	467	0.1326	V	512	0	C
423	0	C	468	0.1828	V	513	0	C
424	0.0166	C	469	0.3007	V	514	0	C
425	0	C	470	0.6451	AV	515	0	C
426	0	C	471	0.0719	C	516	0	C
427	0.0166	C	472	0.565	AV	517	0	C
428	0	C	473	0.5511	AV	518	0	C
429	0.1413	V	474	0	C	519	0	C
430	0	C	475	0.1286	V	520	0	C
431	0	C	476	0.1201	V	521	0	C
432	0	C	477	0.042	C	522	0.0166	C
433	0	C	478	0.6067	AV	523	0	C
434	0	C	479	0.0166	C	524	0	C
435	0	C	480	0.4767	AV	525	0	C
436	0.1413	V	481	0.2632	V	526	0	C
437	0	C	482	0	C	527	0.0299	C
438	0	C	483	0.6187	AV	528	0	C
439	0	C	484	0.5484	AV	529	0	C

530	0	C	575	0.0299	C	620	0.0166	C
531	0	C	576	0	C	621	0	C
532	0	C	577	0	C	622	0	C
533	0.0167	C	578	0	C	623	0.0166	C
534	0	C	579	0.0166	C	624	0	C
535	0	C	580	0	C	625	0.1115	V
536	0	C	581	0	C	626	0.1533	V
537	0	C	582	0	C	627	0	C
538	0.0166	C	583	0	C	628	0	C
539	0	C	584	0.0332	C	629	0	C
540	0	C	585	0.0166	C	630	0.0166	C
541	0	C	586	0	C	631	0.0166	C
542	0.0166	C	587	0	C	632	0.1527	V
543	0.0166	C	588	0	C	633	0.0166	C
544	0.0166	C	589	0	C	634	0.0166	C
545	0	C	590	0	C	635	0	C
546	0	C	591	0	C	636	0	C
547	0	C	592	0.0166	C	637	0	C
548	0	C	593	0	C	638	0.1451	V
549	0.0299	C	594	0.0166	C	639	0.1115	V
550	0	C	595	0	C	640	0.0166	C
551	0	C	596	0.0166	C	641	0	C
552	0.0167	C	597	0.0332	C	642	0.1288	V
553	0.0166	C	598	0	C	643	0	C
554	0	C	599	0.0166	C	644	0.1413	V
555	0	C	600	0.0299	C	645	0.1286	V
556	0.0166	C	601	0	C	646	0	C
557	0.0166	C	602	0.064	C	647	0.0332	C
558	0.0466	C	603	0.1201	V	648	0	C
559	0.0166	C	604	0.2302	V	649	0.1201	V
560	0.0166	C	605	0	C	650	0	C
561	0.0166	C	606	0	C	651	0.0166	C
562	0.0166	C	607	0	C	652	0.0166	C
563	0.0332	C	608	0	C	653	0	C
564	0	C	609	0.0166	C	654	0.0332	C
565	0.0421	C	610	0	C	655	0	C
566	0	C	611	0	C	656	0	C
567	0.0166	C	612	0	C	657	0	C
568	0.0166	C	613	0.0166	C	658	0	C
569	0.0166	C	614	0	C	659	0.0166	C
570	0	C	615	0	C	660	0.1201	V
571	0	C	616	0	C	661	0.3328	V
572	0	C	617	0.0166	C	662	0	C
573	0	C	618	0	C	663	0	C
574	0	C	619	0	C	664	0	C

665	0.0166	C	710	0	C	755	0	C
666	0	C	711	0	C	756	0	C
667	0	C	712	0	C	757	0	C
668	0	C	713	0	C	758	0	C
669	0	C	714	0	C	759	0.0166	C
670	0	C	715	0.0498	C	760	0	C
671	0	C	716	0.0166	C	761	0	C
672	0	C	717	0	C	762	0	C
673	0.042	C	718	0	C	763	0	C
674	0	C	719	0.0166	C	764	0	C
675	0	C	720	0	C	765	0	C
676	0	C	721	0	C	766	0	C
677	0	C	722	0	C	767	0	C
678	0.0166	C	723	0	C	768	0	C
679	0	C	724	0	C	769	0	C
680	0	C	725	0.1453	V	770	0	C
681	0.1583	V	726	0.0166	C	771	0.0167	C
682	0.0166	C	727	0	C	772	0.0907	V
683	0	C	728	0	C	773	0	C
684	0.0299	C	729	0	C	774	0.0166	C
685	0	C	730	0.0166	C	775	0	C
686	0	C	731	0	C	776	0	C
687	0.0166	C	732	0.0166	C	777	0.0166	C
688	0	C	733	0.0166	C	778	0	C
689	0.1286	V	734	0.0299	C	779	0	C
690	0	C	735	0	C	780	0.0166	C
691	0	C	736	0	C	781	0.0166	C
692	0	C	737	0	C	782	0.1026	V
693	0	C	738	0	C	783	0.1451	V
694	0	C	739	0.0166	C	784	0	C
695	0	C	740	0	C	785	0.0166	C
696	0	C	741	0.0166	C	786	0.0167	C
697	0	C	742	0.0742	C	787	0	C
698	0	C	743	0.0299	C	788	0.1286	V
699	0	C	744	0	C	789	0	C
700	0.0299	C	745	0.0299	C	790	0	C
701	0	C	746	0.1604	V	791	0	C
702	0	C	747	0	C	792	0.3121	V
703	0.0166	C	748	0	C	793	0	C
704	0.0332	C	749	0.1704	V	794	0	C
705	0.0166	C	750	0	C	795	0	C
706	0.0166	C	751	0	C	796	0.0299	C
707	0	C	752	0	C	797	0.0166	C
708	0	C	753	0	C	798	0.0166	C
709	0	C	754	0	C	799	0	C

800	0	C	845	0.0166	C	890	0.1451	V
801	0.0166	C	846	0	C	891	0	C
802	0	C	847	0	C	892	0	C
803	0.0166	C	848	0.0299	C	893	0	C
804	0	C	849	0	C	894	0.0166	C
805	0	C	850	0	C	895	0	C
806	0	C	851	0	C	896	0.0332	C
807	0.0166	C	852	0	C	897	0.0299	C
808	0.0166	C	853	0	C	898	0.0166	C
809	0.042	C	854	0	C	899	0	C
810	0	C	855	0	C	900	0.0166	C
811	0	C	856	0	C	901	0	C
812	0.0166	C	857	0	C	902	0	C
813	0	C	858	0.0332	C	903	0	C
814	0	C	859	0.0299	C	904	0	C
815	0	C	860	0	C	905	0	C
816	0.0332	C	861	0.064	C	906	0	C
817	0.0299	C	862	0.0166	C	907	0	C
818	0	C	863	0	C	908	0.1604	V
819	0.0332	C	864	0.0166	C	909	0	C
820	0.1286	V	865	0	C	910	0	C
821	0.1026	V	866	0.0166	C	911	0.0166	C
822	0	C	867	0.0166	C	912	0.0698	C
823	0	C	868	0	C	913	0.0166	C
824	0	C	869	0.0465	C	914	0.0166	C
825	0.0166	C	870	0.042	C	915	0.0299	C
826	0	C	871	0	C	916	0.0166	C
827	0.0166	C	872	0	C	917	0.1604	V
828	0.0166	C	873	0	C	918	0.0332	C
829	0	C	874	0	C	919	0.0166	C
830	0	C	875	0	C	920	0	C
831	0	C	876	0.6343	AV	921	0	C
832	0.0166	C	877	0	C	922	0.0166	C
833	0	C	878	0	C	923	0.0332	C
834	0	C	879	0	C	924	0	C
835	0.1367	V	880	0	C	925	0.0166	C
836	0	C	881	0	C	926	0.0166	C
837	0.0166	C	882	0.0166	C	927	0	C
838	0.0166	C	883	0	C	928	0.0466	C
839	0	C	884	0	C	929	0.0166	C
840	0	C	885	0	C	930	0	C
841	0.0465	C	886	0.0166	C	931	0	C
842	0	C	887	0	C	932	0.03	C
843	0	C	888	0.0166	C	933	0	C
844	0	C	889	0.0166	C	934	0.0299	C

935	0	C	980	0	C	1025	0	C
936	0	C	981	0.1201	V	1026	0.0299	C
937	0	C	982	0.0166	C	1027	0	C
938	0.0166	C	983	0	C	1028	0.0166	C
939	0	C	984	0.0299	C	1029	0	C
940	0	C	985	0.042	C	1030	0	C
941	0.0166	C	986	0	C	1031	0.1853	V
942	0.0299	C	987	0	C	1032	0.1604	V
943	0	C	988	0	C	1033	0.1527	V
944	0	C	989	0.0166	C	1034	0.1368	V
945	0.0166	C	990	0.0166	C	1035	0.1005	V
946	0.0166	C	991	0.0166	C	1036	0.1413	V
947	0	C	992	0	C	1037	0	C
948	0.0299	C	993	0.0166	C	1038	0.6309	AV
949	0	C	994	0.0166	C	1039	0	C
950	0	C	995	0	C	1040	0	C
951	0	C	996	0	C	1041	0.03	C
952	0.0166	C	997	0	C	1042	0.0332	C
953	0	C	998	0	C	1043	0.0698	C
954	0.0167	C	999	0	C	1044	0.0167	C
955	0	C	1000	0	C	1045	0	C
956	0	C	1001	0	C	1046	0	C
957	0	C	1002	0	C	1047	0	C
958	0	C	1003	0	C	1048	0.03	C
959	0	C	1004	0	C	1049	0.6567	AV
960	0.0166	C	1005	0	C	1050	0.0299	C
961	0	C	1006	0	C	1051	0	C
962	0.0166	C	1007	0.0166	C	1052	0.1191	V
963	0	C	1008	0	C	1053	0.1026	V
964	0	C	1009	0	C	1054	0.064	C
965	0	C	1010	0.0166	C	1055	0.1026	V
966	0	C	1011	0.1191	V	1056	0.1026	V
967	0	C	1012	0.0533	C	1057	0.0166	C
968	0.0299	C	1013	0	C	1058	0.0166	C
969	0	C	1014	0.1193	V	1059	0	C
970	0	C	1015	0	C	1060	0	C
971	0.03	C	1016	0.0166	C	1061	0	C
972	0	C	1017	0.1026	V	1062	0	C
973	0	C	1018	0.1533	V	1063	0.0166	C
974	0	C	1019	0.1115	V	1064	0	C
975	0.1118	V	1020	0.084	V	1065	0.0166	C
976	0.0166	C	1021	0.1448	V	1066	0.0166	C
977	0	C	1022	0.1604	V	1067	0	C
978	0	C	1023	0.1604	V	1068	0	C
979	0	C	1024	0.1853	V	1069	0	C

1070	0	C	1115	0	C	1160	0.0166	C
1071	0	C	1116	0.0299	C	1161	0.0166	C
1072	0.042	C	1117	0	C	1162	0.0166	C
1073	0.0166	C	1118	0	C	1163	0.1115	V
1074	0	C	1119	0	C	1164	0.0332	C
1075	0.0166	C	1120	0.0299	C	1165	0.1368	V
1076	0.0598	C	1121	0	C	1166	0.6901	AV
1077	0.0166	C	1122	0	C	1167	0	C
1078	0	C	1123	0	C	1168	0	C
1079	0.0166	C	1124	0	C	1169	0	C
1080	0	C	1125	0	C	1170	0.0166	C
1081	0	C	1126	0	C	1171	0	C
1082	0	C	1127	0	C	1172	0	C
1083	0.0166	C	1128	0.0299	C	1173	0	C
1084	0.0166	C	1129	0	C	1174	0	C
1085	0	C	1130	0	C	1175	0	C
1086	0.0166	C	1131	0	C	1176	0.0299	C
1087	0	C	1132	0.042	C	1177	0.0299	C
1088	0	C	1133	0.137	V	1178	0.0166	C
1089	0	C	1134	0.0805	C	1179	0	C
1090	0	C	1135	0.1115	V	1180	0	C
1091	0	C	1136	0	C	1181	0	C
1092	0	C	1137	0	C	1182	0	C
1093	0.0166	C	1138	0	C	1183	0.0465	C
1094	0.0166	C	1139	0	C	1184	0.0465	C
1095	0	C	1140	0	C	1185	0	C
1096	0	C	1141	0	C	1186	0	C
1097	0	C	1142	0	C	1187	0	C
1098	0.0166	C	1143	0	C	1188	0	C
1099	0.0166	C	1144	0.1898	V	1189	0	C
1100	0	C	1145	0	C	1190	0	C
1101	0	C	1146	0	C	1191	0	C
1102	0.0166	C	1147	0	C	1192	0	C
1103	0	C	1148	0	C	1193	0	C
1104	0.0166	C	1149	0.7848	AV	1194	0	C
1105	0	C	1150	0.0166	C	1195	0	C
1106	0	C	1151	0	C	1196	0	C
1107	0.0166	C	1152	0	C	1197	0	C
1108	0	C	1153	0	C	1198	0	C
1109	0.0166	C	1154	0	C	1199	0	C
1110	0	C	1155	0	C	1200	0	C
1111	0	C	1156	0.0166	C	1201	0.0332	C
1112	0.0166	C	1157	0	C	1202	0	C
1113	0.0166	C	1158	0.0166	C	1203	0	C
1114	0.0166	C	1159	0.042	C	1204	0	C

1205	0	C	1250	0	C	1295	0	C
1206	0.0166	C	1251	0	C	1296	0.0421	C
1207	0	C	1252	0	C	1297	0.0332	C
1208	0	C	1253	0	C	1298	0	C
1209	0	C	1254	0	C	1299	0	C
1210	0	C	1255	0.1604	V	1300	0	C
1211	0	C	1256	0.1869	V	1301	0	C
1212	0	C	1257	0.1448	V	1302	0.1026	V
1213	0.0166	C	1258	0	C	1303	0	C
1214	0.0166	C	1259	0.0465	C	1304	0.1368	V
1215	0	C	1260	0	C	1305	0.1869	V
1216	0	C	1261	0	C	1306	0.1898	V
1217	0	C	1262	0.0166	C	1307	0	C
1218	0.0166	C	1263	0.0166	C	1308	0.042	C
1219	0	C	1264	0.0166	C	1309	0	C
1220	0	C	1265	0	C	1310	0	C
1221	0	C	1266	0.0166	C	1311	0	C
1222	0	C	1267	0	C	1312	0	C
1223	0	C	1268	0.0299	C	1313	0.0166	C
1224	0	C	1269	0.6901	AV	1314	0	C
1225	0	C	1270	0	C	1315	0.0166	C
1226	0	C	1271	0.1026	V	1316	0	C
1227	0.0166	C	1272	0.775	AV	1317	0	C
1228	0.0166	C	1273	0	C	1318	0	C
1229	0	C	1274	0.5743	AV	1319	0.0533	C
1230	0	C	1275	0.6666	AV	1320	0	C
1231	0	C	1276	0.5791	AV	1321	0	C
1232	0	C	1277	0.0166	C	1322	0.6807	AV
1233	0	C	1278	0	C	1323	0.0166	C
1234	0	C	1279	0.1273	V	1324	0	C
1235	0	C	1280	0	C	1325	0	C
1236	0	C	1281	0.0166	C	1326	0.0299	C
1237	0	C	1282	0	C	1327	0	C
1238	0	C	1283	0.1026	V	1328	0	C
1239	0.0299	C	1284	0.7137	AV	1329	0	C
1240	0	C	1285	0.5743	AV	1330	0	C
1241	0.0332	C	1286	0.0166	C	1331	0	C
1242	0	C	1287	0.5668	AV	1332	0.042	C
1243	0.0166	C	1288	0	C	1333	0	C
1244	0	C	1289	0.042	C	1334	0.042	C
1245	0	C	1290	0.6234	AV	1335	0	C
1246	0.0839	V	1291	0.0166	C	1336	0	C
1247	0	C	1292	0	C	1337	0	C
1248	0.064	C	1293	0	C	1338	0	C
1249	0	C	1294	0.084	V	1339	0.6807	AV

1340	0.0166	C	1385	0	C	1430	0.0428	C
1341	0	C	1386	0	C	1431	0.017	C
1342	0.0587	C	1387	0.0166	C	1432	0	C
1343	0	C	1388	0.0166	C	1433	0	C
1344	0	C	1389	0	C	1434	0.017	C
1345	0	C	1390	0.0166	C	1435	0.0306	C
1346	0	C	1391	0	C	1436	0.017	C
1347	0.1527	V	1392	0	C	1437	0.0599	C
1348	0.0166	C	1393	0	C	1438	0.0306	C
1349	0	C	1394	0	C	1439	0.017	C
1350	0	C	1395	0	C	1440	0.0544	C
1351	0	C	1396	0	C	1441	0	C
1352	0	C	1397	0	C	1442	0.017	C
1353	0	C	1398	0	C	1443	0.0171	C
1354	0	C	1399	0.0909	V	1444	0.1406	V
1355	0	C	1400	0.03	C	1445	0	C
1356	0.0586	C	1401	0	C	1446	0.0172	C
1357	0	C	1402	0	C	1447	0.0483	C
1358	0	C	1403	0	C	1448	0.0871	V
1359	0	C	1404	0	C	1449	0	C
1360	0.0299	C	1405	0	C	1450	0.0636	C
1361	0	C	1406	0.03	C	1451	0.1127	V
1362	0	C	1407	0	C	1452	0.0855	V
1363	0	C	1408	0.0167	C	1453	0.1246	V
1364	0	C	1409	0.0167	C	1454	0.1875	V
1365	0	C	1410	0.1142	V	1455	0.077	C
1366	0.0166	C	1411	0.0167	C	1456	0	C
1367	0	C	1412	0.0167	C	1457	0.0183	C
1368	0.0166	C	1413	0.0633	C	1458	0.0186	C
1369	0	C	1414	0.0754	C	1459	0.0472	C
1370	0	C	1415	0.0334	C	1460	0	C
1371	0	C	1416	0.0703	C	1461	0.0475	C
1372	0	C	1417	0.0746	C	1462	0.2007	V
1373	0.0587	C	1418	0	C	1463	0.1346	V
1374	0	C	1419	0.0303	C	1464	0.7065	AV
1375	0.0166	C	1420	0	C	1465	0.1925	V
1376	0.1368	V	1421	0	C	1466	0.2213	V
1377	0	C	1422	0.0168	C	1467	0.0879	V
1378	0.0166	C	1423	0.0168	C	1468	0.0398	C
1379	0.0166	C	1424	0.0169	C	1469	0.0766	C
1380	0.0166	C	1425	0.0854	V	1470	0.1342	V
1381	0.0166	C	1426	0.1559	V	1471	0.0375	C
1382	0.0166	C	1427	0	C	1472	0.0755	C
1383	0.0166	C	1428	0.0305	C	1473	0.1053	V
1384	0.0299	C	1429	0.1633	V	1474	0.1616	V

1475	0.0929	V	1520	0	C
1476	0.0528	C	1521	0.3802	V
1477	0.2034	V	1522	0	C
1478	0.2564	V	1523	0	C
1479	0.1843	V	1524	0	C
1480	0.136	V	1525	0	C
1481	0.0424	C	1526	0	C
1482	0.0667	C			
1483	0.0785	C			
1484	0.0291	C			
1485	0.3704	V			
1486	0.2686	V			
1487	0.201	V			
1488	0.169	V			
1489	0.0597	C			
1490	0.1109	V			
1491	0.0739	C			
1492	0.0909	V			
1493	0.037	C			
1494	0.0372	C			
1495	0.0374	C			
1496	0.0376	C			
1497	0.1084	V			
1498	0.2946	V			
1499	0	C			
1500	0	C			
1501	0.0407	C			
1502	0.0853	V			
1503	0.0984	V			
1504	0.1139	V			
1505	0.0479	C			
1506	0.0842	V			
1507	0.0482	C			
1508	0	C			
1509	0	C			
1510	0.1404	V			
1511	0.1119	V			
1512	0.0565	C			
1513	0	C			
1514	0.2392	V			
1515	0.1047	V			
1516	0.0605	C			
1517	0.0659	C			
1518	0.0693	C			
1519	0.0701	C			

En el cuadro se enlistan las 1,526 posiciones del conjunto de Greengenes con su respectiva entropía y la clasificación a la que cada posición corresponde; C: Conservado, V: Variante y AV: Altamente Variante; H(i): entropía, Pos: posición.

13.7. Clasificación mediante K-medias conjunto Fusionado

Pos	H(i)	Grupo						
1	0.152	V	40	0.0417	C	81	0	C
2	0.0871	V	41	0.1454	V	82	0.0254	C
3	0.0837	V	42	0	C	83	0.1513	V
4	0.0815	V	43	0.0366	C	84	0.1766	V
5	0.0785	C	44	0.031	C	85	0.6642	AV
6	0	C	45	0.031	C	86	0.2133	V
7	0	C	46	0.0306	C	87	0.5232	AV
8	0.0724	C	47	0	C	88	0.5836	AV
9	0.0708	C	48	0.0297	C	89	0.6227	AV
10	0.1599	V	49	0	C	90	0.6073	AV
11	0.4905	AV	50	0	C	91	0.6627	AV
12	0.7585	AV	51	0	C	92	0.9265	AV
13	0.1169	V	52	0.0289	C	93	0.5634	AV
14	0	C	53	0.0302	C	94	0.5554	AV
15	0.064	C	54	0.0289	C	95	0.4472	AV
16	0.064	C	55	0.0289	C	96	0.0958	V
17	0	C	56	0.0805	V	97	0.105	V
18	0.2509	V	57	0.0805	V	98	0.4743	AV
19	0	C	58	0.6715	AV	99	0.1378	V
20	0	C	59	0.0802	V	100	0.9866	AV
21	0.0584	C	60	0.0802	V	101	0.0602	C
22	0.0574	C	61	0.132	V	102	0	C
23	0.0565	C	62	0.1369	V	103	0.017	C
24	0.0555	C	63	0.0805	V	104	0	C
25	0.1502	V	64	0.2842	V	105	0.114	V
26	0	C	65	0.5947	AV	106	0.0508	C
27	0.0503	C	66	0.1621	V	107	0.0428	C
28	0	C	67	0.0905	V	108	0.0304	C
29	0.0482	C	68	0.5007	AV	109	0.583	AV
30	0.0469	C	69	0.6628	AV	110	0	C
31	0	C	70	0.6519	AV	111	0.0168	C
32	0	C	71	0.7297	AV	112	0.1034	V
33	0.0448	C	72	0.3997	V	113	0	C
34	0.0434	C	73	0.7379	AV	114	0.0167	C
35	0.0422	C	74	0.2076	V	115	0	C
36	0.0417	C	75	0.6077	AV	116	0	C
37	0.0414	C	76	0.5236	AV	117	0	C
38	0.0407	C	77	0.0254	C	118	0.0302	C
39	0	C	78	0.5516	AV	119	0	C
			79	0.164	V	120	0.0167	C
			80	0	C	121	0.0167	C

122	0	C	167	0	C	212	0	C
123	0	C	168	0.137	V	213	0	C
124	0	C	169	0.0166	C	214	0.0299	C
125	0.0333	C	170	0	C	215	0	C
126	0.0167	C	171	0	C	216	0.2593	V
127	0	C	172	0	C	217	0.1286	V
128	0	C	173	0	C	218	0	C
129	0.0333	C	174	0	C	219	0.1368	V
130	0.1196	V	175	0	C	220	0.1583	V
131	0	C	176	0.0166	C	221	0.0598	C
132	0.0167	C	177	0	C	222	0.0299	C
133	0	C	178	0	C	223	0.6633	AV
134	0.0167	C	179	0	C	224	0.2261	V
135	0.0167	C	180	0.6403	AV	225	0.4963	AV
136	0.5942	AV	181	0	C	226	0.1203	V
137	0.0167	C	182	0.0166	C	227	0.0332	C
138	0	C	183	0.0166	C	228	0.1503	V
139	0.0167	C	184	0.042	C	229	0.2151	V
140	0.1206	V	185	0.2653	V	230	0.0332	C
141	0.186	V	186	0.1669	V	231	0.5343	AV
142	0.2306	V	187	0.9667	AV	232	0.042	C
143	0.0332	C	188	0.2395	V	233	0	C
144	0.0587	C	189	0.9421	AV	234	0.2186	V
145	0.3261	V	190	0.9412	AV	235	0.2295	V
146	0.6549	AV	191	0.5088	AV	236	0.11	V
147	0.0332	C	192	0.8568	AV	237	0	C
148	0	C	193	0.9436	AV	238	0	C
149	0.0166	C	194	0.217	V	239	0	C
150	0.0166	C	195	0.0166	C	240	0	C
151	0	C	196	0.1557	V	241	0	C
152	0	C	197	0	C	242	0	C
153	0.0166	C	198	0.0719	C	243	0.128	V
154	0.0166	C	199	0	C	244	0.0166	C
155	0	C	200	0.0166	C	245	0	C
156	0	C	201	0.2097	V	246	0.0166	C
157	0.0743	C	202	0.9459	AV	247	0.0166	C
158	0.1282	V	203	0.8383	AV	248	0	C
159	0.0332	C	204	0.543	AV	249	0	C
160	0.0166	C	205	1.0291	AV	250	0.064	C
161	0.0499	C	206	0.8243	AV	251	0.1288	V
162	0.0466	C	207	0.8029	AV	252	0	C
163	0.0166	C	208	0.9788	AV	253	0.1286	V
164	0	C	209	0.6887	AV	254	0	C
165	0	C	210	0.0166	C	255	0	C
166	0.0166	C	211	0.0166	C	256	0	C

257	0.0166	C	302	0	C	347	0	C
258	0.0631	C	303	0	C	348	0.6684	AV
259	0.0166	C	304	0	C	349	0.0166	C
260	0	C	305	0.0299	C	350	0	C
261	0	C	306	0.0166	C	351	0	C
262	0.0166	C	307	0	C	352	0.0166	C
263	0.1191	V	308	0.0166	C	353	0.0167	C
264	0	C	309	0	C	354	0	C
265	0	C	310	0.0166	C	355	0	C
266	0	C	311	0.0299	C	356	0	C
267	0.0465	C	312	0	C	357	0.0166	C
268	0.1115	V	313	0	C	358	0	C
269	0.2105	V	314	0.0299	C	359	0	C
270	0	C	315	0	C	360	0	C
271	0	C	316	0.1115	V	361	0	C
272	0	C	317	0	C	362	0	C
273	0	C	318	0	C	363	0	C
274	0	C	319	0	C	364	0	C
275	0.0465	C	320	0	C	365	0	C
276	0	C	321	0	C	366	0	C
277	0	C	322	0	C	367	0	C
278	0	C	323	0	C	368	0.0166	C
279	0.1201	V	324	0	C	369	0	C
280	0.1201	V	325	0	C	370	0	C
281	0	C	326	0	C	371	0	C
282	0	C	327	0.668	AV	372	0	C
283	0	C	328	0	C	373	0	C
284	0	C	329	0	C	374	0	C
285	0.1026	V	330	0	C	375	0	C
286	0	C	331	0	C	376	0	C
287	0.0166	C	332	0	C	377	0	C
288	0	C	333	0	C	378	0	C
289	0.777	AV	334	0	C	379	0	C
290	0.0586	C	335	0	C	380	0	C
291	0.1826	V	336	0	C	381	0.1583	V
292	0.0166	C	337	0.0166	C	382	0.1286	V
293	0.0166	C	338	0	C	383	0	C
294	0	C	339	0	C	384	0	C
295	0.1286	V	340	0	C	385	0	C
296	0	C	341	0	C	386	0.0166	C
297	0.1286	V	342	0	C	387	0	C
298	0.0166	C	343	0	C	388	0	C
299	0	C	344	0.0332	C	389	0	C
300	0.0166	C	345	0	C	390	0	C
301	0	C	346	0	C	391	0.0166	C

392	0.2314	V	437	0	C	482	0	C
393	0.0166	C	438	0	C	483	0.6187	AV
394	0	C	439	0	C	484	0.5484	AV
395	0	C	440	0	C	485	0.0598	C
396	0	C	441	0	C	486	0.6393	AV
397	0	C	442	0	C	487	0.2842	V
398	0	C	443	0	C	488	0.2104	V
399	0	C	444	0.0166	C	489	0.1324	V
400	0	C	445	0	C	490	0.128	V
401	0.1286	V	446	0.0166	C	491	0.615	AV
402	0.1286	V	447	0.0166	C	492	0	C
403	0	C	448	0.042	C	493	0	C
404	0	C	449	0	C	494	0	C
405	0	C	450	0	C	495	0	C
406	0	C	451	0	C	496	0	C
407	0	C	452	0	C	497	0	C
408	0	C	453	0.0166	C	498	0	C
409	0	C	454	0.0166	C	499	0.0166	C
410	0	C	455	0.0299	C	500	0	C
411	0	C	456	0.128	V	501	0.128	V
412	0	C	457	0	C	502	0.042	C
413	0	C	458	0	C	503	0.0166	C
414	0	C	459	0	C	504	0	C
415	0	C	460	0	C	505	0	C
416	0	C	461	0.0166	C	506	0	C
417	0	C	462	0	C	507	0	C
418	0	C	463	0	C	508	0.0166	C
419	0.0332	C	464	0.1115	V	509	0	C
420	0	C	465	0.0586	C	510	0.0166	C
421	0	C	466	0.1645	V	511	0	C
422	0.0166	C	467	0.1326	V	512	0	C
423	0	C	468	0.1828	V	513	0	C
424	0.0166	C	469	0.3007	V	514	0	C
425	0	C	470	0.6451	AV	515	0	C
426	0	C	471	0.0719	C	516	0	C
427	0.0166	C	472	0.565	AV	517	0	C
428	0	C	473	0.5511	AV	518	0	C
429	0.1413	V	474	0	C	519	0	C
430	0	C	475	0.1286	V	520	0	C
431	0	C	476	0.1201	V	521	0	C
432	0	C	477	0.042	C	522	0.0166	C
433	0	C	478	0.5503	AV	523	0	C
434	0	C	479	0.0166	C	524	0	C
435	0	C	480	0.4767	AV	525	0	C
436	0.1413	V	481	0.2632	V	526	0	C

527	0.0299	C	572	0	C	617	0.0166	C
528	0	C	573	0	C	618	0	C
529	0	C	574	0	C	619	0	C
530	0	C	575	0.0299	C	620	0.0166	C
531	0	C	576	0	C	621	0	C
532	0	C	577	0	C	622	0	C
533	0.0167	C	578	0	C	623	0.0166	C
534	0	C	579	0.0166	C	624	0	C
535	0	C	580	0	C	625	0.1115	V
536	0	C	581	0	C	626	0.1533	V
537	0	C	582	0	C	627	0	C
538	0.0166	C	583	0	C	628	0	C
539	0	C	584	0.0332	C	629	0	C
540	0	C	585	0.0166	C	630	0.0166	C
541	0	C	586	0	C	631	0.0166	C
542	0.0166	C	587	0	C	632	0.1527	V
543	0.0166	C	588	0	C	633	0.0166	C
544	0.0166	C	589	0	C	634	0.0166	C
545	0	C	590	0	C	635	0	C
546	0	C	591	0	C	636	0	C
547	0	C	592	0.0166	C	637	0	C
548	0	C	593	0	C	638	0.1451	V
549	0.0299	C	594	0.0166	C	639	0.1115	V
550	0	C	595	0	C	640	0.0166	C
551	0	C	596	0.0166	C	641	0	C
552	0.0167	C	597	0.0332	C	642	0.1288	V
553	0.0166	C	598	0	C	643	0	C
554	0	C	599	0.0166	C	644	0.1413	V
555	0	C	600	0.0299	C	645	0.1286	V
556	0.0166	C	601	0	C	646	0	C
557	0.0166	C	602	0.064	C	647	0.0332	C
558	0.0466	C	603	0.1201	V	648	0	C
559	0.0166	C	604	0.2302	V	649	0.1201	V
560	0.0166	C	605	0	C	650	0	C
561	0.0166	C	606	0	C	651	0.0166	C
562	0.0166	C	607	0	C	652	0.0166	C
563	0.0332	C	608	0	C	653	0	C
564	0	C	609	0.0166	C	654	0.0332	C
565	0.0421	C	610	0	C	655	0	C
566	0	C	611	0	C	656	0	C
567	0.0166	C	612	0	C	657	0	C
568	0.0166	C	613	0.0166	C	658	0	C
569	0.0166	C	614	0	C	659	0.0166	C
570	0	C	615	0	C	660	0.1201	V
571	0	C	616	0	C	661	0.3328	V

662	0	C	707	0	C	752	0	C
663	0	C	708	0	C	753	0	C
664	0	C	709	0	C	754	0	C
665	0.0166	C	710	0	C	755	0	C
666	0	C	711	0	C	756	0	C
667	0	C	712	0	C	757	0	C
668	0	C	713	0	C	758	0	C
669	0	C	714	0	C	759	0.0166	C
670	0	C	715	0.0498	C	760	0	C
671	0	C	716	0.0166	C	761	0	C
672	0	C	717	0	C	762	0	C
673	0.042	C	718	0	C	763	0	C
674	0	C	719	0.0166	C	764	0	C
675	0	C	720	0	C	765	0	C
676	0	C	721	0	C	766	0	C
677	0	C	722	0	C	767	0	C
678	0.0166	C	723	0	C	768	0	C
679	0	C	724	0	C	769	0	C
680	0	C	725	0.1453	V	770	0	C
681	0.1583	V	726	0.0166	C	771	0.0167	C
682	0.0166	C	727	0	C	772	0.0907	V
683	0	C	728	0	C	773	0	C
684	0.0299	C	729	0	C	774	0.0166	C
685	0	C	730	0.0166	C	775	0	C
686	0	C	731	0	C	776	0	C
687	0.0166	C	732	0.0166	C	777	0.0166	C
688	0	C	733	0.0166	C	778	0	C
689	0.1286	V	734	0.0299	C	779	0	C
690	0	C	735	0	C	780	0.0166	C
691	0	C	736	0	C	781	0.0166	C
692	0	C	737	0	C	782	0.1026	V
693	0	C	738	0	C	783	0.1451	V
694	0	C	739	0.0166	C	784	0	C
695	0	C	740	0	C	785	0.0166	C
696	0	C	741	0.0166	C	786	0.0167	C
697	0	C	742	0.0742	C	787	0	C
698	0	C	743	0.0299	C	788	0.1286	V
699	0	C	744	0	C	789	0	C
700	0.0299	C	745	0.0299	C	790	0	C
701	0	C	746	0.1604	V	791	0	C
702	0	C	747	0	C	792	0.3121	V
703	0.0166	C	748	0	C	793	0	C
704	0.0332	C	749	0.1704	V	794	0	C
705	0.0166	C	750	0	C	795	0	C
706	0.0166	C	751	0	C	796	0.0299	C

797	0.0166	C	842	0	C	887	0	C
798	0.0166	C	843	0	C	888	0.0166	C
799	0	C	844	0	C	889	0.0166	C
800	0	C	845	0.0166	C	890	0.1451	V
801	0.0166	C	846	0	C	891	0	C
802	0	C	847	0	C	892	0	C
803	0.0166	C	848	0.0299	C	893	0	C
804	0	C	849	0	C	894	0.0166	C
805	0	C	850	0	C	895	0	C
806	0	C	851	0	C	896	0.0332	C
807	0.0166	C	852	0	C	897	0.0299	C
808	0.0166	C	853	0	C	898	0.0166	C
809	0.042	C	854	0	C	899	0	C
810	0	C	855	0	C	900	0.0166	C
811	0	C	856	0	C	901	0	C
812	0.0166	C	857	0	C	902	0	C
813	0	C	858	0.0332	C	903	0	C
814	0	C	859	0.0299	C	904	0	C
815	0	C	860	0	C	905	0	C
816	0.0332	C	861	0.064	C	906	0	C
817	0.0299	C	862	0.0166	C	907	0	C
818	0	C	863	0	C	908	0.1604	V
819	0.0332	C	864	0.0166	C	909	0	C
820	0.1286	V	865	0	C	910	0	C
821	0.1026	V	866	0.0166	C	911	0.0166	C
822	0	C	867	0.0166	C	912	0.0698	C
823	0	C	868	0	C	913	0.0166	C
824	0	C	869	0.0465	C	914	0.0166	C
825	0.0166	C	870	0.042	C	915	0.0299	C
826	0	C	871	0	C	916	0.0166	C
827	0.0166	C	872	0	C	917	0.1604	V
828	0.0166	C	873	0	C	918	0.0332	C
829	0	C	874	0	C	919	0.0166	C
830	0	C	875	0	C	920	0	C
831	0	C	876	0.6343	AV	921	0	C
832	0.0166	C	877	0	C	922	0.0166	C
833	0	C	878	0	C	923	0.0332	C
834	0	C	879	0	C	924	0	C
835	0.1367	V	880	0	C	925	0.0166	C
836	0	C	881	0	C	926	0.0166	C
837	0.0166	C	882	0.0166	C	927	0	C
838	0.0166	C	883	0	C	928	0.0466	C
839	0	C	884	0	C	929	0.0166	C
840	0	C	885	0	C	930	0	C
841	0.0465	C	886	0.0166	C	931	0	C

932	0.03	C	977	0	C	1022	0.1604	V
933	0	C	978	0	C	1023	0.1604	V
934	0.0299	C	979	0	C	1024	0.1853	V
935	0	C	980	0	C	1025	0	C
936	0	C	981	0.1201	V	1026	0.0299	C
937	0	C	982	0.0166	C	1027	0	C
938	0.0166	C	983	0	C	1028	0.0166	C
939	0	C	984	0.0299	C	1029	0	C
940	0	C	985	0.042	C	1030	0	C
941	0.0166	C	986	0	C	1031	0.1853	V
942	0.0299	C	987	0	C	1032	0.1604	V
943	0	C	988	0	C	1033	0.1527	V
944	0	C	989	0.0166	C	1034	0.1368	V
945	0.0166	C	990	0.0166	C	1035	0.1005	V
946	0.0166	C	991	0.0166	C	1036	0.1413	V
947	0	C	992	0	C	1037	0	C
948	0.0299	C	993	0.0166	C	1038	0.6309	AV
949	0	C	994	0.0166	C	1039	0	C
950	0	C	995	0	C	1040	0	C
951	0	C	996	0	C	1041	0.03	C
952	0.0166	C	997	0	C	1042	0.0332	C
953	0	C	998	0	C	1043	0.0698	C
954	0.0167	C	999	0	C	1044	0.0167	C
955	0	C	1000	0	C	1045	0	C
956	0	C	1001	0	C	1046	0	C
957	0	C	1002	0	C	1047	0	C
958	0	C	1003	0	C	1048	0.03	C
959	0	C	1004	0	C	1049	0.6567	AV
960	0.0166	C	1005	0	C	1050	0.0299	C
961	0	C	1006	0	C	1051	0	C
962	0.0166	C	1007	0.0166	C	1052	0.1191	V
963	0	C	1008	0	C	1053	0.1026	V
964	0	C	1009	0	C	1054	0.064	C
965	0	C	1010	0.0166	C	1055	0.1026	V
966	0	C	1011	0.1191	V	1056	0.1026	V
967	0	C	1012	0.0533	C	1057	0.0166	C
968	0.0299	C	1013	0	C	1058	0.0166	C
969	0	C	1014	0.1193	V	1059	0	C
970	0	C	1015	0	C	1060	0	C
971	0.03	C	1016	0.0166	C	1061	0	C
972	0	C	1017	0.1026	V	1062	0	C
973	0	C	1018	0.1533	V	1063	0.0166	C
974	0	C	1019	0.1115	V	1064	0	C
975	0.1118	V	1020	0.084	V	1065	0.0166	C
976	0.0166	C	1021	0.1448	V	1066	0.0166	C

1067	0	C	1112	0.0166	C	1157	0	C
1068	0	C	1113	0.0166	C	1158	0.0166	C
1069	0	C	1114	0.0166	C	1159	0.042	C
1070	0	C	1115	0	C	1160	0.0166	C
1071	0	C	1116	0.0299	C	1161	0.0166	C
1072	0.042	C	1117	0	C	1162	0.0166	C
1073	0.0166	C	1118	0	C	1163	0.1115	V
1074	0	C	1119	0	C	1164	0.0332	C
1075	0.0166	C	1120	0.0299	C	1165	0.1368	V
1076	0.0598	C	1121	0	C	1166	0.6901	AV
1077	0.0166	C	1122	0	C	1167	0	C
1078	0	C	1123	0	C	1168	0	C
1079	0.0166	C	1124	0	C	1169	0	C
1080	0	C	1125	0	C	1170	0.0166	C
1081	0	C	1126	0	C	1171	0	C
1082	0	C	1127	0	C	1172	0	C
1083	0.0166	C	1128	0.0299	C	1173	0	C
1084	0.0166	C	1129	0	C	1174	0	C
1085	0	C	1130	0	C	1175	0	C
1086	0.0166	C	1131	0	C	1176	0.0299	C
1087	0	C	1132	0.042	C	1177	0.0299	C
1088	0	C	1133	0.137	V	1178	0.0166	C
1089	0	C	1134	0.0805	V	1179	0	C
1090	0	C	1135	0.1115	V	1180	0	C
1091	0	C	1136	0	C	1181	0	C
1092	0	C	1137	0	C	1182	0	C
1093	0.0166	C	1138	0	C	1183	0.0465	C
1094	0.0166	C	1139	0	C	1184	0.0465	C
1095	0	C	1140	0	C	1185	0	C
1096	0	C	1141	0	C	1186	0	C
1097	0	C	1142	0	C	1187	0	C
1098	0.0166	C	1143	0	C	1188	0	C
1099	0.0166	C	1144	0.1898	V	1189	0	C
1100	0	C	1145	0	C	1190	0	C
1101	0	C	1146	0	C	1191	0	C
1102	0.0166	C	1147	0	C	1192	0	C
1103	0	C	1148	0	C	1193	0	C
1104	0.0166	C	1149	0.7848	AV	1194	0	C
1105	0	C	1150	0.0166	C	1195	0	C
1106	0	C	1151	0	C	1196	0	C
1107	0.0166	C	1152	0	C	1197	0	C
1108	0	C	1153	0	C	1198	0	C
1109	0.0166	C	1154	0	C	1199	0	C
1110	0	C	1155	0	C	1200	0	C
1111	0	C	1156	0.0166	C	1201	0.0332	C

1202	0	C	1247	0	C	1292	0	C
1203	0	C	1248	0.064	C	1293	0	C
1204	0	C	1249	0	C	1294	0.084	V
1205	0	C	1250	0	C	1295	0	C
1206	0.0166	C	1251	0	C	1296	0.0421	C
1207	0	C	1252	0	C	1297	0.0332	C
1208	0	C	1253	0	C	1298	0	C
1209	0	C	1254	0	C	1299	0	C
1210	0	C	1255	0.1604	V	1300	0	C
1211	0	C	1256	0.1869	V	1301	0	C
1212	0	C	1257	0.1448	V	1302	0.1026	V
1213	0.0166	C	1258	0	C	1303	0	C
1214	0.0166	C	1259	0.0465	C	1304	0.1368	V
1215	0	C	1260	0	C	1305	0.1869	V
1216	0	C	1261	0	C	1306	0.1898	V
1217	0	C	1262	0.0166	C	1307	0	C
1218	0.0166	C	1263	0.0166	C	1308	0.042	C
1219	0	C	1264	0.0166	C	1309	0	C
1220	0	C	1265	0	C	1310	0	C
1221	0	C	1266	0.0166	C	1311	0	C
1222	0	C	1267	0	C	1312	0	C
1223	0	C	1268	0.0299	C	1313	0.0166	C
1224	0	C	1269	0.6901	AV	1314	0	C
1225	0	C	1270	0	C	1315	0.0166	C
1226	0	C	1271	0.1026	V	1316	0	C
1227	0.0166	C	1272	0.775	AV	1317	0	C
1228	0.0166	C	1273	0	C	1318	0	C
1229	0	C	1274	0.5743	AV	1319	0.0533	C
1230	0	C	1275	0.6666	AV	1320	0	C
1231	0	C	1276	0.5791	AV	1321	0	C
1232	0	C	1277	0.0166	C	1322	0.6807	AV
1233	0	C	1278	0	C	1323	0.0166	C
1234	0	C	1279	0.1273	V	1324	0	C
1235	0	C	1280	0	C	1325	0	C
1236	0	C	1281	0.0166	C	1326	0.0299	C
1237	0	C	1282	0	C	1327	0	C
1238	0	C	1283	0.1026	V	1328	0	C
1239	0.0299	C	1284	0.7137	AV	1329	0	C
1240	0	C	1285	0.5743	AV	1330	0	C
1241	0.0332	C	1286	0.0166	C	1331	0	C
1242	0	C	1287	0.5668	AV	1332	0.042	C
1243	0.0166	C	1288	0	C	1333	0	C
1244	0	C	1289	0.042	C	1334	0.042	C
1245	0	C	1290	0.6234	AV	1335	0	C
1246	0.0839	V	1291	0.0166	C	1336	0	C

1337	0	C	1382	0.0166	C	1427	0	C
1338	0	C	1383	0.0166	C	1428	0.0305	C
1339	0.6807	AV	1384	0.0299	C	1429	0.1633	V
1340	0.0166	C	1385	0	C	1430	0.0428	C
1341	0	C	1386	0	C	1431	0.017	C
1342	0.0587	C	1387	0.0166	C	1432	0	C
1343	0	C	1388	0.0166	C	1433	0	C
1344	0	C	1389	0	C	1434	0.017	C
1345	0	C	1390	0.0166	C	1435	0.0306	C
1346	0	C	1391	0	C	1436	0.017	C
1347	0.1527	V	1392	0	C	1437	0.0599	C
1348	0.0166	C	1393	0	C	1438	0.0306	C
1349	0	C	1394	0	C	1439	0.017	C
1350	0	C	1395	0	C	1440	0.0544	C
1351	0	C	1396	0	C	1441	0	C
1352	0	C	1397	0	C	1442	0.017	C
1353	0	C	1398	0	C	1443	0.0171	C
1354	0	C	1399	0.0909	V	1444	0.1406	V
1355	0	C	1400	0.03	C	1445	0	C
1356	0.0586	C	1401	0	C	1446	0.0172	C
1357	0	C	1402	0	C	1447	0.0483	C
1358	0	C	1403	0	C	1448	0.0871	V
1359	0	C	1404	0	C	1449	0	C
1360	0.0299	C	1405	0	C	1450	0.0636	C
1361	0	C	1406	0.03	C	1451	0.1127	V
1362	0	C	1407	0	C	1452	0.0855	V
1363	0	C	1408	0.0167	C	1453	0.1246	V
1364	0	C	1409	0.0167	C	1454	0.1875	V
1365	0	C	1410	0.1142	V	1455	0.077	C
1366	0.0166	C	1411	0.0167	C	1456	0	C
1367	0	C	1412	0.0167	C	1457	0.0183	C
1368	0.0166	C	1413	0.0633	C	1458	0.0186	C
1369	0	C	1414	0.0754	C	1459	0.0472	C
1370	0	C	1415	0.0334	C	1460	0	C
1371	0	C	1416	0.0703	C	1461	0.0475	C
1372	0	C	1417	0.0746	C	1462	0.2007	V
1373	0.0587	C	1418	0	C	1463	0.1346	V
1374	0	C	1419	0.0303	C	1464	0.0613	C
1375	0.0166	C	1420	0	C	1465	0.1925	V
1376	0.1368	V	1421	0	C	1466	0.2213	V
1377	0	C	1422	0.0168	C	1467	0.0879	V
1378	0.0166	C	1423	0.0168	C	1468	0.0398	C
1379	0.0166	C	1424	0.0169	C	1469	0.0766	C
1380	0.0166	C	1425	0.0854	V	1470	0.1342	V
1381	0.0166	C	1426	0.1559	V	1471	0.0375	C

1472	0.0755	C
1473	0.1053	V
1474	0.1616	V
1475	0.0929	V
1476	0.0528	C
1477	0.2034	V
1478	0.2564	V
1479	0.1843	V
1480	0.136	V
1481	0.0424	C
1482	0.0667	C
1483	0.0785	C
1484	0.0291	C
1485	0.3704	V
1486	0.2686	V
1487	0.201	V
1488	0.169	V
1489	0.0597	C
1490	0.1109	V
1491	0.0739	C
1492	0.0909	V
1493	0.037	C
1494	0.0372	C
1495	0.0374	C
1496	0.0376	C
1497	0.1084	V
1498	0.2946	V
1499	0	C
1500	0	C
1501	0.0407	C
1502	0.0853	V
1503	0.0984	V
1504	0.1139	V
1505	0.0479	C
1506	0.0842	V
1507	0.0482	C
1508	0	C
1509	0	C
1510	0.1404	V
1511	0.1119	V
1512	0.0565	C
1513	0	C

En el cuadro se enlistan las 1,513 posiciones del conjunto Fusionado con su respectiva entropía y la clasificación a la que cada posición corresponde; C: Conservado, V: Variante y AV: Altamente Variante; H(i): entropía, Pos: posición.

13.8. Lista de modelos ordenados por puntajes CIB árbol ARNr 16S

Modelo	-LnL	df	AIC	AICc	CIB
GTR+F	140719.432	4459	290356.865	40064636.9	314130.982
GTR+F+I	139873.313	4460	288666.627	40080786.6	312446.076
GTR+F+G4	115363.464	4460	239646.928	40031766.9	263426.377
GTR+F+I+G4	115298.161	4461	239518.323	40049482.3	263303.103
SYM+I+G4	115251.36	4458	239418.721	39995862.7	263187.506
TVM+F+I+G4	115358.614	4460	239637.229	40031757.2	263416.677
TVMe+I+G4	115274.882	4457	239463.763	39978075.8	263227.217
TIM3+F+I+G4	115301.813	4459	239521.626	40013801.6	263295.743
TIM3e+I+G4	115253.094	4456	239418.187	39960202.2	263176.309
TIM2+F+I+G4	115371.142	4459	239660.284	40013940.3	263434.401
TIM2e+I+G4	115268.215	4456	239448.43	39960232.4	263206.552
TIM+F+I+G4	115374.685	4459	239667.37	40013947.4	263441.487
TIMe+I+G4	115269.859	4456	239451.717	39960235.7	263209.839
TPM3u+F+I+G4	115362.572	4458	239641.143	39996085.1	263409.929
TPM3+I+G4	115276.56	4455	239463.119	39942423.1	263215.91
TPM2u+F+I+G4	115431.172	4458	239778.344	39996222.3	263547.129
TPM2+I+G4	115295.488	4455	239500.976	39942461	263253.767
K3Pu+F+I+G4	115434.946	4458	239785.892	39996229.9	263554.678
K3P+I+G4	115296.956	4455	239503.912	39942463.9	263256.702
TN+F+I+G4	115374.812	4458	239665.623	39996109.6	263434.409
TNe+I+G4	115269.931	4455	239449.862	39942409.9	263202.652
HKY+F+I+G4	115435.067	4457	239784.134	39978396.1	263547.588
K2P+I+G4	115297.026	4454	239502.052	39924642.1	263249.511
F81+F+I+G4	116371.56	4456	241655.12	39962439.1	265413.242
JC+I+G4	116211.743	4453	241329.486	39908653.5	265071.613

En el cuadro se enlistan todos los modelos probados para la construcción del árbol por Máxima Verosimilitud, ordenado con base en el mejor puntaje CIB. **AIC**: **P**riterio de información de Akaike puntajes, **AICc**: AIC corregido puntajes. **CIB**: Criterio de información Bayesiano puntajes.

13.9. Lista de modelos ordenados por puntajes CIB árbol de marcadores Similares

Modelo	-LnL	df	AIC	AICc	CIB
GTR+F	40228.986	2957	86371.972	17579984	99601.267
GTR+F+I	40056.185	2958	86028.371	17591472.4	99262.14
GTR+F+G4	35002.884	2958	75921.767	17581365.8	89155.536
GTR+F+I+G4	35002.881	2959	75923.762	17593203.8	89162.005
SYM+G4	35001.339	2955	75912.679	17545872.7	89133.026
SYM+I+G4	35001.334	2956	75914.668	17557698.7	89139.489
TVM+F+G4	35003.359	2957	75920.717	17569532.7	89150.012
TVM+F+I+G4	35003.36	2958	75922.721	17581366.7	89156.489
TVMe+G4	35006.204	2954	75920.408	17534060.4	89136.281
TVMe+I+G4	35006.211	2955	75922.422	17545882.4	89142.769
TIM3+F+G4	35009.775	2956	75931.551	17557715.6	89156.371
TIM3+F+I+G4	35009.777	2957	75933.554	17569545.6	89162.849
TIM3e+G4	35005.63	2953	75917.26	17522241.3	89128.659
TIM3e+I+G4	35005.625	2954	75919.251	17534059.3	89135.124
TIM2+F+G4	35031.036	2956	75974.072	17557758.1	89198.893
TIM2+F+I+G4	35031.038	2957	75976.076	17569588.1	89205.37
TIM2e+G4	35054.132	2953	76014.263	17522338.3	89225.662
TIM2e+I+G4	35054.132	2954	76016.264	17534156.3	89232.137
TIM+F+G4	35033.296	2956	75978.592	17557762.6	89203.413
TIM+F+I+G4	35033.297	2957	75980.595	17569592.6	89209.889
TIMe+G4	35054.044	2953	76014.088	17522338.1	89225.487
TIMe+I+G4	35054.046	2954	76016.092	17534156.1	89231.965
TPM3u+F+G4	35010.177	2955	75930.354	17545890.4	89150.701
TPM3u+F+I+G4	35010.178	2956	75932.357	17557716.4	89157.178
TPM3+G4	35010.292	2952	75924.583	17510436.6	89131.508
TPM3+I+G4	35010.284	2953	75926.568	17522250.6	89137.968
TPM2u+F+G4	35031.673	2955	75973.346	17545933.3	89193.693
TPM2u+F+I+G4	35031.673	2956	75975.346	17557759.3	89200.167
TPM2+G4	35060.15	2952	76024.3	17510536.3	89231.225
TPM2+I+G4	35060.151	2953	76026.302	17522350.3	89237.702
K3Pu+F+G4	35033.839	2955	75977.678	17545937.7	89198.025
K3Pu+F+I+G4	35033.84	2956	75979.681	17557763.7	89204.501
K3P+G4	35059.851	2952	76023.702	17510535.7	89230.627
K3P+I+G4	35059.852	2953	76025.705	17522349.7	89237.104
TN+F+G4	35036.053	2955	75982.107	17545942.1	89202.454
TN+F+I+G4	35036.055	2956	75984.11	17557768.1	89208.931
TNe+G4	35056.527	2952	76017.053	17510529.1	89223.979
TNe+I+G4	35056.528	2953	76019.056	17522343.1	89230.456
HKY+F+G4	35036.591	2954	75981.182	17534121.2	89197.055
HKY+F+I+G4	35036.593	2955	75983.185	17545943.2	89203.532

K2P+G4	35062.338	2951	76026.675	17498730.7	89229.127
K2P+I+G4	35062.339	2952	76028.678	17510540.7	89235.603
F81+F+G4	35251.982	2953	76409.964	17522734	89621.363
F81+F+I+G4	35251.983	2954	76411.965	17534552	89627.838
JC+G4	35275.663	2950	76451.327	17487351.3	89649.304
JC+I+G4	35275.664	2951	76453.327	17499157.3	89655.778

En el cuadro se enlistan todos los modelos probados para la construcción del árbol por Máxima Verosimilitud, ordenado con base en el mejor puntaje CIB. **AIC**: Puntaje de Criterio de información de Akaike, **AICc**: Puntaje de AIC corregido. **CIB**: Puntaje de Criterio de información Bayesiano.

13.10. Lista de modelos ordenados por puntajes CIB árbol de marcadores Disimilares

Modelo	-LnL	df	AIC	AICc	CIB
GTR+F	81837.968	2909	169493.935	17099873.9	183402.032
GTR+F+I	81531.528	2910	168883.056	17110903.1	182795.934
GTR+F+G4	66495.655	2910	138811.309	17080831.3	152724.187
GTR+F+I+G4	66481.197	2911	138784.393	17092448.4	152702.052
SYM+I+G4	66446.706	2908	138709.411	17057453.4	152612.727
TVM+F+I+G4	66481.952	2910	138783.904	17080803.9	152696.781
TVMe+I+G4	66466.21	2907	138746.419	17045858.4	152644.954
TIM3+F+I+G4	66480.6	2909	138779.199	17069159.2	152687.296
TIM3e+I+G4	66447.301	2906	138706.603	17034190.6	152600.356
TIM2+F+I+G4	66481.678	2909	138781.357	17069161.4	152689.454
TIM2e+I+G4	66447.182	2906	138706.364	17034190.4	152600.118
TIM+F+I+G4	66482.335	2909	138782.671	17069162.7	152690.768
TIMe+I+G4	66447.78	2906	138707.56	17034191.6	152601.314
TPM3u+F+I+G4	66482.607	2908	138781.214	17057525.2	152684.529
TPM3+I+G4	66466.786	2905	138743.572	17022603.6	152632.544
TPM2u+F+I+G4	66483.847	2908	138783.694	17057527.7	152687.01
TPM2+I+G4	66466.441	2905	138742.881	17022602.9	152631.854
K3Pu+F+I+G4	66484.498	2908	138784.997	17057529	152688.313
K3P+I+G4	66467.012	2905	138744.025	17022604	152632.997
TN+F+I+G4	66482.329	2908	138780.657	17057524.7	152683.973
TNe+I+G4	66447.78	2905	138705.561	17022565.6	152594.533
HKY+F+I+G4	66484.498	2907	138782.996	17045895	152681.53
K2P+I+G4	66467.013	2904	138742.026	17010982	152626.218
F81+F+I+G4	66903.65	2906	139619.3	17035103.3	153513.053
JC+I+G4	66878.552	2903	139563.103	17000187.1	153442.513

En el cuadro se enlistan todos los modelos probados para la construcción del árbol por Máxima Verosimilitud, ordenado con base en el mejor puntaje CIB. **AIC**: Puntaje de Criterio de información de Akaike, **AICc**: Puntaje de AIC corregido. **CIB**: Puntaje de Criterio de información Bayesiano.