

DETERMINACIÓN Y EVALUACIÓN DE LA MICROBIOTA INTESTINAL COMÚN A LA
POBLACIÓN COLOMBIANA

TRABAJO DE GRADO PRESENTADO COMO REQUISITO PARA OBTENER EL
TÍTULO DE BIÓLOGA

MARGARITA MARÍA VALDERRAMA MAYA

ASESOR: JUAN SEBASTIÁN ESCOBAR RESTREPO, PhD.
CO-ASESOR: NICOLÁS PINEL PELÁEZ, PhD.

UNIVERSIDAD EAFIT
ESCUELA DE CIENCIAS
DEPARTAMENTO DE CIENCIAS BIOLÓGICAS
MEDELLÍN - ANTIOQUIA
2017

Nota de Aceptación

Presidente del Jurado

Jurado

Jurado

Medellín 20, 10, 2017

CONTENIDO

	Pág.
INTRODUCCIÓN	9
JUSTIFICACIÓN	11
OBJETIVOS	12
OBJETIVO GENERAL	12
OBJETIVOS ESPECÍFICOS.....	12
MARCO TEÓRICO.....	13
METODOLOGÍA	18
PROCEDIMIENTO EXPERIMENTAL.....	18
<i>Población de estudio</i>	18
<i>Extracción de ADN</i>	18
<i>Amplificación del gen 16S ADNr</i>	19
<i>Análisis de secuencias y detección de secuencias representativas de OTUs</i>	19
DETERMINACIÓN DEL NÚCLEO.....	20
<i>Análisis de datos</i>	20
RESULTADOS	22
CARACTERÍSTICAS GENERALES DE LA MICROBIOTA	22
DETERMINACIÓN DE LA MICROBIOTA NÚCLEO	22
POTENCIAL FUNCIONAL DEL NÚCLEO DE MEMBRESÍA-COMPOSICIÓN	26
NÚCLEO DE MEMBRESÍA EN SUBCONJUNTOS POBLACIONALES.....	28
DISCUSIÓN	31
CONCLUSIONES Y RECOMENDACIONES	37
DECLARACIONES.....	38
CONSIDERACIONES ÉTICAS.....	38
DISPONIBILIDAD DE LOS DATOS Y MATERIALES.....	38
FINANCIAMIENTO	38
REFERENCIAS.....	39
ANEXOS	45

LISTA DE FIGURAS

	Pág.
Figura 1. Definición de los distintos núcleos microbianos..	16
Figura 2. Paradoja actual sobre la microbiota intestinal..	23
Figura 3. Diagrama de caja y bigotes con la abundancia relativa (en escala logarítmica) de los phyla principales detectados en la población estudiada.	24
Figura 4. Número de OTUs que conforman el núcleo de membresía para diferentes valores de prevalencia poblacional.	24
Figura 5. Predicción del potencial funcional de las OTUs representantes de distintos núcleos de membresía-composición.....	27
Figura 6. Núcleo de membresía con parámetro de prevalencia del 95% para cada subconjunto poblacional.....	29

LISTA DE TABLAS

	Pág.
Tabla 1. Criterios para definir el núcleo microbiano implementado por otros estudios. 17	17
Tabla 2. Variación del núcleo de membresía-composición.....	25
Tabla 3. Núcleo filogenético.....	26

LISTA DE ANEXOS

Anexo 1. Núcleo de membresía con prevalencia del 95%.	45
Anexo 2. Módulos metabólicos predichos en el núcleo funcional.....	48
Anexo 3. Núcleo de membresía por ciudad considerando una prevalencia del 95%. ..	49
Anexo 4. Núcleo de membresía por categorías de índice de masa corporal considerando una prevalencia del 95%.....	50
Anexo 5. Núcleo de membresía por sexo considerando una prevalencia del 95%.	51

GLOSARIO

Biomarcador: característica (ej: conjunto de genes o de especies) que de forma objetiva logra ser indicadora de procesos biológicos normales, procesos patogénicos, o respuestas farmacológicas a una intervención terapéutica.

Clado: es una agrupación que contiene un antepasado común y todos sus descendientes, ya sean vivos o extintos. Es sinónimo de grupo monofilético.

Coevolución es el fenómeno de adaptación evolutiva que ocurre cuando dos o más especies se afectan recíprocamente de distintas formas.

Comensal: organismo que saca provecho de otro sin provocar algún tipo de perjuicio.

Contig: *reads* que han sido ensamblados juntos.

Disbiosis: desbalance del equilibrio normal en la microbiota intestinal que puede conducir a una enfermedad.

Especie molecular: conjunto de organismos que se agrupan únicamente en base a la identidad de sus secuencias de ADN. Este concepto se prefiere sobre el de especie biológica ya que, en organismos procariontes, este último es difuso por la transferencia horizontal de genes entre grupos filogenéticamente distantes, y debido a que no se puede hablar de aislamiento reproductivo en organismos que tienen reproducción asexual. En consecuencia, se ha optado por la identificación molecular y las “especies” (mejor llamadas unidades taxonómicas operativas, por sus siglas en inglés OTUs) se definen a partir de un porcentaje de identidad en las secuencias de ADN (usualmente, mayor o igual al 97%).

Índice de masa corporal: es una medida del estado nutricional de una persona; se define como el peso en kg/estatura en m al cuadrado (kg/m^2). En otras palabras, indica la cantidad de masa de una persona para una estatura dada. Cuando este índice está entre 18.5 y 24.9 kg/m^2 se considera que una persona tiene un peso adecuado; entre 25.0 y 29.9 kg/m^2 sobrepeso y $>30 \text{ kg/m}^2$ obesidad.

Parafilético: es un grupo que incluye al ancestro común de sus miembros, pero no a todos los descendientes de este.

Polifilético: un grupo constituido por la unión artificial de ramas dispersas del árbol evolutivo.

Secuencia quimérica: secuencia formada a partir de dos o más secuencias unidas entre sí. Las secuencias quiméricas se forman durante la PCR.

Read: secuencia de ADN que se obtiene al final del proceso de secuenciación. Esto, en última instancia, es la secuencia de una sección de un fragmento único de ADN.

Valor selectivo: representación cuantitativa de la selección natural dentro de un contexto evolutivo. Esto puede estar dado respecto a un genotipo o a un fenotipo en un entorno dado. En cualquier caso, describe el éxito reproductivo individual y es igual a la contribución media al acervo genético de la siguiente generación que es hecha por individuos del genotipo o fenotipo especificado.

RESUMEN

La paradoja existente entre la altísima variación entre individuos en la composición microbiana del intestino humano y la homogeneidad en el potencial funcional de tal variedad de especies conduce a explorar la existencia de un núcleo microbiano común a un grupo de individuos. Dicho concepto permitiría entender la fracción mínima de especies y/o funciones que tienen que estar presentes en la microbiota intestinal para que el hospedero funcione adecuadamente. Con este trabajo se hace una exploración del núcleo microbiano a partir de datos de la caracterización de la microbiota intestinal de 441 colombianos. Este análisis está basado en conceptos de la ecología tradicional, como lo son la composición, las relaciones filogenéticas y la membresía de las especies presentes en la comunidad. En esta tesis se encuentra que en la cohorte muestreada no existen unidades taxonómicas operativas (OTUs) compartidas por el 100% de los individuos. Sin embargo, se detectaron cuatro OTUs en el 99% (*Escherichia coli*, *Faecalibacterium prausnitzii*, *Roseburia faecis* y *Blautia* sp.) y 19 OTUs en el 95% de los individuos. A partir de los datos de abundancia se encontró que las OTUs que más prevalecen en la microbiota intestinal de los colombianos son de taxa dominantes. Respecto al potencial funcional hallado bajo distintos parámetros de prevalencia y abundancia, los análisis hechos exponen una serie de módulos metabólicos sobre-representados a medida que se consideran parámetros más estrictos. En conclusión, a pesar de la alta riqueza de especies y la alta variabilidad entre individuos, el análisis de los diferentes núcleos de membresía, composición, filogenético y funcional indican que las 19 OTUs compartidas por el 95% de la cohorte de estudio, junto con sus funciones metabólicas, representarían el núcleo microbiano de la población colombiana.

PALABRAS CLAVE: Microbiota intestinal, población colombiana, biomarcadores, núcleo microbiano.

INTRODUCCIÓN

Nuestro cuerpo alberga un amplio y poco explorado mundo microbiano, conocido como la microbiota humana (The Human Microbiome Project Consortium, 2012). Este complejo de microorganismos simbioses tiene un rol integral en aspectos de la salud y la fisiología humana. En particular, la microbiota del intestino es un eficiente y complejo biorreactor natural, el cual es reconocido como un componente clave para la homeostasis del tracto gastrointestinal (de Vos & de Vos, 2012; Sonnenburg, Angenent, & Gordon, 2004; Tap et al., 2009). Esto, dado que la vasta mayoría de microorganismos que habitan allí codifican funciones metabólicas que la especie humana no ha tenido que evolucionar por sí misma (Bäckhed, Ley, Sonnenburg, Peterson, & Gordon, 2005; Ley, Peterson, & Gordon, 2006; Simpson & Campbell, 2015).

Diversos estudios han establecido que el cambio en la composición y abundancia de este conjunto de microorganismos en el intestino de un individuo puede generar susceptibilidad a determinadas patologías o algunas infecciones (The Human Microbiome Project Consortium, 2012). Por tanto, la estabilidad de las comunidades microbianas que habitan en el intestino puede ser pensada como una propiedad descriptiva de la salud humana. Dicha estabilidad ecológica puede estar dada por la composición o el perfil funcional de estas comunidades, las cuales van a favorecer al hospedero (Bäckhed et al., 2012). Consecuentemente, es preciso sugerir posibles **biomarcadores**¹ de la salud humana, los cuales, en esencia, pueden ser representados por un grupo de **especies moleculares** comunes en un grupo de individuos aparentemente saludables.

Aunque el concepto de núcleo microbiano parece intuitivo (las especies moleculares compartidas por un grupo de individuos), no existe una definición formal y universal para éste (Salonen, Salojärvi, Lahti, & de Vos, 2012). Es por ello que Shade & Handelsman (2012) presentan un marco conceptual novedoso para la identificación de especies comunes a un hábitat, el cual se encuentra basado en los conceptos teóricos de la ecología tradicional. Estos conceptos permiten comprender la complejidad de la microbiota-núcleo y avanzar en la comprensión del papel de estos microorganismos y sus funciones dentro del intestino (Shade & Handelsman, 2012). Adicionalmente, la búsqueda de un núcleo microbiano se ve sustentada por la paradoja existente entre la altísima variación inter-individual en la composición microbiana y la homogeneidad en el potencial funcional de tal variedad de especies (Tap et al., 2009; The Human Microbiome Project Consortium, 2012). Esto conduce a explorar si, a pesar de dicha diversidad filogenética, existe un conjunto de especies comunes a todos los individuos de una

¹ Los términos que aparecen en el glosario se señalan por primera vez con el formato negrita.

población, lo que demostraría que, aunque exista una redundancia a nivel funcional, habría un grupo de especies esenciales para el funcionamiento adecuado del organismo y la mantención de la vida misma del hospedero. Este trabajo tiene como objetivo hacer una búsqueda del conjunto de grupos bacterianos comunes a la población colombiana, empleando conceptos novedosos del núcleo microbiano.

JUSTIFICACIÓN

El tracto gastrointestinal de los seres humanos es un ambiente rico y habitado hasta por 100 billones (10^{14}) de microorganismos que, en conjunto, conforman la microbiota intestinal (Whitman, Coleman, & Wiebe, 1998). En consecuencia, actualmente se reconoce el microbioma intestinal (es decir, el conjunto de microorganismos y sus genes) como el segundo genoma de la especie humana, el cual es determinante para moldear algunos rasgos fenotípicos y aspectos de salud y enfermedad que, incluso, se creían que estaban directamente asociados con la carga genética humana (Grice et al., 2012; The Human Microbiome Project Consortium, 2012; Turnbaugh & Gordon, 2009).

La microbiota se puede entender, entonces, como un biomarcador de salud y enfermedad; así, podrían plantearse estrategias de modulación para contrarrestar una afección a través de la intervención sobre esta comunidad microbiana. Sin embargo, tales estrategias no son evidentes puesto que existe una paradoja entre composición (qué especies hay) y función (qué hacen esas especies): la composición microbiana entre personas saludables es altamente variable mientras que existe una redundancia en los rasgos funcionales de estos microorganismos (The Human Microbiome Project Consortium, 2012). Por tanto, la exploración de un núcleo de especies común a un grupo de individuos sanos es de interés para entender la fracción de especies que tienen que estar presentes en la microbiota intestinal para que el organismo funcione adecuadamente (Tap et al., 2009; The Human Microbiome Project Consortium, 2012).

Otra razón que impulsa esta investigación es que la mayor parte de estudios sobre la microbiota humana se han centrado en poblaciones estadounidenses o europeas (Arumugam et al., 2011; Qin et al., 2010; The Human Microbiome Project Consortium, 2012; Turnbaugh & Gordon, 2009), dando así una visión limitada y sesgada de la microbiota intestinal humana; más aún, considerando que existen múltiples estudios que evidencian cómo el núcleo filogenético y la composición de la microbiota difiere entre individuos y poblaciones conforme a factores dietarios, el contenido genético del hospedero, la edad, el origen étnico y la geografía (Falony et al., 2016; Qin et al., 2010; Tap et al., 2009; Yatsunenکو et al., 2012). En particular, la investigación llevada a cabo por Escobar et al., (2014) sobre la microbiota intestinal de los colombianos permite evidenciar cómo la composición de las comunidades microbianas de esta población difiere significativamente de los americanos, europeos y asiáticos. En consecuencia, resta por conocer qué comparten estas comunidades con otras poblaciones de hospederos al incluir esta nueva población de estudio. Todo ello, con el fin de contribuir al conocimiento de base sobre el papel que cumplen estos microorganismos en la salud de los colombianos para futuras intervenciones clínicas con moduladores de la microbiota.

OBJETIVOS

Objetivo general

Determinar la microbiota intestinal común a la población colombiana a partir de datos de secuenciamiento de una unidad variable del gen 16S ADNr.

Objetivos específicos

Evaluar los criterios de inclusión para la determinación de la microbiota núcleo a través de una búsqueda exhaustiva de información bibliográfica.

Implementar el método de inclusión para la determinación de la microbiota intestinal núcleo de la población colombiana a partir de una rutina propia creada en la herramienta computacional R.

Determinar el núcleo microbiano considerando distintos subconjuntos de la población de estudio (de acuerdo al índice de masa corporal, al origen geográfico y al género).

MARCO TEÓRICO

Nuestro cuerpo alberga una amplia diversidad microbiana, conocida comúnmente como la "microbiota humana". Estos microorganismos superan en número nuestras propias células y están íntimamente asociados con nuestros tejidos y órganos (Robinson, Bohannon, & Young, 2010). Nuestra relación con los componentes de esta microbiota es generalmente reconocida como **comensal** en lugar de mutualista. Sin embargo, el uso frecuente del término comensal refleja la falta de conocimiento o, al menos, una actitud agnóstica sobre la contribución de la mayoría de los habitantes de este complejo a nuestro propio **valor selectivo** o al de los miembros de esta comunidad microbiana (Bäckhed et al., 2005). A pesar de esto, varios estudios han puesto de manifiesto cómo la microbiota intestinal ha **coevolucionado** con sus hospederos y la manera en la que estos microorganismos modulan y complementan nuestra biología de forma mutuamente beneficiosa (Bäckhed et al., 2005; Moeller et al., 2016)

La vasta mayoría de microorganismos que habitan en el intestino codifican funciones metabólicas que la especie humana no ha tenido que evolucionar por sí misma; éstas incluyen, por ejemplo, la capacidad de hidrolizar fibra dietaria y carbohidratos complejos, la síntesis *de novo* de vitamina K y vitaminas del complejo B, como biotina, cobalamina, folato, ácido nicotínico, ácido pantoténico, riboflavina y tiamina (Bäckhed et al., 2005; Ley et al., 2006; Simpson & Campbell, 2015). De hecho, algunos estudios han logrado evidenciar cómo el cambio en la composición y abundancia de microorganismos en el tracto gastrointestinal se encuentra asociado a enfermedades crónicas como obesidad, alergias, diabetes tipo 1, asma, autismo, aterosclerosis y enfermedades inflamatorias del intestino (Bäckhed et al., 2012; Honda & Littman, 2012; Le Chatelier et al., 2013; Sartor, 2008, 2010). Por tanto, la estabilidad o resiliencia de las comunidades microbianas que habitan en el intestino puede ser pensada como una propiedad descriptiva de la salud humana. Dicha estabilidad ecológica puede estar dada por la composición o el perfil funcional de estas comunidades, las cuales van a favorecer al hospedero (Frederik Bäckhed et al., 2012; Salonen et al., 2012). Consecuentemente, es preciso sugerir posibles biomarcadores de la salud humana, los cuales, en esencia, pueden ser representados por un grupo de especies moleculares comunes en un grupo de individuos aparentemente saludables.

De hecho, la búsqueda de un núcleo microbiano se ve sustentada por la paradoja existente de la altísima variación en la composición microbiana entre individuos acompañada de una homogeneidad en el potencial funcional de tal variedad de especies, lo cual conduce a explorar si, a pesar de dicha diversidad filogenética, existe un conjunto de especies comunes a todos los individuos de una población (Tap et al., 2009; The

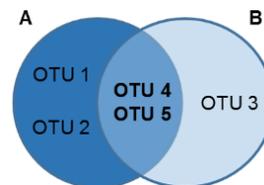
Human Microbiome Project Consortium, 2012); esto demostraría que, aunque exista una redundancia a nivel funcional, habría un grupo de especies esenciales para el funcionamiento adecuado del organismo y la manutención de la vida misma del hospedero.

Aunque el concepto de núcleo microbiano es aparentemente intuitivo (las especies moleculares compartidas por un grupo de individuos), no existe una definición formal y universal para éste (Salonen et al., 2012). En vista de esto, Shade & Handelsman (2012) presentan un marco conceptual novedoso para la identificación de especies comunes a un hábitat, el cual se encuentra enmarcado en los conceptos teóricos de la ecología tradicional. En esencia, estos autores proponen cinco conceptos distintos para explorar el núcleo microbiano, los cuales permitirían tener una visión más compleja de la microbiota-núcleo y avanzar en la comprensión del papel de estos microorganismos y sus funciones clave dentro del intestino (Shade & Handelsman, 2012).

Entre los conceptos sugeridos por Shade & Handelsman (2012) se encuentra el *núcleo de membresía* (Fig. 1A), el cual se define de la siguiente forma: *una OTU pertenece al núcleo de membresía de prevalencia p (%) si se encuentra en al menos p (%) de la población de estudio*. Aunque este análisis pondera todas la OTUs de manera similar, ignora la representatividad de cada OTU dentro de la comunidad. Sin embargo, la abundancia relativa de una especie dada se ha asociado con la predisposición a padecer ciertas patologías. Adicionalmente, existe una ardua discusión sobre el papel que tienen las distintas especies dentro del ecosistema. Esto, dado que algunos estudios ponen de manifiesto que la conservación de un número relativamente pequeño de especies, generalmente dominantes, es suficiente para mantener la mayoría de los procesos en un ecosistema (Grime, 1998), mientras que otros estudios han logrado demostrar que las especies menos comunes y raras contribuyen significativamente al funcionamiento del ecosistema (Lyons et al., 2005). En vista de esto, se propone el concepto de *núcleo de composición* (Fig. 1B), el cual se define así: *una OTU pertenece al núcleo de composición con parámetro de abundancia a si tiene al menos una abundancia relativa mediana de a* . En esencia, este análisis permite resaltar el rol que juegan las especies raras, comunes y dominantes en el ecosistema de interés. Adicionalmente, se entenderá como *núcleo de membresía-composición al conjunto de OTUs que cumplen tanto la restricción de prevalencia (p) como de abundancia (a)*.

A. Núcleo de membresía

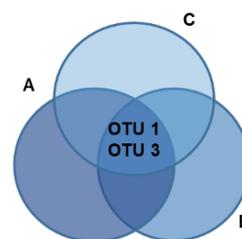
	Muestra A	Muestra B	Compartida por A & B
OTU 1	1	0	
OTU 2	1	0	
OTU 3	0	1	
OTU 4	1	1	X
OTU 5	1	1	X



Las OTUs que pertenecen al núcleo de membresía con prevalencia del 100% son la OTU 4 y la OTU 5.

B. Núcleo de composición

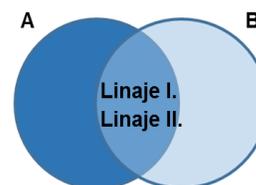
	Muestra A	Muestra B	Muestra C	Mediana
OTU 1	0,01	0,036	0,002	0,01
OTU 2	0,004	0,002	0,05	0,004
OTU 3	0,005	0,02	0,01	0,01
OTU 4	0	0,004	0,02	0,004
OTU 5	0,003	0	0,02	0,003



Las OTUs que pertenecen al núcleo de composición con una abundancia mediana de 0,01 son: la OTU 1 y la OTU 3. Nótese en este ejemplo que la OTU 2 hace parte del núcleo de membresía con prevalencia del 100%, pero no del núcleo de composición con abundancia de 0,01.

C. Núcleo filogenético

Linaje		Muestra A	Muestra B	Linaje compartido por A & B
I.	OTU 1	0,01	0	Linaje I
	OTU 2	0	0,002	
II.	OTU 3	0,005	0,02	Linaje II
	OTU 4	0	0,004	
III.	OTU 5	0,003	0	



Los linajes que pertenecen al núcleo filogenético son el linaje I y el linaje II.

D. Núcleo taxonómico

	Muestra A	Muestra B	Taxa compartida por A & B
<i>Prevotella copri</i>	0,01	0	Prevotella
<i>Prevotella</i> sp.	0	0,002	
<i>Streptococcus</i> sp.	0	0,001	
<i>Ruminococcus bromii</i>	0,005	0	Clostridiales
<i>Ruminococcus gnavus</i>	0	0,004	



Las taxa que pertenecen al núcleo taxonómico son el género *Prevotella* y el orden Clostridiales. Cabe señalar que el género *Ruminococcus* es parafilético y polifilético.

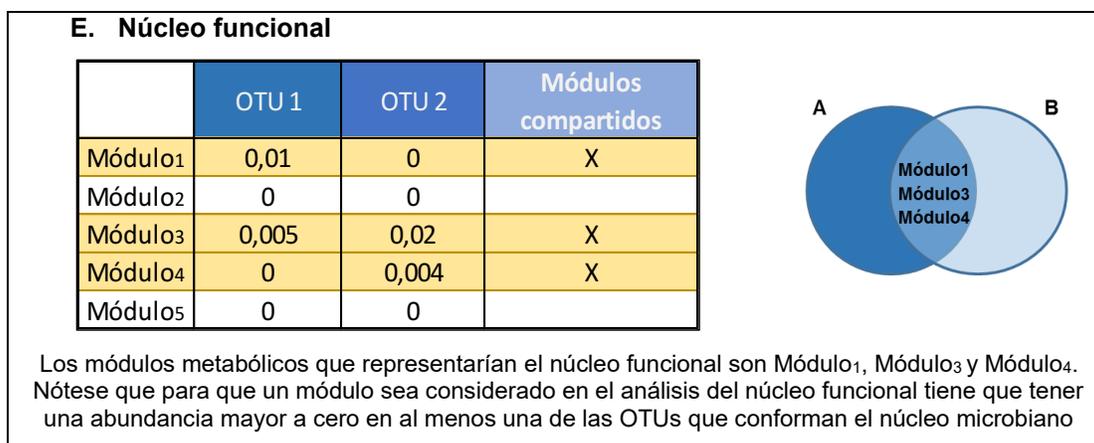


Figura 1. Definición de los distintos núcleos microbianos. **(A)** Representación del núcleo de membresía. **(B)** Representación de núcleo de composición. **(C)** Representación del núcleo filogenético. **(D)** Representación del núcleo taxonómico. **(D)** Representación del núcleo funcional. Tomado y adaptado de Shade & Handelsman (2012).

Por otra parte, Shade & Handelsman (2012) proponen el concepto de núcleo filogenético (núcleo taxonómico), el cual está basado en la existencia de una coherencia ecológica a niveles evolutivos superiores. Esto significa que los miembros de un **clado** comparten estrategias de vida generales o rasgos que los distinguen de los miembros de otros clados (Lozupone & Knight, 2007; Ren & Wu, 2016). En este sentido, *los núcleos microbianos podrían existir a niveles evolutivos diferentes a las especies (OTUs)* (Fig. 1C-D). Sin embargo, los intentos de atribuir rasgos ecológicos a grupos bacterianos que se definen en rangos taxonómicos superiores a las especies son raros, y la idea de que las ramas profundas de los árboles filogenéticos bacterianos podrían ser ecológicamente coherentes rara vez se ha considerado (Doolittle, 2009). Esto, dado que se ha observado que el porcentaje de genes conservados entre las cepas de una misma especie bacteriana (por ejemplo, *Ralstonia solanacearum*) puede ser tan bajo como 68%; en comparación, el 75% de los genes humanos tienen homólogos en el genoma del pez *Takifugu rubripes* (Aparicio et al., 2002; Guidot et al., 2007). Asimismo, puede haber una diversidad fisiológica sustancial dentro de los clados bacterianos; por ejemplo, el phylum Proteobacteria incluye organismos heterótrofos, litótrofos y fotótrofos (Johnson et al., 2006). En las bacterias, rasgos ecológicos importantes pueden ser específicos de la cepa, y bacterias estrechamente relacionadas pueden ocupar nichos distintos. Además, la transferencia horizontal de genes puede ocurrir incluso entre organismos lejanamente relacionados (Hunt et al., 2008)

Finalmente, debido a que la estabilidad ecológica del hospedero puede estar dada no sólo por los microorganismos que componen la microbiota del intestino sino también por el perfil funcional de ésta, se entenderá cómo *núcleo funcional al conjunto de funciones metabólicas codificadas por los genomas de los microorganismos que*

conforman un determinado núcleo de membresía-composición (Fig. 1E). En esencia, este análisis permite identificar las posibles funciones que permanecen a medida que se asumen parámetros más restrictivos de prevalencia (**p**) y abundancia (**a**).

Aunque todos estos conceptos empleados para identificar el núcleo microbiano son de gran interés para resaltar el rol ecológico de estos organismos dentro de la comunidad, la mayor parte de los estudios han centrado sus análisis en considerar como núcleo microbiano únicamente al núcleo de membresía. Aun restringiéndonos a este último, es importante señalar que los criterios de inclusión sobre la prevalencia requerida para incluir especies a este núcleo han variado considerablemente entre los distintos estudios (Salonen et al., 2012; Shade & Handelsman, 2012). En donde sí concuerdan las investigaciones es en señalar que no existe un núcleo microbiano con especies moleculares que prevalezcan en el 100% de la población de estudio (Falony et al., 2016; Salonen et al., 2012; Tap et al., 2009; Turnbaugh et al., 2009), promoviendo así una búsqueda del núcleo microbiano con umbrales de prevalencia que van desde 50% al 100% de los individuos (Tabla 1).

Tabla 1. Criterios para definir el núcleo microbiano implementado por distintos estudios.

Concepto de núcleo empleado	Parámetros	Referencia
Núcleo de membresía	Prevalencia del 95%	Falony et al., 2016
	Prevalencia del 95%	Huse et al., 2012
	Prevalencia del 50%	Tap et al., 2009
	Prevalencia del 70%	Cheng et al., 2016
Núcleo de composición	Prevalencia: 1-100% Abundancia/Intensidad: 1.93–4.98	Jalanka-Tuovinen et al., 2011
	Prevalencia: 1-100% Abundancia/Intensidad: 1.93–4.98	Salonen et al., 2012
Núcleo filogenético	Prevalencia 97%	Zhang et al., 2015

METODOLOGÍA

Procedimiento experimental

El desarrollo del presente estudio se encuentra basado en un conjunto de datos obtenido previamente por el grupo de investigación Vidarium (de la Cuesta-Zuluaga et al., 2017). A continuación, se detallará sucintamente el proceso experimental.

Población de estudio

En la población de estudio se incluyeron 441 individuos, tanto hombres como mujeres que se encontraban viviendo en las ciudades de Barranquilla, Bogotá, Medellín, Bucaramanga o Cali; estos individuos tenían un rango de edad de 18 a 62 años, y un **índice de masa corporal (IMC)** $\geq 18.5 \text{ kg/m}^2$. Para el estudio se excluyeron personas que habían sido diagnosticadas con cualquiera de las siguientes enfermedades: Alzheimer, Parkinson o cualquier otra enfermedad neurodegenerativa; cáncer, y enfermedades gastrointestinales (enfermedad de Crohn, colitis ulcerativa, síndrome de intestino corto, diverticulosis o enfermedad celíaca). Tampoco se consideraron mujeres en embarazo ni personas que consumieron antibióticos o antiparasitarios en los 3 meses anteriores al muestreo.

Extracción de ADN

Cada individuo muestreado era responsable de coleccionar una muestra de material fecal en un recipiente hermético y estéril provisto por el grupo de investigación. Dicho material fue inmediatamente almacenado en el refrigerador y llevado dentro de las siguientes 12 horas a una de las instalaciones de la EPS SURA; allí las muestras eran almacenadas en hielo seco para luego ser enviadas a un laboratorio central.

El ADN total fue extraído usando el kit comercial QIAamp DNA Stool Mini Kit (Qiagen, Hilden, Alemania) siguiendo las instrucciones del fabricante, con una ligera modificación en la etapa de homogenización con perlas con el tampón de lisis (20 segundos a 15 Hz usando perlas de acero inoxidable de 5 mm de diámetro). Posterior al proceso de extracción, se cuantificó la concentración de ADN utilizando un Nanodrop (Nyxor Biotech, París, Francia).

Amplificación del gen 16S ADNr

Las muestras de ADN total fueron enviadas al laboratorio de Biología Molecular de Sistemas Microbianos de la Universidad de Michigan, Escuela de Medicina (Ann Arbor, MI, USA). Allí se amplificó la región hipervariable V4 del gen 16S ADNr para cada una de las muestras, haciendo uso de los cebadores F515 (5'-CACGGTTCGKCGGCCATT-3') y R806 (5'-GGACTACHVGGTWTCTAAT-3') (Caporaso et al., 2011; Kozich, Westcott, Baxter, Highlander, & Schloss, 2013) y, secuenciado con la plataforma Illumina MiSeq con el kit de secuenciación MiSeq V2 (Kozich et al., 2013).

Análisis de secuencias y detección de secuencias representativas de OTUs

Los amplicones de las secuencias del gen 16S ADNr fueron procesados con el programa Mothur v.1.36 (Schloss, 2009) siguiendo el procedimiento de operación estándar para secuencias obtenidas con Illumina MiSeq. En primer lugar, se extrajeron las secuencias y los datos de puntuación de calidad de los archivos *fastq* emparejados para ensamblar las lecturas y formar **contigs**. Posteriormente, se eliminaron las secuencias que contenían bases con una puntuación de calidad (phred) inferior a 20. Secuencias con bases ambiguas o de longitud menor a 275 pb fueron eliminadas. Las secuencias fueron alineadas utilizando la base de datos de referencia Silva v.1.2.3 (Quast et al., 2013) y se eliminaron los homopolímeros de 8 o más nucleótidos, así como las secuencias que no se superponían a la región de alineación V4. Subsiguientemente, se llevó a cabo una etapa de pre-agrupamiento, en la cual se combinaron secuencias con una identidad $\geq 99\%$ (es decir, secuencias que difirieron en 2 nucleótidos o menos). Adicionalmente, las **secuencias quiméricas** fueron detectadas y descartadas con el programa UCHIME (Edgar, Haas, Clemente, Quince, & Knight, 2011). Posterior a ello, se asignó la clasificación taxonómica a las distintas secuencias usando como base de referencia Greengenes 13_8_99 (DeSantis et al., 2006) y se eliminaron las secuencias clasificadas como mitocondrias, eucariotas o desconocidas. Se implementó el algoritmo de vecino más cercano promedio para generar las distintas unidades taxonómicas operacionales (OTUs) delimitadas con un 97% de identidad y clasificadas taxonómicamente por consenso utilizando la base de referencia Greengenes 13_8_99. Finalmente, se calcularon distancias pareadas no corregidas entre secuencias alineadas (una secuencia representativa por OTU) para obtener un árbol de distancias utilizando el algoritmo “relaxed neighbor joining” implementado en el programa Clearcut (Evans et al., 2006).

Determinación del núcleo

A continuación, se detallará la metodología implementada para la determinación de los núcleos microbianos.

Análisis de datos

Tanto para el concepto de núcleo de membresía como para el de composición se crearon rutinas propias en la herramienta computacional R x64 3.1.1 (R Development Core Team, 2011) que permiten determinar los taxones compartidos en la población de estudio con una variación en los parámetros de inclusión (p y a). En primer lugar, para el núcleo de membresía se hizo una búsqueda exploratoria bajo diferentes umbrales de prevalencia, los cuales variaron entre 0% y 100% de los individuos de estudio. En análisis posteriores, únicamente se consideraron los valores de prevalencia del 50, 75, 90, 95 y 99%. Para el núcleo de composición, se consideraron las siguientes abundancias relativas medianas: 10^{-6} , 10^{-5} , 10^{-4} , 10^{-3} , 10^{-2} , y 2×10^{-2} .

Para identificar el núcleo filogenético, se implementó la herramienta PhyloCore (Ren & Wu, 2016) considerando prevalencias del 50, 75, 90, 95, 99 y 100%. Para ello, en primer lugar, se hizo una exploración de este concepto empleando el árbol de distancias antes mencionado. Cabe señalar que dicho árbol carece de valores de soporte de ramas y no es filogenético, en sentido estricto. Adicionalmente, se determinó el núcleo filogenético a partir de un análisis por filotipos², es decir, considerando únicamente la clasificación taxonómica como fuente de información sobre las relaciones evolutivas entre las especies moleculares.

El potencial funcional, es decir, la aproximación del perfil metagenómico de la microbiota intestinal, fue inferido con el paquete de R Tax4Fun (Aßhauer, Wemheuer, Daniel, & Meinicke, 2015). Esta herramienta, a partir de datos del gen 16S ADNr y basado en una clasificación taxonómica usando como base de datos de referencia Silva v.123 (Quast et al., 2013), permite hacer una transformación lineal del perfil taxonómico de los organismos procariotas empleando la base de datos KEGG (*Kyoto Encyclopedia of Genes and Genomes*). Posterior a ello, las abundancias estimadas de los genes ortólogos reportados en KEGG (KOs) se normalizaron de acuerdo al número de copias del gen 16S ADNr presentes en una especie dada, número obtenido a partir de las anotaciones del genoma de dicha especie en el repositorio del NCBI (Aßhauer et al., 2015). Finalmente, las abundancias taxonómicas normalizadas se utilizaron para

² Estrictamente hablando este no es un núcleo filogenético sino un núcleo taxonómico.

combinar linealmente los KOs y realizar la predicción del perfil funcional de la comunidad microbiana (Aßhauer et al., 2015).

Una vez obtenida la información de salida del paquete Tax4Fun, los KOs se condensaron en módulos metabólicos usando la plataforma GOMixer v.1.7.3 (Darzi, Falony, Vieira-silva, & Raes, 2015); estos módulos se pueden entender como conjuntos de funciones enzimáticas estrechamente relacionadas que representan un proceso celular con metabolitos de entrada y salida definidos (Darzi et al., 2015). En general, el flujo de trabajo de GOMixer comienza por cuantificar los módulos de la vía metabólica del intestino humano para cada muestra mediante el mapeo de la abundancia de genes en una base de datos de módulos predefinidos específicos del intestino (Darzi et al., 2015). Para la inferencia de los módulos se tendrá en cuenta la cobertura mínima de los KOs para un módulo (porcentaje de pasos metabólicos presentes). Por defecto, este valor se infiere de la predicción del módulo. La abundancia del módulo se definirá como la abundancia media de los pasos metabólicos cubiertos para esta vía (Darzi et al., 2015).

El potencial funcional de la microbiota fue determinado para todo el juego de datos (n=441 individuos). Adicionalmente, dicha predicción fue hallada para casos particulares del núcleo de membresía-composición, entre los cuales se incluyen las siguientes configuraciones de parámetros (prevalencia-abundancia): 99%-10⁻², 99%-10⁻⁴, 95%-10⁻⁴, 90%-10⁻⁴, 75%-10⁻⁴, 50%-10⁻⁴ y 50%-10⁻⁵. Para visualizar los módulos por núcleo de membresía-composición se hizo un mapa de calor usando R x64 3.1.1 (R Development Core Team, 2011); para ello se obtuvo el valor medio de cada módulo considerando el valor de todas las muestras de cada núcleo de membresía-composición.

RESULTADOS

Características generales de la microbiota

Del análisis bioinformático global, 14.750.448 secuencias (**reads**) pasaron los filtros de calidad y fueron usadas para tratamientos subsiguientes. A partir del análisis de agrupamiento jerárquico de vecino cercano promedio se obtuvieron 4720 OTUs con una identidad del 97% y una correlación de Matthew³ de 0,79, lo que indica un agrupamiento de alta calidad. En esencia, la composición de la microbiota intestinal de la cohorte muestreada (441 colombianos) comprende una gran variación taxonómica a nivel de phylum (Fig. 2A). Los phyla Firmicutes y Bacteroidetes fueron los más abundantes (con una abundancia media \pm DS = $48,9 \pm 19,6\%$ y $16,1 \pm 17,6\%$, respectivamente), seguido de Proteobacteria ($12,1 \pm 17,2\%$), Actinobacteria ($9,4 \pm 9,2\%$), Verrucomicrobia ($7,8 \pm 14,8\%$), Euryarchaeota ($2,6 \pm 5,0\%$) y otros phyla presentes en frecuencias menores (Fig. 3). No obstante, a partir de la predicción del potencial funcional de la microbiota intestinal de esta cohorte, se obtuvo una gran homogeneidad en la abundancia relativa de los módulos metabólicos predichos (Fig. 2B). Esto pone de manifiesto la paradoja antes mencionada acerca de la altísima variación en la composición microbiana entre individuos acompañada de una homogeneidad en el potencial funcional de tal variedad de especies y conduce a explorar los microorganismos que se han conservado durante la coevolución mutua del ser humano y sus microbios intestinales.

Determinación de la microbiota núcleo

Con la obtención del núcleo de membresía se encontró que de las 4720 OTUs asociadas con la microbiota intestinal de la población colombiana, ninguna fue detectada en el 100% de los individuos muestreados. Sin embargo, con una prevalencia del 99% se detectaron las siguientes cuatro especies bacterianas: *Escherichia coli*, *Faecalibacterium prausnitzii*, *Roseburia faecis*, y *Blautia* sp. Bajo un criterio de inclusión más laxo (las OTUs compartidas por el 95% de la población de estudio) se observó un aumento en la cantidad de OTUs que se comparten entre los individuos estudiados (Fig. 4), incrementando el número de linajes bacterianos a 19, los cuales son: *Akkermansia muciniphila*, *Bacteroides* sp., *Bifidobacterium adolescentis*, *Blautia obeum*, *Blautia* spp. (3), *Collinsella aerofaciens*, *Escherichia coli*, *Faecalibacterium prausnitzii* (2), *Gemmiger formicilis*, *Oscillospira* sp., *Roseburia faecis*, *Ruminococcus gnavus*, *Ruminococcus lactaris*, *SMB53* sp. (familia Clostridiaceae), *Streptococcus* spp. (2) (Anexo 1).

³ El coeficiente de correlación de Matthew es una medida de la asociación entre dos variables binarias. Esta medida es similar al coeficiente de correlación de Pearson en su interpretación.

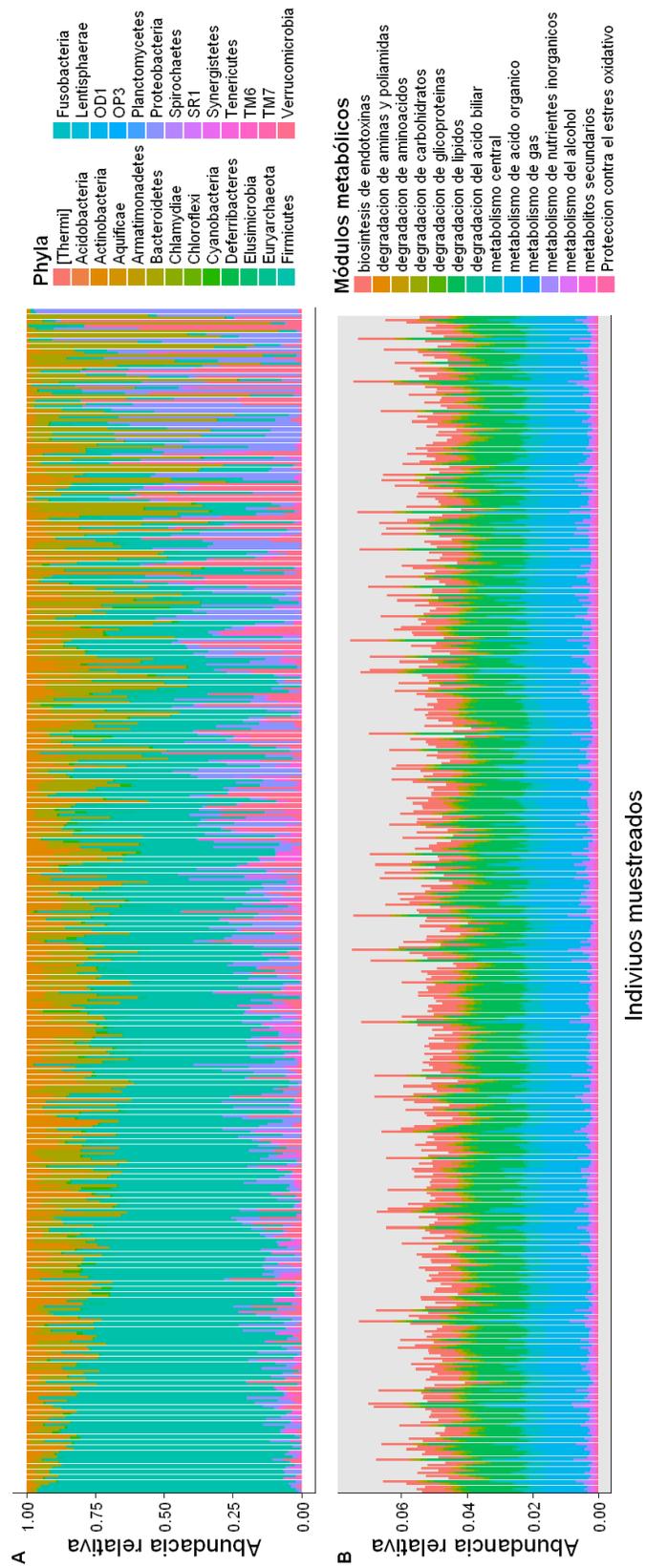


Figura 2. Paradoja actual sobre la microbiota intestinal. **(A)** Representación de la composición taxonómica (phyla) por individuos muestreados. **(B)** Representación de la abundancia de los módulos metabólicos predichos. Nótese que la escala de abundancia en el panel B no suma 1 porque se relativiza por individuo y no por módulo (la suma de un solo módulo en todos los individuos es igual a 1).

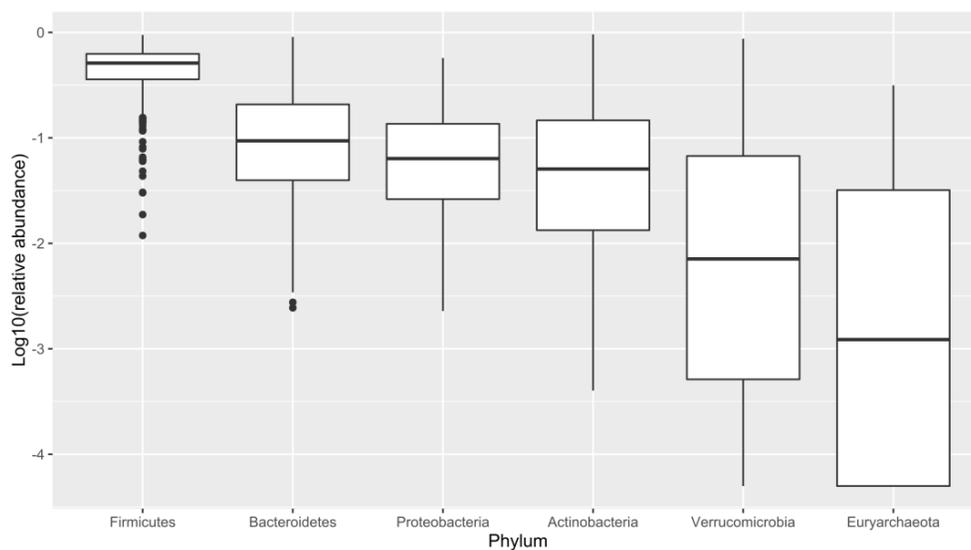


Figura 3. Diagrama de caja y bigotes con la abundancia relativa (en escala logarítmica) de los phyla principales detectados en la población estudiada.

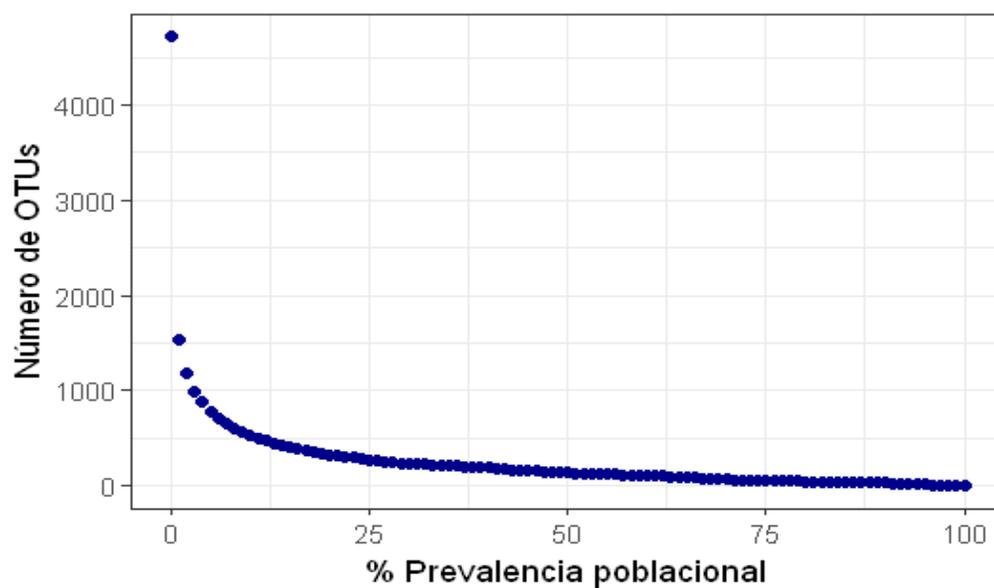


Figura 4. Número de OTUs que conforman el núcleo de membresía para diferentes valores de prevalencia poblacional.

Como se explicó en apartados anteriores, el núcleo de membresía pondera todas la OTUs de manera similar, ignorando la representatividad de cada OTU dentro de la comunidad. Por tanto, al determinar el *núcleo de composición* se resalta el rol que tiene la abundancia de las distintas especies dentro de un ecosistema de interés. En esencia, respecto al núcleo de composición, se puede observar que las OTUs que más prevalecen

en la microbiota intestinal de los colombianos son especies dominantes, es decir, especies con alta abundancia respecto a las demás. Esto se ejemplifica claramente con el núcleo de membresía-composición (Tabla 2). Considérese la columna de prevalencia del 95% y hágase variar la abundancia (de mayor a menor): se observa que por debajo de una abundancia de 0,001 no se incorporan nuevas OTUs al núcleo, demostrando así que las 19 OTUs que conforman el núcleo membresía-composición 95%-0,001 tienen una abundancia mediana relativamente alta ($0,001 \leq x \leq 0,02$). Adicionalmente, se puede observar que las restricciones de abundancia no afectan en gran medida el número de especies en el núcleo microbiano de membresía-composición, como se puede observar para valores razonables de abundancia (10^{-3}) y valores altos de prevalencia.

Tabla 2. Variación del núcleo de membresía-composición. En la primera columna se encuentran los parámetros de abundancia, en la primera fila los valores de prevalencia considerados. En cada casilla se encuentra el número de OTUs que cumplen con los criterios tanto de abundancia como de prevalencia.

		Prevalencia					
		0%	50%	75%	90%	95%	99%
Abundancia	0	4720					
	$1 \cdot 10^{-6}$	137		60	32	19	4
	$1 \cdot 10^{-5}$						
	$1 \cdot 10^{-4}$	100					
	$1 \cdot 10^{-3}$	38		34	28		
	$1 \cdot 10^{-2}$	6				5	2
	$2 \cdot 10^{-2}$	2				1	

El núcleo filogenético obtenido a partir del árbol de distancias de las distintas OTUs incorporó los mismos grupos microbianos en este núcleo que los que se incluyeron en el núcleo de membresía, considerando distintos valores de prevalencia (Tabla 3). No obstante, al ignorar las relaciones filogenéticas entre las OTUs y considerar únicamente su clasificación taxonómica como fuente de información sobre las relaciones evolutivas entre ellas (núcleo taxonómico), se detectaron varios taxa en el 100% de los individuos (Tabla 3).

Tabla 3. Núcleo filogenético (segunda columna) y núcleo taxonómico (tercera columna). En la primera columna se especifican los valores de prevalencia utilizados.

	Núcleo filogenético	Núcleo taxonómico
Prevalencia del 100%	Ningún taxón.	Una clase (Actinobacteria), un orden (Bacteroidales), cuatro familias (Clostridiaceae, Coriobacteriaceae, Enterobacteriaceae, Veillonellaceae), dos géneros (<i>Clostridium</i> , <i>Streptococcus</i>) y, dos especies (<i>Faecalibacterium prausnitzii</i> , <i>Blautia</i> sp.)
Prevalencia del 99%	<i>Escherichia coli</i> (Otu00004); <i>Faecalibacterium prausnitzii</i> (Otu00005); <i>Roseburia faecis</i> (Otu00014); <i>Blautia</i> sp. (Otu00037).	Tres familias (Coriobacteriaceae, Erysipelotrichaceae, Veillonellaceae), cinco géneros (<i>Bacteroides</i> , <i>Clostridium</i> , <i>Dorea</i> , <i>Prevotella</i> , <i>Actinomycetales</i>), dos especies (<i>Ruminococcus gnavus</i> , <i>Streptococcus</i> sp.) y, cuatro OTUs (Otu00004, Otu00005, Otu00014, Otu00037)

Potencial funcional del núcleo de membresía-composición

Con el fin de predecir el potencial funcional de las OTUs representantes de los núcleos de membresía-composición y filogenético, se recurre a realizar una predicción de metagenomas a partir de las secuencias del gen 16S ADNr. A diferencia de los núcleos hasta ahora presentados, el núcleo funcional no busca hallar el conjunto de OTUs que lo conforman sino identificar las posibles funciones metabólicas que permanecen o desaparecen a medida que se excluyen microorganismos dados ciertos parámetros de abundancia y prevalencia. Las configuraciones de parámetros que se probaron fueron las siguientes (% prevalencia poblacional-abundancia relativa mediana): 99%-10⁻², 99%-10⁻⁴, 95%-10⁻⁴, 90%-10⁻⁴, 75%-10⁻⁴, 50%-10⁻⁴ y 50%-10⁻⁵, 0%-0.

En el mapa de calor presentado en la Figura 5 se puede observar cómo el potencial funcional de las 4720 OTUs que conforman la microbiota intestinal de los colombianos es semejante al representado por los núcleos de membresía-composición 50%-10⁻⁴ (100 OTUs) y 50%-10⁻⁵ (137 OTUs). Esto de por sí es un hecho notable, ya que un porcentaje muy bajo de OTUs (entre 2,1% y 2,9%) da cuenta de todo el perfil funcional de la microbiota intestinal de esta población. Sin embargo, cuando disminuimos el número de OTUs haciendo más restrictivos los parámetros de prevalencia y abundancia, es decir, aproximándonos a lo que en esencia sería un verdadero núcleo microbiano, observamos que el potencial funcional varía, mostrando sobre-representación de ciertos módulos metabólicos. Claramente se observa que el metagenoma de 60 (75%-10⁻⁴), 32 (90%-10⁻⁴) y 19 (95%-10⁻⁴) OTUs es muy similar entre ellos, y sólo vuelve a variar cuando se consideran 4 (99%-10⁻⁴) y 2 (99%-10⁻²) OTUs.

Todo indica que las funciones metabólicas esenciales para el buen funcionamiento del hospedero estarían aseguradas con el conjunto de 19 OTUs que conforman el núcleo 95%-10⁻⁴; si bien su abundancia relativa no es exactamente igual a la del conjunto total de OTUs, la diversidad de funciones metabólicas que abarca sugiere que éste podría ser el núcleo funcional de la comunidad bajo escrutinio, como se detalla a continuación.

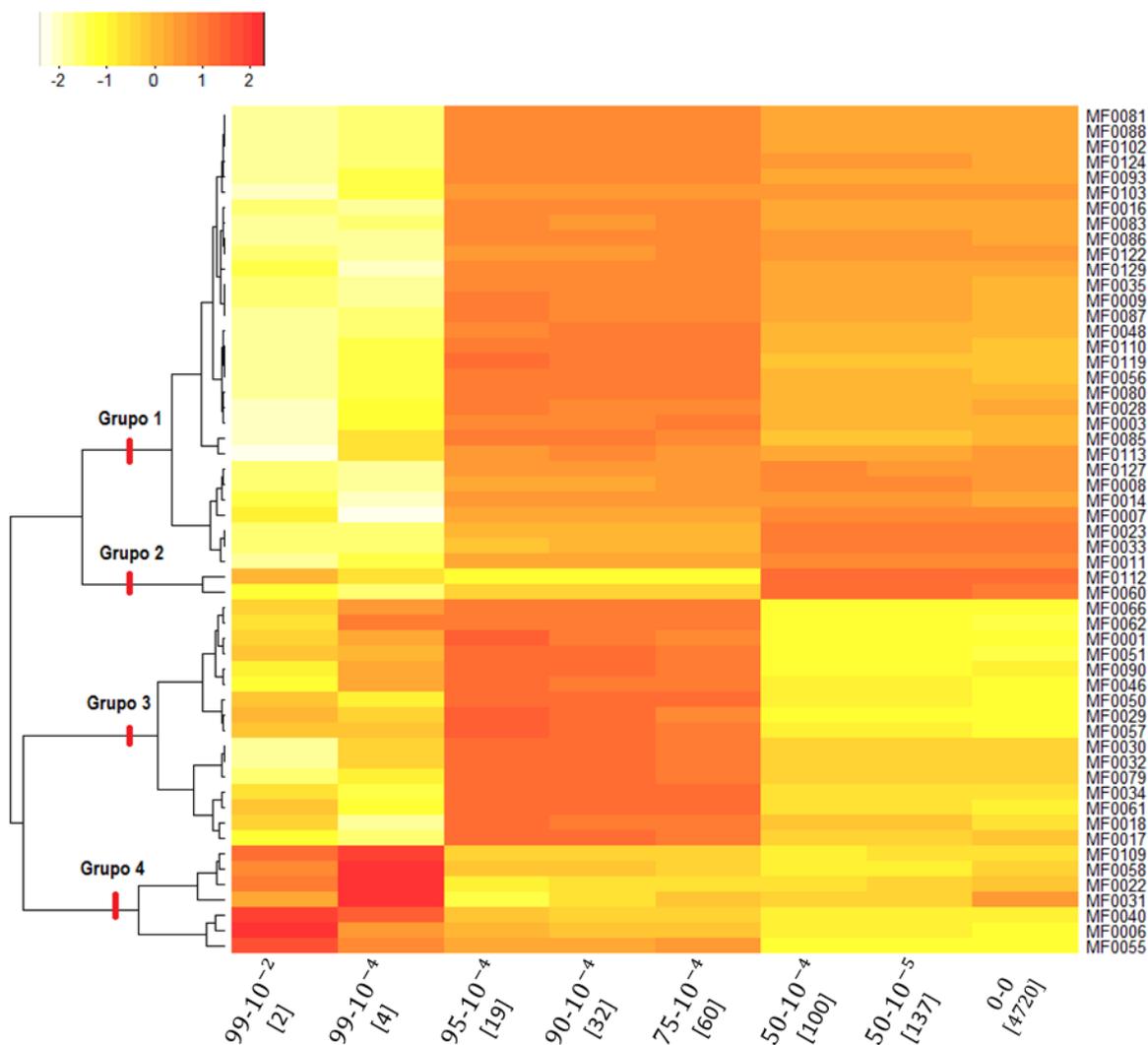


Figura 5. Predicción del potencial funcional de las OTUs representantes de distintos núcleos de membresía-composición. Las filas representan los módulos metabólicos predichos y sintetizados por GOMixer, y las columnas los distintos núcleos de membresía-composición. El dendrograma de la izquierda representa la agrupación de los módulos metabólicos dada su similitud en términos de abundancia, mas no por semejanzas funcionales o de ontología genética. La escala del mapa de calor corresponde a valores z (también llamados unidades tipificadas, estandarizadas o normalizadas). Los tonos rojos representan sobre-representación de los módulos metabólicos, mientras que los colores más claros indican sub-representación de éstos.

A partir del mapa de calor obtenido y la construcción del dendrograma basado en la similitud de las abundancias de los distintos módulos metabólicos (Fig. 5), se logran evidenciar cuatro grandes agrupaciones de funciones metabólicas. La primera de estas

agrupaciones está compuesta por un conjunto de funciones sobre-representadas por el conjunto de 4720, 137, 100, 60, 32 y 19 OTUs. Entre éstas se encuentran funciones involucradas con el metabolismo central (ej. ciclo del ácido cítrico, glicólisis), la degradación de aminoácidos aromáticos cargados de forma negativa, la degradación de carbohidratos complejos y simples (ej. monosacáridos y disacáridos), la degradación de mucinas, la producción de ácidos grasos de cadena corta (ej. acetogénesis) y de catalasas. En esencia, este conjunto de módulos metabólicos está compuesto por una alta diversidad de funciones involucradas en la extracción de energía de los alimentos de la dieta del hospedero (Anexo 2). Por otro lado, el segundo grupo de funciones metabólicas se encuentra sobre-representado únicamente por el conjunto de 4720, 137 y 100 OTUs, y están asociadas con la degradación de la ribosa y el metabolismo del acetato. Adicionalmente, el tercer grupo de módulos metabólicos está sobre-representado únicamente en los núcleos $75\%-10^{-4}$ (60 OTUs), $90\%-10^{-4}$ (32 OTUs) y $95\%-10^{-4}$ (19 OTUs); dichas funciones se encuentran relacionadas a la extracción de energía de los alimentos de la dieta, como lo son las funciones involucradas con el metabolismo central, el almacenamiento de energía y la degradación de polisacáridos y aminoácidos polares, alifáticos y no cargados. Finalmente, el último grupo de funciones que se encuentra sobre-representado por el conjunto de 2 y 4 OTUs está relacionado con la degradación de carbohidratos, aminoácidos y lípidos estructuralmente simples (Anexo 2).

Núcleo de membresía en subconjuntos poblacionales

Considerando que existen algunos estudios que evidencian cómo el núcleo microbiano y la composición de la microbiota difiere entre individuos y poblaciones conforme a factores dietarios, a la composición genética del hospedero, la edad, el origen étnico y la geografía (Falony et al., 2016; Qin et al., 2010; Tap et al., 2009), se procede a realizar un análisis del núcleo de membresía en tres subconjuntos poblacionales: de acuerdo con el sitio de origen de los individuos muestreados, al índice de masa corporal y al género del individuo muestreado (Fig. 6). Para estos análisis sólo se consideró el núcleo de membresía-composición $95\%-10^{-3}$ dado que, como se mostró en secciones anteriores, es el que considera que representa mejor el núcleo microbiano, ya que asegura una alta prevalencia poblacional y asegura alta diversidad funcional.

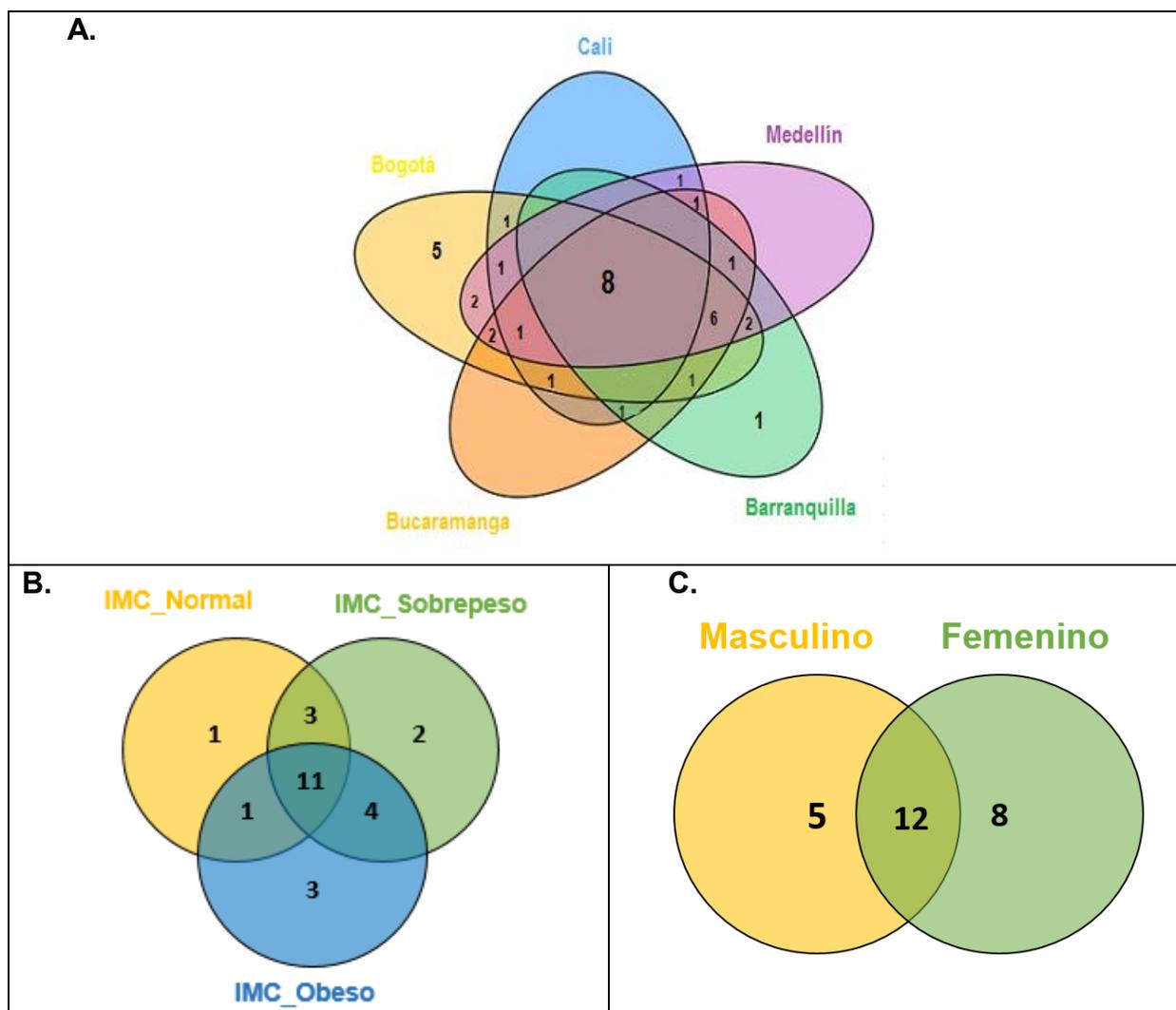


Figura 6. Núcleo de membresía con parámetro de prevalencia del 95% para cada subconjunto poblacional. **(A)** Núcleo de membresía por ciudad de origen de los participantes del estudio. **(B)** Núcleo de membresía por estado nutricional (IMC). **(C)** Núcleo de membresía por género.

La identificación de la microbiota núcleo considerando como variable categórica la ciudad de origen refleja parcialmente las diferencias en estilos de vida y dieta entre regiones geográficas (De Filippo et al., 2010; Yatsunenکو et al., 2012). En términos generales, con la obtención del núcleo considerando la ciudad (Anexo 3) se encontró que los individuos muestreados en Bogotá tienen en su núcleo cinco especies que no se encuentran en los núcleos de las demás ciudades (*Dorea* sp., *Mogibacteriaceae* no clasificada, *Roseburia inulinivorans*, *Subdoligranulum variabile*, y *Bacteroides fragilis*); lo mismo ocurre en Barranquilla, cuyo núcleo contiene una OTU que no se encuentra en ningún otro núcleo (*Veillonella dispar*). Es de anotar, sin embargo, que la mayoría de las OTUs de los núcleos por ciudad se comparten (Fig. 6A), demostrando, de otra manera, la existencia de una microbiota núcleo a una escala geográfica superior (Colombia).

La comparación del perfil microbiano entre sujetos con distintos estados nutricionales (evaluados a través del IMC) permite la identificación de una microbiota núcleo que se podría convertir en una herramienta sencilla para la evaluación de la predisposición a la obesidad (Turnbaugh et al., 2007). De hecho, en el núcleo por IMC (Fig. 6B) se detectan especies moleculares únicas a cada categoría, e incluso se logra apreciar una transición desde un estado normal a uno obeso. Esto, dado que las categorías sobrepeso y obeso comparten más OTUs entre ellas que con el estado normal; a su vez, el estado normal comparte más OTUs con el estado sobrepeso que con el estado obeso. No obstante, entre todas las categorías se comparte la gran mayoría de las OTUs identificadas (Anexo 4). Para el IMC normal se detectó una especie molecular única (*Bacteroides fragilis*), para sobrepeso dos (*Clostridium celatum* y *Ruminococcus bromii*) y para obesidad tres (*Blautia* sp., *Prevotella copri* y *Dorea formicigenerans*).

Estudios recientes han demostrado diferencias marcadas en la composición de la microbiota intestinal entre mellizos de sexo opuesto (Taneja, 2017) así como entre hombres y mujeres, de manera general (Haro et al., 2016; Wallis et al., 2016, 2017). Sin embargo, no se sabe si tales diferencias se traducen en núcleos microbianos diferentes. En este trabajo se encuentran taxones microbianos clave compartidos exclusivamente entre hombres o entre mujeres (Fig. 6C). En el núcleo de hombres se encontraron las siguientes cinco especies moleculares únicas: *Prevotella copri*, *Bifidobacterium adolescentis*, *Enterobacter hormaechei*, *Ruminococcus lactaris* y *Dorea formicigenerans*. En el de mujeres se encontraron ocho especies moleculares únicas: *Akkermansia muciniphila*, *Clostridium celatum*, *Ruminococcus bromii*, *Blautia* sp., *Bacteroides* sp., *Clostridium hathewayi*, *SMB53* sp. (familia Clostridiaceae) y *Blautia obeum*. No obstante, cabe resaltar que la mayor parte de las OTUs identificadas son compartidas por ambos géneros (Anexo 5).

DISCUSIÓN

Uno de los principales objetivos del *International Human Microbiome Project Consortium* es determinar si existe un "núcleo microbiano" identificable de organismos, genes o capacidades funcionales compartidos, que se encuentran en un hábitat corporal de todos o de la gran mayoría de los seres humanos (Turnbaugh et al., 2007). De hecho, distintos estudios han sugerido que, a pesar de los perfiles altamente específicos de la microbiota de cada individuo, las personas comparten un subconjunto común de grupos bacterianos relativamente abundantes, denominado núcleo microbiano (Huse et al., 2012; Salonen et al., 2012; Tap et al., 2009). Adicionalmente, otros estudios han puesto de manifiesto la existencia de un núcleo microbiano a nivel de genes (Turnbaugh & Gordon, 2009).

La mayor parte de los esfuerzos por determinar y definir el núcleo microbiano se han centrado en la presencia y ausencia de los grupos bacterianos en una proporción definida de los sujetos del estudio, es decir, en el núcleo de membresía (Salonen et al., 2012; Tap et al., 2009). En este estudio, se hace una exploración del núcleo microbiano aplicando distintos conceptos de la ecología tradicional, entre ellos la prevalencia, la abundancia, las relaciones filogenéticas de las OTUs y el potencial funcional de estos grupos microbianos. En esencia, estos conceptos permiten aprehender la complejidad de la microbiota-núcleo y avanzar en la comprensión del papel de estos microorganismos y sus funciones dentro del intestino (Shade & Handelsman, 2012).

Al igual que en los distintos estudios que han explorado el núcleo microbiano (Falony et al., 2016; Huse et al., 2012; Salonen et al., 2012; Tap et al., 2009), con el presente análisis no se encontró ninguna OTU con una prevalencia del 100%. Este hecho sugeriría, en principio, que no existe un núcleo microbiano en sentido estricto en esta población. No obstante, tal rigidez conceptual podría ocultar resultados interesantes acerca de esta comunidad microbiana. Es por eso que se explora el núcleo microbiano empleando distintos conceptos y diferentes parámetros, para entender el núcleo microbiano de una forma más incluyente que la simple prevalencia al 100%. Además, esta exploración revelaría la existencia de un núcleo de funciones compartidas por la microbiota de los individuos que serían llevadas a cabo por organismos diferentes con funciones redundantes; alternatively, podría existir un núcleo microbiano a niveles taxonómicos superiores a las OTUs.

Como se detallará y explicará más adelante, el núcleo microbiano que se ha encontrado como el más apropiado es el núcleo de membresía-composición con parámetros de prevalencia-abundancia de 95%-10⁻³, es decir, aquel que incluye 19

OTUs. En términos generales, los géneros *Akkermansia*, *Bacteroides*, *Blautia*, *Bifidobacterium*, *Faecalibacterium*, *Roseburia*, y *Ruminococcus*, se han encontrado asociados con estado de buena salud, con una dieta rica en fibra dietaria y con la producción de ácidos grasos de cadena corta. En particular, los géneros *Bifidobacterium*, *Faecalibacterium*, *Roseburia* y *Ruminococcus* proveen altos niveles de butirato a su hospedero, el cual sirve como fuente de energía para las células epiteliales del intestino e incrementa la expresión de las proteínas de unión estrecha, reforzando así la función de la barrera epitelial (Jalanka-Tuovinen et al., 2011; Nam et al., 2011; Schnorr et al., 2014). Por otro lado, *Blautia* y *Bacteroides* se reconocen por ser productores de acetato, en particular, *Bacteroides* se registra como un marcador de poblaciones urbanizadas/industrializadas y algunos miembros de este género son reconocidos por metabolizar polisacáridos complejos (Mancabelli et al., 2017). *Akkermansia* ha sido constantemente vinculada con la salud metabólica y la obesidad; además, se ha encontrado una correlación inversa entre la abundancia de este grupo, el IMC y los niveles de glucosa en ayunas (Shin et al., 2014). Adicionalmente, *Oscillospira* se ha detectado en baja expresión en individuos obesos y se considera como indicador de buena salud (Konikoff & Gophna, 2016). Por el contrario, las bacterias del género *Escherichia* son enterobacterias patobiontes que podrían indicar estados de **disbiosis** y poca integridad epitelial (Litvak et al., 2017).

Estudios del núcleo microbiano en otras poblaciones han sido previamente reportados. Zhang et al., (2015) caracterizaron las comunidades microbianas de los mongoles y detectaron nueve OTUs en la mayor parte de las muestras, pertenecientes principalmente a los géneros *Bacteroides*, *Faecalibacterium*, *Ruminococcus*, *Subdoligranulum* y *Coprococcus*. Por otro lado, Martínez et al., (2013) caracterizaron las comunidades microbianas del intestino de tres jóvenes norteamericanos durante un período de un año para investigar las características temporales de sus comunidades bacterianas. Estos autores detectaron 16 OTUs en el núcleo microbiano, pertenecientes a los siguientes géneros: *Bifidobacterium*, *Bacteroides*, *Faecalibacterium*, *Ruminococcus*, *Blautia*, *Dorea*, *Eubacterium* y *Coprococcus*. Ling et al., (2013) evaluaron el núcleo microbiano de diez estudiantes chinos sanos y encontraron que los grupos taxonómicos prevalecientes en estas muestras fecales fueron: *Faecalibacterium*, *Coprococcus*, *Blautia*, *Bacteroides*, *Roseburia*, *Ruminococcus*, *Subdoligranulum*, *Sporacetigenium*, *Oscillibacter*, *Dorea*, *Phascolarctobacterium* y *Prevotella*. Finalmente, Huse et al., (2012) exploraron la microbiota intestinal-núcleo en una población de más de 200 individuos y lograron identificar siete OTUs representadas por los géneros *Faecalibacterium*, *Oscillibacter* y *Bacteroides*. Al contraponer nuestro núcleo microbiano con las demás poblaciones de estudio, encontramos tres géneros comunes a todos: *Bacteroides*, *Faecalibacterium* y *Ruminococcus*. Como bien se dijo anteriormente, los

miembros de estos géneros se han visto involucrados en la producción de butirato y la degradación de polisacáridos complejos.

Estos hallazgos sugieren una consistencia sorprendente en la microbiota intestinal-núcleo entre distintas poblaciones. De hecho, la prevalencia de los grupos bacterianos en una población es un indicativo de la habilidad de estos microorganismos de persistir en el tracto gastrointestinal, y varios factores ecológicos podrían explicar esto (Tap et al., 2009); entre ellos, factores ligados a condiciones del ecosistema, la adhesión a algunos alimentos, resistencia a cambios de pH y al estrés provisto por fuerzas mecánicas del movimiento peristáltico, podrían impedir el lavado de estas especies. No obstante, algunos estudios señalan que buena parte de los géneros anteriormente nombrados desempeñan un papel clave en la síntesis de metabolitos básicos en el tracto gastrointestinal humano (Zhang et al., 2015). Estos hallazgos, a grosso modo, sugieren que la microbiota intestinal-núcleo en todos los seres humanos puede variar dentro de un rango limitado.

Curiosamente, el núcleo de membresía hallado en la población colombiana fue equivalente al obtenido con el núcleo filogenético al tener en cuenta un árbol de distancias. No obstante, las circunstancias que explican estos dos núcleos (membresía y filogenético) son disímiles, pero no excluyentes. En primera instancia, este resultado sugiere que aun tomando en cuenta las relaciones de ancestría entre las OTUs no se logra identificar un núcleo con prevalencia 100%, es decir, ni siquiera los ancestros de estas especies moleculares son compartidos por el 100% de los individuos de estudio; esta situación sugiere que la gran variabilidad inter-individual tiene un origen lejano, de manera que no se logran detectar estas taxa. Adicionalmente, cabe resaltar que, en un sentido estricto, no contamos con un árbol filogenético sino con un árbol de distancias que carece de una resolución robusta de las relaciones evolutivas, es decir, no cuenta con una topología fiable; por tanto, resulta un tanto abusivo determinar el núcleo filogenético a partir de esta información. Lo anterior, puesto que algunos phyla (ej. Firmicutes) presentes en el árbol de distancias son **polifiléticos**. Sin embargo, queda como una perspectiva a futuro validar este núcleo utilizando un verdadero árbol filogenético, así como herramientas computacionales diferentes.

Si ignoramos las relaciones evolutivas entre las especies y consideramos únicamente su clasificación taxonómica como fuente de información sobre tales relaciones (núcleo taxonómico), vemos que se comparten 10 taxa en el 100% de los individuos: una clase (Actinobacteria), un orden (Bacteroidales), cuatro familias (Clostridiaceae, Coriobacteriaceae, Enterobacteriaceae y Veillonellaceae), dos géneros (*Clostridium* y *Streptococcus*) y dos especies (*Faecalibacterium prausnitzii*, *Blautia* sp.). Si bien este resultado es aparentemente interesante y expone grupos frecuentemente

encontrados en otros núcleos microbianos (Tap et al., 2009; Zhang et al., 2015), este análisis cuenta con restricciones serias. La clasificación por filotipos se limita a los perfiles taxonómicos que han sido depositados en las distintas bases de datos, es decir, a los microorganismos que han sido cultivados (Schloss & Westcott, 2011). Por tanto, un análisis por filotipos puede carecer de una taxonomía bien definida y la existente no es del todo confiable, dado que reúne grupos **parafiléticos** y polifiléticos (ej: *Ruminococcus*, *Clostridium*) (DeSantis et al., 2006; Quast et al., 2013). En segundo lugar, algunos estudios han demostrado que especies bacterianas cercanamente relacionadas no tienen un solapamiento completo en sus funciones (Cordero et al., 2012; Youngblut et al., 2013). Por tanto, aunque estos 10 taxones sean compartidos por el 100% de la población, sería más conservador considerar como los únicos miembros de este núcleo a las taxa que comprenden niveles taxonómicos inferiores a la familia: *Faecalibacterium prausnitzii*, *Blautia* sp., *Streptococcus* y *Clostridium*, dado que es arduo asegurar una redundancia ecológica en rangos taxonómicos superiores al género. Sin embargo, en vista que *Clostridium* es un género polifilético, no se podría asumir una redundancia funcional entre sus descendientes. Al final, el núcleo taxonómico con prevalencia del 100% estaría, a lo sumo, conformado únicamente por *Faecalibacterium prausnitzii*, *Blautia* sp. y *Streptococcus*.

Dejando a un lado los problemas taxonómicos, a partir de los resultados obtenidos con el núcleo de membresía-composición, se puede observar que las OTUs que más prevalecen en la población colombiana son especies dominantes, es decir, con una alta abundancia. Esto, en esencia, conduce a pensar que estas especies bacterianas dominantes son suficientes para la manutención de la mayor parte de los procesos del ecosistema (el intestino grueso) (Grime, 1998). Por el contrario Tschöp et al., (2009), sugieren que el núcleo microbiano se compone de poblaciones de baja abundancia (raras), a prueba de fallas y que contribuirían con funciones críticas en caso de que la población dominante que llena ese nicho colapse. No obstante, la mayor parte de los estudios del núcleo microbiano sugieren que éste se compone de grupos bacterianos dominantes (Huse et al., 2012; Tap et al., 2009; Zhang et al., 2015), lo cual se confirma con estos análisis. Esto es consistente con estudios que han demostrado la alta resiliencia de la microbiota intestinal humana dominante bajo una exposición corta a antibióticos (Cochetière et al., 2005).

De hecho, la idea sobre las capacidades que tienen las especies dominantes sobre el ecosistema se ve soportada en el mapa de calor de los módulos metabólicos, construido con los distintos núcleos de membresía-composición. Primeramente, se puede observar que al considerar parámetros más restrictivos de prevalencia-abundancia ($99\%-10^{-2}$, $99\%-10^{-4}$), donde se hallaron 2 y 4 OTUs, se encuentran funciones encargadas de la degradación de monosacáridos, aminoácidos y glicerol, es

decir las formas estructuralmente más simples de las macromoléculas que componen nuestra dieta y de la que extraemos energía. En esencia, estas funciones representan vías metabólicas de los degradadores finales de la cadena trófica. Por otro lado, al considerar unos parámetros más laxos ($95\%-10^{-4}$), es decir, al obtener las posibles funciones del conjunto de 19 especies bacterianas, se encuentra un enriquecimiento funcional importante, entre los cuales se incluyen módulos encargados de la extracción de energía de la dieta del hospedero, incluyendo aquellos para el metabolismo de macromoléculas no digeribles por las enzimas humanas, es decir, las vías metabólicas presentes en los degradadores iniciales de la cadena trófica. De hecho, esta diversidad de funciones metabólicas se mantiene estable al relajar aún más los parámetros de prevalencia-abundancia ($75\%-10^{-4}$, $90\%-10^{-4}$); no obstante, cuando se obtienen las posibles funciones de las 100, 137 y 4720 OTUs presentes en los núcleos de membresía-composición $50\%-10^{-4}$, $50\%-10^{-5}$ y $0\%-0$ se obtiene una cantidad menor de funciones, pero siguen estando relacionadas con la degradación de estructuras moleculares complejas.

De forma similar a la arriba descrita, a partir de la reconstrucción del núcleo microbiano basado en las vías metabólicas, Turnbaugh et al. (2009) revelaron que la representación de los grupos funcionales fue altamente consistente en todas las muestras. Entre las funciones conservadas se incluyen varias vías que son probablemente importantes para la subsistencia en el intestino, como la degradación de carbohidratos y metabolismo de aminoácidos (ej. metabolismo de fructosa/manosa, el metabolismo del amino-azúcar y la degradación de N-glicano). Además, se incluyen sistemas de secreción y transporte de membrana (ej. sistemas de fosfotransferasa implicados en el transporte de nutrientes). Aunque muy pocos estudios han llevado a cabo la búsqueda de un núcleo funcional, el estudio de Turnbaugh et al. (2009) es bastante consistente con el núcleo funcional obtenido para la población colombiana. Basada en estos hallazgos y su comparación con estudios previos, se sugiere que el mínimo conjunto de especies bacterianas (es decir, el “verdadero” núcleo microbiano) capaz de realizar la mayor parte de funciones metabólicas dentro del ecosistema y mantener el buen funcionamiento del hospedero no es el núcleo más restrictivo, sino el núcleo de membresía del 95%. A pesar de que el concepto teórico del núcleo microbiano sugiera un parámetro de prevalencia más estricto, en términos biológicos, considerar únicamente a las 4 OTUs que componen el núcleo de membresía al 99% no sería adecuado dado que su potencial funcional es reducido (degradación de estructuras simples de compuestos previamente degradados por otros microorganismos) y, a lo mejor, no lograrían por sí solas sobrevivir dentro del intestino.

Por tal motivo, se considera que el núcleo más apropiado debería ser el conformado por 19 especies (membresía-composición: $95\%-10^{-3}$). No obstante, cabe

resaltar que llegar a esta conclusión requirió la exploración del núcleo microbiano bajo distintas ópticas, es decir, considerando la membresía, la abundancia y el perfil funcional de la microbiota de estudio. Por tanto, es el análisis conjunto de estos tres conceptos el que guía esta determinación y no una escogencia arbitraria, como se propone en otros estudios (Huse et al., 2012; Tap et al., 2009; Zhang et al., 2015).

Teniendo en cuenta que el núcleo de la población colombiana se ve bien representado por una prevalencia del 95%, se hizo un análisis de este núcleo microbiano en diferentes subconjuntos poblacionales (subdividido por ciudad de origen, IMC y género). Al determinar dichos núcleos, se encontró que, a pesar de que el núcleo microbiano tenga variaciones menores por subconjunto considerado (Fig. 6), la mayor parte de las OTUs identificadas se comparte entre las distintas categorías. En consecuencia, estos hallazgos refuerzan la idea de la existencia de un núcleo microbiano común a la población colombiana, capaz de sufrir pequeñas variaciones según la subpoblación de interés.

CONCLUSIONES Y RECOMENDACIONES

El interés específico hacia las bacterias universalmente compartidas surge del hecho de que, a diferencia de los habitantes intestinales transitorios que fluctúan dependiendo de la dieta y otros factores ambientales, las bacterias del núcleo común se conservan durante la coevolución mutua del hombre y sus microbios intestinales. En consecuencia, se prevé que el núcleo microbiano represente un conjunto seleccionado de simbioses. Una vez catalogada, la caracterización dirigida de las bacterias nucleares proporcionaría una estrategia científicamente sólida y económicamente relevante para acceder a los microbios gastrointestinales que son los más relevantes para la salud humana y que pueden tener potencial diagnóstico o terapéutico.

El presente estudio concibe un acercamiento novedoso al concepto del núcleo microbiano empleando nociones de la ecología tradicional, entre ellos la prevalencia, la abundancia, las relaciones filogenéticas de las OTUs y el potencial funcional de estos grupos microbianos. Este trabajo de grado demuestra que la búsqueda de un núcleo microbiano es más rica y compleja que la simple noción intuitiva sobre el conjunto de especies compartidas por todos los miembros de una población. Adicionalmente, se sugiere que el mínimo conjunto de microorganismos capaz de realizar la mayor parte de funciones metabólicas dentro del ecosistema y mantener el buen funcionamiento del hospedero, no es el núcleo más restrictivo (núcleo de membresía al 99%), sino un núcleo de membresía del 95%. Esto, dado que, en términos biológicos, al considerar las 19 OTUs que componen este núcleo se asegura una amplia gama de funciones asociadas tanto con el inicio como el final de la red trófica del intestino, caso contrario a lo que ocurre en el núcleo de membresía al 99%.

En este sentido, las investigaciones futuras podrían encaminarse en, por lo menos, dos vías: (1) explorar el concepto de núcleo del mismo modo que se hace aquí en otras poblaciones ya estudiadas, con el ánimo de determinar la existencia o no de un núcleo microbiano a escala humana. (2) Validar experimentalmente el o los núcleos microbianos, por ejemplo, haciendo uso de modelos animales libres de microorganismos (ej: ratones gnotobióticos). En este último caso, es de esperar que la colonización intestinal por los miembros del núcleo (en el caso de la población aquí estudiada, las 19 OTUs del núcleo 95%- 10^{-3}) permita el adecuado desarrollo y funcionamiento de los hospederos, lo que en últimas validaría experimentalmente la existencia del núcleo.

DECLARACIONES

Consideraciones éticas

El proyecto fue realizado en conformidad con los Principios de la Declaración de Helsinki revisado en 2008 y tiene un mínimo de riesgo acorde al Ministerio de Salud (Resolución 8430 de 1993). Todos los participantes del estudio estuvieron bien informados sobre el proyecto y los procedimientos antes de firmar el consentimiento informado. De hecho, se garantiza el anonimato y la confidencialidad de todos los participantes. El protocolo y el consentimiento informado del proyecto fueron aprobados por el Comité de Bioética de la SIU, Universidad de Antioquia (acta de aprobación 14-24-588, 28 mayo del 2014).

Disponibilidad de los datos y materiales

Las secuencias crudas fueron depositadas en el *Sequence Read Archive* (SRA) (número de acceso pendiente). El análisis bioinformático implementado en este artículo está disponible en <https://github.com/MargaritaValderrama>.

Financiamiento

Este trabajo fue financiado por el Grupo Empresarial Nutresa. Los financiadores no han tenido ningún papel en el diseño o realización del estudio; en la recopilación, gestión, análisis o interpretación de los datos; o en la preparación, revisión o aprobación del manuscrito.

REFERENCIAS

- Aparicio, S., Chapman, J., Stupka, E., Putnam, N., Chia, J., Dehal, P., ... Brenner, S. (2002). Whole-Genome shotgun assembly and analysis of the genome of *Fugu rubripes*.; *Science*, 297(5585), 1301 LP-1310. JOUR. Retrieved from <http://science.sciencemag.org/content/297/5585/1301.abstract>
- Arumugam, M., Raes, J., Pelletier, E., Le, P., Yamada, T., Mende, D. R., ... Bork, P. (2011). Enterotypes of the human gut microbiome. *Nature*, 473(7346), 174–180. <https://doi.org/10.1038/nature09944>
- Aßhauer, K. P., Wemheuer, B., Daniel, R., & Meinicke, P. (2015). Sequence analysis Tax4Fun : predicting functional profiles from metagenomic 16S rRNA data, *Bioinformatics*, 31(17), 2882–2884. <https://doi.org/10.1093/bioinformatics/btv287>
- Backhed, F. (2005). Host-Bacterial Mutualism in the Human Intestine. *Science*, 307(5717), 1915–1920. <https://doi.org/10.1126/science.1104816>
- Bäckhed, F., Fraser, C. M., Ringel, Y., Sanders, M. E., Sartor, R. B., Sherman, P. M., ... Finlay, B. B. (2012). Defining a healthy Human Gut Microbiome: current concepts, future directions, and clinical applications. *Cell Host & Microbe*, 12, 611–622. <https://doi.org/10.1016/j.chom.2012.10.012>
- Bäckhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A., & Gordon, J. I. (2005). Host-Bacterial Mutualism in the Human Intestine. *Science*, 307(5717), 1915 LP-1920. JOUR. Retrieved from <http://science.sciencemag.org/content/307/5717/1915.abstract>
- Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., ... Knight, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences*, 108(Supplement 1), 4516-4522. JOUR. <https://doi.org/10.1073/pnas.1000080107>
- Cheng, J., Ringel-kulka, T., Jong, I. H., Ringel, Y., Carroll, I., Vos, W. M. De, ... Satokari, R. (2016). Discordant temporal development of bacterial phyla and the emergence of core in the fecal microbiota of young children, *Nature*, 10, 1002–1014. <https://doi.org/10.1038/ismej.2015.177>
- Cochetière, M. F. D. La, Durand, T., Lepage, P., Bourreille, a, Galmiche, J. P., Doré, J., & Cochetie, M. F. D. La. (2005). Resilience of the Dominant Human Fecal Microbiota upon Short-Course Antibiotic Challenge Resilience on the Dominant Human Fecal Microbiota upon. *Journal of Clinical Microbiology*, 43(11), 5588. <https://doi.org/10.1128/JCM.43.11.5588>
- Cordero, O. X., Wildschutte, H., Kirkup, B., Proehl, S., Ngo, L., Hussain, F., ... Polz, M. F. (2012). Ecological Populations of Bacteria Act as Socially Cohesive Units of Antibiotic Production and Resistance. *Science*, 337(6099), 1228 LP-1231. JOUR. Retrieved from <http://science.sciencemag.org/content/337/6099/1228.abstract>
- Darzi, Y., Falony, G., Vieira-silva, S., & Raes, J. (2015). Towards biome-specific analysis of meta-

- omics data, *Nature*, 10(5), 1025–1028. <https://doi.org/10.1038/ismej.2015.188>
- De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J. B., Massart, S., ... Lionetti, P. (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proceedings of the National Academy of Sciences of the United States of America*, 107(33), 14691–6. <https://doi.org/10.1073/pnas.1005963107>
- de la Cuesta-Zuluaga, J., Mueller, N. T., Corrales-Agudelo, V., Velásquez-Mejía, E. P., Carmona, J. A., Abad, J. M., & Escobar, J. S. (2016). Metformin Is Associated With Higher Relative Abundance of Mucin-Degrading *Akkermansia muciniphila*; and Several Short-Chain Fatty Acid–Producing Microbiota in the Gut. *Diabetes Care*, 40(1), 54 LP-62. JOUR. Retrieved from <http://care.diabetesjournals.org/content/40/1/54.abstract>
- de Vos, W. M., & de Vos, E. A. J. (2012). Role of the intestinal microbiome in health and disease: from correlation to causation. *Nutrition Reviews*, 70(suppl_1), S45–S56. JOUR. Retrieved from <http://dx.doi.org/10.1111/j.1753-4887.2012.00505.x>
- DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., ... 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. JOUR. <https://doi.org/10.1128/AEM.03006-05>
- Doolittle, W. F. (2009). The practice of classification and the theory of evolution, and what the demise of Charles Darwin's tree of life hypothesis means for both of them. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1527), 2221–2228. JOUR. <https://doi.org/10.1098/rstb.2009.0032>
- Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., & Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics*, 27(16), 2194–2200. JOUR. <https://doi.org/10.1093/bioinformatics/btr381>
- Escobar, J. S., Klotz, B., Valdes, B. E., & Agudelo, G. M. (2014). The gut microbiota of Colombians differs from that of Americans, Europeans and Asians. *BMC Microbiology*, 14(1), 311. <https://doi.org/10.1186/s12866-014-0311-6>
- Evans, J., Sheneman, L., & Foster, J. (2006). Relaxed Neighbor Joining: A Fast Distance-Based Phylogenetic Tree Construction Method. *Journal of Molecular Evolution*, 62(6), 785–792. JOUR. <https://doi.org/10.1007/s00239-005-0176-2>
- Falony, G., Joossens, M., Vieira-silva, S., Wang, J., Darzi, Y., Faust, K., ... Raes, J. (2016). Population-level analysis of gut microbiome variation. *Nature*, 352(6285), 560–564.
- Grice, Elizabeth, Julia A, S. (2012). The Human Microbiome: Our Second Genome*. *Annu Rev Genomics Hum Genet.*, 13(10), 151–170. <https://doi.org/10.1146/annurev-genom-090711-163814>.
- Grime, J. P. (1998). Benefits of plant diversity to ecosystems: immediate, filter and founder effects. *Journal of Ecology*, 86(6), 902–910. <https://doi.org/10.1046/j.1365-2745.1998.00306.x>
- Guidot, A., Prior, P., Schoenfeld, J., Carrère, S., Genin, S., & Boucher, C. (2007). Genomic

- structure and phylogeny of the plant pathogen *Ralstonia solanacearum* inferred from gene distribution analysis. *Journal of Bacteriology*, 189(2), 377–387. <https://doi.org/10.1128/JB.00999-06>
- Haro, C., Rangel-Zúñiga, O. A., Alcalá-Díaz, J. F., Gómez-Delgado, F., Pérez-Martínez, P., Delgado-Lista, J., ... Camargo, A. (2016). Intestinal Microbiota Is Influenced by Gender and Body Mass Index. *PLOS ONE*, 11(5), e0154090. JOUR. Retrieved from <https://doi.org/10.1371/journal.pone.0154090>
- Honda, K., & Littman, D. R. (2012). The Microbiome in Infectious Disease and Inflammation. *Annual Review of Immunology*, 30, 759–795. JOUR. <https://doi.org/10.1146/annurev-immunol-020711-074937>
- Hunt, D. E., David, L. A., Gevers, D., Preheim, S. P., Alm, E. J., & Polz, M. F. (2008). Resource Partitioning and Sympatric Differentiation Among Closely Related Bacterioplankton. *Science*, 320(5879), 1081 LP-1085. JOUR. Retrieved from <http://science.sciencemag.org/content/320/5879/1081.abstract>
- Huse, S. M., Ye, Y., Zhou, Y., & Fodor, A. A. (2012). A core human microbiome as viewed through 16S rRNA sequence clusters. *PLoS ONE*, 7(6), 1–12. <https://doi.org/10.1371/journal.pone.0034242>
- Jalanka-Tuovinen, J., Salonen, A., Nikkilä, J., Immonen, O., Kekkonen, R., Lahti, L., ... de Vos, W. M. (2011). Intestinal microbiota in healthy adults: Temporal analysis reveals individual and common core and relation to intestinal symptoms. *PLOS ONE*, 6(7). <https://doi.org/10.1371/journal.pone.0023035>
- Johnson, Z. I., Zinser, E. R., Coe, A., McNulty, N. P., Woodward, E. M. S., & Chisholm, S. W. (2006). Niche Partitioning Among; Prochlorococcus Ecotypes Along Ocean-Scale Environmental Gradients. *Science*, 311(5768), 1737 LP-1740. JOUR. Retrieved from <http://science.sciencemag.org/content/311/5768/1737.abstract>
- Konikoff, T., & Gophna, U. (2016). Oscillospira : a Central , Enigmatic Component of the Human Gut Microbiota. *Trends in Microbiology*, 24(7), 523–524. <https://doi.org/10.1016/j.tim.2016.02.015>
- Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. *Applied and Environmental Microbiology*, 79(17), 5112–5120. <https://doi.org/10.1128/AEM.01043-13>
- Le Chatelier, E., Nielsen, T., Qin, J., Prifti, E., Hildebrand, F., Falony, G., ... Pedersen, O. (2013). Richness of human gut microbiome correlates with metabolic markers. *Nature*, 500(7464), 541–546. <https://doi.org/10.1038/nature12506>
- Ley, R. E., Peterson, D. A., & Gordon, J. I. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell*, 124(4), 837–848. <https://doi.org/10.1016/j.cell.2006.02.017>
- Ling, Z., Liu, X., Luo, Y., Yuan, L., Nelson, K. E., Wang, Y., ... Li, L. (2013). Pyrosequencing analysis of the human microbiota of healthy Chinese undergraduates. *BMC Genomics*, 14,

390. JOUR. <https://doi.org/10.1186/1471-2164-14-390>
- Litvak, Y., Byndloss, M. X., Tsohis, R. M., & Bäumlér, A. J. (2017). Dysbiotic Proteobacteria expansion: a microbial signature of epithelial dysfunction. *Current Opinion in Microbiology*, 39, 1–6. <https://doi.org/10.1016/j.mib.2017.07.003>
- Lozupone, C. A., & Knight, R. (2007). Global patterns in bacterial diversity. *Proceedings of the National Academy of Sciences of the United States of America*, 104(27), 11436–11440. JOUR. <https://doi.org/10.1073/pnas.0611525104>
- LYONS, K. G., BRIGHAM, C. A., TRAUT, B. H., & SCHWARTZ, M. W. (2005). Rare Species and Ecosystem Functioning. *Conservation Biology*, 19(4), 1019–1024. <https://doi.org/10.1111/j.1523-1739.2005.00106.x>
- Mancabelli, L., Milani, C., Lugli, G. A., Turróni, F., Ferrario, C., & Sinderen, D. Van. (2017). Minireview Meta-analysis of the human gut microbiome from urbanized and pre-agricultural populations, *Environmental Microbiology* 00(00), 00–00 0. <https://doi.org/10.1111/1462-2920.13692>
- Martínez, I., Muller, C. E., & Walter, J. (2013). Long-Term Temporal Analysis of the Human Fecal Microbiota Revealed a Stable Core of Dominant Bacterial Species. *PLoS ONE*, 8(7), e69621. JOUR. <https://doi.org/10.1371/journal.pone.0069621>
- Moeller, A. H., Caro-Quintero, A., Mjungu, D., Georgiev, A. V., Lonsdorf, E. V., Muller, M. N., ... Ochman, H. (2016). Cospeciation of gut microbiota with hominids. *Science (New York, N. Y.)*, 353(6297), 380–382. JOUR. <https://doi.org/10.1126/science.aaf3951>
- Nam, Y., Jung, M., Roh, S. W., Kim, M., & Bae, J. (2011). Comparative Analysis of Korean Human Gut Microbiota by Barcoded Pyrosequencing, 6(7). <https://doi.org/10.1371/journal.pone.0022109>
- Qin, J, R, Li, J, Raes, et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing: Commentary. *Nature*, 11(1), 28. <https://doi.org/10.1038/nature08821>
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., ... 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. JOUR. <https://doi.org/10.1093/nar/gks1219>
- R Development Core Team, R. (2011). R: A Language and Environment for Statistical Computing. (R. D. C. Team, Ed.), *R Foundation for Statistical Computing*. R Foundation for Statistical Computing. <https://doi.org/10.1007/978-3-540-74686-7>
- Ren, T., & Wu, M. (2016). PhyloCore: A phylogenetic approach to identifying core taxa in microbial communities. *Gene*, 593(2), 330–333. JOUR. <https://doi.org/https://doi.org/10.1016/j.gene.2016.08.032>
- Robinson, C. J., Bohannan, B. J. M., & Young, V. B. (2010). From structure to function: the ecology of host-associated microbial communities. *Microbiology and Molecular Biology Reviews: MMBR*, 74(3), 453–76. <https://doi.org/10.1128/MMBR.00014-10>

- Salonen, A., Salojärvi, J., Lahti, L., & de Vos, W. M. (2012). The adult intestinal core microbiota is determined by analysis depth and health status. *Clinical Microbiology and Infection : The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, 18 Suppl 4, 16–20. <https://doi.org/10.1111/j.1469-0691.2012.03855.x>
- Sartor, R. B. (2008). Microbial Influences in Inflammatory Bowel Diseases. *Gastroenterology*, 134(2), 577–594. JOUR. <https://doi.org/10.1053/j.gastro.2007.11.059>
- Sartor, R. B. (2010, December). Genetics and environmental interactions shape the intestinal microbiome to promote inflammatory bowel disease versus mucosal homeostasis . *Gastroenterology*, 139 (6), 1816-1819 . JOUR . <https://doi.org/10.1053/j.gastro.2010.10.036>
- Schloss, P. D. (2009). A high-throughput DNA sequence aligner for microbial ecology studies. *PLOS ONE*, 4(12): e8230.. <https://doi.org/10.1371/journal.pone.0008230>
- Schloss, P. D., & Westcott, S. L. (2011). Assessing and Improving Methods Used in Operational Taxonomic Unit-Based Approaches for 16S rRNA Gene Sequence Analysis . *Applied and Environmental Microbiology*, 77(10), 3219–3226. JOUR. <https://doi.org/10.1128/AEM.02810-10>
- Schnorr, S. L., Candela, M., Rampelli, S., Centanni, M., Consolandi, C., Basaglia, G., ... Crittenden, A. N. (2014). Gut microbiome of the Hadza hunter-gatherers. *Nat Commun*, 5, 3654. <https://doi.org/10.1038/ncomms4654>
- Shade, A., & Handelsman, J. (2012). Beyond the Venn diagram: The hunt for a core microbiome. *Environmental Microbiology*, 14(1), 4–12. <https://doi.org/10.1111/j.1462-2920.2011.02585.x>
- Shin, N.-R., Lee, J.-C., Lee, H.-Y., Kim, M.-S., Whon, T. W., Lee, M.-S., & Bae, J.-W. (2014). An increase in the *Akkermansia* spp. population induced by metformin treatment improves glucose homeostasis in diet-induced obese mice. *Gut*, 63(5), 727 LP-735. JOUR. Retrieved from <http://gut.bmj.com/content/63/5/727.abstract>
- Simpson, H. L., & Campbell, B. J. (2015). Review article: Dietary fibre-microbiota interactions. *Alimentary Pharmacology and Therapeutics*, 42(2), 158–179. <https://doi.org/10.1111/apt.13248>
- Sonnenburg, J. L., Angenent, L. T., & Gordon, J. I. (2004). Getting a grip on things: how do communities of bacterial symbionts become established in our intestine? *Nat Immunol*, 5(6), 569–573. JOUR. Retrieved from <http://dx.doi.org/10.1038/ni1079>
- Tap, J., Mondot, S., Levenez, F., Pelletier, E., Caron, C., Furet, J. P., ... Leclerc, M. (2009). Towards the human intestinal microbiota phylogenetic core. *Environmental Microbiology*, 11(10), 2574–2584. <https://doi.org/10.1111/j.1462-2920.2009.01982.x>
- The Human Microbiome Project Consortium. (2012). Structure, function and diversity of the healthy human microbiome. *Nature*, 486(7402), 207–214. JOUR. Retrieved from <http://dx.doi.org/10.1038/nature11234>
- Tschöp, M. H., Hugenholtz, P., & Karp, C. L. (2009). Getting to the core of the gut microbiome. *Nature Biotechnology*, 27(4), 344–346. <https://doi.org/10.1038/nbt0409-344>

- Taneja V. (2017). Microbiome: Impact of Gender on Functions & Characteristics of Gut Microbiome. In: Legato, MJ (editor). Principles of Gender-Specific Medicine: Gender in the Genomic Era. 3rd edition. Academic Press. 792 pp.
- Turnbaugh, P. J., & Gordon, J. I. (2009). The core gut microbiome, energy balance and obesity. *The Journal of Physiology*, 587(Pt 17), 4153–8. <https://doi.org/10.1113/jphysiol.2009.174136>
- Turnbaugh, P. J., Hamady, M., Yatsunencko, T., Cantarel, B. L., Duncan, A., Ley, R. E., ... Gordon, J. I. (2009). A core gut microbiome in obese and lean twins. *Nature*, 457(7228), 480–484. <https://doi.org/10.1038/nature07540>
- Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-liggett, C., Knight, R., & Gordon, J. I. (2007). The human microbiome project: exploring the microbial part of ourselves in a changing world. *Nature*, 449(7164), 804–810. <https://doi.org/10.1038/nature06244>.The
- Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., & Gordon, J. I. (2007). The Human Microbiome Project. *Nature*, 449(7164), 804–810. JOUR. Retrieved from <http://dx.doi.org/10.1038/nature06244>
- Wallis, A., Butt, H., Ball, M., Lewis, D. P., & Bruck, D. (2016). Support for the Microgenderome: Associations in a Human Clinical Population, 6, 19171. JOUR. Retrieved from <http://dx.doi.org/10.1038/srep19171>
- Wallis, A., Butt, H., Ball, M., Lewis, D. P., & Bruck, D. (2017). Support for the microgenderome invites enquiry into sex differences. *Gut Microbes*, 8(1), 46–52. JOUR. <https://doi.org/10.1080/19490976.2016.1256524>
- Whitman, W. B., Coleman, D. C., & Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *Proc Natl Acad Sci U S A*, 95(12), 6578–6583. <https://doi.org/10.1073/pnas.95.12.6578>
- Yatsunencko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., ... Gordon, J. I. (2012). Human gut microbiome viewed across age and geography. *Nature*, 486(7402), 222–227. <https://doi.org/10.1038/nature11053>
- Youngblut, N. D., Shade, A., Read, J. S., McMahan, K. D., & Whitaker, R. J. (2013). Lineage-specific responses of microbial communities to environmental change. *Applied and Environmental Microbiology*, 79(1), 39–47. <https://doi.org/10.1128/AEM.02226-12>
- Zhang, J., Guo, Z., Lim, A. A. Q., Zheng, Y., Koh, E. Y., Ho, D., ... Zhang, H. (2015). Mongolians core gut microbiota and its correlation with seasonal dietary changes. *Scientific Reports*, 4(1), 5001. <https://doi.org/10.1038/srep05001>
- Zhang, J., Guo, Z., Xue, Z., Sun, Z., Zhang, M., Wang, L., ... Zhang, H. (2015). A phylo-functional core of gut microbiota in healthy young Chinese cohorts across lifestyles, geography and ethnicities. *The ISME Journal*, 9(9), 1–12. <https://doi.org/10.1038/ismej.2015.11>

ANEXOS

OTU id	Clasificación taxonómica a nivel de especie	Clasificación taxonómica a nivel de phylum
Otu00006	<i>Akkermansia muciniphila</i>	Verrucomicrobia
Otu00025	<i>Bacteroides</i> sp.	Bacteroidetes
Otu00007	<i>Bifidobacterium adolescentis</i>	Actinobacteria
Otu00030	<i>Blautia obeum</i>	Firmicutes
Otu00037	<i>Blautia</i> sp.	Firmicutes
Otu00046	<i>Blautia</i> sp.	Firmicutes
Otu00059	<i>Blautia</i> sp.	Firmicutes
Otu00017	<i>Collinsella aerofaciens</i>	Actinobacteria
Otu00004	<i>Escherichia coli</i>	Proteobacteria
Otu00005	<i>Faecalibacterium prausnitzii</i>	Firmicutes
Otu00008	<i>Faecalibacterium prausnitzii</i>	Firmicutes
Otu00003	<i>Gemmiger formicilis</i>	Firmicutes
Otu00012	<i>Oscillospira</i> sp.	Firmicutes
Otu00014	<i>Roseburia faecis</i>	Firmicutes
Otu00024	<i>Ruminococcus gnavus</i>	Firmicutes
Otu00035	<i>Ruminococcus lactaris</i>	Firmicutes
Otu00029	<i>SMB53</i> sp. (familia Clostridiaceae)	Firmicutes
Otu00020	<i>Streptococcus</i> sp.	Firmicutes
Otu00022	<i>Streptococcus</i> sp.	Firmicutes

Anexo 1. Núcleo de membresía con prevalencia del 95%.

Módulos (Nivel I)	Módulos (Nivel II)	Módulos (Nivel III)	Módulos id	Grupo
Degradación de acetil glucosamina	Degradación de aminas y poliamidas	Degradación de acetil glucosamina	MF0003	1
Degradación de fenilalanina	Degradación de aminoácidos	Degradación de aminoácidos aromáticos	MF0007	1
Degradación de tirosina	Degradación de aminoácidos	Degradación de aminoácidos aromáticos	MF0008	1
Degradación de triptófano	Degradación de aminoácidos	Degradación de aminoácidos aromáticos	MF0009	1
Degradación de aspartato	Degradación de aminoácidos	Degradación de aminoácidos cargados de forma negativa	MF0011	1
Degradación de glutamato	Degradación de aminoácidos	Degradación de aminoácidos cargados de forma negativa	MF0014	1
Degradación de glicina	Degradación de aminoácidos	Degradación de aminoácidos no-polares, alifáticos	MF0016	1
Degradación de metionina	Degradación de aminoácidos	Degradación de aminoácidos no-polares, alifáticos	MF0023	1
Degradación de serina	Degradación de aminoácidos	Degradación de aminoácidos no-polares, alifáticos	MF0028	1
Degradación de cisteína	Degradación de aminoácidos	Degradación de aminoácidos no-polares, alifáticos	MF0033	1
Degradación de arginina	Degradación de aminoácidos	Degradación de aminoácidos cargados de forma positiva	MF0035	1
Degradación de lactosa	Degradación de carbohidratos	Degradación de disacáridos	MF0048	1
Degradación de galactosa	Degradación de carbohidratos	Degradación de monosacáridos	MF0056	1
Glicólisis (Fase preparatoria)	Metabolismo central	Metabolismo energético	MF0080	1
Glicólisis	Metabolismo central	Metabolismo energético	MF0081	1
Complejo de piruvato deshidrogenasa	Metabolismo central	Metabolismo energético	MF0083	1
Piruvato: formato liasa	Metabolismo central	Metabolismo energético	MF0085	1
Ciclo de ácido cítrico	Metabolismo central	Metabolismo energético	MF0086	1
Ciclo de ácido cítrico	Metabolismo central	Metabolismo energético	MF0087	1
Ciclo de ácido cítrico	Metabolismo central	Metabolismo energético	MF0088	1
Homoacetogénesis	Metabolismo de gas	Acetogénesis	MF0093	1
Degradación de mucinas	Degradación de glicoproteínas	Degradación de mucina	MF0102	1
Reducción de nitrato	Metabolismos de nutrientes inorgánicos	Nitrógeno	MF0103	1
Derivación de glioxilato	Degradación de lípidos	Derivación de glioxilato	MF0110	1
De acetil-CoA a acetato	Metabolismo de ácido orgánico	Metabolismo del acetato	MF0113	1

Módulos (Nivel I)	Módulos (Nivel II)	Módulos (Nivel III)	Módulos id	Grupo
Producción de lactato	Metabolismo de ácido orgánico	Metabolismo de lactato	MF0119	1
Producción de propionato	Metabolismo de ácido orgánico	Metabolismo de propionato	MF0122	1
Degradación de fucosa	Metabolismo de ácido orgánico	Metabolismo de propionato	MF0124	1
Producción de succinato	Metabolismo de ácido orgánico	Metabolismo de succinato	MF0127	1
Catalasa	Protección contra el estrés oxidativo	Catalasa	MF0129	1
Degradación de ribosa	Degradación de carbohidratos	Degradación de monosacáridos	MF0060	2
Desde acetato a acetil-CoA	Metabolismo de ácido orgánico	Metabolismo del acetato	MF0112	2
Degradación de alpha-D-glucosa y alfa-D-glucosa 1-fosfato	Degradación de carbohidratos	Degradación de monosacáridos	MF0057	3
Producción de etanol	Metabolismo del alcohol	Metabolismo de etanol	MF0001	3
Degradación de alanina	Degradación de aminoácidos	Degradación de aminoácidos no-polares, alifáticos	MF0017	3
Degradación de prolina	Degradación de aminoácidos	Degradación de aminoácidos no-polares, alifáticos	MF0018	3
Degradación de treonina	Degradación de aminoácidos	Degradación de aminoácidos polar, no-cargados	MF0029	3
Degradación de treonina	Degradación de aminoácidos	Degradación de aminoácidos polar, no-cargados	MF0030	3
Degradación glutamina	Degradación de aminoácidos	Degradación de aminoácidos polar, no-cargados	MF0032	3
Degradación de arginina	Degradación de aminoácidos	Degradación de aminoácidos polar, no-cargados	MF0034	3
Degradación de sucrosa	Degradación de carbohidratos	Degradación de disacáridos	MF0046	3
Degradación de melibiose	Degradación de carbohidratos	Degradación de disacáridos	MF0050	3
Degradación de sucrosa	Degradación de carbohidratos	Degradación de disacáridos	MF0051	3
Degradación de manosa	Degradación de carbohidratos	Degradación de monosacáridos	MF0061	3
Degradación de azúcar	Degradación de carbohidratos	Degradación de polisacáridos	MF0062	3
Metabolismo de glicógeno	Degradación de carbohidratos	Almacenaje	MF0066	3
bifidobacterium shunt	Metabolismo central	Metabolismo energético	MF0079	3

Módulos (Nivel I)	Módulos (Nivel II)	Módulos (Nivel III)	Módulos id	Grupo
Vía pentosa fosfato	Metabolismo central	Metabolismo energético	MF0090	3
Degradación de urea	Degradación de aminas y poliamidas	Degradación de urea	MF0006	4
Degradación de isoleucina	Degradación de aminoácidos	Degradación de aminoácidos no-polares, alifáticos	MF0022	4
Degradación de asparagina	Degradación de aminoácidos	Degradación de aminoácidos polar, no-cargados	MF0031	4
Degradación de lisina	Degradación de aminoácidos	Degradación de aminoácidos cargados de forma positiva	MF0040	4
Degradación de xilosa	Degradación de carbohidratos	Degradación de monosacáridos	MF0055	4
Degradación de fructosa	Degradación de carbohidratos	Degradación de monosacáridos	MF0058	4
Degradación de glicerol	Degradación de lípidos	Degradación de glicerol	MF0109	4

Anexo 2. Módulos metabólicos predichos en el núcleo funcional.

OTU id	Clasificación taxonómica inferior	Clasificación a nivel de phylum	Barranquilla	Bogotá	Bucaramanga	Cali	Medellín
Otu00058	<i>Actinomyces</i> sp.	Actinobacteria		○			○
Otu00006	<i>Akkermansia muciniphila</i>	Verrucomicrobia		○	○	○	
Otu00027	<i>Bacteroides fragilis</i>	Bacteroidetes		○			
Otu00025	<i>Bacteroides</i> sp.	Bacteroidetes	○	○	○		○
Otu00007	<i>Bifidobacterium adolescentis</i>	Firmicutes	○	○	○		○
Otu00030	<i>Blautia obeum</i>	Firmicutes	○	○			○
Otu00037	<i>Blautia</i> sp.	Firmicutes	○	○	○	○	○
Otu00046	<i>Blautia</i> sp.	Firmicutes	○	○			○
Otu00059	<i>Blautia</i> sp.	Firmicutes	○	○	○		○
Otu00016	<i>Cellulosibacter alkalithermophilus</i>	Firmicutes		○	○		○
Otu00010	<i>Clostridium celatum</i>	Firmicutes	○	○	○		
Otu00026	<i>Clostridium hathewayi</i>	Firmicutes		○		○	○
Otu00017	<i>Collinsella aerofaciens</i>	Actinobacteria	○	○	○		○
Otu00083	<i>Dorea formicigenerans</i>	Firmicutes		○	○		○
Otu00066	<i>Dorea</i> sp.	Firmicutes		○			
Otu00009	<i>Enterobacter hormaechei</i>	Proteobacteria	○		○	○	
Otu00004	<i>Escherichia coli</i>	Proteobacteria	○	○	○	○	○
Otu00005	<i>Faecalibacterium prausnitzii</i>	Firmicutes	○	○	○	○	○
Otu00008	<i>Faecalibacterium prausnitzii</i>	Firmicutes	○	○	○	○	○
Otu00003	<i>Gemmiger formicilis</i>	Firmicutes	○	○	○	○	○
Otu00097	Mogibacteriaceae	Firmicutes		○			
Otu00045	<i>Streptophyta</i>	Cyanobacteria		○			○
Otu00012	<i>Oscillospira</i> sp.	Firmicutes			○	○	○
Otu00002	<i>Prevotella copri</i>	Bacteroidetes		○		○	○
Otu00014	<i>Roseburia faecis</i>	Firmicutes	○	○	○	○	○
Otu00064	<i>Roseburia inulinivorans</i>	Firmicutes		○			
Otu00013	<i>Ruminococcus bromii</i>	Firmicutes		○		○	
Otu00024	<i>Ruminococcus gnavus</i>	Firmicutes	○	○	○	○	○
Otu00035	<i>Ruminococcus lactaris</i>	Firmicutes	○	○	○		○
Otu00029	SMB53 sp. (familia Clostridiaceae)	Firmicutes	○		○		○
Otu00020	<i>Streptococcus</i> sp.	Firmicutes	○	○	○		○
Otu00022	<i>Streptococcus</i> sp.	Firmicutes	○	○	○	○	○
Otu00056	<i>Subdoligranulum variable</i>	Firmicutes		○			
Otu00044	<i>Veillonella dispar</i>	Firmicutes	○				

Anexo 3. Núcleo de membresía por ciudad considerando una prevalencia del 95%.

OTU id	Clasificación taxonómica a nivel de especie	Clasificación a nivel de phylum	Normal	Sobrepeso	Obesidad
Otu00006	<i>Akkermansia muciniphila</i>	Verrucomicrobia		○	○
Otu00027	<i>Bacteroides fragilis</i>	Bacteroidetes	○		
Otu00025	<i>Bacteroides</i> sp.	Bacteroidetes	○	○	
Otu00007	<i>Bifidobacterium adolescentis</i>	Firmicutes		○	○
Otu00030	<i>Blautia obeum</i>	Firmicutes	○		○
Otu00037	<i>Blautia</i> sp.	Firmicutes	○	○	○
Otu00046	<i>Blautia</i> sp.	Firmicutes			○
Otu00059	<i>Blautia</i> sp.	Firmicutes	○	○	○
Otu00010	<i>Clostridium celatum</i>	Firmicutes		○	
Otu00026	<i>Clostridium hathewayi</i>	Firmicutes	○	○	
Otu00017	<i>Collinsella aerofaciens</i>	Actinobacteria	○	○	○
Otu00083	<i>Dorea formicigenerans</i>	Firmicutes			○
Otu00004	<i>Escherichia coli</i>	Proteobacteria	○	○	○
Otu00005	<i>Faecalibacterium prausnitzii</i>	Firmicutes	○	○	○
Otu00008	<i>Faecalibacterium prausnitzii</i>	Firmicutes	○	○	○
Otu00003	<i>Gemmiger formicilis</i>	Firmicutes	○	○	○
Otu00012	<i>Oscillospira</i> sp.	Firmicutes	○	○	○
Otu00002	<i>Prevotella copri</i>	Bacteroidetes			○
Otu00024	<i>Ruminococcus gnavus</i>	Firmicutes	○	○	○
Otu00014	<i>Roseburia faecis</i>	Firmicutes	○	○	○
Otu00013	<i>Ruminococcus bromii</i>	Firmicutes		○	
Otu00035	<i>Ruminococcus lactaris</i>	Firmicutes		○	○
Otu00029	<i>SMB53</i> sp.	Firmicutes		○	○
Otu00020	<i>Streptococcus</i> sp.	Firmicutes	○	○	
Otu00022	<i>Streptococcus</i> sp.	Firmicutes	○	○	○

Anexo 4. Núcleo de membresía por categorías de índice de masa corporal considerando una prevalencia del 95%.

OTU id	Clasificación taxonómica a nivel de especie	Clasificación a nivel de phylum	Masculino	Femenino
Otu00006	<i>Akkermansia muciniphila</i>	Verrucomicrobia		○
Otu00025	<i>Bacteroides</i> sp.	Bacteroidetes		○
Otu00007	<i>Bifidobacterium adolescentis</i>	Actinobacteria	○	
Otu00030	<i>Blautia obeum</i>	Firmicutes		○
Otu00037	<i>Blautia</i> sp.	Firmicutes	○	○
Otu00046	<i>Blautia</i> sp.	Firmicutes		○
Otu00059	<i>Blautia</i> sp.	Firmicutes	○	○
Otu00010	<i>Clostridium celatum</i>	Firmicutes		○
Otu00026	<i>Clostridium hathewayi</i>	Firmicutes		○
Otu00017	<i>Collinsella aerofaciens</i>	Actinobacteria	○	○
Otu00083	<i>Dorea formicigenerans</i>	Firmicutes	○	
Otu00009	<i>Enterobacter hormaechei</i>	Proteobacteria	○	
Otu00004	<i>Escherichia coli</i>	Proteobacteria	○	○
Otu00005	<i>Faecalibacterium prausnitzii</i>	Firmicutes	○	○
Otu00008	<i>Faecalibacterium prausnitzii</i>	Firmicutes	○	○
Otu00003	<i>Gemmiger formicilis</i>	Firmicutes	○	○
Otu00012	<i>Oscillospira</i> sp.	Firmicutes	○	○
Otu00002	<i>Prevotella copri</i>	Bacteroidetes	○	
Otu00014	<i>Roseburia faecis</i>	Firmicutes	○	○
Otu00013	<i>Ruminococcus bromii</i>	Firmicutes		○
Otu00024	<i>Ruminococcus gnavus</i>	Firmicutes	○	○
Otu00035	<i>Ruminococcus lactaris</i>	Firmicutes	○	
Otu00029	<i>SMB53</i> sp.	Firmicutes		○
Otu00020	<i>Streptococcus</i> sp.	Firmicutes	○	○
Otu00022	<i>Streptococcus</i> sp.	Firmicutes	○	○

Anexo 5. Núcleo de membresía por sexo considerando una prevalencia del 95%.