



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



Escola Tècnica Superior  
d'Enginyeria Agronòmica i del Medi Natural

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Técnica Superior de Ingeniería Agronómica  
y del Medio Natural

Secuenciación, ensamblaje y evaluación funcional del  
genoma de la especie *Blautia* spp., aislado del tracto  
intestinal humano y con potencial anti-inflamatorio

Trabajo Fin de Grado

Grado en Biotecnología

AUTOR/A: Casado García-Consuegra, Jaime

Tutor/a: Forment Millet, José Javier

Cotutor/a externo: BENITEZ PAEZ, AFONSO

CURSO ACADÉMICO: 2021/2022



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



Escola Tècnica Superior  
d'Enginyeria Agrònoma i del Medi Natural

# UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Escuela Técnica Superior de Ingeniería Agronómica  
y del Medio Natural

Secuenciación, ensamblaje y evaluación funcional del  
genoma de la especie *Blautia* spp., aislado del tracto  
intestinal humano y con potencial anti-inflamatorio

Trabajo Fin de Grado  
Grado en Biotecnología

AUTOR/A: Casado García-Consuegra, Jaime

Tutor/a: Forment Millet, José Javier

Cotutor/a externo: BENITEZ PAEZ, AFONSO

CURSO ACADÉMICO: Biotecnología

**Título**

*Secuenciación, ensamblaje y evaluación funcional del genoma de la especie *Blautia spp.*, aislado del tracto intestinal humano y con potencial anti-inflamatorio*

**Resumen**

La investigación en el campo del microbioma humano ha permitido dilucidar el gran impacto que tienen en nuestra salud los microbios que habitan en nuestro tracto digestivo. Las interacciones dieta-hospedador microbiota también han permitido concluir que existe un gran potencial biomédico/biotecnológico sobre ciertas especies predominantes en nuestro intestino y que podrían contribuir de manera importante en el mantenimiento del mismo, con un amplio impacto a nivel sistémico. Por ello, el conocimiento en profundidad de la fisiología, genética y las funciones metabólicas codificadas en los genomas de dichos microorganismos representa un aspecto relevante en la caracterización tanto de cepas con potencial probiótico como de biomarcadores de enfermedades no comunicables. El presente proyecto tiene como objetivo ensamblar y evaluar desde un punto de vista funcional el genoma de un aislado de la especie *Blautia spp.*, que mediante estudios clínicos previos ha demostrado tener un posible papel en la obesidad infantil y en la prevención del tono inflamatorio intestinal subyacente a dicha condición mórbida.

**Palabras clave:** microbioma; *Blautia*; obesidad infantil; bioinformática; ensamblaje; anotación funcional; probiótico; biomarcadores; filogenia; secuenciación

**Alumno/a:** Jaime Casado García-Consuegra

**Tutor experimental:** Alfonso Benítez Páez

**Tutor UPV:** Forment Millet, José Javier

València, 30 de agosto de 2022

**Title**

*Sequencing, assembly and functional evaluation of the genome of Blautia spp. isolated from the human intestinal tract with anti-inflammatory potential*

**Abstract**

Research in the field of the human microbiome has made possible to elucidate the great impact that the microbes that inhabit our digestive tract have on our health. Diet-host-microbiota interactions have also led to the conclusion that there is a big biomedical/biotechnological potential for certain predominant species in our intestine and that they could contribute significantly to its maintenance, with a broad impact at the systemic level. Therefore, in-depth knowledge of the physiology, genetics and metabolic functions encoded in the genomes of these microorganisms represents a relevant aspect in the characterization of both strains with probiotic potential and biomarkers of non-communicable diseases. The objective of this project is to assemble and evaluate from a functional point of view the genome of an isolate of a *Blautia spp.* species, which through previous clinical studies has been shown to have a possible role in childhood obesity and in the prevention of tone intestinal inflammation underlying said morbid condition.

**Key words:** microbiome; Blautia; childhood obesity; bioinformatics; assembly; functional annotation; probiotic; biomarkers; phylogeny; sequencing

**Student:** Jaime Casado García-Consuegra

**Experimental tutor:** Alfonso Benítez Páez

**UPV tutor:** José Javier Forment Millet

València, 30th August, 2022.



# ÍNDICE

## **1. Introducción.**

### 1.1 Obesidad.

#### 1.1.1 Generalidades.

#### 1.1.2 Problemas asociados

#### 1.1.3 Obesidad infantil.

### 1.2 Inflamación.

#### 1.2.1 Generalidades.

#### 1.2.2 Causas y consecuencias.

#### 1.2.3 Relación con el microbiota.

### 1.3 Microbiota.

#### 1.3.1 Funciones.

#### 1.3.2 Importancia en la obesidad.

#### 1.3.3 Géneros que influyen en la obesidad.

### 1.4 Blautia.

#### 1.4.1 Generalidades.

#### 1.4.2 Estudios previos.

## **2. Objetivos**

## **3. Materiales y métodos**

### 3.1 Pretratamiento de muestras.

#### 3.1.1 Calidad de muestras.

#### 3.1.2 Recorte de lecturas.

### 3.2 Ensamblaje.

#### 3.2.1 Generalidades.

3.2.2 Abyss.

3.2.3 Spades.

3.2.4 Quast.

3.2.5 Repetición del proceso.

3.3 Anotación estructural y funcional.

3.3.1 Quast.

3.3.2 Prodigal.

3.3.3 Egnog.

3.4 Alineación de secuencias y filogenia.

## **4. Resultado y Discusión**

4.1 Ensamblaje.

4.2 Identificación taxonómica.

4.2.1 Resultados Blast

4.2.2 Comparación de genomas *Blautia massiliensis*

4.2.3 Comparación de filogenias

4.3 Resultados análisis funcional.

## **5. Conclusiones**

## **6. Bibliografía**

## **7. Anexos**

## ÍNDICE DE TABLAS

*Tabla 1.* Comparación de resultados del ensamblaje en diferentes momentos del estudio

*Tabla 2.* Resultados de Blast al realizarlo con la secuencia 16S.

*Tabla 3.* Resultados de Blast al realizarlo con la secuencia 23S.

*Tabla 4.* Comparación de ensamblajes de *Blautia massiliensis*.

*Tabla 5.* Clasificación de la composición genética del ensamblaje mediante categorías COG.

## ÍNDICE DE FIGURAS

*Figura 1.* Filogenias de la secuencias ribosomales 16S.

*Figura 2.* Filogenias de la secuencias ribosomales 23S.

*Figura 3.* Comparación de los patrones génicos de diferentes cepas de *Blautia massiliensis*.



# 1. Introducción.

## 1.1 Obesidad

### 1.1.1 Generalidades.

En las últimas décadas como consecuencia del gran aumento de casos en el mundo, la obesidad ha comenzado a constituir un grave problema para la sanidad global, adquiriendo el grado de epidemia según la OMS (1). La obesidad se define como un desbalance entre la ingesta de energía y el gasto de la misma, que puede desembocar en una acumulación excesiva de tejido adiposo que afecta de manera negativa a la salud (2). Aunque en grandes rasgos se acepta la definición, es en los criterios de diagnóstico donde existen las diferencias.

Existen diversas formas de clasificación de la obesidad siendo la más común la dependiente del índice de masa corporal (IMC). El IMC es un parámetro que correlaciona el peso y la altura de un individuo para poder determinar problemas en su desarrollo (3). Según la OMS se considera que un adulto sufre de sobrepeso si su IMC es igual o superior a 25 (kg/m<sup>2</sup>) y de obesidad si su IMC es igual o superior a 30 (kg/m<sup>2</sup>). La consideración cambia si se refiere a niños y adolescentes. Para estos casos, se define la obesidad para una determinada altura de acuerdo con la cantidad de desviaciones típicas que presente con respecto a la mediana de los patrones de crecimiento infantil de la OMS (2).

Aun así, otras organizaciones han acordado diferentes criterios. Para gobiernos de países asiáticos, aunque se sigue utilizando el IMC como medida indicativa, los valores que determinan el sobrepeso son sensiblemente menores. Por ejemplo, Japón define la obesidad cuando el IMC es superior a 25 (kg/m<sup>2</sup>) (4) y China cuando es superior a 28 (kg/m<sup>2</sup>). (5)

De acuerdo con esta clasificación la OMS estima que en 2016 existían más de 1900 millones de adultos con sobrepeso, de los cuales 650 millones eran obesos -más del 10% de la población adulta del mundo. (2) Desde 1975 la prevalencia mundial de la obesidad se ha triplicado y se espera que para el 2025 alcance cuotas de alrededor del 20% en la población adulta mundial. (6)

Los datos aumentan cuando el foco se centra en la población infantil. En 2016 habían más de 100 millones de niños y adolescentes que sufrían de obesidad (7). Comparado con 40 años atrás, la prevalencia se ha multiplicado por 8 (8). Además, también existe un cambio de tendencia en cuanto a las zonas afectadas. Si históricamente se asociaba la obesidad con

países de altos ingresos, de un tiempo a esta parte este trastorno también ha aumentado en zonas de medios y bajos ingresos, sobre todo en entornos urbanos (9). Como muestra de ello, en África se han duplicado el número de niños con sobrepeso desde el año 2000 y alrededor de la mitad de los niños obesos menores de 5 años vivían en Asia (2).

Los informes referentes a la obesidad infantil son excepcionalmente importantes ya que, aunque se espera que el 60% de los niños que sufren de esta enfermedad sigan desarrollándola en edades adultas tempranas (10), es en esta etapa donde mayores serán las oportunidades de cambio y más decisivos serán éstos para revertir los principales problemas asociados que trae consigo el desarrollo de la obesidad (11).

Existen diferentes enfermedades asociadas al diagnóstico de la obesidad que dificultan y empeoran el desarrollo de la misma. Conforme menor sea la persona que la sufre, más posibilidades tendrá de padecer estas complicaciones. Estudios han mostrado que sufrir de sobrepeso en edades tempranas aumenta la posibilidad de mortalidad adulta por enfermedades sistémicas (12). Según la OMS, casi 3 millones de personas fallecen anualmente por complicaciones relacionadas con el sobrepeso, mayoritariamente por problemas cardiovasculares (13). Aun así, estos problemas no son los únicos a los que un individuo que padece esta enfermedad debe hacer frente a lo largo de su vida, la obesidad infantil puede afectar a casi todos los sistemas de órganos de manera grave (14).

### 1.1.2 Problemas asociados

Se han reportado en numerosas ocasiones que el exceso de tejido adiposo en individuos disminuye la capacidad de su sistema inmune, tanto innato como adquirido. En modelos murinos se detectó que la obesidad, independientemente de su origen, producía una alteración en mecanismos relacionados con la leptina y como consecuencia, una pérdida en la funcionalidad de mecanismos inmunológicos (15).

En cuanto a humanos, la obesidad se asocia a una mayor posibilidad de padecer infecciones nosocomiales después de cirugías (16). Además, se relacionó el sobrepeso con una disminución en la efectividad de la vacunación (17). Esta deficiencia en el sistema inmune produce un efecto en el que se amplifican las comorbilidades, ya que las posibles infecciones aumentan la inflamación crónica que, como se explicará después, tienen, a su vez, un efecto negativo en el desarrollo de obesidad (18).

Otros problemas que genera el desarrollo de obesidad son los de tipo metabólico. Se observó que la mitad de los niños cuyo IMC se encontraba en el percentil 97 sufrían de algún tipo de

síndrome metabólico (19). Entre este tipo de enfermedades, las que más posibilidades tienen estos niños de desarrollar son la diabetes de tipo II, si se producía una resistencia prolongada a la insulina (20) o enfermedades cardiovasculares de un amplio rango (21).

A todas estas enfermedades se les añaden otras condiciones que, si no se tratan de la manera adecuada, pueden llegar a convertirse en un grave problema. Entre éstas, se han observado trastornos pulmonares (22), daño renal (23), potencialidad a diferentes tipos de cánceres (24) y problemas en el desarrollo óseo o complicaciones ortopédicas, tales como fracturas o molestias musculares (25). Finalmente, cabe destacar otras disposiciones que puede acarrear la obesidad como el avance prematuro de la pubertad mediante la aceleración de la telarquia y menarquia en niñas (26) o efectos adversos en el ámbito psicosocial (27).

Como ya se ha explicado, es en la etapa de la niñez donde más posibilidades habrá para erradicar todas estas dolencias. Se demostró que un cambio, por mínimo que sea, hacia un peso correcto en las etapas iniciales del desarrollo disminuían de forma considerable el riesgo a enfermedades metabólicas en posteriores etapas de la vida (28).

### 1.1.3 Obesidad infantil.

Conseguir reducir el peso en niños con obesidad no es sencillo pues existe una gran cantidad de causas para el desarrollo de esta problemática. En la base, existe un rasgo histórico que explica el auge de la enfermedad. Durante miles de años el alimento fue escaso, lo que favoreció que los individuos con un metabolismo “calmado” prosperasen y fijasen este rasgo en su descendencia, lo que se conoce como hipótesis del gen ahorrativo (29). Conforme el alimento fue asegurándose, resultado de un avance tecnológico y el gasto energético requerido para llevar a cabo los trabajos diarios fue reduciéndose, se fue produciendo un aumento en la prevalencia de la obesidad (30).

Estos hechos produjeron una serie de cambios sociales, biológicos y genéticos que explicarían la gran cantidad de causas que pueden llegar a provocar la obesidad. De estas causas destacan las de tipo genético. Se han detectado genes, muchos de ellos relacionados con el metabolismo de la leptina, cuyo fallo generan anomalías que aumentan la probabilidad de obesidad (31). Aparte de este hecho, es cierto que existen ciertas variantes genéticas asociadas con el aumento de la adiposidad en el cuerpo (32). Muchas de estas variantes están relacionadas con enfermedades endocrinas como el hipotiroidismo, la falta de la hormona del crecimiento o el síndrome de ovario poliquístico, entre otros (33).

Además de la componente genética, (en los últimos años) se ha investigado como afecta el ambiente a la expresión de la enfermedad, es decir la relación entre la obesidad y la epigenética. Se descubrió que las condiciones en el útero eran capaces de regular la expresión de determinados genes mediante la metilación del ADN o la modificación de histonas (34). Así, una exposición a una mayor cantidad de grasa (35) o a diabetes gestacional (36) en el útero afectó a la futura potencialidad de desarrollar obesidad. Otro elemento que influyó en el riesgo a la obesidad fue el peso del gestante al nacer. Se advirtió que bebés pequeños para su edad podrían estar en peligro de obesidad al recuperar su talla (37).

A estos factores biológicos se le añaden otros de carácter social, como la posición socioeconómica (38); carácter conductual, como los hábitos al dormir (39) o la dieta (40) y de carácter ambiental como el país o etnia de origen (41); todos los factores anteriormente nombrados, influyen en otra variable que en la última década ha comenzado a estudiarse como posible determinante en la obesidad infantil, la microbiota humana y su posible implicación en la obesidad mediante procesos inflamatorios.

## **1.2 Inflamación**

### **1.2.1 Generalidades.**

La inflamación se define como una respuesta adaptativa a ciertos factores malignos, tales como lesiones o infecciones (42). Esta respuesta será positiva siempre y cuando sea controlada y no se perpetúe en el tiempo más allá de lo necesario. Si es así, podría desembocar en una inflamación crónica con consecuencias negativas para el individuo como enfermedades autoinmunes u obesidad (43). De hecho, se observó la presencia de marcadores inflamatorios, como citocinas y quimiocinas, en las deposiciones de niños con resistencia a la insulina y obesidad (44).

### **1.2.2 Causas y consecuencias.**

A nivel molecular, el conocimiento sobre la inflamación ha avanzado mucho en las últimas décadas. Aunque existen diferentes vías para la activación y desarrollo de la condición, esencialmente la inflamación se basa en la coordinación de sistemas del organismo para conseguir la llegada de distintas células y componentes sanguíneos al lugar de la infección.

Para que se inicie el proceso inflamatorio se necesita la activación de los receptores del sistema inmunitario presentes en los macrófagos y mastocitos del tejido afectado. La activación de receptores como los de tipo TOLL (TLR) o los similares a NOD (NLRs) generarán la

producción, por parte de las células anteriormente nombradas, de diferentes mediadores inflamatorios. El efecto de estos mediadores será el de facilitar a las proteínas plasmáticas y leucocitos la llegada al lugar afectado. Una vez en el tejido dañado, los neutrófilos, un tipo de leucocito, se activarán por contacto con determinadas moléculas y secretarán productos reactivos (ROS, nitrógeno, elastasa, catepsina G...) (45) para intentar acabar con los patógenos infiltrados. Finalmente, cuando el peligro ya ha sido eliminado, la producción de mediadores proinflamatorios tornará a mediadores antiinflamatorios como las lipoxinas, resolvinas o las protectinas. Estas, evitarán la activación de neutrófilos y favorecerán la de monocitos que se desharán de los residuos y comenzarán la remodelación celular (46).

Es importante remarcar que los productos reactivos que los macrófagos liberan para intentar eliminar a los patógenos también son tóxicos para las células del organismo. Es por ello que, si la inflamación perdura, puede tener efectos negativos. Además, conforme el proceso persiste va pasando por nuevas fases que pueden degenerar en inflamación crónica con efectos más graves para el individuo (46).

La inflamación crónica es, por tanto, consecuencia de una regulación deficiente en el control del proceso inflamatorio. Aunque los mecanismos moleculares que causan la dolencia se siguen investigando, el origen es muy diverso. La condición puede ser consecuencia desde por un fracaso en la eliminación del patógeno (y por lo tanto la prolongación de la fase de inflamación aguda) a trastornos autoinmunes como el lupus o la artritis reumatoide, donde el sistema inmunitario reconoce como dañinos elementos del organismo atacándolos durante el proceso (47). Otros motivos para el desarrollo de la enfermedad son la manifestación de episodios recurrentes de inflamación aguda, la exposición a alérgenos o efectos asociados a la presencia de determinados microorganismos de la microbiota (47).

Todos estos factores pueden llegar a desencadenar un estado fisiológico que tiene graves consecuencias para la salud. Así, se asocia la resistencia a la insulina con tejidos adiposos ricos en receptores proinflamatorios tipo Toll (TLR), que aparte de ser regulables por elementos microbianos, también lo son por lípidos presentes en la dieta (48). Además, también se ha relacionado el aumento de citocinas proinflamatorias como IL-6, IL-1 o IL-18 con una disrupción en la señalización de la insulina (49).

Otra secuela que puede tener la inflamación crónica es la regulación del sistema inmune de la persona. La activación del sistema inmunitario dependerá de los estímulos que capte el organismo. Éstos, al unirse a sus respectivos receptores desencadenarán una respuesta que conllevará la diferenciación de células T y macrófagos para acudir a la zona afectada (18) (50).

Es por ello, que dependiendo del tipo de receptor que se active, aumentará la concentración de un tipo concreto de célula inmune u otro.

A todos estos problemas, subyace uno de mayor importancia, la influencia de la inflamación en la obesidad. Se ha encontrado asociaciones entre la inflamación hipotalámica y la obesidad (51). Esto es debido a que es en el hipotálamo donde se produce el control del equilibrio de la ingesta y gasto energético mediante neuroreguladores como la leptina o la insulina. Cuando el hipotálamo se encuentra inflamado se producen resistencias a los neuroreguladores anteriores, creando desajustes en este control y disminuyendo la sensación de saciedad del organismo. Es por ello, que en animales se encontró que eliminando receptores proinflamatorios como el TLR4, se reducía la inflamación hipotalámica y con ello mejoraban la señalización de los reguladores (18).

### 1.2.3 Relación con la microbiota.

Todos estos efectos, como se ha explicado, se generarán según las vías de inflamación que se active. La mayoría de las veces, la activación dependerá de factores como el tipo de inductor que promueve el inicio del proceso. Existen dos tipos de inductores: los de tipo endógeno, que provienen de estímulos creados por las propias células del organismo y los de tipo exógenos, originarios del exterior.

Cuando se habla de la relación de la microbiota con la inflamación se refiere a inductores de tipo exógeno, más concretamente de tipo bacteriano. Es importante destacar que cualquier inductor del que se hable nunca tendrá una respuesta fija en todas las circunstancias, sino que dependerá del contexto en que se encuentre (52). Un mismo estímulo puede tener características inflamatorias y antiinflamatorias según el lugar en que se encuentre.

Para determinar qué clase de respuesta generará un componente es muy importante las primeras etapas de la vida, ya que serán en éstas donde se determinará la microbiota del individuo. Cada pequeña acción influirá en la composición. Por ejemplo, se observaron diferencias en las especies que componían la microbiota de una persona dependiendo la forma en la que nacían, si mediante parto vaginal (en las que se encontró mayor variabilidad en las especies) o cesárea (53).

Durante estas primeras épocas la respuesta inmune del individuo es menor por la inmadurez de su sistema inmunitario. Este hecho, facilitará el asentamiento de las especies que conformarán la microbiota (54). Además, se promoverán la colonización de los microorganismos comensales mediante una serie de mecanismos. El individuo neonatal

atenuará, mediante mediadores inflamatorios o células eritroides, las respuestas defensoras incluso si se activan los receptores que en la etapa adulta generarían respuestas inflamatorias (54)(55). Se teoriza con que estos procesos de colonización determinarán el futuro de la composición de la microbiota del organismo a lo largo de su vida por cambios epigenéticos en las células epiteliales que generarán tolerancia hacia algunos microorganismos (56).

Aun así, en el futuro, para evitar posibles efectos negativos se llevarán a cabo estrategias que minimicen la posibilidad de contacto entre los receptores proinflamatorios y los microorganismos. Éstas se basarán en la presencia de células epiteliales, moco, IgA y péptidos antimicrobianos en los lugares en donde la existencia de microbiota se puede considerar como densa (57). Además de todo ello, si se produjese una traslocación de estos organismos comensales se ocasionaría una activación de macrófagos que los eliminarían sin tener que activar la respuesta inflamatoria (58). Se ha probado que la modificación de la microbiota intestinal a través de la dieta influye en la migración de los macrófagos en los tejidos asociados (59). En el mismo estudio, el cambio en la composición de los microorganismos produjo una disminución de algunas citoquinas pro y antiinflamatorias (59).

Incluso con todas estas precauciones la microbiota puede generar respuestas inflamatorias en el huésped. Aparte de esto, el estado más probable si un componente del microbiota genera una inflamación es el crónico, dado que, al estar en el organismo en gran densidad, el componente seguirá existiendo, aunque se elimine una pequeña proporción.

Los mecanismos que tienen las bacterias comensales para activar la inflamación son variados. Seguramente el desencadenante prioritario se encuentre en los lipopolisacáridos (LPS) que forman parte de la pared celular de las bacterias Gramnegativas. Se ha asociado una alta concentración de este componente en personas obesas y con síndromes metabólicos (60). Su mecanismo molecular es doble. En primer lugar, es capaz de unirse a receptores de tipo TLR activando vías de señalización proinflamatorias y además su disociación de la barrera intestinal permite la entrada de macrófagos a la zona que posteriormente provocarán la inflamación (61).

Otra forma de controlar la inflamación es mediante los ácidos grasos de cadena corta (SCFA) (62). Éstos son productos metabólicos que las bacterias generan por la digestión de ciertas moléculas, como la fibra, que no son capaces de digerir el huésped. Estos ácidos grasos jugarán diferentes papeles según su composición y concentración. Por ejemplo, mientras el butirato será un factor antiinflamatorio que inhibirá la síntesis de citoquinas proinflamatorias (62), el acetato, en ciertas condiciones, puede promover la generación de colesterol contribuyendo a

la aparición de obesidad o controlar el metabolismo de los lípidos ayudando a mantener un equilibrio lipídico (63).

Las citoquinas son otro de los principales factores con los cuales el sistema inmunológico controla la aparición de la inflamación, tanto positiva como negativamente. Estas moléculas interactuarán con la microbiota creando una relación de retroalimentación con muchas de ellas. Por ejemplo, se observó una correlación entre un bajo nivel de la citocina proinflamatoria TNF-alfa y un nivel abundante de la bacteria *Bifidobacterium adolescentis* (64). Otro hecho remarcable fue la asociación positiva entre grandes concentraciones de *Lactobacillus* y la IL-6, una citocina proinflamatoria relacionada con la aparición de resistencia a la insulina (65)(66).

Por último, también existen otros elementos microbianos que, de una forma más minoritaria regulan la inflamación. Entre ellos, los más importantes son los ácidos biliares secundarios que han sido metabolizados previamente por la microbiota y que podrán activar algunos receptores proinflamatorios (67) o el N-óxido de trimetilamina (TMAO) que son metabolitos de bacterias comensales que generan a partir de compuestos de la dieta, y que facilitarán la aparición de inflamación y enfermedades cardiovasculares (68).

## **1.3 Microbiota**

### **1.3.1 Funciones.**

Como se ha explicado, la microbiota juega un papel fundamental en la activación de algunas vías de inflamación, por lo que también tendrá gran influencia en otras áreas de la vida del organismo. Desde el punto de vista cuantitativo la microbiota humana, es decir, el conjunto de microorganismo que viven en las personas, representa una proporción mayor que el número total de células que tiene un adulto (69). Se calcula que un adulto promedio alberga alrededor de 100 billones de microorganismos en su cuerpo (69). Además, éstos tendrán una alta variabilidad entre ellos. En un estudio realizado entre 14 adultos se observó que se obtuvieron más de 18000 secuencias distintas del gen del RNA ribosomal 16S (70).

Otro rasgo básico para entender la importancia de la microbiota es la relativa facilidad para colonizar casi todas las partes del cuerpo humano. Los microorganismos tendrán una presencia fundamental en zonas esenciales del huésped. Aunque se pueda creer que su rol principal se encuentra en el sistema digestivo, en un estudio el 53% de especies microbianas que se consiguieron secuenciar fueron a partir del intestino (71), existen otras áreas donde su presencia es notable. Por ejemplo, la microbiota juega un papel trascendental en la boca por



su actividad proinflamatoria a la hora de activar rutas inmunológicas (72). Además, se ha acreditado su existencia en zonas como la pulmonar, vaginal o renal entre otras (52). Incluso en órganos donde a priori sería muy difícil su implantación por las duras condiciones en que se encuentra como es la piel, se han detectado hasta 1000 millones de bacterias por centímetro cuadrado (73).

Así, al encontrarse la microbiota en tantos tejidos tienen un papel trascendental en multitud de funciones. Una de las funciones primordiales son las de tipo inmunitario. Se ha reportado la utilidad de algunas especies para la lucha contra infecciones de patógenos multirresistentes mediante la transferencia de heces entre donantes (74). En otro estudio similar se relacionó la disminución de organismos probióticos por el uso de antibióticos con una reducción en la eficacia de las células T en infecciones nasales (75).

La microbiota también ha sido relacionada con el correcto funcionamiento cerebral. Se confirmó una red de comunicación intestino-cerebro en la cual los comensales tenían bastante influencia (76). Se obtuvo que la química cerebral cambiaba según la composición de la microbiota en el intestino (77). En otros informes, además de demostrar que la exposición murina a estresores, como la separación materna en las primeras etapas de vida, afectaba al desarrollo cerebral a través de cambios en la microbiota, también se informó que la administración temprana de combinaciones de microorganismos conseguía, hasta cierto punto revertir los cambios causados por los estresores (78)(79).

Al haber una influencia en el eje intestino-cerebro los microorganismos también median en el comportamiento de los huéspedes. Se informó de fallos en acciones como el nado en crías de ratón que habían sido sometidas a estrés al producirse una separación de la madre. Además, estas crías presentaban algunos rasgos tanto cerebrales como conductuales de depresión y ansiedad. Curiosamente, este efecto se apaciguaba cuando se administraba *Bifidobacterium infantis* incluso en la adultez (80).

Otro ámbito en que se sabe de la relación con la microbiota es el de las enfermedades metabólicas y como se comentó anteriormente la inflamación. En la misma investigación que la anterior también se expuso una disminución de IL-6, una citocina proinflamatoria en el momento en que se iniciaba el tratamiento con *Bifidobacterium infantis* (80). En otro estudio se informó que cambios en la microbiota a través de la dieta, provocaban un aumento en la permeabilidad de los tejidos. Este hecho permitía la interacción de determinadas moléculas y receptores proinflamatorios pudiendo provocar finalmente resistencia a la insulina (81). Adicionalmente se reconoció que tanto las bacterias grampositivas como las negativas eran

capaces de activar algunas vías de señalización proinflamatoria pudiendo ayudar a la generación de resistencia a la insulina (82).

### 1.3.2 Relación con la obesidad

Finalmente, si hay una circunstancia con la cual se asocia a la microbiota, y el objetivo de este trabajo, es la relación de los microorganismos comensales y la obesidad. Numerosos estudios han conseguido relacionar el control de microorganismos mediante antibióticos con la regulación de la cantidad de grasa corporal. Mientras que en ratones adultos la administración de floxacina y ampicilina consiguió un mayor control glucémico (83), tanto en ratones jóvenes como en niños se asoció el suministro de antibióticos con la obesidad (83)(84).

Otra evidencia que parece confirmar la relación microbiota-obesidad se encontró cuando al trasplantar restos fecales de ratones gemelos con diferente fenotipo obeso/delgado, los ratones a los que eran trasplantados adquirían el mismo fenotipo que de los cuales se trasplantaba (85). A este estudio se le suman otros que indican que incluso en animales con la misma dieta e idéntica genética pueden existir diferencia en el fenotipo de obesidad dependiendo de la composición de su microbiota (86). Aparte de estas investigaciones también se ha visto que la composición en la microbiota afecta a la cantidad de energía obtenida de los alimentos, y por lo tanto al apetito y obesidad (87). También se detectó que cambios en la dieta que reducían el nivel de masa corporal, como una dieta baja en carbohidratos, produjo una disminución de algunas especies de bacterias productoras de SCFA (88).

### 1.3.3 Géneros que influyen en la obesidad

Demostrada la relación entre la microbiota y la obesidad, falta determinar qué microorganismos son los que favorecen la obesidad y cuáles son los que previenen su aparición. Como se ha establecido anteriormente, muchas veces será el contexto en el que se encuentre la bacteria el que determinará el efecto último de la misma. Por ello, en este ámbito es normal encontrar investigaciones cuyas conclusiones se contradicen con otras. Por ejemplo, mientras que algunos estudios relacionaron un aumento en la concentración de Proteobacteria con aumentos en marcadores de enfermedades metabólicas e inflamatorias (89), otros no encontraron asociación alguna (90).

Se relacionó el aumento de microorganismos del filo *Bacteroidetes* y *Firmicutes*, por medio de dietas ricas en grasas, con un aumento en el peso corporal, producción de marcadores proinflamatorios y deficiencias metabólicas (91). Además, en otro estudio, el trasplante de microbiota de heces rico en *Bacteroidetes* y *Firmicutes* a ratones libres de gérmenes y delgados, dio como resultado el aumento de la masa corporal de estos ratones (92). Estas carencias se consiguieron remediar, o bien, cuando se administraron antibióticos que acababan con dichas bacterias (93), o cuando se introducían otros microorganismos, en este caso *Bifidobacterium pseudocatenuatum*, que desplazaban y reducían el número de *Firmicutes* y *Bacteroidetes* (94), demostrando que sí que existen microorganismos que regulan de forma beneficiosa la aparición de obesidad. Aun así, algunos reportes no han encontrado asociación alguna entre estos filos y el aumento en la masa corporal, pero sí con otros microorganismos como las *Actinobacterias* (95)(96).

Se han encontrado otras asociaciones con bacterias comensales que aumentan o disminuyen la probabilidad de desarrollar obesidad. Se informó de un aumento de especies de la familia Enterobacteriaceae en pacientes obesos (97). En cambio, otra familia de microorganismos, Christensenellaceae, se relacionó en pacientes con un buen nivel de masa corporal (98). Algunas publicaciones, aunque evidenciaron un aumento en la concentración de *Lactobacillus* en niños que sufrieron una pérdida de peso (99), también se encontraron gran cantidad de estas bacterias en niños con obesidad (100).

Diferentes géneros de microorganismos se encontraron de manera diferencial en individuos cuyo nivel de masa corporal era saludable. Entre éstos se encuentran los géneros *Erwinia*, *Succinivibrio* (101) o *Alistipes* (102). En otro estudio metagenómico japonés que intentaba analizar las diferencias en la composición de la microbiota entre individuos obesos y normo peso se hallaron muchas especies determinantes. Entre las que promovieron un aumento de peso se encontraron especies del género *Ruminococcus* (*R. obeum* y *R. bromii*), *Eubacterium ventriosum* o *Blautia hydrogenotrophica* entre otras (103). Curiosamente, en ese mismo estudio entre las especies que se encontraban significativamente diferenciadas entre individuos normo peso respecto a los obesos, se informó de la presencia de otras especies del género *Blautia* (*B. wexlerae*) entre otras como especies del género *Bacteroides* (*B. faecichinchillae* y *B. thetaiotaomicron*) o *Clostridium boltea* (103).

## **1.4 Género *Blautia***

### **1.4.1 Generalidades.**

Las especies del género *Blautia* juega un papel ambiguo en la obesidad. Cuando se habla de *Blautia* se refiere a todas las especies que se engloban dentro de un género encuadrado en el cluster *Clostridium XIVa* (104). Estas bacterias tienen ciertas características comunes. La gran mayoría son grampositivas (105), inmóviles, en forma de cocobacilo y no producen esporas, aunque hay especies que son capaces de producirlas. Metabólicamente se consideran anaerobias estrictas, catalasa negativas y obtienen su energía fermentando los nutrientes hasta producir acetato y succinato (104) principalmente, aunque también otros productos como etanol y lactato (105). Su aislamiento se consiguió a partir de restos fecales, en donde se encontraban en abundancia (106), surgiendo la hipótesis de que debían ser una parte importante de la microbiota. Es por todo esto en los últimos años ha sido objeto de investigación por su influencia metabólica en la vida del individuo.

#### 1.4.2 Estudios previos.

Por ejemplo, se obtuvo una relación significativa entre la presencia de dos especies de *Blautia* (*B. luti* y *B. wexleare*) en niños con un peso normal respecto a los obesos (107). En el mismo estudio, al compararse marcadores proinflamatorios en restos fecales de niños obesos respecto a los delgados se detectó una mayor cantidad de éstos en las heces de los niños con obesidad, demostrando la capacidad antiinflamatoria de estas bacterias y su relación con el buen metabolismo de la insulina (107). En esa misma línea, se han detectado efectos positivos en la contención del tono inflamatorio como consecuencia de la producción de SCFA como el butirato (108). En investigaciones paralelas se ha conseguido beneficios en pacientes con diabetes de tipo 2 cuando se administraban ciertas combinaciones que aumentaban las especies de *Blautia* (109), correlacionando el metabolismo de la glucosa con las especies que conforman el género *Blautia* (110). Aun así, existen otros reportes que atribuyen a la especie de *Blautia coccooides* actividad proinflamatoria (111), por lo es necesaria más investigación sobre el género.

En otros ámbitos, una investigación se informó sobre una menor mortalidad en pacientes con abundancia de *Blautia* cuando se realizaban trasplantes alogénicos (112). Además, se ha conseguido regular y aumentar la concentración de *Blautia* mediante la ingesta de extracto de salvado de trigo enriquecido en arabinoxilanos (113), hecho que permitiría el uso de probióticos concretos para aumentar los efectos positivos de *Blautia*.

Es observable como de un tiempo a esta parte se ha ido incrementando el conocimiento previo sobre este género. Tanto es así que desde el momento en el cual se acotaron las

especies y características que comprendían el género, se han ido descubriendo otros microorganismos que englobándose en *Blautia*, aportan variabilidad.

## **2. Objetivos**

El objetivo del siguiente trabajo será el ensamblaje, anotación funcional y evaluación del genoma de una especie de *Blautia* aislada a partir de las heces de un individuo metabólicamente sano y con fenotipo delgado, con el propósito de hacer una caracterización preliminar de su genoma, una clasificación taxonómica fiable a nivel de especie y determinar rasgos metabólicos de importancia para su hospedador.

## **3. Materiales y métodos**

### **3.1 Pretratamiento de muestras.**

#### **3.1.1 Calidad de muestras.**

Las secuencias se obtuvieron tras la caracterización de aislados en el laboratorio de Interacciones Microbio-Hospedador en salud metabólica del Centro Investigador Príncipe Felipe. Inicialmente se caracterizó al microorganismo, obteniendo que eran bacterias cocoides grampositivas, que se encontraban en agrupaciones de pares de células y cuyas condiciones óptimas de crecimiento eran de anaerobiosis estricta, a una temperatura de 37º C en medio Wilkins Chalgren. La secuenciación de su genoma se realizó a partir de una extracción de ADN usando el kit MasterPure Gram Positive DNA Purification (Epicentre) y secuenciando dicho ADN mediante la plataforma Illumina MiSeq generando un total de 2,8 millones de secuencias bajo una configuración de 2x300bp en formato Fastq. Paralelamente, se obtuvieron un total de aproximadamente 12mil lecturas de ADN a partir de la secuenciación del mismo material genómico mediante la tecnología Nanopore y química R9.4.1.

Una vez obtenidas las secuencias, se realiza un proceso de pretratamiento en el que se controla la calidad de las muestras obtenidas para que posteriormente, las conclusiones que se obtengan puedan ser reproducibles y tengan alto grado de fiabilidad.

Para ello, se aplicarán una serie de análisis, realizables mediante diferentes programas informáticos, que calcularán parámetros como el grado de calidad o la concentración de las bases en la secuencia y actuarán en consecuencia si algunos de estos valores no cumplen con unos estándares mínimos.

El software utilizado para el control de calidad fue fastqc (v 0.11.9) (114). La principal ventaja del programa, a parte de la presentación de los resultados en un entorno interactivo mucho más accesible, es la gran variedad de módulos que se comprueban de manera paralela. Además, la forma de cálculo de los parámetros de fastqc en la cual se evalúan una cantidad fija de secuencias lo suficientemente representativa y se extrapolan los resultados a las demás, reduce en gran medida tanto el costo como el tiempo de operación (115).

Introducidos los datos en el programa se observó resultados anómalos en 2 de los parámetros calculados. En primer lugar, se evidenció la presencia de adaptadores propios de Illumina en la zona (3') de las secuencias. Al mismo tiempo, se observó una reducción en la calidad de la secuenciación conforme las posiciones de la secuencia avanzaban. Este hecho es debido a la degradación en la calidad de secuenciación con el aumento en la longitud de la lectura.

### 3.1.2 Recorte de lecturas.

Es por ello por lo que se optó por el recorte de las secuencias mediante el software Cutadapt (v 3.7) (116). Éste es un programa capaz de hacer la poda tanto de manera automática (determinando cual es el adaptador para eliminar según unas condiciones predefinidas) o de manera manual (especificando las secuencias a eliminar o eligiendo la cantidad de lectura que se quiere conservar).

Como además del adaptador se querían eliminar las zonas de las lecturas donde se comenzaba a producir una reducción de la calidad, se eligió la opción de eliminar una cantidad fija de las secuencias obteniendo como output las primeras 90 bases de cada lectura y como línea de comandos:

```
cutadapt -u 90 -o /home/jaicaga1/nuevo/ABL231_1_90.fastq /home/jaicaga1/nuevo/ABL231_1.fastq  
cutadapt -u 90 -o /home/jaicaga1/nuevo/ABL231_2_90.fastq /home/jaicaga1/nuevo/ABL231_2.fastq
```

Finalmente, las reads modificadas después de la aplicación de Cutadapt se volvieron a analizar con fastqc y se observó que el conjunto de datos era óptimo para el posterior ensamblaje.

## 3.2 Ensamblaje.

### 3.2.1 Generalidades.

Asegurada las calidades de la muestras se procedió a la fase siguiente, el ensamblaje. El ensamblaje es un proceso en el cual se agrupan y ordenan las reads obtenidas para conseguir una secuencia consenso. Normalmente nunca se obtendrá la secuencia completa al finalizar el procedimiento, sino que por el contrario alcanzarán niveles menores formados por grupos de

reads que solapan entre ellos. Estos niveles podrán ser “*scaffolds*”, la entidad mínima del ensamblaje constituida por las lecturas o “*contigs*”, unidades mayores que se basan en agrupaciones de “*scaffolds*”.

Como pasaba en el control de calidad de las lecturas, el ensamblaje también presenta una serie de características que lo definen y que ayudan a discernir la confiabilidad del resultado. En concreto, entre todos los parámetros que se calculan destacan el grado de fragmentación de las secuencias, medible mediante el N50 (longitud de los contigs tal que usando contigs de igual o mayor tamaño produce la mitad de las bases del genoma) y L50 (el número más pequeño de contigs cuya suma de longitud constituye la mitad del tamaño del genoma), o la composición del resultado, que se evalúa mediante métricas como el porcentaje del contenido de guaninas y citocinas (GC%).

Al no saber el organismo al cual pertenecían las secuencias obtenidas, se optó por realizar un ensamblaje de *novo*, es decir, que a diferencia del mapeo no utiliza un genoma de referencia para ubicar los datos. Conforme esta premisa y para poder comparar y contrastar los resultados se utilizaron dos ensambladores que utilizaban diferentes métodos computacionales y posteriormente se analizaron mediante un software que determinaba la calidad del procedimiento.

### 3.2.2 ABySS.

En primer lugar, se utilizó el software ABySS (v. 2.3.2) (117) para realizar la primera aproximación al ensamblaje. Éste es un programa de código abierto, implementado en C++ que se especializa en ensamblajes de lectura corta de *novo*. Su algoritmo está basado en la construcción de grafos ‘*de Bruijn*’, es decir, un método que representa los solapamientos de una extensión determinada entre diferentes fragmentos de lecturas de longitud K (k-mers). En otras palabras, el funcionamiento de la operación consta de 2 fases. En la primera se generan todas las posibles subcadenas de longitud k (k-mers) que forman las lecturas y en la segunda se superponen los k-mers para solucionar problemas de ambigüedades y extender los *contigs*.

Además, para especificar más, de las distintas opciones elegibles se indicó que el máximo tamaño de k-mer que se generase debería ser de 96 bases (k = 96) y que el ensamblaje final no podría superar las 2G de bases (B=2G). Así, la secuencia de la pipeline en Linux fue:

```
abyss-pe name=/home/jaicaga1/nuevo/blautia k=96 B=2G  
in='/home/jaicaga1/nuevo/ABL231_1_90.fastq /1 /home/jaicaga1/nuevo/ABL231_2_90.fastq /2'
```

### 3.2.3 Spades.

Para el segundo ensamblaje el algoritmo elegido fue *Spades* (v. 3.13.1) (118) un ensamblador que se suele aplicar para casos concretos en los que solo se dispone de una única célula del microorganismo. La diferencia de este ensamblador con respecto a los demás radica en su técnica, que, aunque basada en los grafos de *Bruijn*, está modificada en ciertos aspectos que aumentan su precisión y disminuyen el tiempo y esfuerzo de computación. En concreto, el enfoque innovador es el llamado PDBG (Grafos Emparejados de Bruijn) que consigue calcular las desviaciones en la distancia entre todos los dos k-mers posibles de cada lectura (k-bimer) y ajustarla para conseguir el mejor resultado posible.

De todas las especificaciones posibles se escogieron las opciones de `--sc` (single cell), que indica al programa que los datos recogidos son de una única procedencia y `--careful`, que es una forma de reducción del número de errores mediante la ejecución de la herramienta integrada *MismatchCorrect* y que además es recomendable en genomas pequeños como era el caso del estudio. Con todo esto, se construyó la siguiente pipeline:

```
spades -1 /home/jaicaga1/spadesresult/ABL231_1_90.fastq -2  
/home/jaicaga1/spadesresult/ABL231_2_90.fastq --sc --careful -o  
/home/jaicaga1/spadesresult/ensamblaje.fastq
```

### 3.2.4 Quast.

Finalmente, para evaluar los ensamblajes obtenidos se utilizó *Quast* (v. 5.0.2) (119), una herramienta que compara y valora la calidad en los ensamblajes, incluso sin disponer de un genoma de referencia con el que contrastar los datos. Esta tarea la realiza principalmente con el cálculo de distintos módulos que generan una idea de lo confiable que puede ser el resultado de los ensambladores. La elección de dichos módulos dependerá de las necesidades de la investigación. En este caso se establecieron los siguientes parámetros. En primer lugar, se determinó en 10 el número máximo de subprocesos que se podían realizar de manera paralela (`-t 10`) para que se equilibrase el tiempo necesario por la computadora para llevar a cabo los procesos y aun así no produjese un gran esfuerzo computacional. Se especificó que los datos introducidos se encontraban en *scaffolds* en lugar de *contigs* (`-split-scaffolds`). Además, se introdujeron variables para el cómputo de las calidades de las agrupaciones mediante k-mers (`--k-mer-stats`). Finalmente, para poder comparar el ensamblaje obtenidos con otros realizados de la misma especie y para analizar de manera superficial la calidad del resultado. Se añadieron opciones que estimaban tanto la posición de las diferentes estructuras genómicas



como CDS, ORFs o genes (*--glimmer*), las secuencias de ADN ribosomal (*--rna-finding*) necesario para la determinación del microorganismo y la presencia de ciertos genes ortólogos en otras especies (*--conserved-genes-finding*). Con la elección de todas estas medidas, la línea de comandos fue:

Para Spades (v. 3.13.1):

```
python2 ./quast.py -o quastresult -t 10 --glimmer --rna-finding --split-scaffolds --k-mer-stats --est-insert-size 130 --conserved-genes-finding -1 /home/jaicaga1/nuevo/ABL231_1_90.fastq -2 /home/jaicaga1/nuevo/ABL231_2_90.fastq /home/jaicaga1/spadesresult/scaffolds.fasta
```

### 3.2.5 Repetición del proceso.

Cuando se hubo analizado los ensamblajes con *quast*, aunque se observó que los productos elaborados por *Spades* tenían una mejor calidad con respecto a los de *Abyss*, sobre todo en ciertos ámbitos como el de la fragmentación, no existió la suficiente disposición como para que se diesen por confiables. Es por ello, que el microorganismo se volvió a secuenciar, en este caso mediante la plataforma Nanopore (que generó lecturas mucho más largas) y se repitió el proceso, pero combinando las lecturas pasadas de Illumina con las de Nanopore. En esta ocasión, teniendo presente el anterior análisis de los ensambladores por parte de *quast*, se prefirió realizar el ensamblaje únicamente con *Spades*. La pipeline de cada una de las aplicaciones usadas quedaron:

Spades (v. 3.13.1):

```
spades -1 /home/jaicaga1/spadesresult/ABL231_1_90.fastq -2 /home/jaicaga1/spadesresult/ABL231_2_90.fastq --nanopore /home/jaicaga1/spadesresult/ABL231_ONT_R941.fastq --sc --careful -o /home/jaicaga1/spadesresult/emsamblaje.fastq
```

Para Abyss (v. 2.3.2):

```
python2 ./quast.py -o quastresult -t 10 --glimmer --rna-finding --split-scaffolds --k-mer-stats --est-insert-size 130 --conserved-genes-finding -1 /home/jaicaga1/nuevo/ABL231_1_90.fastq -2 /home/jaicaga1/nuevo/ABL231_2_90.fastq /home/jaicaga1/abysss/scaffolds.fasta
```

Para Quast (v. 5.0.2):

```
python2 ./quast.py -o quastresult -t 10 --glimmer --rna-finding --split-scaffolds --k-mer-stats --est-insert-size 130 --conserved-genes-finding -1 /home/jaicaga1/nuevo/ABL231_1_90.fastq -2 /home/jaicaga1/nuevo/ABL231_2_90.fastq --nanopore /home/jaicaga1/nuevo/ABL231_ONT_R941.fastq /home/jaicaga1/nuevo/scaffolds.fasta
```

### 3.3 Anotación estructural y funcional.

#### 3.3.1 Quast

El siguiente paso, después de ensamblar las lecturas obtenidas, fue la anotación del genoma ensamblado. Esta etapa consiste en la caracterización de los rasgos distintivos de las secuencias. Se conseguirá mediante dos formas, desde un punto de vista estructural, es decir, indicando la disposición de los distintos elementos genómicos (exones, ORFs, transposones ...) y desde un punto de vista funcional revelando la función de los genes que forman la secuencia.

Para llevar a cabo esta tarea y a su vez poder comparar resultados, se confió en tres programas informáticos especializados con diferentes algoritmos y bases de datos. Como se ha explicado en el apartado anterior quast proporciona algunos módulos que realizan estas deducciones. Es por ello, que se utilizaron ciertos apartados para confirmar la información de los demás softwares. En concreto se usó la opción *Glimmer*, que utiliza la herramienta GlimmerHMM (120) de *GeneMark* basada en modelos probabilísticos de *Markov* (GHMM) y que a partir de ciertos supuestos es capaz de generar un mapa con las regiones de las secuencias. Además, también se empleó la herramienta de código abierto *Barrnap* (121), mediante la opción *rna-finding*, para la búsqueda de ARN ribosómico a través de su propia base de datos. Finalmente, la búsqueda de ortólogos de genes en otras especies (*conserved-genes-finding*) se ejecutó con el programa *BUSCO* (122), que a partir de los datos generados por medio de los modelos ocultos de *Markov* (GHMM) y utilizando una base de datos propia (*OrthoDB*) es capaz de predecir la funcionalidad de los genes propuestos.

#### 3.3.2 Prodigal

Los dos otros programas que se gastaron fueron *Prodigal* y *Eggnog Mapper*. El primero, *Prodigal* (v. 2.6.3) (123), es una aproximación para análisis funcionales de genomas creada con la intención de ser lo más verídica posible. Así, el algoritmo del programa se encuentra en constante entrenamiento para aumentar su sensibilidad en la detección de ORFs y disminuir la cantidad de falsos positivos. Para este trabajo se utiliza una base de datos creada a partir de un centenar de genomas del *GenBank*. La línea de comando empleada para el proceso fue:

```
prodigal -d /home/jaicaga1/prediccion/ABL231.fna -f gbk -i /home/jaicaga1/spadesresult/scaffolds.fasta  
-n -o /home/jaicaga1/prediccion/ABL231.gbk -a /home/jaicaga1/prediccion/proteina.fnn
```

### 3.3.3 Egnog Mapper

Finalmente, también se dispuso de la plataforma online *Egnog Mapper* (v 2.17) (124). La principal ventaja con respecto a los demás programas es su amplia base de datos propia (*eggNOG*) con registros muy completos de los genes en cualquiera de los ámbitos y que permite al algoritmo de la plataforma buscar coincidencias significativas en colecciones precalculadas de modelos de *Markov* (HMM). Otra razón que explica su uso es que al ser online y realizar los cálculos en servidores con mayor potencia que un ordenador convencional, puede filtrar y evaluar mucha más información para conseguir mayor sensibilidad. Finalmente, la utilización de términos jerarquizados que son conocidos en el mundo de la bioinformática, como las vías *KEGG* o las categorías *COG*, facilitan en gran medida el análisis.

### 3.4 Alineación de secuencias y filogenia.

Las secuencias del RNA ribosómico obtenidas en el anterior paso se utilizaron para cotejar la especie de organismos a la que pertenecían mediante *BLAST*. *BLAST* (125) es un software gratuito del NCBI que a través de la evaluación de alineamientos de secuencias locales que se encuentran en la base de datos de la institución, es capaz de obtener las secuencias que más se asemejan a la secuencia problema. Esta labor la conseguirá mediante 3 pasos: el *seeding*, donde el algoritmo creará k-mers de la secuencia problema y la alineará contra secuencias de la base de datos; la extensión del alineamiento y la evaluación de dicha extensión con un sistema de puntuación.

Obtenida la presunta especie a la que pertenecía el genoma secuenciado, se buscó las secuencias de ribosómicas 16S y 5S de especies relacionadas en la base de datos de *GenBank* y se llevó a cabo las acciones necesarias para crear un árbol filogenético entre todas las secuencias obtenidas. Además, también se captaron genomas tanto de otras cepas de la misma especie, como de especies relacionadas para su posterior análisis.

Aunque los árboles filogenéticos fueron creados mediante el programa *Mega11* (v. 11.0.11) (126), se realizaron diversos alineamientos múltiples con diferentes algoritmos para confirmar la distribución de los árboles. En primer lugar, se utilizó la forma web de *Muscle* (127) como alineador. Este procesador se basa en la idea de la creación de sucesivas matrices de relación entre todas las secuencias que se quieren alinear y después mediante el método de

jerarquización *UPGMA* establecer arboles guía que ayuden a la generación de los alineamientos múltiples.

Además, con el propio Mega también se alinearon las secuencias mediante el algoritmo de *Muscle* que tiene integrado, pero en vez de utilizar el método *UPGMA* se usó el de *Neighbour-Joining*, que se ha estimado óptimo para árboles de tipo evolutivo (127). Este hecho es así ya que el árbol mediante esta manera se irá creando desde las secuencias con mayor relación hasta las de menor.

Con los distintos alineamientos múltiples construidos se realizó la filogenia mediante Mega11 (126). Se decidió que se efectuarían árboles de máxima verosimilitud (*Maximum Likelihood*) y como forma de comprobación se utilizó la técnica del *Bootstrap* con un número de 1000 iteraciones distintas de árboles para producir el árbol consenso.

## 4. Resultados y discusión

### 4.1 Ensamblaje.

A lo largo del presente trabajo, el principal objetivo del estudio ha sido la elaboración y posterior confirmación de un ensamblaje, a partir de las reads obtenidas, con el fin de evaluar la relación entre el microorganismo al cual pertenecía el genoma y su potencial rol en la inflamación. Para ello, como se ha expuesto anteriormente, se eligió *Spades* como una primera aproximación preliminar que se completaría con una evaluación y perfeccionamiento de los resultados mediante *Quast*.

Una vez obtenidos los posibles ensamblajes se escogieron una serie de parámetros para valorar la calidad de los resultados (Tabla 1). Para poder realizar una comparación significativa se optó por 3 de los estados más representativos del proceso. En primer lugar, se eligieron los datos provenientes del procedimiento completo (realización preliminar con *Spades* y corrección mediante *Quast*) pero esta vez sin los datos genómicos de la secuenciación por tecnología *Nanopore*, es decir únicamente con las reads proporcionadas por *Illumina*. También se escogieron los resultados obtenidos exclusivamente por *Spades* con las reads tanto de *Illumina* como de *Nanopore*. Finalmente, también se optó por los resultados procedentes del proceso completo de ensamblaje con las reads resultantes de las dos plataformas de secuenciación.

Como se puede observar, existen grandes diferencias entre los ensamblajes. Si bien es cierto que el tamaño total es parecido en todos los casos, es en el fraccionamiento de éstos donde se aprecian las principales discordancias. En primer lugar, el número de *contigs* será un gran

indicador para estimar la calidad. Cuanto menos *contigs* haya, mejor será el ensamblaje, pues más compacto será el resultado (128). Como se observa el uso de *Quast* para el perfeccionamiento es esencial, además de que se hace patente la necesidad de complementar los datos iniciales con reads de *Nanopore* para un resultado óptimo. Estos valores se correlacionarán tanto con el N50, el tamaño del *contig* más largo y el L50. El N50 es el valor de la longitud del *contig* más pequeño a partir del cual sumando las longitudes de *contigs* mayores e iguales da como resultado más del 50% de la longitud total del ensamblaje. En cambio, el L50 es la cantidad mínima de *contigs* que suman más del 50% de la longitud del ensamblaje. Así, cuanto más grandes sean los valores de N50 y la longitud del *contig* más largo, y más pequeño sea el valor de L50, el ensamblaje estará menos fragmentado, tendrá menos zonas de incógnita y por tanto se asemejará más a la realidad. Se observa que con estos datos se repite la misma situación que con el número de *contigs*. Por un lado, se reitera la necesidad de usar *Quast* para mejorar el ensamblaje resultante de *Spades* y por otro, se apunta a las limitaciones que hubiesen ocurrido si solo se hubiese usado las lecturas de *Illumina*. Es por todo ello, que se decidió que dentro del contexto del estudio el mejor ensamblaje sería el procedentes del proceso completo de ensamblaje (*Spades* y *Quast*) con las reads resultantes tanto de *Illumina* como de *Nanopore*.

*Tabla 1.* Comparación de resultados del ensamblaje en diferentes momentos del estudio

	<b>Resultados Quast (sin dataset Nanopore)</b>	<b>Resultado Spades (sin corrección de errores)</b>	<b>Resultado Quast</b>
<b>Tamaño (pb)</b>	3.772.749	3.798.197	3.836.144
<b>Nº Contings</b>	137	393	5
<b>N50</b>	198.427	72.660	1.006.803
<b>L50</b>	8	37	2
<b>Conting más largo</b>	1.045.274	248.563	1.635.515
<b>Profundidad de cobertura</b>	—	23,2x	143x
<b>GC (%)</b>	44,44	44,12	44,48

## 4.2 Identificación taxonómica.

### 4.2.1 Resultados Blast.

Una vez determinado cual es el mejor ensamblaje dentro del ensayo, se pasó a averiguar cuál era la especie a la que pertenecía el genoma ensamblado obtenido. Para este propósito se utilizó el predictor de genes ribosomales incorporado dentro del paquete de *Quast*,

obteniendo como resultado tanto las secuencias de ARN ribosomal de la subunidad 16S como de la 23S del ensamblaje. Se eligieron estos operones ribosomales puesto que son secuencias muy conservadas que permiten identificar fácilmente las especies incluso con cambios mínimos entre las bases de diferentes organismos (129).

Estas secuencias sirvieron para crear emparejamiento con otras secuencias parecidas de la base de datos del *NCBI* mediante *BLAST*. Atendiendo a la similitud en los resultados es posible especificar qué especie es la más probable que sea. Así, se obtuvo: (Tabla 2) (Tabla 3)

*Tabla 2.* Resultados de Blast al realizarlo con la secuencia 16S

<b>16S</b>			
<b>Organismo</b>	<b>Score</b>	<b>Porcentaje identidad</b>	<b>Nº Acceso (GenBank)</b>
Bacteria no identificada	2645	99,59%	GQ897153.1
Bacteria no identificada	2639	99,52%	GQ897209.1
<i>Blautia massiliensis</i>	2636	99,45%	LN890282.1
Bacteria no identificada	2619	99,24%	GQ897517.1
<i>Blautia sp.</i>	2608	99,10%	CP040518.1

*Tabla 3.* Resultados de Blast al realizarlo con la secuencia 23S

<b>23S</b>			
<b>Organismo</b>	<b>Score</b>	<b>Porcentaje identidad</b>	<b>Nº Acceso (GenBank)</b>
<i>Blautia massiliensis</i>	5137	99,79%	CP085976.1
<i>Blautia sp.</i>	5110	99,36%	CP040518.1
<i>Blautia sp.</i>	4253	93,89%	LR994469.1
<i>Blautia argi</i>	3938	91,91%	LR994468.1
<i>Blautia argi</i>	3928	91,84%	CP030280.1

Se detecta que, sin lugar a duda, como se sospechaba por la evaluación morfológica preliminar, el organismo del cual provienen las lecturas es del género *Blautia*. Atendiendo al porcentaje de identidad, y teniendo en cuenta la gran conservación de esta porción del genoma, existen grandes probabilidades que en concreto se trate de *Blautia massiliensis*, una especie del género recientemente descrita y que, según algunos informes anteriores, podría formar parte dentro del proceso de regulación de la obesidad mediante la inflamación (105).

#### 4.2.2 Comparación de genomas *Blautia massiliensis*.

Aun así, para mayor fiabilidad se realizaron una serie de comprobaciones posteriores que se describirán a continuación. En primer lugar, se efectuó una búsqueda en distintas bases de datos (*GeneBank* y *ENA*) con la finalidad de encontrar ensamblajes existentes de la misma especie y así poder compararlos con el generado en este trabajo. Los criterios de elección de los ensamblajes existentes se basaron en valores óptimos de características que indicaban un resultado no fragmentado, es decir una estimación alta de la N50 y un valor bajo del número de contigs. Teniendo en cuenta estas apreciaciones se eligieron las siguientes cepas: *B. massiliensis* GD9 (SAMEA3676640), *B. massiliensis* MSK.22.60 (SAMN14067542), *B. massiliensis* 1001254J\_160919\_F9 (SAMN15533053), y *B. massiliensis* DFI.1\_1 (SAMN24725948) y se compararon con el resultado del ensamblaje de la cepa del estudio *Blautia* spp. ABL231. (Tabla 4)

Como se puede observar el ensamblaje generado en este trabajo tiene muy buenas prestaciones en las características que determinan su calidad con respecto a los demás ensamblajes. Aunque no es el que mayor N50 tiene, sí que es mayor que la mayoría de los genomas registrados. Además, sí que es el que menos fragmentado está si atendemos únicamente a la cantidad de *contigs* formados. Estos rasgos, se deberán seguramente a la combinación de plataformas de secuenciación que forman mejores resultados.

Una vez comprobada la calidad del ensamblaje con respecto a otros, se comparan los rasgos fundamentales para asegurarse de que concuerdan entre sí y por lo tanto pueden ser considerados de la misma especie, es decir que el resultado del estudio es *Blautia massiliensis*. (Tabla 4)

Se advierte como se asemejan bastante la mayoría de las características fundamentales del ensamblaje problema con los demás. En primer lugar, el tamaño del genoma, aunque es ligeramente mayor el obtenido en la investigación, es del mismo orden de magnitud que el resto, siendo al que más se parece la cepa GD9, que es justo la que mejor calidad de ensamblaje tenía. Con lo que respecta a la cantidad de genes predichos, se encuentra una alta similitud entre todas las opciones consultada, ya que solamente en una cepa el valor de este rasgo se encuentra en un rango diferente al de 3350-3300. Posteriormente también se analizarán a grandes rasgos las categoría COG de cada uno de los ensamblajes para averiguar si los datos del estudio son coherentes con los registros existentes. Finalmente, otro parámetro que se examinó fue el porcentaje de GC en el genoma. Éste, una vez más, expresó una alta afinidad entre cepas reforzando la idea de que la especie a estudiar fue *Blautia massiliensis*.

Tabla 4. Comparación de ensamblajes de *Blautia massiliensis*

	<b>B. massiliensis GD9</b>	<b>B. massiliensis MSK.22.60</b>	<b>Blautia spp. ABL231</b>	<b>B. massiliensis 1001254J_160919_F 9</b>	<b>B. massiliensis DFI.1_1</b>
Plataforma de secuenciación	Illumina MiSeq	Illumina HiSeq	IlluminaMiSeq/Nanopore-MinION	Illumina HiSeq	Illumina MiSeq
Tamaño (Mb)	3,72	3,61	3,8	3,41	3,59
Contigs	9	51	5	47	59
N50	2.528.264	170.554	1.006.803	130.922	158.108
Genes predichos	3.346	3.311	3.321	3.171	3.337
Operones de rRNA	16	14	15	2	5
GC%	43,98	43,90	44,48	44,40	44,20

Los datos del ensamblaje de las cepas *B. massiliensis* GD9, *B. massiliensis* MSK.22.60, *B. massiliensis* 1001254J\_160919\_F9 y *B. massiliensis* DFI.1\_1 fueron obtenidos mediante información disponible en GenBank y ENA.

#### 4.2.2 Comparación filogenias

Una vez comprobada la similitud de los datos mediante la comparación entre el ensamblaje generado en el laboratorio y diferentes resultados bibliográficos, se pasó a realizar arboles filogenéticos con la intención de clarificar la relación de la cepa investigada con las demás especies del género. Es por ello, que se eligieron secuencias de ADN de especies del género *Blautia*, además de una especie no relacionada, como es el *Lactobacillus Casei*, para enraizar los árboles.

Las secuencias elegidas fueron de las subunidades ribosomales 16S (Figura 1) y 23S (Figura 2) que como se ha especificado anteriormente están altamente conservadas. Es por ello, que cualquier mínimo cambio en el ADN se podrá identificar de manera más sencilla y generará un rastro mayor para los algoritmos de creación de árboles filogenéticos. Además, normalmente este tipo de secuencias son de las más estudiadas ya que permiten la identificación de las especies, por lo que es más fácil encontrar datos de los organismos que se quieren comparar. Aun así, es importante especificar que el género *Blautia* comprende, por el momento, una veintena de especies distintas (130), por lo que en el caso de la subunidad 23S no fue posible encontrar secuencias de todas las especies en las bases de datos biológicas. Todas las secuencias que se emplearon para la construcción de los árboles filogenéticos se pueden encontrar en la base de datos de *GenBank*, por lo que en los árboles también se detallaron el número de identificación de las secuencias entre paréntesis. Con todas estas especificaciones, se generaron los siguientes dos árboles.



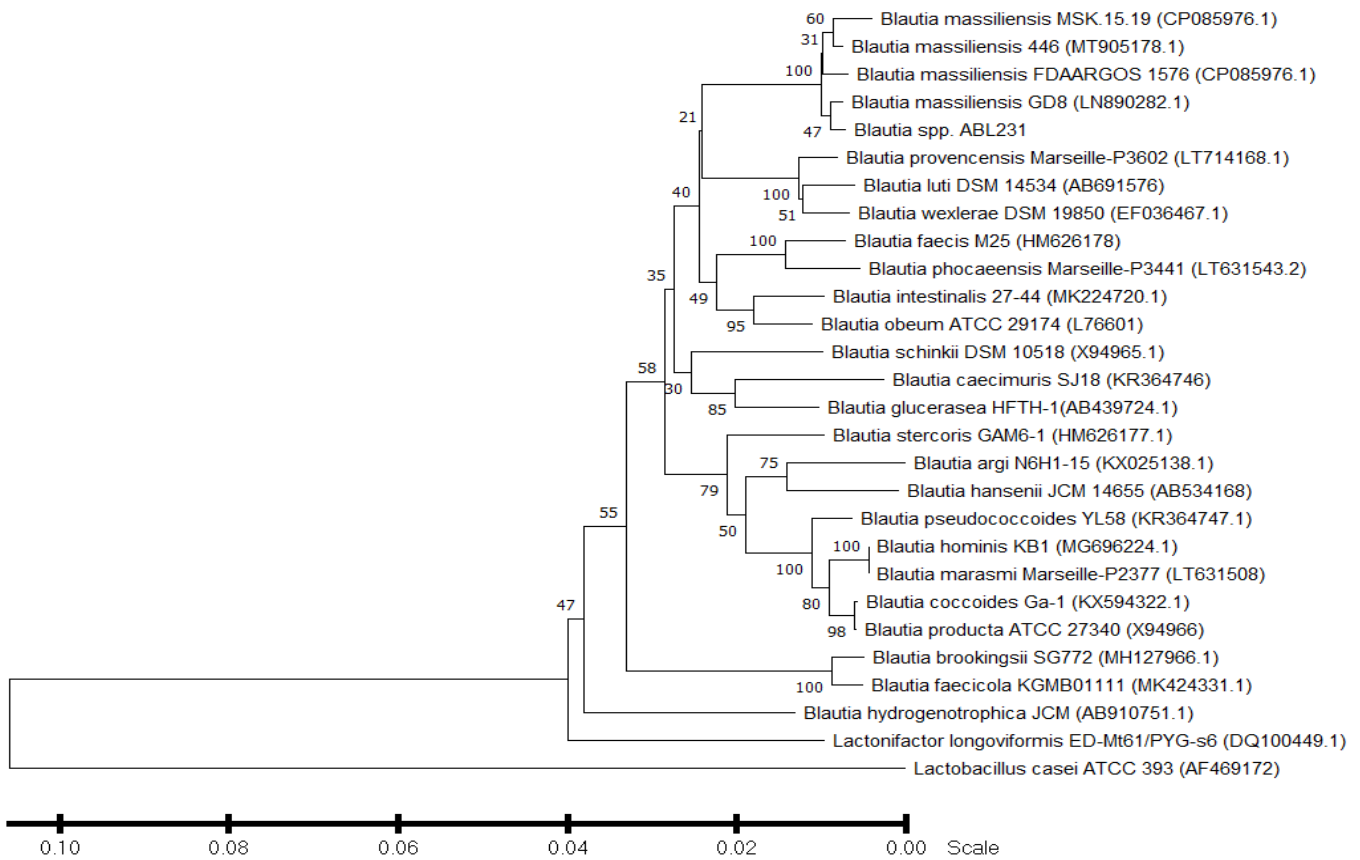


Figura 1. Filogenias de la secuencias ribosomales 16S

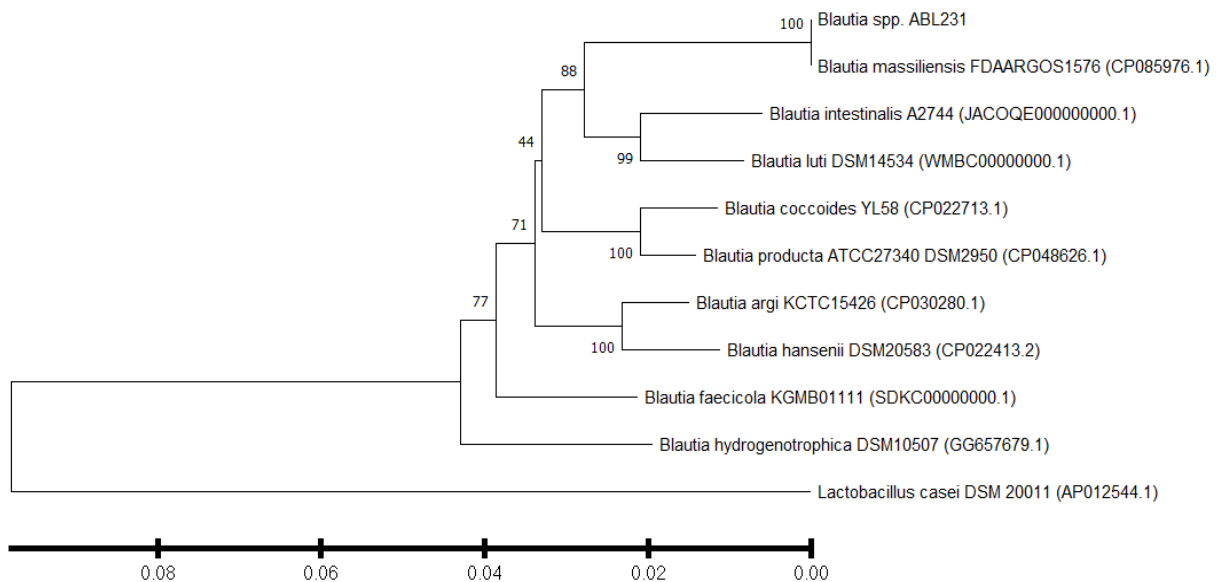


Figura 2. Filogenias de la secuencias ribosomales 23S

A grandes rasgos, sin analizar las filogenias de forma pormenorizada, se distingue en los dos árboles un clado perteneciente de manera exclusiva a cepas de la especie de *Blautia massiliensis*, donde, de manera conveniente se encuadra el organismo problema del estudio. Además, para mayor seguridad, los valores de *bootstrap* (re-muestreo aleatorio) obtenidos en ambos árboles para el clado perteneciente a *Blautia massiliensis* son de 100, es decir en todos los supuestos arboles contruidos por el algoritmo para confirmar las distribuciones de las especies, se creaba un subgrupo en que se agrupaban las especies de *Blautia massiliensis*, entre ellas el ensamblaje del trabajo.

Si se comenta más detalles de los resultados, lo primero que habrá que especificar es, que como se aprecia, los valores de *bootstrap* de la mayoría de los nodos son inferiores a 70, por lo que habrá que tener cuidado a la hora de sacar cualquier tipo de conclusión. Aun así, se puede recurrir a otros estudios bibliográficos para poder discernir si los árboles están bien contruidos y por lo tanto se confirma la especie a la que pertenece el organismo que se quiere estudiar.

En primer lugar, se observa que la especie más alejada filogenéticamente del género, en los dos casos, es la correspondiente a *Blautia hydrogenotrophica*. Si se compara con los datos bibliográficos consultados (130) se observa que esta apreciaciones concuerdan con lo publicado anteriormente. En estos mismos estudios (130) también se advierten ciertos subgrupos con alto valor de *bootstrap* y que, por lo tanto, tienen alta posibilidades de ser cercanos a nivel filogenético. Entre estos nodos, destacan los formados por dos binomios, el de *Blautia coccooides* y *Blautia producta* y el de *Blautia marasmi* y *Blautia hominis*. Además, estos nodos, normalmente están conectados tanto con las especies *Blautia hanseii* como con *Blautia argi*. Si se compara el clado anteriormente nombrado con los árboles generados, se observa que están presentes en los dos, si bien es cierto que en el árbol de las secuencias 23S de algunas de las especies nombradas no existen datos.

Finalmente, para verificar las filogenias, el análisis se centra en los nodos que se encuentran junto a la especie *Blautia massiliensis*. En la bibliografía consultada se aprecia como esta especie se encuentra estrechamente relacionada, en particular con *Blautia luti* y de forma más alejada con *Blautia wexlerae* y *Blautia provencensis*. Si el foco se centra en el árbol de 16S, ya que es el único en el que se encontraron datos de todas estas especies, se detecta una relación bastante más estrecha respecto a las demás especies, entre las 3 especies anteriormente nombradas y el clado en el que se encuadran todos los organismos pertenecientes a *Blautia massiliensis*.

Acorde a todas estas evaluaciones se puede determinar con alto grado de seguridad, que la especie que se encontró de manera diferencial en personas con un índice de masa corporal correcto es *Blautia massiliensis*. Aun así, se encontraron algunas discrepancias morfológicas cuando se comparó la evaluación que se realizó en el laboratorio de este estudio con respecto al realizado en las referencias bibliográficas encontradas.

Según el estudio que descubrió la bacteria (105) la caracterización de la especie dio como resultado bacterias ovaladas Gram-negativas, mientras que en este estudio se observaron bacterias cocoides Gram-positivas, que se encontraban en agrupaciones de pares de células. Esta discrepancia es llamativa ya que todas las especies pertenecientes al género *Blautia* que se han detectado hasta el momento se encontraron Gram-positivas. Además, cuando se comparó el ensamblaje resultado de la investigación con el que se obtuvo en el artículo científico que se referencia, se observó que eran parecidos en la gran mayoría de sus características, por lo que la explicación más razonable seguramente sea una equivocación involuntaria.

### **4.3 Resultado análisis funcional**

Para acabar, una vez determinada la especie a la que pertenecía el microorganismo problema se realizó la anotación estructural y funcional del ensamblaje generado. Para ello, se acudió a la clasificación mediante categorías COG. Estas son un tipo de categorización que se fundamenta en bases de datos de agrupamiento de genes ortólogos, es decir, cada gen se clasifica dependiendo de su función (131). Existen 26 grupos diferentes en los que se pueden clasificar a los genes dependiendo de su función, así, según las necesidades de los microorganismo tendrá más o menos cantidad de genes de una categoría o de otra. Es por ello, que analizando la composición de tipos de genes se puede averiguar el contexto y las funciones para el sistema que realiza la bacteria (131). Teniendo esto en cuenta, al analizar funcionalmente a la bacteria se obtuvo la siguiente tabla (Tabla 5).

Lo que primero llama la atención es la gran cantidad de genes cuya función se desconoce o no están registrados. Este hecho podría sesgar los resultados, ya que es posible que justo los genes que no se conocen sean los que aumenten la proporción de genes de una determinada función, pero se acepta que lo más probable es que, habiendo interpretado casi el 70% de los genes predichos, los restantes sigan el mismo patrón.

Tabla 5. Clasificación de la composición genética del ensamblaje mediante categorías COG

Categoría	Porcentaje	Nº Genes	Explicación
[A]	-	-	Procesamiento y modificación de ARN
[J]	4,5	151	Translación
[K]	9,3	310	Transcripción
[L]	7,9	261	Replicación y reparación
[D]	1,6	54	Control del ciclo celular y mitosis
[B]	-	-	Estructura y dinámica de la cromatina
[Y]	-	-	Estructura nuclear
[V]	3,2	107	Mecanismos de defensa
[T]	4,4	145	Transducción
[M]	5,1	170	Biogénesis de pared/membrana/envoltura celular
[N]	0,5	17	Motilidad celular
[Z]	0,0	1	Citoesqueleto
[W]	-	-	Estructuras extracelulares
[U]	1,8	59	Tráfico intracelular y secreción
[O]	1,9	63	Modificación post-traducciona l y funciones de chaperona
[X]	-	-	Transposones
[C]	6,1	201	Producción y conversión de energía
[G]	6,4	211	Metabolismo y transporte de carbohidratos
[E]	5,7	189	Metabolismo y transporte de aminoácidos
[F]	2,5	82	Metabolismo y transporte de nucleótidos
[H]	5,7	189	Metabolismo coenzimático
[I]	1,6	53	Metabolismo lipídico
[P]	4,3	143	Transporte inorgánico de iones y metabolismo
[Q]	0,9	31	Estructura secundaria
[R]	-	-	Predicción funcional general solamente
[S]	22,2	738	Función desconocida
-	8,3	275	

Otro hecho que sobresale sobre el resto es que en ciertos grupos de genes no se ha detectado ninguno. Este es el caso de, por ejemplo, las categorías B e Y correspondientes a genes que regulan la estructura de la cromatina y que controlan la organización nuclear respectivamente. Es normal que no existan genes de este tipo, puesto que los organismos procariontes carecen de este tipo de estructuras (132). Lo que sí puede llamar más la atención es la baja cantidad de genes correspondientes a los tipos X y Z, responsables tanto de la regulación de la función de los transposones como del citoesqueleto. En lo que respecta a los transposones, posteriormente se comparará con otras especies del mismo género para dilucidar la razón del resultado, pero sí que es llamativa la no presencia de ningún gen que regule el citoesqueleto

procariota cuando desde la última década se han ido descubriendo diferentes moléculas que sí están implicadas en esta función (133). Este hecho hace patente lo escrito anteriormente, la gran cantidad de genes cuya función se desconoce puede alterar el resultado de la discusión y conforme se vayan descubriendo nuevas funcionalidades, los resultados serán más precisos.

En las demás categorías, su abundancia varía dependiendo su utilidad de ese ámbito en su contexto. Es por ello, y como última verificación de la calidad de los datos, se comparó las predicciones de las categorías COG de los ensamblajes de *Blautia massiliensis* que se utilizaron anteriormente. (Figura 3)

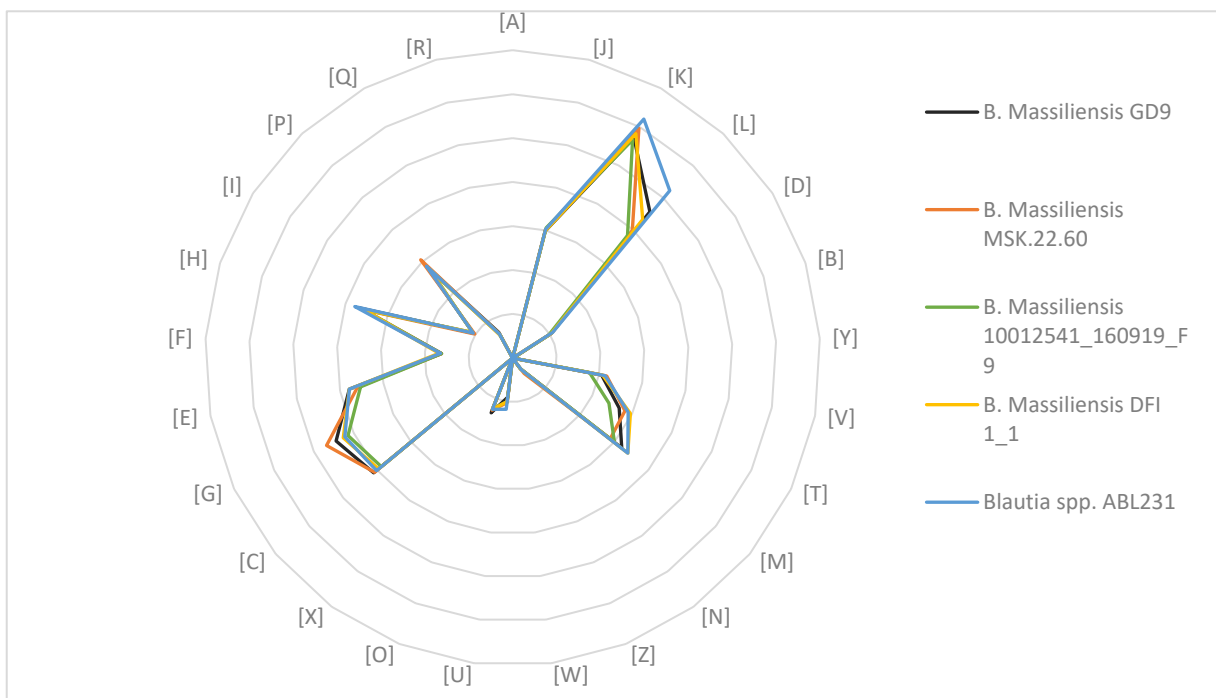


Figura 3. Comparación de los patrones génicos de diferentes cepas de *Blautia Massiliensis*.

Como se puede apreciar los patrones de los resultados en la comparativa son bastante similares con pequeñas divergencias atribuidas a las diferencias entre los ensamblajes, pero a nivel de forma, encajan casi a la perfección. Sí que destaca que en ciertos aspectos el ensamblaje del estudio presenta mayor cantidad de genes comparado con los demás, sobre todo en la regulación de la transcripción y factores relacionados con la replicación y reparación del genoma. A priori, no debería existir razón alguna para estas diferencias por lo que la hipótesis más acertada sea otra vez discrepancias en la generación de ensamblajes.

Un rasgo que sí que podría ser bastante relevante es el alto porcentaje en todos los ensamblajes de las categorías relacionadas con el metabolismo de biomoléculas y generación

de energía. Se ha demostrado que en circunstancias de inanición uno de los géneros de bacterias que más disminuyen es precisamente *Blautia* (134). Este hecho, además de la hipótesis de otros estudios que encontraron especies de *Blautia* sobrerrepresentadas en heces de humanos normopesos (107) y que detectaron relaciones entre la presencia de especies de *Blautia* y beneficios en pacientes con diabetes tipo 2 (109), se podría explicar por la alta capacidad de estos microorganismos en catabolizar grasas y azúcares.

Otra característica que podría tener importancia en la falta de desarrollo del proceso inflamatorio es la baja cantidad de genes que provienen a la bacteria de motilidad celular. Así, este factor dificultaría a la bacteria su difusión al torrente sanguíneo o lugares en los cuales se encuentran células inmunitarias que desencadenan inflamación (47).

## 5. Conclusiones

Como conclusión, en este trabajo se realizó un ensamblaje de un genoma a partir de las lecturas generadas mediante dos plataformas de secuenciación distintas, *Illumina* y *Nanopore*. Después de realizar una serie de comprobaciones se determinó que el microorganismo que se encontraba en abundancia en la microbiota de personas con un buen índice de masa corporal era *Blautia massiliensis*.

Posteriormente se realizó un estudio funcional del genoma mediante la clasificación de los genes que resultaron en la predicción estructural en categorías COG. Con la comparación de resultados de diferentes especies se llegó a la conclusión que podría existir un mecanismo común entre especies estrechamente relacionadas de *Blautia* para realizar un control del desarrollo de la respuesta inflamatoria.

Si se quiere entender la razón y la importancia de esta especie en la composición de la microbiota y sus consecuencias para la regulación de la inflamación y posterior desarrollo de ciertas enfermedades, una investigación más detallada y basada en el conocimiento de sus mecanismos metabólicos es absolutamente necesaria.

Finalmente, a vista de los resultados y la bibliografía existente, es indudable que esta especie, *Blautia massiliensis*, juega un papel muy importante dentro de la composición de la microbiota. Aun así, todos los estudios referentes a especies de *Blautia* relacionadas, e incluso a otras que no tienen nada que ver con *Blautia*, indican que el comportamiento de cualquier microorganismo está mediado por su contexto y entorno, por lo que aun siendo optimista, no hay que olvidar que se necesita más información para conocer la naturaleza real de la bacteria.

## 6. Bibliografía

- (1) Obesity and overweight. (2021, 9 junio). WHO. Recuperado 5 de mayo de 2022, de <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>
- (2) Di Angelantonio, E., Bhupathiraju, S. N., Wormser, D., Gao, P., Kaptoge, S., de Gonzalez, A. B., ... & Hu, F. B. (2016). Body-mass index and all-cause mortality: individual-participant-data meta-analysis of 239 prospective studies in four continents. *The Lancet*, 388(10046), 776-786.
- (3) Kanazawa, M., Yoshiike, N., Osaka, T., Numba, Y., Zimmet, P., & Inoue, S. (2002). Criteria and classification of obesity in Japan and Asia-Oceania. *Asia Pacific journal of clinical nutrition*, 11, S732-S737.
- (4) Zhou B; Cooperative Meta-Analysis Group Of China Obesity Task Force. [Predictive values of body mass index and waist circumference to risk factors of related diseases in Chinese adult population]. *Zhonghua Liu Xing Bing Xue Za Zhi*. 2002 Feb;23(1):5-10. Chinese. PMID: 12015100.
- (5) NCD Risk Factor Collaboration. (2016). Trends in adult body-mass index in 200 countries from 1975 to 2014: a pooled analysis of 1698 population-based measurement studies with 19· 2 million participants. *The lancet*, 387(10026), 1377-1396.
- (6) Forouzanfar, M. H., Afshin, A., Alexander, L. T., Anderson, H. R., Bhutta, Z. A., Biryukov, S., ... & Carrero, J. J. (2016). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The lancet*, 388(10053), 1659-1724.
- (7) Weihrauch-Blüher, S., & Wiegand, S. (2018). Risk factors and implications of childhood obesity. *Current obesity reports*, 7(4), 254-259.
- (8) Wang Y, Lobstein T. Worldwide trends in childhood overweight and obesity. *Int J Pediatr Obes* 2006; 1: 11–25.
- (9) Benítez-Páez, A., Gómez del Pugar, E. M., López-Almela, I., Moya-Pérez, Á., Codoñer-Franch, P., & Sanz, Y. (2020). Depletion of Blautia species in the microbiota of obese children relates to intestinal inflammation and metabolic phenotype worsening. *Msystems*, 5(2), e00857-19.
- (10) Kim J, Lee I, Lim S. 2017. Overweight or obesity in children aged 0 to 6 and the risk of adult metabolic syndrome: a systematic review and meta-analysis. *J Clin Nurs* 26:3869–3880.
- (11) Bjørge, T., Engeland, A., Tverdal, A., & Smith, G. D. (2008). Body mass index in adolescence in relation to cause-specific mortality: a follow-up of 230,000 Norwegian adolescents. *American journal of epidemiology*, 168(1), 30-37.
- (12) Obesity. (2021, 9 junio). WHO. Recuperado 5 de mayo de 2022, de <https://www.who.int/news-room/facts-in-pictures/detail/6-facts-on-obesity>
- (13) Daniels, S. R. (2009). Complications of obesity in children and adolescents. *International journal of obesity*, 33(1), S60-S65.
- (14) Gauffin Cano, P., Santacruz, A., Moya, Á., & Sanz, Y. (2012). *Bacteroides uniformis* CECT 7771 ameliorates metabolic and immunological dysfunction in mice with high-fat-diet induced obesity.

- (15) Milner, J. J., & Beck, M. A. (2012). The impact of obesity on the immune response to infection. *Proceedings of the Nutrition Society*, 71(2), 298-306.
- (16) Painter, S. D., Ovsyannikova, I. G., & Poland, G. A. (2015). The weight of obesity on the human immune response to vaccination. *Vaccine*, 33(36), 4422-4429.
- (17) Sanz, Y., & Moya-Pérez, A. (2014). Microbiota, inflammation and obesity. *Microbial endocrinology: The microbiota-gut-brain axis in health and disease*, 291-317.
- (18) Calcaterra, V., Klersy, C., Muratori, T., Telli, S., Caramagna, C., Scaglia, F., ... & Larizza, D. (2008). Prevalence of metabolic syndrome (MS) in children and adolescents with varying degrees of obesity. *Clinical endocrinology*, 68(6), 868-872.
- (19) Bjerregaard, L. G., Jensen, B. W., Ängquist, L., Osler, M., Sørensen, T. I., & Baker, J. L. (2018). Change in overweight from childhood to early adulthood and risk of type 2 diabetes. *New England Journal of Medicine*.
- (20) GBD 2015 Obesity Collaborators. (2017). Health effects of overweight and obesity in 195 countries over 25 years. *New England Journal of Medicine*, 377(1), 13-27.
- (21) Gilliland, F. D., Berhane, K., Islam, T., McConnell, R., Gauderman, W. J., Gilliland, S. S., ... & Peters, J. M. (2003). Obesity and the risk of newly diagnosed asthma in school-age children. *American journal of epidemiology*, 158(5), 406-415.
- (22) Ding, W., & Mak, R. H. (2015). Early markers of obesity-related renal injury in childhood. *Pediatric nephrology*, 30(1), 1-4.
- (23) De Pergola, G., & Silvestris, F. (2013). Obesity as a major risk factor for cancer. *Journal of obesity*, 2013.
- (24) Taylor, E. D., Theim, K. R., Mirch, M. C., Ghorbani, S., Tanofsky-Kraff, M., Adler-Wailes, D. C., ... & Yanovski, J. A. (2006). Orthopedic complications of overweight in children and adolescents. *Pediatrics*, 117(6), 2167-2174.
- (25) Rosenfield, R. L., Lipton, R. B., & Drum, M. L. (2009). Thelarche, pubarche, and menarche attainment in children with normal and elevated body mass index. *Pediatrics*, 123(1), 84-88.
- (26) Pizzi, M. A., & Vroman, K. (2013). Childhood obesity: effects on children's participation, mental health, and psychosocial development. *Occupational therapy in health care*, 27(2), 99-112.
- (27) Bjerregaard, L. G., Jensen, B. W., Ängquist, L., Osler, M., Sørensen, T. I., & Baker, J. L. (2018). Change in overweight from childhood to early adulthood and risk of type 2 diabetes. *New England Journal of Medicine*.
- (28) Neel, J. V. (1962). Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"? *American journal of human genetics*, 14(4), 353.
- (29) Han, J. C., Lawlor, D. A., & Kimm, S. Y. (2010). Childhood obesity. *The lancet*, 375(9727), 1737-1748.
- (30) Montague, C. T., Farooqi, I. S., Whitehead, J. P., Soos, M. A., Rau, H., Wareham, N. J., ... & O'Rahilly, S. (1997). Congenital leptin deficiency is associated with severe early-onset obesity in humans. *Nature*, 387(6636), 903-908.
- (31) Willer, C. J., Speliotes, E. K., Loos, R. J., Li, S., Lindgren, C. M., Heid, I. M., ... & Wellcome Trust Case Control Consortium. (2009). Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nature genetics*, 41(1), 25.
- (32) Weaver, J. U. (2008). Classical endocrine diseases causing obesity. *Obesity and Metabolism*, 36, 212-228.
- (33) Campion, J., Milagro, F. I., & Martinez, J. A. (2009). Individuality and epigenetics in obesity. *Obesity reviews*, 10(4), 383-392.



- (34) Kral, J. G., Biron, S., Simard, S., Hould, F. S., Lebel, S., Marceau, S., & Marceau, P. (2006). Large maternal weight loss from obesity surgery prevents transmission of obesity to children who were followed for 2 to 18 years. *Pediatrics*, *118*(6), e1644-e1649.
- (35) Dabelea, D. (2007). The predisposition to obesity and diabetes in offspring of diabetic mothers. *Diabetes care*, *30*(Supplement\_2), S169-S174.
- (36) Marchi, J., Berg, M., Dencker, A., Olander, E. K., & Begley, C. (2015). Risks associated with obesity in pregnancy, for the mother and baby: a systematic review of reviews. *Obesity Reviews*, *16*(8), 621-638.
- (37) Iguacel, I., Michels, N., Ahrens, W., Bammann, K., Eiben, G., Fernández-Alvira, J. M., ... & IDEFICS consortium. (2018). Prospective associations between socioeconomically disadvantaged groups and metabolic syndrome risk in European children. Results from the IDEFICS study. *International Journal of Cardiology*, *272*, 333-340.
- (38) Cespedes Feliciano, E. M., Quante, M., Rifas-Shiman, S. L., Redline, S., Oken, E., & Taveras, E. M. (2018). Objective sleep characteristics and cardiometabolic health in young adolescents. *Pediatrics*, *142*(1).
- (39) Nead, K. G., Halterman, J. S., Kaczorowski, J. M., Auinger, P., & Weitzman, M. (2004). Overweight children and adolescents: a risk group for iron deficiency. *Pediatrics*, *114*(1), 104-108.
- (40) Freedman, D. S., Khan, L. K., Serdula, M. K., Ogden, C. L., & Dietz, W. H. (2006). Racial and ethnic differences in secular trends for childhood BMI, weight, and height. *Obesity*, *14*(2), 301-308.
- (41) Majno, G., & Joris, I. (2004). *Cells, tissues, and disease: principles of general pathology*. Oxford University Press.
- (42) Pahwa, R., Goyal, A., Bansal, P., & Jialal, I. (2018). Chronic inflammation.
- (43) Benítez-Páez, A., Gómez del Pugar, E. M., López-Almela, I., Moya-Pérez, Á., Codoñer-Franch, P., & Sanz, Y. (2020). Depletion of *Blautia* species in the microbiota of obese children relates to intestinal inflammation and metabolic phenotype worsening. *Msystems*, *5*(2), e00857-19.
- (44) Nathan, C. (2006). Neutrophils and immunity: challenges and opportunities. *Nature reviews immunology*, *6*(3), 173-182.
- (45) Medzhitov, R. (2008). Origin and physiological roles of inflammation. *Nature*, *454*(7203), 428-435.
- (46) Pahwa, R., Goyal, A., Bansal, P., & Jialal, I. (2018). Chronic inflammation.
- (47) Shi, H., Kokoeva, M. V., Inouye, K., Tzamelis, I., Yin, H., & Flier, J. S. (2006). TLR4 links innate immunity and fatty acid-induced insulin resistance. *The Journal of clinical investigation*, *116*(11), 3015-3025.
- (48) Osborn, O., & Olefsky, J. M. (2012). The cellular and signaling networks linking the immune system and metabolism in disease. *Nature medicine*, *18*(3), 363-374.
- (49) Lumeng, C. N., Bodzin, J. L., & Saltiel, A. R. (2007). Obesity induces a phenotypic switch in adipose tissue macrophage polarization. *The Journal of clinical investigation*, *117*(1), 175-184.
- (50) Milanski, M., Arruda, A. P., Coope, A., Ignacio-Souza, L. M., Nunez, C. E., Roman, E. A., ... & Velloso, L. A. (2012). Inhibition of hypothalamic inflammation reverses diet-induced insulin resistance in the liver. *Diabetes*, *61*(6), 1455-1462.
- (51) Belkaid, Y., & Hand, T. W. (2014). Role of the microbiota in immunity and inflammation. *Cell*, *157*(1), 121-141.
- (52) Dominguez-Bello, M. G., Costello, E. K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., & Knight, R. (2010). Delivery mode shapes the acquisition and structure of the

- initial microbiota across multiple body habitats in newborns. *Proceedings of the National Academy of Sciences*, 107(26), 11971-11975.
- (53) Elahi, S., Ertelt, J. M., Kinder, J. M., Jiang, T. T., Zhang, X., Xin, L., ... & Way, S. S. (2013). Immunosuppressive CD71+ erythroid cells compromise neonatal host defence against infection. *Nature*, 504(7478), 158-162.
- (54) Kollmann, T. R., Levy, O., Montgomery, R. R., & Goriely, S. (2012). Innate immune function by Toll-like receptors: distinct responses in newborns and the elderly. *Immunity*, 37(5), 771-783.
- (55) Alenghat, T., Osborne, L. C., Saenz, S. A., Kobuley, D., Ziegler, C. G., Mullican, S. E., ... & Artis, D. (2013). Histone deacetylase 3 coordinates commensal-bacteria-dependent intestinal homeostasis. *Nature*, 504(7478), 153-157.
- (56) Hooper, L. V., & Macpherson, A. J. (2010). Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nature Reviews Immunology*, 10(3), 159-169.
- (57) Kelsall, B. (2008). Recent progress in understanding the phenotype and function of intestinal dendritic cells and macrophages. *Mucosal immunology*, 1(6), 460-469.
- (58) Cani, P. D., Possemiers, S., Van de Wiele, T., Guiot, Y., Everard, A., Rottier, O., ... & Delzenne, N. M. (2009). Changes in gut microbiota control inflammation in obese mice through a mechanism involving GLP-2-driven improvement of gut permeability. *Gut*, 58(8), 1091-1103.
- (59) Davis, C. D. (2016). The gut microbiome and its role in obesity. *Nutrition today*, 51(4), 167.
- (60) Ghosh, S. S., Wang, J., Yannie, P. J., & Ghosh, S. (2020). Intestinal barrier dysfunction, LPS translocation, and disease development. *Journal of the Endocrine Society*, 4(2), bvz039.
- (61) Segain, J. P., De La Blétière, D. R., Bourreille, A., Leray, V., Gervois, N., Rosales, C., ... & Galmiche, J. P. (2000). Butyrate inhibits inflammatory responses through NFκB inhibition: implications for Crohn's disease. *Gut*, 47(3), 397-403.
- (62) Schwartz, A., Taras, D., Schäfer, K., Beijer, S., Bos, N. A., Donus, C., & Hardt, P. D. (2010). Microbiota and SCFA in lean and overweight healthy subjects. *Obesity*, 18(1), 190-195.
- (63) González Hernández, M. A., Canfora, E. E., Jocken, J. W., & Blaak, E. E. (2019). The short-chain fatty acid acetate in body weight control and insulin sensitivity. *Nutrients*, 11(8), 1943.
- (64) Pradhan, A. D., Manson, J. E., Rifai, N., Buring, J. E., & Ridker, P. M. (2001). C-reactive protein, interleukin 6, and risk of developing type 2 diabetes mellitus. *Jama*, 286(3), 327-334.
- (65) Cooper, D., Kim, E. B., Marco, M., Rust, B., Welch, L., Horn, W., ... & Keim, N. (2016). Relationship between human gut microbiota and interleukin 6 levels in overweight and obese adults. *The FASEB Journal*, 30, 146-4.
- (66) Just, S., Mondot, S., Ecker, J., Wegner, K., Rath, E., Gau, L., ... & Clavel, T. (2018). The gut microbiota drives the impact of bile acids and fat source in diet on mouse metabolism. *Microbiome*, 6(1), 1-18.
- (67) Fennema, D., Phillips, I. R., & Shephard, E. A. (2016). Trimethylamine and trimethylamine N-oxide, a flavin-containing monooxygenase 3 (FMO3)-mediated host-microbiome metabolic axis implicated in health and disease. *Drug Metabolism and Disposition*, 44(11), 1839-1850.

- (68) 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.
- (69) Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Microbial ecology: Human gut microbes associated with obesity. *Nature* *444*, 1022–1023 (2006).
- (70) Hugon, P., Dufour, J. C., Colson, P., Fournier, P. E., Sallah, K., & Raoult, D. (2015). A comprehensive repertoire of prokaryotic species identified in human beings. *The Lancet Infectious Diseases*, *15*(10), 1211-1219.
- (71) Dixon, D. R., Reife, R. A., Cebra, J. J., & Darveau, R. P. (2004). Commensal bacteria influence innate status within gingival tissues: a pilot study. *Journal of periodontology*, *75*(11), 1486-1492.
- (72) Grice, E. A., Kong, H. H., Conlan, S., Deming, C. B., Davis, J., Young, A. C., ... & Segre, J. A. (2009). Topographical and temporal diversity of the human skin microbiome. *science*, *324*(5931), 1190-1192.
- (73) Van Nood, E., Vrieze, A., Nieuwdorp, M., Fuentes, S., Zoetendal, E. G., de Vos, W. M., ... & Keller, J. J. (2013). Duodenal infusion of donor feces for recurrent *Clostridium difficile*. *New England Journal of Medicine*, *368*(5), 407-415.
- (74) Ichinohe, T., Pang, I. K., Kumamoto, Y., Peaper, D. R., Ho, J. H., Murray, T. S., & Iwasaki, A. (2011). Microbiota regulates immune defense against respiratory tract influenza A virus infection. *Proceedings of the National Academy of Sciences*, *108*(13), 5354-5359.
- (75) Collins, S. M., & Bercik, P. (2009). The relationship between intestinal microbiota and the central nervous system in normal gastrointestinal function and disease. *Gastroenterology*, *136*(6), 2003-2014.
- (76) Nishino, R., Mikami, K., Takahashi, H., Tomonaga, S., Furuse, M., Hiramoto, T., ... & Sudo, N. (2013). Commensal microbiota modulate murine behaviors in a strictly contamination-free environment confirmed by culture-based methods. *Neurogastroenterology & Motility*, *25*(6), 521-e371.
- (77) Distrutti, E., Cipriani, S., Mencarelli, A., Renga, B., & Fiorucci, S. (2013). Probiotics VSL# 3 protect against development of visceral pain in murine model of irritable bowel syndrome. *PLoS one*, *8*(5), e63893.
- (78) Moya-Perez, A., Perez-Villalba, A., Benitez-Paez, A., Campillo, I., & Sanz, Y. (2017). *Bifidobacterium* CECT 7765 modulates early stress-induced immune, neuroendocrine and behavioral alterations in mice. *Brain, behavior, and immunity*, *65*, 43-56.
- (79) Desbonnet, L., Garrett, L., Clarke, G., Kiely, B., Cryan, J. F., & Dinan, T. (2010). Effects of the probiotic *Bifidobacterium infantis* in the maternal separation model of depression. *Neuroscience*, *170*(4), 1179-1188.
- (80) Cani, P. D., Bibiloni, R., Knauf, C., Waget, A., Neyrinck, A. M., Delzenne, N. M., & Burcelin, R. (2008). Changes in gut microbiota control metabolic endotoxemia-induced inflammation in high-fat diet–induced obesity and diabetes in mice. *Diabetes*, *57*(6), 1470-1481.
- (81) Laflamme, N., Echchannaoui, H., Landmann, R., & Rivest, S. (2003). Cooperation between toll-like receptor 2 and 4 in the brain of mice challenged with cell wall components derived from gram-negative and gram-positive bacteria. *European journal of immunology*, *33*(4), 1127-1138.
- (82) Membrez, M., Blancher, F., Jaquet, M., Bibiloni, R., Cani, P. D., Burcelin, R. G., ... & Chou, C. J. (2008). Gut microbiota modulation with norfloxacin and ampicillin enhances glucose tolerance in mice. *The FASEB Journal*, *22*(7), 2416-2426.

- (83) Bailey, L. C., Forrest, C. B., Zhang, P., Richards, T. M., Livshits, A., & DeRusso, P. A. (2014). Association of antibiotics in infancy with early childhood obesity. *JAMA pediatrics*, *168*(11), 1063-1069.
- (84) Ridaura, V. K., Faith, J. J., Rey, F. E., Cheng, J., Duncan, A. E., Kau, A. L., ... & Gordon, J. I. (2013). Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science*, *341*(6150), 1241-1244.
- (85) Fleissner, C. K., Huebel, N., Abd El-Bary, M. M., Loh, G., Klaus, S., & Blaut, M. (2010). Absence of intestinal microbiota does not protect mice from diet-induced obesity. *British Journal of Nutrition*, *104*(6), 919-929.
- (86) Jumpertz, R., Le, D. S., Turnbaugh, P. J., Trinidad, C., Bogardus, C., Gordon, J. I., & Krakoff, J. (2011). Energy-balance studies reveal associations between gut microbes, caloric load, and nutrient absorption in humans. *The American journal of clinical nutrition*, *94*(1), 58-65.
- (87) Duncan, S. H., Belongue, A., Holtrop, G., Johnstone, A. M., Flint, H. J., & Lopley, G. E. (2007). Reduced dietary intake of carbohydrates by obese subjects results in decreased concentrations of butyrate and butyrate-producing bacteria in feces. *Applied and environmental microbiology*, *73*(4), 1073-1078.
- (88) Verdum, F. J., Fuentes, S., de Jonge, C., Zoetendal, E. G., Erbil, R., Greve, J. W., ... & Rensen, S. S. (2013). Human intestinal microbiota composition is associated with local and systemic inflammation in obesity. *Obesity*, *21*(12), E607-E615.
- (89) Sanz, Y., Rastmanesh, R., & Agostonic, C. (2013). Understanding the role of gut microbes and probiotics in obesity: how far are we?. *Pharmacological Research*, *69*(1), 144-155.
- (90) Kim, K. A., Gu, W., Lee, I. A., Joh, E. H., & Kim, D. H. (2012). High fat diet-induced gut microbiota exacerbates inflammation and obesity in mice via the TLR4 signaling pathway.
- (91) Turnbaugh, P. J., Ley, R. E., Mahowald, M. A., Magrini, V., Mardis, E. R., & Gordon, J. I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. *nature*, *444*(7122), 1027-1031.
- (92) Murphy, E. F., Cotter, P. D., Hogan, A., O'Sullivan, O., Joyce, A., Fouhy, F., ... & Shanahan, F. (2013). Divergent metabolic outcomes arising from targeted manipulation of the gut microbiota in diet-induced obesity. *Gut*, *62*(2), 220-226.
- (93) Cano, P. G., Santacruz, A., Trejo, F. M., & Sanz, Y. (2013). Bifidobacterium CECT 7765 improves metabolic and immunological alterations associated with obesity in high-fat diet-fed mice. *Obesity*, *21*(11), 2310-2321.
- (94) 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.
- (95) Walters, W. A., Xu, Z., & Knight, R. (2014). Meta-analyses of human gut microbes associated with obesity and IBD. *FEBS letters*, *588*(22), 4223-4233.
- (96) Fei, N., & Zhao, L. (2013). An opportunistic pathogen isolated from the gut of an obese human causes obesity in germfree mice. *The ISME journal*, *7*(4), 880-884.
- (97) Goodrich, J. K., Waters, J. L., Poole, A. C., Sutter, J. L., Koren, O., Blekhnman, R., ... & Ley, R. E. (2014). Human genetics shape the gut microbiome. *Cell*, *159*(4), 789-799.
- (98) Santacruz, A., Marcos, A., Wärnberg, J., Martí, A., Martín-Matillas, M., Campoy, C., ... & EVASYON Study Group. (2009). Interplay between weight loss and gut microbiota composition in overweight adolescents. *Obesity*, *17*(10), 1906-1915.

- (99) Ignacio, A., Fernandes, M. R., Rodrigues, V. A. A., Groppo, F. C., Cardoso, A. L., Avila-Campos, M. J., & Nakano, V. (2016). Correlation between body mass index and faecal microbiota from children. *Clinical microbiology and infection*, *22*(3), 258-e1.
- (100) Murugesan, S., Ulloa-Martínez, M., Martínez-Rojano, H., Galván-Rodríguez, F. M., Miranda-Brito, C., Romano, M. C., ... & García-Mena, J. J. E. J. O. C. M. (2015). Study of the diversity and short-chain fatty acids production by the bacterial community in overweight and obese Mexican children. *European Journal of Clinical Microbiology & Infectious Diseases*, *34*(7), 1337-1346.
- (101) Verdam, F. J., Fuentes, S., de Jonge, C., Zoetendal, E. G., Erbil, R., Greve, J. W., ... & Rensen, S. S. (2013). Human intestinal microbiota composition is associated with local and systemic inflammation in obesity. *Obesity*, *21*(12), E607-E615.
- (102) Kasai, C., Sugimoto, K., Moritani, I., Tanaka, J., Oya, Y., Inoue, H., ... & Takase, K. (2015). Comparison of the gut microbiota composition between obese and non-obese individuals in a Japanese population, as analyzed by terminal restriction fragment length polymorphism and next-generation sequencing. *BMC gastroenterology*, *15*(1), 1-10.
- (103) Liu, C., Finegold, S. M., Song, Y., & Lawson, P. A. (2008). Reclassification of *Clostridium coccoides*, *Ruminococcus hansenii*, *Ruminococcus hydrogenotrophicus*, *Ruminococcus luti*, *Ruminococcus productus* and *Ruminococcus schinkii* as *Blautia coccoides* gen. nov., comb. nov., *Blautia hansenii* comb. nov., *Blautia hydrogenotrophica* comb. nov., *Blautia luti* comb. nov., *Blautia producta* comb. nov., *Blautia schinkii* comb. nov. and description of *Blautia wexlerae* sp. nov., isolated from human faeces. *International Journal of Systematic and Evolutionary Microbiology*, *58*(8), 1896-1902.
- (104) Durand, G. A., Pham, T., Ndongo, S., Traore, S. I., Dubourg, G., Lagier, J. C., ... & Million, M. (2017). *Blautia massiliensis* sp. nov., isolated from a fresh human fecal sample and emended description of the genus *Blautia*. *Anaerobe*, *43*, 47-55.
- (105) Touyama, M., Jin, J. S., Kibe, R., Hayashi, H., & Benno, Y. (2015). Quantification of *Blautia wexlerae* and *Blautia luti* in human faeces by real-time PCR using specific primers. *Beneficial Microbes*, *6*(4), 583-590.
- (106) Benítez-Páez, A., Gómez del Pugar, E. M., López-Almela, I., Moya-Pérez, Á., Codoñer-Franch, P., & Sanz, Y. (2020). Depletion of *Blautia* species in the microbiota of obese children relates to intestinal inflammation and metabolic phenotype worsening. *Msystems*, *5*(2), e00857-19.
- (107) Zhang, Q., Xiao, X., Li, M., Yu, M., Ping, F., Zheng, J., ... & Wang, X. (2017). Vildagliptin increases butyrate-producing bacteria in the gut of diabetic rats. *PLoS one*, *12*(10), e0184735.
- (108) Tong, X., Xu, J., Lian, F., Yu, X., Zhao, Y., Xu, L., ... & Zhao, L. (2018). Structural alteration of gut microbiota during the amelioration of human type 2 diabetes with hyperlipidemia by metformin and a traditional Chinese herbal formula: a multicenter, randomized, open label clinical trial. *MBio*, *9*(3), e02392-17.
- (109) Egshatyan, L., Kashtanova, D., Popenko, A., Tkacheva, O., Tyakht, A., Alexeev, D., ... & Boytsov, S. (2016). Gut microbiota and diet in patients with different glucose tolerance. *Endocrine Connections*, *5*(1), 1-9.
- (110) Tuovinen, E., Keto, J., Nikkilä, J., Mättö, J., & Lähteenmäki, K. (2013). Cytokine response of human mononuclear cells induced by intestinal *Clostridium* species. *Anaerobe*, *19*, 70-76.

- (111) Jenq, R. R., Taur, Y., Devlin, S. M., Ponce, D. M., Goldberg, J. D., Ahr, K. F., ... & van den Brink, M. R. (2015). Intestinal *Blautia* is associated with reduced death from graft-versus-host disease. *Biology of Blood and Marrow Transplantation*, *21*(8), 1373-1383.
- (112) Kjølbaek, L., Benítez-Páez, A., Del Pulgar, E. M. G., Brahe, L. K., Liebisch, G., Matysik, S., ... & Sanz, Y. (2020). Arabinoxylan oligosaccharides and polyunsaturated fatty acid effects on gut microbiota and metabolic markers in overweight individuals with signs of metabolic syndrome: a randomized cross-over trial. *Clinical Nutrition*, *39*(1), 67-79.
- (113) Andrew, S. (s. f.). Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. [www.bioinformatics.babraham.ac.uk](http://www.bioinformatics.babraham.ac.uk). Recuperado 3 de mayo de 2022, de <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- (114) Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884-i890.
- (115) Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal*, *17*(1), 10-12.
- (116) Simpson, J. T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J., & Birol, I. (2009). ABySS: a parallel assembler for short read sequence data. *Genome research*, *19*(6), 1117-1123.
- (117) Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., ... & Pevzner, P. A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology*, *19*(5), 455-477.
- (118) Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, *29*(8), 1072-1075.
- (119) Majoros, W. H., Pertea, M., & Salzberg, S. L. (2004). TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics*, *20*(16), 2878-2879.
- (120) Seemann, T., & Booth, T. (2018). Barrnap: BAsic Rapid Ribosomal RNA Predictor.
- (121) Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210-3212.
- (122) Hyatt, D., Chen, G. L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics*, *11*(1), 1-11.
- (123) Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., Von Mering, C., & Bork, P. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Molecular biology and evolution*, *34*(8), 2115-2122.
- (124) Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), 2460-2461.
- (125) Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Molecular biology and evolution*, *38*(7), 3022-3027.
- (126) Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC bioinformatics*, *5*(1), 1-19.
- (127) Goodwin, S., Gurtowski, J., Ethe-Sayers, S., Deshpande, P., Schatz, M. C., & McCombie, W. R. (2015). Oxford Nanopore sequencing, hybrid error correction, and de novo assembly of a eukaryotic genome. *Genome research*, *25*(11), 1750-1756.

- (128) Patel, J. B. (2001). 16S rRNA gene sequencing for bacterial pathogen identification in the clinical laboratory. *Molecular diagnosis*, 6(4), 313-321.
- (129) Liu, X., Mao, B., Gu, J., Wu, J., Cui, S., Wang, G., ... & Chen, W. (2021). Blautia—a new functional genus with potential probiotic properties?. *Gut microbes*, 13(1), 1875796.
- (130) Tatusov, R. L., Galperin, M. Y., Natale, D. A., & Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic acids research*, 28(1), 33-36.
- (131) Selker, E. U. (1990). DNA methylation and chromatin structure: a view from below. *Trends in biochemical sciences*, 15(3), 103-107.
- (132) Erickson, H. P. (2017). The discovery of the prokaryotic cytoskeleton: 25th anniversary. *Molecular Biology of the Cell*, 28(3), 357-358.
- (133) Million, M., Diallo, A., & Raoult, D. (2017). Gut microbiota and malnutrition. *Microbial pathogenesis*, 106, 127-138.
- (134) Chapat, L., Chemin, K., Dubois, B., Bourdet-Sicard, R., & Kaiserlian, D. (2004). Lactobacillus casei reduces CD8+ T cell-mediated skin inflammation. *European journal of immunology*, 34(9), 2520-2528.