



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE  
ESCUELA DE INGENIERÍA

**MODELAMIENTO A ESCALA GENÓMICA  
DE LAS INTERACCIONES METABÓLICAS  
ENTRE MICROORGANISMOS  
COMENSALES Y PRODUCTORES DE  
BUTIRATO**

**GUILLERMO ORELLANA BECERRA**

Tesis para optar al grado de  
Magíster en Ciencias de la Ingeniería

Profesor Supervisor:  
DANIEL GARRIDO CORTÉS

Santiago de Chile, Enero, 2022

© 2022, GUILLERMO ORELLANA BECERRA



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE  
ESCUELA DE INGENIERÍA

**MODELAMIENTO A ESCALA GENÓMICA  
DE LAS INTERACCIONES METABÓLICAS  
ENTRE MICROORGANISMOS  
COMENSALES Y PRODUCTORES DE  
BUTIRATO**

**GUILLERMO ORELLANA BECERRA**

Miembros del Comité:

DANIEL GARRIDO CORTÉS 

PEDRO SAA HIGUERA 

CAROLINA SÁNCHEZ DOÑAS 

SERGIO MATURANA VALDERRAMA 

Tesis para optar al grado de  
Magíster en Ciencias de la Ingeniería

Santiago de Chile, Enero, 2022

© 2022, GUILLERMO ORELLANA BECERRA

*A mi familia y a Belén, con todo mi  
aprecio*

## **AGRADECIMIENTOS**

Creo que una página no es suficiente para dar todas las gracias. Quisiera agradecer, en primer lugar, a mi supervisor Daniel Garrido por su apoyo durante el desarrollo de esta tesis y por sus enseñanzas a lo largo de todos estos años. Agradezco que haya confiado en mis habilidades y que me haya dado la oportunidad de investigar y aprender en este tema. También le agradezco al profesor Pedro Saa por su apoyo y confianza durante este postgrado y en mi paso por Ingeniería UC.

Mis agradecimientos más profundos son para mi familia por su increíble fuerza y persistencia. Ha sido un camino largo y lleno de peñascos, pero ya está terminado. No habría podido llegar sin ustedes, Guadalupe, Juan, María y Guillermo, esta tesis es de ustedes.

Finalmente, mi más sincero cariño y agradecimiento a Belén Hirmas por su apoyo incondicional durante estos años. La vida es mucho mejor a tu lado.

## ÍNDICE DE CONTENIDOS

AGRADECIMIENTOS	iv
ÍNDICE DE FIGURAS	viii
ÍNDICE DE TABLAS	x
ABSTRACT	xii
RESUMEN	xiii
1. INTRODUCCIÓN	1
1.1. La microbiota intestinal . . . . .	1
1.2. Fermentación de fibras alimenticias y rol en la salud humana . . . . .	1
1.3. Fibras alimenticias como prebióticos . . . . .	2
1.4. Nociones generales del modelamiento a escala genómica . . . . .	3
1.4.1. GSMMs para modelar la microbiota intestinal . . . . .	5
1.4.2. Enfoques de modelamiento de comunidades microbianas . . . . .	6
1.5. Motivación . . . . .	7
1.6. Hipótesis . . . . .	8
1.7. Objetivos . . . . .	8
2. MATERIALES Y MÉTODOS	10
2.1. Procedimiento general . . . . .	10
2.2. Genómica comparativa y análisis bioinformático . . . . .	11
2.2.1. Selección de los genomas de referencia para el modelamiento . . . . .	11
2.2.2. Búsqueda de modelos de referencia en <i>Bacteroides</i> para el modelamiento de <i>Phocaeicola dorei</i> . . . . .	12
2.2.3. Búsqueda de especies filogenéticamente cercanas a <i>Lachnoclostridium</i> <i>symbiosum</i> . . . . .	13
2.2.4. Anotación de glicosil hidrolasas y PULs . . . . .	13

2.2.5.	Búsqueda de genes ortólogos . . . . .	14
2.2.6.	Potencial probiótico de las especies de interés . . . . .	14
2.3.	Reconstrucción y modelamiento a escala genómica . . . . .	14
2.3.1.	Reconstrucción de <i>Phocaeicola dorei</i> 5_1_36/D4 . . . . .	15
2.3.2.	Reconstrucción de <i>Lachnoclostridium symbiosum</i> WAL-14673 . . . . .	16
2.3.3.	Determinación de la reacción de biomasa y requerimiento de ATP . . . . .	16
2.3.4.	Transportadores de membrana, fuentes de carbono y productos de fermentación . . . . .	18
2.4.	Simulación . . . . .	18
2.4.1.	Representación matemática de un GSMM . . . . .	19
2.4.2.	Simulación de crecimiento en monocultivo . . . . .	19
2.4.3.	Validación de las reconstrucciones . . . . .	22
2.4.4.	Simulación de crecimiento en cocultivo y producción de SCFA . . . . .	23
2.4.5.	Identificación de interacciones por muestreo aleatorio . . . . .	25
3.	RESULTADOS . . . . .	26
3.1.	Reconstrucción de <i>Phocaeicola dorei</i> 5_1_36/D4 . . . . .	26
3.1.1.	Elección de las especies de referencia por genómica comparativa . . . . .	26
3.1.2.	Resultados del proceso de reconstrucción . . . . .	28
3.1.3.	Consumo de las fibras xilano e inulina . . . . .	28
3.1.4.	Validación de la reconstrucción de <i>Phocaeicola dorei</i> 5_1_36/D4 . . . . .	31
3.2.	Genómica comparativa y reconstrucción de <i>Lachnoclostridium symbiosum</i> . . . . .	32
3.2.1.	Consumo de las fibras xilano e inulina . . . . .	34
3.2.2.	Validación de la reconstrucción de <i>Lachnoclostridium symbiosum</i> WAL-14673 . . . . .	35
3.3.	Simulaciones de monocultivo . . . . .	35
3.4.	Simulaciones en cocultivo . . . . .	36
3.4.1.	Búsqueda de interacciones metabólicas . . . . .	39
3.4.2.	Interacciones metabólicas en inulina . . . . .	40
3.4.3.	Interacciones metabólicas en xilano . . . . .	43

4. DISCUSIÓN	47
4.1. Reconstrucciones metabólicas . . . . .	47
4.1.1. Validación de las reconstrucciones . . . . .	48
4.2. Modelamiento de fibras alimenticias y su consumo en bacterias . . . . .	48
4.3. Simulaciones en monocultivo . . . . .	49
4.4. Simulaciones en cocultivo . . . . .	50
5. CONCLUSIONES	53
REFERENCIAS	54
ANEXO	68
A. Información complementaria sobre genómica comparativa y reconstrucciones	70
A.1. <i>Phocaeicola dorei</i> . . . . .	70
A.2. <i>Lachnoclostridium symbiosum</i> . . . . .	73
B. Información complementaria sobre las simulaciones . . . . .	75
B.1. Validación de los modelos . . . . .	75

## ÍNDICE DE FIGURAS

1.1	Pasos para generar un modelo a escala genómica . . . . .	4
3.1	Árbol filogenético de las especies de <i>Bacteroides</i> con una reconstrucción disponible en AGORA v3.0 . . . . .	27
3.2	Mapa de calor con las enzimas identificadas en DBCAN para la familia <i>Lachnospiraceae</i> . . . . .	33
3.3	Abundancia relativa de las dos especies de interés respecto a la tasa de crecimiento de la comunidad y producción de butirato total en inulina . . . . .	37
3.4	Abundancia relativa de las dos especies de interés respecto a la tasa de crecimiento de la comunidad y producción de butirato total en xilano . . . . .	38
3.5	Distribuciones de flujo de las reacciones de intercambio identificadas como interacciones en las simulaciones de cocultivo en medio mZBM suplementado con inulina . . . . .	41
3.6	Diagrama simplificado de las interacciones en cocultivo en medio mZMB suplementado con inulina . . . . .	43
3.7	Distribuciones de flujo de las reacciones de intercambio identificadas como interacciones en las simulaciones de cocultivo en medio mZBM suplementado con xilano . . . . .	44
3.8	Diagrama simplificado de las interacciones en cocultivo en medio mZBM suplementado con xilano . . . . .	46
A.1	Mapa de calor del porcentaje de identidad media nucleotídica (ANI) entre las especies de <i>Bacteroides</i> modeladas en AGORA . . . . .	70

A.2	Mapa de calor del porcentaje de ortólogos respecto al total de regiones codificantes entre los genomas de referencia para la reconstrucción de <i>Phocaeicola dorei</i> 5_1_36/D4 . . . . .	71
A.3	Árbol filogenético con los miembros de la familia <i>Lachnospiraceae</i> basado en el gen <i>rpoB</i> . . . . .	73

## ÍNDICE DE TABLAS

2.1	Características genómicas de los ensamblados de las cepas de estudio . . . . .	11
3.1	Identidad media nucleotídica de las especies más cercanas a <i>Phocaeicola dorei</i> 5_1_36/D4 modeladas en AGORA . . . . .	26
3.2	Lista de familias de enzimas glicosil hidrolasas relacionadas al consumo de inulina y xilano . . . . .	29
3.3	Resultados de la validación para la reconstrucción de <i>Phocaeicola dorei</i> . . .	31
3.4	Resultados de la validación para la reconstrucción de <i>Lachnoclostridium</i> <i>symbiosum</i> . . . . .	35
3.5	Tasas de crecimiento predichas por FBA en medio mZMB con inulina o xilano	36
3.6	Interacciones metabólicas en cocultivo con mZMB e inulina como fuente de carbono predichas por SteadyCom. . . . .	42
3.7	Interacciones metabólicas en cocultivo con mZMB y xilano como fuente de carbono predichas por SteadyCom. . . . .	45
A.1	Fuentes de carbono con evidencia experimental para <i>Phocaeicola dorei</i> . . .	72
A.2	Ortólogos del PUL0425 para el consumo de inulina encontrados en <i>Phocaeicola</i> <i>dorei</i> 5_1_36/D4 . . . . .	72
A.3	Ortólogos del PUL0346 para el consumo de xilano encontrados en <i>Phocaeicola</i> <i>dorei</i> 5_1_36/D4 . . . . .	73
A.4	Genomas seleccionados para el análisis de genómica comparativa en las especies de la familia <i>Lachnospiraceae</i> . . . . .	74

A.5	Fuentes de carbono con evidencia experimental para <i>Lachnoclostridium symbiosum</i> . . . . .	75
B.1	Composición del medio mZMB. Parte 1. . . . .	76
B.2	Composición del medio mZMB. Parte 2. . . . .	77
B.3	Interacciones metabólicas en cocultivo con mZMB y fructosa como fuente de carbono predichas por SteadyCom. . . . .	77

## ABSTRACT

The gut microbiota is composed of a great diversity of microorganisms capable of degrading complex fibers, such as inulin and xylan, and of producing compounds beneficial to human health. These compounds include butyrate, a short-chain fatty acid (SCFA) that stimulates the anti-inflammatory response and the protection of the intestinal epithelium.

*Lachnospirillum symbiosum* is a butyrate producer that benefits from the presence of the primary fermenter *Phocaeicola dorei*. In this work, the metabolic interactions between both bacteria were studied when growing in cocultures with inulin and xylan, using a comparative genomics approach and metabolic modeling. Specific reconstructions were generated and complex fiber and carbohydrate metabolism was studied. Using statistical metrics, it was concluded that the predictive capacity of the reconstructions was lower than others published models, mainly due to the lack of information on fiber metabolism. Simulations with the SteadyCom method and random sampling analysis predicted the existence of a lactate and succinate interaction in both substrates. Additionally, interactions involving amino acids and acetate were found, although these interactions require experimental validation. Finally, simulations predicted that the production of butyrate occurs only in cocultures with xylan, not in inulin.

The building process of a genomic-scale model is iterative, and the reconstructions in this work require further refinement. However, the models were able to identify the relevant metabolic interactions in the coculture of the two bacteria. This study provides key insights to guide the development of therapeutic treatments based on diet and the gut microbiota.

**Keywords:** butyrate production, genome scale modelling, gut microbiota

## RESUMEN

La microbiota intestinal está compuesta por una gran diversidad de microorganismos capaces de degradar fibras complejas, como la inulina y el xilano, y de producir compuestos benéficos para la salud humana. Entre estos compuestos destaca butirato, un ácido graso de cadena corta que estimula la respuesta antiinflamatoria y la protección del epitelio intestinal.

*Lachnoclostridium symbiosum* es un productor de butirato que se ve beneficiado por la presencia del fermentador primario *Phocaeicola dorei*. En este trabajo se estudiaron las interacciones metabólicas entre ambas bacterias al crecer en cocultivos con inulina y xilano, mediante un enfoque de genómica comparativa y modelamiento metabólico. Se generaron reconstrucciones específicas y se estudió el metabolismo de carbohidratos y fibras complejas. A través de métricas estadísticas se concluyó que la capacidad predictiva de las reconstrucciones era inferior a la de otros modelos publicados, principalmente por la falta de información sobre el metabolismo de fibras. Simulaciones con el método SteadyCom y análisis de muestreo aleatorio predijeron la existencia de una interacción de lactato y succinato en ambos sustratos. Adicionalmente se encontraron interacciones que involucran aminoácidos y acetato, aunque dichas interacciones requieren validación experimental. Finalmente, las simulaciones predijeron que la producción de butirato ocurre solo en los cocultivos con xilano.

El proceso de construcción de un modelo a escala genómica es iterativo y las reconstrucciones de este trabajo requieren más refinamiento. Sin embargo, los modelos fueron capaces de encontrar las interacciones metabólicas relevantes en el cocultivo de las dos bacterias. Este estudio entrega claves para guiar el desarrollo de tratamientos terapéuticos basados en la dieta y el microbioma intestinal.

**Palabras Claves:** producción de butirato, modelamiento a escala genómica, microbiota intestinal.

## **1. INTRODUCCIÓN**

### **1.1. La microbiota intestinal**

La microbiota intestinal humana es una compleja comunidad de microorganismos compuesta por una gran diversidad de especies, en su mayoría miembros de los filos Bacteroidetes, Actinobacteria y Firmicutes (Xu & Knight, 2015). Se ha observado que su composición está correlacionada con la edad (Claesson et al., 2011), el tipo de dieta del huésped (Xu & Knight, 2015), su estado de salud y sus condiciones genéticas (Schmidt, Raes, & Bork, 2018), entre otros factores. Esto da paso a una gran variabilidad en la composición de la microbiota intestinal entre personas.

La abundancia y diversidad de las especies de la microbiota intestinal juegan un papel fundamental en la estabilidad de esta comunidad (Thursby & Juge, 2017). Se ha observado que existe redundancia funcional y perfiles metabólicos que se mantienen constantes, pese a los cambios en la composición de la microbiota (Moya & Ferrer, 2016). Lo anterior es de vital importancia, pues la microbiota intestinal realiza una serie de funciones esenciales para el huésped. Estas incluyen la síntesis de vitaminas, la eliminación de compuestos tóxicos, el fortalecimiento de la barrera intestinal, la estimulación y regulación del sistema inmune y en particular, la fermentación de macromoléculas de la dieta que no son degradadas por el intestino (Heintz-Buschart & Wilmes, 2018).

### **1.2. Fermentación de fibras alimenticias y rol en la salud humana**

Fibras alimenticias como el almidón, celulosa, inulina, xilano y otros polisacáridos derivados de estructuras vegetales, no pueden ser hidrolizadas por las enzimas digestivas del ser humano debido a su complejidad estructural (Singh, Banerjee, & Arora, 2015). En su lugar, la fibra llega hasta el intestino grueso y es degradada por bacterias en cadenas más pequeñas y/o monosacáridos, los cuales posteriormente son absorbidos por el huésped o consumidos por otras bacterias de la microbiota (Sen & Orešič, 2019; Graf et al., 2015).

Esto le entrega ventajas competitivas a aquellas especies con la capacidad de fermentar sustratos específicos, pues se convierten en los fermentadores primarios de los recursos disponibles en el intestino, y los productos de su metabolismo pasan a formar parte de interacciones de alimentación cruzada con otras especies.

Como resultado de la fermentación de fibras, las bacterias del intestino producen gases (hidrógeno y dióxido de carbono) y ácidos grasos de cadena corta (SCFA por sus siglas en inglés): acetato, propionato, butirato, lactato, entre otros. Estos hidrocarburos son utilizados como fuente de energía y poseen propiedades benéficas para el huésped. En el caso particular del butirato, se ha documentado que induce la expresión de genes relacionados con la producción de mucina, lo que fortalece al epitelio intestinal frente a la acción de otras bacterias (Burger-van Paassen et al., 2009). Otras de sus propiedades están relacionadas con la restauración y modulación de las proteínas de unión estrecha entre células epiteliales (Davis-Richardson & Triplett, 2015) y la capacidad de reducir la respuesta inflamatoria (Thomson, Medina, Ortúzar, Gotteland, & Garrido, 2018). Por otro lado, conocidos productores de butirato como *Roseburia sp.*, *Faecalibacterium prausnitzii* y *Eubacterium rectale*, utilizan acetato y lactato como fuente de energía (Louis & Flint, 2017). Gracias a estas propiedades, el butirato se ha convertido en un importante objetivo de estudio para desarrollar nuevos tratamientos basados en la microbiota intestinal.

### **1.3. Fibras alimenticias como prebióticos**

Diversas fibras han sido consideradas como prebióticos debido a su capacidad para favorecer a grupos de bacterias productoras de SCFA. La inulina, por ejemplo, es una fibra de origen vegetal compuesta por fructanos (cadenas lineales de fructosa) unidos por enlaces  $\beta(2,1)$  (Niness, 1999). Esta fibra promueve el crecimiento del filo *Bifidobacterium* e inhibe la colonización de patógenos en experimentos *in vivo* (Bosscher, Loo, & Franck, 2006; Vandeputte et al., 2017), además de incrementar la producción de propionato, acetato y butirato en cultivos fecales *batch* (Kaur, Rose, Rumpagaporn, Patterson,

& Hamaker, 2011) y mejorar el metabolismo de azúcares y lípidos en pacientes con obesidad sometidos a dietas suplementadas con inulina (van der Beek et al., 2018).

Por otro lado, se ha observado que los oligosacáridos derivados de la hidrólisis de xilano, un polímero de xilosas (pentosas) unidas por enlaces  $\beta(1,4)$ , promueven el crecimiento de bifidobacterias en estudios *in vivo* (Li, Summanen, Komoriya, & Finegold, 2015) e incrementan la producción de butirato (Lecerf et al., 2012). En particular, pocas bacterias poseen los mecanismos para degradar xilano, mientras que los miembros del filo *Bacteroidetes* sí poseen la capacidad de hidrolizarlo. Un reciente estudio de cocultivos microbianos sugiere que *Bacteroides*, el filo más abundante en la microbiota intestinal, es el responsable de los efectos prebióticos del xilano y su acción bifidogénica (Zeybek, Rastall, & Buyukkileci, 2020). Sin embargo, no se ha logrado describir el mecanismo detrás de esa interacción.

Debido al potencial terapéutico de los prebióticos y sus posibles aplicaciones clínicas, es necesario profundizar en los mecanismos de interacción producidos a nivel de especie. Las técnicas convencionales usadas en el estudio del microbioma, como por ejemplo la secuenciación de 16S rRNA y *whole genome shotgun metagenomics*, no permiten obtener este nivel de detalle (Sen & Orešič, 2019). Nuevas aproximaciones se hacen necesarias para el estudio del microbioma intestinal, como por ejemplo, el modelamiento metabólico a escala genómica.

#### **1.4. Nociones generales del modelamiento a escala genómica**

Una reconstrucción metabólica (GENRE, por su sigla en inglés), corresponde a una red que consolida las características de un organismo desde su información genética, bioquímica y genómica, gracias a un extenso proceso de curación de información (Palsson, 2015). La representación matemática de la red y sus características corresponde a un modelo a escala genómica (GSMM, por sus siglas en inglés). A partir de una serie de supuestos

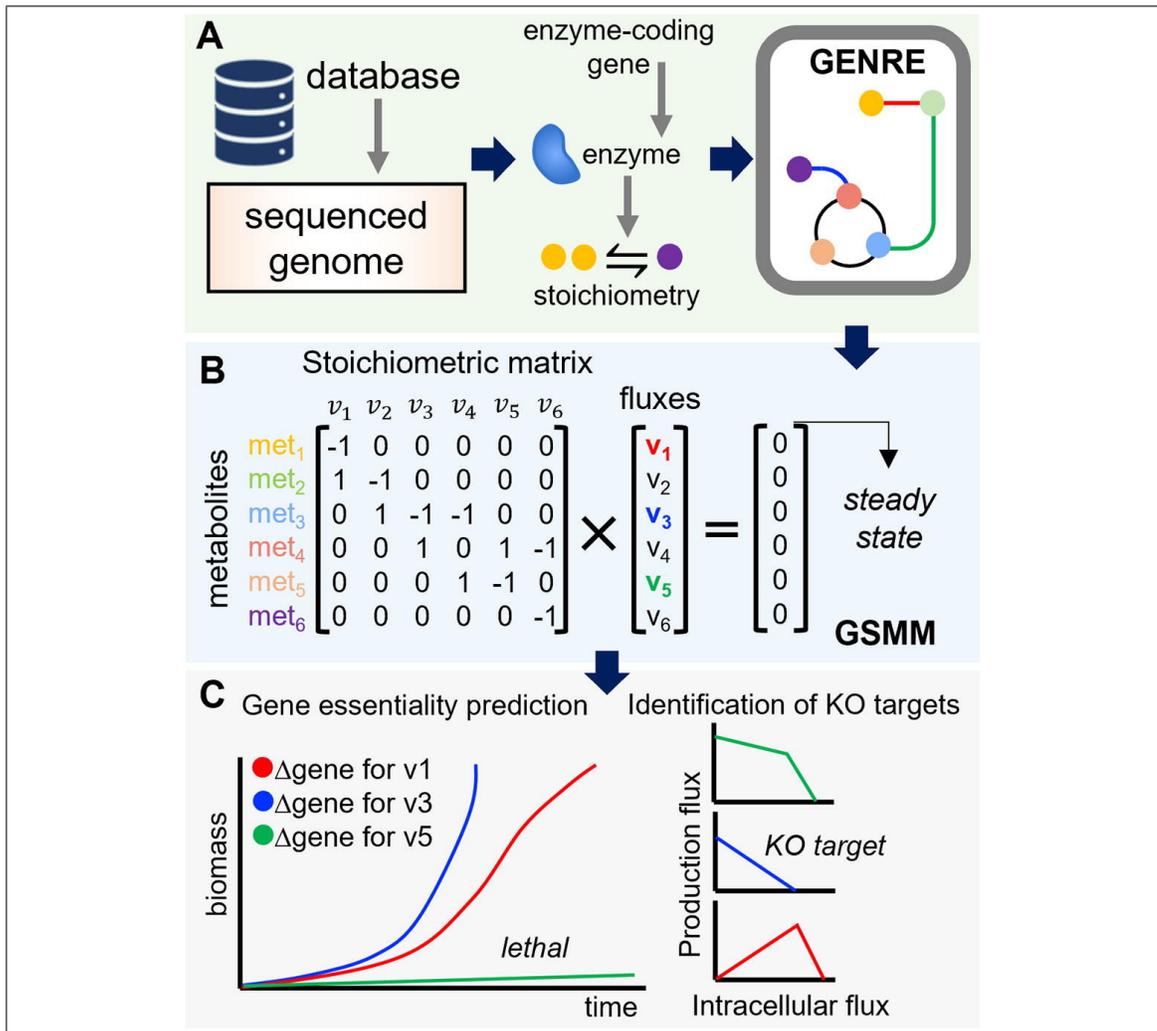


Figura 1.1. Pasos para generar un modelo a escala genómica. A partir del genoma secuenciado y la información de bases de datos de reacciones químicas se genera la reconstrucción de la red metabólica del organismo de interés (Panel A). La representación matemática de esta red permite generar un GSM (Panel B), el que es usado para realizar predicciones sobre esencialidad de genes, identificación de objetivos de silenciamiento, entre otros (Panel C). Imagen obtenida de (Altamirano et al., 2020).

y la aplicación de restricciones y del acoplamiento de información experimental (generalmente de cultivos en biorreactores y tecnologías ómicas), un GSM puede ser usado para predecir comportamientos biológicos y establecer relaciones entre genotipo y fenotipo (Altamirano, Saa, & Garrido, 2020). La figura 1.1 ilustra el proceso de construcción de un GSM.

La generación de una reconstrucción (GENRE) puede ser realizada de forma automática gracias a herramientas como AuReMe (Aite et al., 2018), CarveMe (Machado, Andrejev, Tramontano, & Patil, 2018), ModelSEED (Seaver et al., 2021), entre muchas otras (Mendoza, Olivier, Molenaar, & Teusink, 2019). En general, estas herramientas utilizan una red bioquímica como plantilla y a través del análisis de identidad de secuencia reconstruyen la red del organismo objetivo. Como resultado, se crea una red con metabolitos, reacciones, genes y su estequiometría asociada, codificada de acuerdo a la base de datos usada por la herramienta. La conciliación de las bases de datos de redes bioquímicas es uno de los actuales desafíos del área de modelamiento a escala genómica, aunque se han realizado esfuerzos por generar la base de datos MetaNetX (Moretti et al., 2016), que recopila la asociación entre diversas bases de datos de redes metabólicas, y BiGG (Norsigian et al., 2020), un repositorio que alberga modelos a escala genómica publicados por la comunidad científica y escritos con una codificación común.

Una vez reconstruida la red bioquímica se genera su representación matemática tal como se observa en el panel B de la figura 1.1. Con esta representación se realizan simulaciones y se aplican algoritmos generados por la comunidad científica. El método más usado en modelos a escala genómica es el análisis de balance de flujos o FBA, el cual impone una restricción de estado estacionario sobre la red estequiométrica y optimiza sus flujos en relación a un objetivo metabólico específico, como es la producción de un metabolito o de biomasa (Orth, Thiele, & Palsson, 2010).

#### **1.4.1. GSMMs para modelar la microbiota intestinal**

Aunque tradicionalmente los GSMMs han sido utilizados para estudiar un solo organismo o célula, el desarrollo de nuevas aplicaciones y métodos han expandido el modelamiento al contexto de comunidades microbianas (Sen & Orešič, 2019). Por ejemplo, simulaciones con modelos de 58 especies han logrado predecir el perfil de metabolitos en niños con desnutrición en comparación a un grupo control saludable (Kumar et al., 2018).

Por otro lado, la simulación de una comunidad de 4 GSMMs permitió establecer relaciones de competencia y mutualismo en pacientes tratados con metformina, un fármaco utilizado para el tratamiento de la diabetes. (Rosario et al., 2018).

Adicionalmente, el avance de la metagenómica y la secuenciación masiva ha permitido la generación de un gran número de GENREs de especies del microbioma intestinal disponibles para su uso y curación en la comunidad científica. En esta línea cabe destacar al proyecto AGORA (Magnúsdóttir et al., 2017), un repositorio que contiene 808 reconstrucciones generadas por un método semiautomático basado en genómica comparativa. Las reconstrucciones iniciales de AGORA fueron generadas con ModelSEED (Seaver et al., 2021) y sus reacciones y metabolitos se encuentran codificados en una base de datos exclusiva de AGORA, diseñada para conciliar con el modelo RECON, una reconstrucción del metabolismo humano (Swainston et al., 2016).

#### **1.4.2. Enfoques de modelamiento de comunidades microbianas**

Existen diversos enfoques para simular comunidades microbianas y sus interacciones. MICOM (Diener, Gibbons, & Resendis-Antonio, 2020), por ejemplo, integra información de abundancia relativa en muestras de metagenómica para generar un modelo global de la comunidad. Este método emplea un enfoque de modelamiento compartimentalizado, en donde cada especie es modelada en su propio espacio físico o compartimento, y se define un espacio común que sirve para el intercambio de metabolitos entre cada uno de sus integrantes.

Otras aproximaciones como BacArena, COMETS y d-OptCom permiten simular la distribución espacial y temporal de los miembros de una comunidad gracias a la integración de ecuaciones diferenciales en la modelación, pero presentan dificultades para escalar a comunidades de gran tamaño por la cantidad de parámetros y cómputo requerido para la simulación (Altamirano et al., 2020).

SteadyCom (Chan, Simons, & Maranas, 2017), por otro lado, es un método que extiende las capacidades del FBA a comunidades microbianas a través de la adición de la abundancia relativa y el escalamiento de flujos. La resolución del problema de optimización propuesto por SteadyCom escala linealmente respecto al número de organismos de la comunidad simulada y las aplicaciones del método han predicho con éxito la predominancia de los filos Bacteroides y Firmicutes. Esto último constituye una de las principales ventajas del método para realizar simulaciones de grandes comunidades microbianas. El algoritmo emplea un enfoque de modelamiento compartimentalizado, al igual que MICOM, por lo que es posible extraer información sobre la interacción de las especies modeladas en él.

### 1.5. Motivación

Existen antecedentes que indican que dos bacterias del intestino, *Phocaeicola dorei* y *Lachnospirillum symbiosum*, son esenciales para la producción de butirato durante la utilización de las fibras prebióticas inulina y xilano. Cultivos en inulina de un consorcio de 15 bacterias representativas de la microbiota intestinal (Gutiérrez & Garrido, 2019) demostraron que la delección de *Phocaeicola dorei* tenía un impacto negativo en la comunidad y su presencia beneficiaba el crecimiento del productor de butirato *Lachnospirillum symbiosum* (Vital, Howe, & Tiedje, 2014). Adicionalmente, se ha comprobado que el sobrenadante de cocultivos de ambas bacterias en xilano posee efectos antiinflamatorios en experimentos con células epiteliales, reduciendo la inflamación producida por la citoquina TNF- $\alpha$  (Thomson et al., 2018). Por último, resultados en proceso de publicación indican que el sobrenadante de estos cultivos es capaz de promover la generación de una barrera celular en células Caco2 expuestas a la toxina TcdB (unpublished), resultado que concuerda con otros reportes sobre los beneficios del butirato en la salud del huésped (Davis-Richardson & Triplett, 2015; Burger-van Paassen et al., 2009).

Los antecedentes previos demuestran que esta pareja es un buen modelo para estudiar la relación entre bacterias comensales y productores de butirato, y que es necesario

entender las interacciones metabólicas que dan paso a la producción de butirato cuando crecen en cocultivo. Sin embargo, existen limitaciones para profundizar en la naturaleza de las interacciones. Aunque *Phocaeicola dorei* pertenece al género *Bacteroidetes*, uno de los más abundantes en la microbiota intestinal, no existe información en literatura que detalle el metabolismo de esta especie. Por otro lado, *Lachnoclostridium symbiosum* es una bacteria que pertenece a la recientemente descrita familia *Lachnospiraceae* y pese a haber sido aislada en 1976 (Kaneuchi, Watanabe, Terada, Benno, & Mitsuoka, 1976), es una de las especies menos estudiadas del intestino. Debido a la complejidad de los organismos mencionados, la falta de evidencia experimental y su importancia en el desarrollo de tratamientos antiinflamatorios, existe una gran oportunidad para profundizar en la reconstrucción del metabolismo de estas especies y comprender sus interacciones metabólicas a través del modelamiento metabólico.

## 1.6. Hipótesis

Dadas las capacidades de los modelos a escala genómica y su poder predictivo para simular comunidades bacterianas, se propone la siguiente hipótesis para este trabajo de investigación:

**El desarrollo de modelos a escala genómica de las especies *Phocaeicola dorei* y *Lachnoclostridium symbiosum* permite identificar sus interacciones metabólicas *in silico* durante su crecimiento en inulina y xilano.**

## 1.7. Objetivos

El principal objetivo de este estudio es determinar las interacciones entre las bacterias propuestas a través de experimentos de crecimiento *in silico*. Los objetivos específicos son:

- (i) Caracterizar las rutas metabólicas de las especies *Phocaeicola dorei* y *Lachnoclostridium symbiosum* a través de genómica comparativa y herramientas de predicción bioinformática.
- (ii) Generar reconstrucciones de *Phocaeicola dorei* y *Lachnoclostridium symbiosum* a nivel de cepa bacteriana.
- (iii) Realizar simulaciones en monocultivo y cocultivo para predecir la producción de butirato y las interacciones metabólicas entre las bacterias.
- (iv) Validar la predicción de los modelos con información experimental disponible en literatura.

## 2. MATERIALES Y MÉTODOS

### 2.1. Procedimiento general

Para cumplir con los objetivos de este trabajo se tomaron las siguientes decisiones. En primer lugar, debido a la escasez de información disponible sobre las bacterias *Phocaeicola dorei* y *Lachnoclostridium symbiosum*, se complementó su estudio con diversos análisis de genómica comparativa, que incluyen análisis filogenético, de distancia génica, anotación de enzimas relacionadas con el consumo de fuentes de carbono, búsqueda de locis verificados experimentalmente para el consumo de polisacáridos, entre otros. El análisis fue realizado sobre diversos conjuntos de genomas para obtener pistas sobre las capacidades de las bacterias seleccionadas en este estudio.

En segundo lugar, se decidió utilizar como punto de partida las reconstrucciones de las especies de interés publicadas en el proyecto AGORA (Magnúsdóttir et al., 2017) y utilizar el mismo sistema de codificación para reacciones y metabolitos. Esto con el fin de mantener la consistencia con otros modelos a escala genómica de la microbiota intestinal y así poder reutilizar las reconstrucciones generadas en este trabajo en futuras simulaciones de comunidades de mayor tamaño dentro del mismo grupo de investigación.

Por último, este trabajo fue realizado en paralelo a una tesis de carácter experimental sobre crecimiento en monocultivo y cocultivo de la pareja de bacterias presentadas en la motivación de este estudio (Hirmas, 2021). Parte de los resultados aquí obtenidos apoyaron el diseño experimental de dicho trabajo, motivo por el cual las reconstrucciones se hicieron específicas a nivel de cepa, tal como se detalla más adelante.

En resumen, este trabajo contempla tres secciones:

- Genómica comparativa y análisis bioinformático para encontrar a las especies más cercanas a las de interés y profundizar en sus capacidades metabólicas.

- Elaboración de las reconstrucciones y validación de las capacidades predictivas de los modelos.
- Simulación de los crecimientos de monocultivo y cocultivo en condiciones similares a las reportadas en trabajos experimentales, en particular, los ensayos bidireccionales de Hirmas (Hirmas, 2021).

## 2.2. Genómica comparativa y análisis bioinformático

### 2.2.1. Selección de los genomas de referencia para el modelamiento

Los modelos a escala genómica parten desde el genoma del organismo a modelar. En el caso de las especies de este trabajo se seleccionó el ensamble de la cepa *Phocaeicola dorei* 5\_1\_36/D4, código de acceso GenBank GCA000158335.2, y el ensamble de *Lachnoclostridium symbiosum* WAL-14673, código de acceso GenBank GCA000189615.1. Estos ensamblajes corresponden a las cepas disponibles en el Laboratorio de Microbiología de Sistemas de la Universidad Católica. Las características de los ensamblajes se resumen en la siguiente tabla:

Tabla 2.1. Características genómicas de los ensamblajes de las cepas de estudio

Cepa	ID PATRIC	Tamaño (Mb)	CDS
<i>Phocaeicola dorei</i> 5_1_36/D4	556260.3	5.53	4601
<i>Lachnoclostridium symbiosum</i> WAL-14673	742741.3	4.91	4704

La anotación de los genomas y las secuencias de los genes codificantes (CDS, por sus siglas en inglés) se obtuvo desde la base de datos PATRIC (Davis et al., 2020). Los identificadores de PATRIC se encuentran en la tabla 2.1.

### 2.2.2. Búsqueda de modelos de referencia en *Bacteroides* para el modelamiento de *Phocaeicola dorei*

La cepa *Phocaeicola dorei* 5\_1\_36/D4 no posee modelos a escala genómica disponibles, motivo por el cual se debió generar una reconstrucción específica para esta cepa. Para encontrar las especies y cepas de *Bacteroides* más similares a la cepa de estudio se realizó un análisis filogenético con los genomas de las especies modeladas en el proyecto AGORA. El identificador de los ensamblajes de los genomas usados en la elaboración de los modelos se obtuvo desde la metadata de los modelos alojados en la plataforma *Virtual Metabolic Human* (Noronha et al., 2019). Los archivos de los proteomas y las secuencias nucleotídicas de los genes codificantes se obtuvieron desde la base de datos de PATRIC con el identificador correspondiente. En total se seleccionaron 55 ensamblajes de *Bacteroides*.

Un árbol filogenético de máxima verosimilitud fue generado a partir de 14 genes marcadores correspondientes a proteínas de las subunidades ribosomales (Hug et al., 2016). Se utilizó como referencia el genoma de *Bacteroides thetaiotaomicron* VPI-5482 y los ortólogos de las secuencias se obtuvieron con Orthofinder v2.5.4 (Emms & Kelly, 2019) en los 55 genomas seleccionados (incluyendo la referencia) y en el genoma de la cepa *Phocaeicola dorei* 5\_1\_36/D4. Los ortólogos fueron alineados con MAFFT v7 (Kato & Standley, 2013) usando la opción G-INS-i para alineamiento global de las secuencias. El árbol fue generado por máxima verosimilitud con raxml-ng v1.0.3 (Kozlov, Darriba, Flouri, Morel, & Stamatakis, 2019) y el alineamiento concatenado de las 14 proteínas mencionadas. Se seleccionó RTREV+G4+F como el mejor modelo evolutivo para el alineamiento a partir de los criterios AICc, AIC y BIC calculados por modeltest-ng v0.1.7 (Darriba et al., 2020). La visualización fue realizada en R v4.1 (R Core Team, 2021) con la librería *ggtree*.

Adicionalmente, se calculó la identidad media nucleotídica (ANI) con pyani v.0.2.8 y su opción ANIb para realizar alineamientos locales con BLAST (Pritchard, Glover, Humphris, Elphinstone, & Toth, 2015). La visualización fue realizada en Python con el paquete *seaborn*.

### 2.2.3. Búsqueda de especies filogenéticamente cercanas a *Lachnoclostridium symbiosum*

Para el estudio de las características de la especie *Lachnoclostridium symbiosum*, se recopilaron 11 genomas de referencia de las especies del cluster XIVa de *Clostridium* documentados en el trabajo de reclasificación de la familia *Lachnospiraceae* publicado por Haas (Haas & Blanchard, 2020). En este trabajo se definen dos géneros que no incluyen a la especie de interés, pero que son los más cercanos a *Lachnoclostridium symbiosum* en los árboles genéticos elaborados con tres conjuntos de marcadores distintos (ver figura A.3 del Anexo 1). Los archivos con los CDS, anotación y secuencias proteicas fueron obtenidos desde PATRIC (Davis et al., 2020).

### 2.2.4. Anotación de glicosil hidrolasas y PULs

Se realizó la anotación de enzimas glicosil hidrolasas, importantes para la degradación de azúcares complejos, en los 56 genomas seleccionados para hacer el análisis filogenético de *Bacteroides* (se incluye a la cepa de interés) y los 12 genomas del cluster XIVa de *Clostridium*. Los proteomas respectivos fueron anotados con DBCAN v2.0.11 (Zhang et al., 2018) y procesados con códigos de Python de elaboración propia usando las librerías pandas y biopython. El código de programación se encuentra en el repositorio de esta tesis disponible en el siguiente enlace: [https://github.com/almabeta/tesis\\_GO](https://github.com/almabeta/tesis_GO).

Por otro lado, se realizó una búsqueda de los loci de utilización de polisacáridos (PUL, por sus siglas en inglés) usando la base de datos DBCAN-PUL (Ausland et al., 2021) y BLASTX a través del servidor web disponible en [https://ccb.unl.edu/dbCAN\\_PUL/blast](https://ccb.unl.edu/dbCAN_PUL/blast). La búsqueda se realizó sobre los CDS de los ensamblajes de las cepas de interés y se escribió código personalizado para el procesamiento de resultados. Solo se consideraron los resultados de BLAST con al menos un 50% de identidad de secuencia y un valor-e menor a  $1e^{-5}$ . Los PULs encontrados fueron filtrados y solo se consideró a aquellos que tenían al menos un 50% de sus genes presentes en los ensamblajes.

### **2.2.5. Búsqueda de genes ortólogos**

Se buscaron los genes ortólogos para hacer la comparación entre las especies de *Bacteroides*, las especies usadas como referencia para el modelamiento de *Phocaeicola dorei* 5\_1\_36/D4 y las especies de la familia *Clostridium*. Todas las búsquedas de genes ortólogos fueron realizadas con Orthofinder v2.5.1 sobre los proteomas de las especies y cepas de interés. Los resultados de Orthofinder fueron procesados por programas de Python de elaboración propia disponibles en el repositorio de este trabajo.

### **2.2.6. Potencial probiótico de las especies de interés**

Para el estudio del potencial probiótico de las cepas se realizó la búsqueda de genes de resistencia a antibióticos con la base de datos CARD y la herramienta *Resistance Gene Identifier* (Alcock et al., 2020). Por otro lado, la información de los factores de virulencia fue obtenida desde *Virulence Factors Database* (Liu, Zheng, Jin, Chen, & Yang, 2019).

Todos los resultados fueron comparados con la información de los ensamblajes disponible en PATRIC.

## **2.3. Reconstrucción y modelamiento a escala genómica**

Para la generación de las reconstrucciones a escala genómica se siguió el protocolo de Thiele para la generación de modelos de alta calidad (Thiele & Palsson, 2010), específicamente (i) se incorporaron nuevas reacciones químicas en base a información bibliográfica y bases de datos, (ii) se asignaron reglas de asociación gen-proteína-reacción (GPR) y (iii) se modificaron reacciones manualmente hasta coincidir la predicción del modelo con la información experimental recopilada.

### 2.3.1. Reconstrucción de *Phocaeicola dorei* 5\_1\_36/D4

Para la generación del modelo inicial de *Phocaeicola dorei* 5\_1\_36/D4 se utilizaron 6 reconstrucciones de *Bacteroides* disponibles en el proyecto AGORA (Magnúsdóttir et al., 2017). Los modelos fueron descargados desde la plataforma de la base de datos Virtual Metabolic Human (Noronha et al., 2019). Estas especies de referencia se seleccionaron de acuerdo a los valores de identidad media nucleotídica y a la relación filogenética respecto a la cepa del estudio. Como punto de partida se usó la reconstrucción de *Phocaeicola dorei* DSM-17855, la única bacteria identificada con la misma especie que la cepa de interés. Se generaron tres versiones del modelo de acuerdo a los siguientes pasos:

- v1: Actualización de las reglas de asociación (GPR) del modelo de *Phocaeicola dorei* DSM-17855 basado en homología con *Phocaeicola dorei* 5\_1\_36/D4. Se mantuvieron las reacciones que no tenían evidencia génica para su posterior revisión.
- v2: Adición de reacciones basada en genómica comparativa de los genomas del resto de las reconstrucciones de referencia. Para cada reconstrucción de referencia se revisaron sus reacciones y las reglas GPR asociadas. Se añadieron aquellas reacciones cuyos genes poseían homólogos en la cepa de interés. Algunas reacciones fueron contrastadas con la información del microorganismo en la base de datos Biocyc (Karp et al., 2017). También se revisaron los transportadores y se añadieron nuevas reacciones de intercambio.
- v3: Reconstrucción refinada con información sobre consumo de polisacáridos, en base a literatura e información experimental. A partir de la evidencia obtenida por genómica comparativa se añadieron las reacciones para el consumo de inulina y xilano, las principales fibras prebióticas de este trabajo.

Respecto a las reglas de regulación y sus reacciones asociadas, estas fueron transferidas a la reconstrucción mediante los siguientes criterios:

- Se seleccionaron todas las reacciones cuyos genes asociados poseían ortólogos en la cepa de interés.
- Se mantuvieron las reglas de asociación para aquellas reacciones que poseían ortólogos de **todos** los genes asociados a su GPR.
- Se simplificaron las reglas de asociación para aquellas reacciones que no poseían ortólogos de todos sus genes asociados y solo se añadieron las reacciones cuyas reglas preservaban el sentido lógico y biológico. Por ejemplo, si una reacción posee una regla de regulación del tipo “*gen1 OR gen2*”, y solo *gen1* posee un ortólogo en la cepa de interés, se añade la misma reacción y se reescribe la regla de regulación como “*gen1*”.

### **2.3.2. Reconstrucción de *Lachnoclostridium symbiosum* WAL-14673**

Respecto al modelo de *Lachnoclostridium symbiosum* WAL-14673, se utilizó como base la reconstrucción disponible en el proyecto AGORA para la misma cepa. La familia *Lachnospiraceae*, a diferencia de *Bacteroides*, ha sido poco estudiada por la comunidad científica y existe poca información experimental sobre la especie de interés. Para esta reconstrucción se generó una sola versión con la información del consumo de polisacáridos y productos de fermentación de acuerdo a los resultados de genómica comparativa propios y aquellos obtenidos por Vital para la producción de butirato (Vital et al., 2014), y de experimentos del laboratorio de Microbiología de Sistemas UC (Hirmas, 2021).

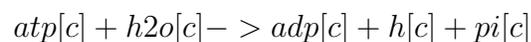
### **2.3.3. Determinación de la reacción de biomasa y requerimiento de ATP**

La reacción de biomasa contiene la información de los requerimientos en macromoléculas y el gasto energético asociado al crecimiento. Los coeficientes y metabolitos de la reacción de biomasa dependen fuertemente de las condiciones de crecimiento del microorganismo, en particular del medio, y son específicas a nivel de cepa (Folsom & Carlson, 2015). En el caso de los modelos de AGORA, la reacción de biomasa fue generada automáticamente con la plataforma de reconstrucciones automáticas ModelSeed

(Seaver et al., 2021) desde una plantilla universal y en base a la información taxonómica de cada especie. A través de una serie de reglas definidas en la plantilla se remueven y/o agregan componentes a la ecuación, por lo que fue necesario examinar las diferencias entre los modelos de referencia.

Para el caso de *Phocaeicola dorei* 5\_1\_36/D4, se realizó la comparación de las reacciones de biomasa en los 6 modelos de referencia seleccionados y la reacción del modelo iKS1119 (K. Kim et al., 2021), el más reciente modelo de *Bacteroides thetatiotaomicron* VPI-5482 publicado a la fecha. No se encontró diferencias entre las reacciones de los modelos de AGORA, salvo en la presencia del metabolito *colipa* que representa a un componente del lipopolisacárido de la membrana de las bacterias Gram negativas. Por otro lado, la reacción del modelo iKS1119 contiene a la mayoría de los componentes de las reacciones de los modelos de AGORA, pero también añade compuestos derivados del metabolismo de folato, fosfolípidos y tiamina. Debido a que no es posible estimar los coeficientes ni los componentes de la reacción de biomasa a nivel de cepa ni extrapolar la información del modelo iKS1119, se decidió mantener en el modelo la reacción de *Phocaeicola dorei* DSM-17855, correspondiente a la reacción de biomasa de una bacteria Gram negativa.

Respecto al requerimiento de ATP asociado y no asociado al crecimiento (GAM y NGAM, por sus siglas en inglés), se decidió mantener los valores presentes en las reconstrucciones de AGORA pues no existe la información experimental necesaria en literatura para su estimación. De esta manera, se mantuvo el requerimiento asociado al crecimiento (GAM) en  $40 \frac{\text{mmol}}{\text{gDW}h}$  y la cota inferior de la reacción de demanda de ATP (NGAM), en  $0 \frac{\text{mmol}}{\text{gDW}h}$ . La reacción correspondiente a la demanda de ATP posee el identificador *DM\_atp\_c\_* y su ecuación estequiométrica es la siguiente:



Los mismos criterios fueron aplicados sobre la reconstrucción de *Lachnoclostridium symbiosum* WAL-14673 y se mantuvo la reacción de biomasa original de la reconstrucción de AGORA.

#### **2.3.4. Transportadores de membrana, fuentes de carbono y productos de fermentación**

Se revisó manualmente la información de los transportadores anotados en Biocyc (Karp et al., 2017) y obtenidos con TransportDB (Elbourne, Tetu, Hassan, & Paulsen, 2017). Por otro lado, se recopiló información sobre las fuentes de carbono fermentadas por las bacterias de interés para la posterior curación y validación de los modelos. Esta información se encuentra detallada en los anexos A.1 y A.5 con sus respectivas referencias. Pese a que todas las fuentes de carbono fueron evaluadas en la validación de los modelos, muchas de ellas son fibras complejas cuyos mecanismos de degradación requieren una revisión más exhaustiva. Esta información se deja disponible para futuras mejoras de la reconstrucción.

### **2.4. Simulación**

Todas las simulaciones de crecimiento realizadas en este trabajo, ya sea en monocultivo o cocultivo, fueron calculadas utilizando el paquete *CO*nstraint-Based *Re*construction and *A*nalysis *T*oolbox (COBRA) disponible en MATLAB y la librería de Python *CO*BRAPy (Ebrahim, Lerman, Palsson, & Hyduke, 2013). El paquete COBRA contiene una colección de algoritmos y métodos usados en el estudio de redes bioquímicas con un enfoque en modelamiento basado en restricciones.

Las simulaciones fueron ejecutadas en Matlab 2019b y Python v3.9 usando COBRAPy v0.22.1, COBRA Toolbox v3.2 y la librería de optimización Gurobi Optimizer v9.0.3.

### 2.4.1. Representación matemática de un GSMM

En un modelo a escala genómica toda la información del organismo o célula es resumida en la matriz estequiométrica  $\mathbf{S}$ , la cual contiene una representación de todas las reacciones codificadas en la red bioquímica. Cada fila de  $\mathbf{S}$  corresponde a un metabolito y cada columna a una reacción de la red. En una red metabólica de  $m$  metabolitos y  $n$  reacciones, el coeficiente estequiométrico del metabolito  $i$  en la reacción  $j$  corresponde al elemento  $S_{ij}$  en la matriz  $S$ , tal como se muestra a continuación:

$$\mathbf{S} = \begin{bmatrix} S_{11} & S_{12} & S_{13} & \dots & S_{1n} \\ S_{21} & S_{22} & S_{23} & \dots & S_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ S_{m1} & S_{m2} & S_{m3} & \dots & S_{mn} \end{bmatrix}_{m \times n} \quad (2.1)$$

Adicionalmente, el flujo que pasa por la reacción  $j$  es representado por una variable continua  $v_j$ . En modelamiento basado en restricciones, se añaden restricciones de capacidad inferior y superior para cada flujo, correspondientes a las constantes  $lb_j$  y  $ub_j$ , respectivamente. Esta representación matemática es la base para realizar análisis de flujos o FBA.

### 2.4.2. Simulación de crecimiento en monocultivo

La tasa de crecimiento de las cepas de interés y la distribución de flujos asociada a cada solución fue calculado mediante el método de análisis de balance de flujos o FBA, por sus siglas en inglés (Orth et al., 2010). Este método resuelve el siguiente problema de optimización lineal:

$$\underset{z}{\text{maximize}} \quad z = \mathbf{c}^\top \mathbf{v} \quad (2.2a)$$

$$\text{subject to} \quad \sum_{j \in J} S_{ij} v_j = 0, \quad \forall i \in I, \quad (2.2b)$$

$$lb_j \leq v_j \leq ub_j, \quad \forall j \in J, \quad (2.2c)$$

$$v_j \in \mathbb{R}, \quad \forall j \in J \quad (2.2d)$$

Donde  $J$  representa el conjunto de reacciones de la red,  $I$  corresponde al conjunto de metabolitos,  $z$  corresponde al valor de la función objetivo,  $lb_j$  y  $ub_j$  corresponden a los límites inferior y superior para cada uno de los flujos metabólicos  $v_j$  y  $S_{ij}$  corresponde al coeficiente estequiométrico del metabolito  $i$  en la reacción  $j$ . En cada optimización el objetivo fue definido como una sola variable de flujo, es decir,  $c_{obj} = 1$  para  $v_{obj}$  y  $c_j = 0 \forall j \neq obj$ . En particular, la variable que representa el flujo de la reacción de biomasa será llamada  $v_{biomass}$  en adelante. Cabe resaltar que los flujos representan tasas específicas para un organismo, pues la reacción de biomasa representa, por convención, el requerimiento para producir 1 gramo de biomasa seca del microorganismo representado y sus unidades corresponden a  $\frac{mmol}{gDWh}$ .

Por otro lado, los valores mínimos y máximos para los flujos de las reacciones de intercambio asociadas a SCFA y otros metabolitos de interés fueron calculados con el método de análisis de variabilidad de flujos o FVA. Mediante este método el espacio de soluciones de la red se restringe a una fracción ( $\alpha$ ) del valor óptimo de crecimiento y se resuelve el siguiente problema de optimización:

$$\begin{aligned}
& \max|\min && v_{obj} \\
\text{subject to} && \sum_{j \in J} S_{ij} v_j = 0, && \forall i \in I, \\
&& v_{biomass} = \alpha \cdot v_{biomass}, && \\
&& lb_j \leq v_j \leq ub_j, && \forall j \in J, \\
&& v_j \in \mathbb{R}, && \forall j \in J
\end{aligned} \tag{2.3}$$

Las condiciones del ambiente en el que crecen los microorganismos son representadas a través de las restricciones de capacidad de los flujos asociados al intercambio de metabolitos específicos entre la bacteria y el medio. Para asegurar el crecimiento de los microorganismos estudiados se empleó una representación del medio mZMB (Medina, Pinto, Ovalle, Thomson, & Garrido, 2017), un medio complejo que incluye vitaminas, iones, aminoácidos, etc. En este trabajo se simularon múltiples condiciones en monocultivo, agrupadas en dos categorías:

- Simulación en medio mZMB con las fuentes de carbono para la validación de las reconstrucciones (se explica en detalle en la siguiente sección). La composición del medio se encuentra en las tablas B.1 y B.2 del Anexo 2. Por convención, el flujo de cada fuente de carbono se fijó en  $-10 \frac{mmol}{gDWh}$ .
- Simulación de los cultivos en medio complejo mZMB suplementado con la fuente de carbono inulina o xilano. La composición del medio mZMB es la misma que la mencionada en el punto anterior, pero los flujos de inulina y xilano fueron fijados en unidades equivalentes respecto al número de carbonos que aportan sus reacciones de hidrólisis. Para inulina se utilizó un flujo de  $1 \frac{mmol}{gDWh}$  mientras que para xilano se utilizó un flujo de  $0.066 \frac{mmol}{gDWh}$ .

Todas las simulaciones fueron realizadas en medio anaeróbico ajustando el límite inferior de la reacción de intercambio de oxígeno, EX\_o2(e), a  $0 \frac{mmol}{gDWh}$ .

### 2.4.3. Validación de las reconstrucciones

Uno de los principales supuestos empleados en el modelamiento realizado en este trabajo es la imposición de un estado estacionario, supuesto que se utiliza tanto en las simulaciones de mono y cocultivo (FBA y SteadyCom, respectivamente). Esta aproximación representa a un sistema de cultivo continuo en biorreactores, de los cuales se obtienen mediciones para refinar las reconstrucciones en términos de rendimientos, productividades volumétricas, tasas de crecimiento y flujos intracelulares obtenidos con técnicas de fluxómica. Para las bacterias empleadas en este estudio no existe esta información experimental disponible en literatura, motivo por el cual se aplicó un enfoque cualitativo para estudiar las capacidades predictivas de las reconstrucciones.

Se realizaron una serie de simulaciones *in silico* para revisar la producción de metabolitos y el consumo de distintas fuentes de carbono documentadas experimentalmente. En el caso de las fuentes de carbono, se comparó el crecimiento del modelo en medio mZMB con el sustrato de interés respecto a su crecimiento en medio mZBM suplementado con glucosa. Si la simulación predice un crecimiento mayor al 20% del crecimiento de referencia, entonces se asume que el microorganismo consume esa fuente de carbono. El resultado fue contrastado con la información experimental y se calculó el número de verdaderos positivos (TP), verdaderos negativos (TN), falsos positivos (FP) y falsos negativos (FN). Estos indicadores se usaron para calcular las siguientes métricas (Mendoza, Cañón, Contreras, Ribbeck, & Agosín, 2017):

- Sensibilidad =  $\frac{TP}{TP+FN}$
- Especificidad =  $\frac{TN}{TN+FP}$
- P.P.V (Valores positivos predichos) =  $\frac{TP}{TP+FP}$
- N.P.V (Valores negativos predichos) =  $\frac{TN}{TN+FN}$
- Precisión =  $\frac{TP+TN}{TP+TN+FP+FN}$
- F-score =  $\frac{2(PPV \cdot Sensibilidad)}{Precision + Sensibilidad}$

Los mismos indicadores fueron calculados para los productos de fermentación descritos en literatura. Las simulaciones se realizaron en medio mZMB suplementado con glucosa y se cambió el objetivo a la reacción de intercambio asociada al producto de fermentación. Se definió un umbral de  $0.001 \frac{mmol}{gDWh}$  para filtrar los artefactos numéricos obtenidos por el optimizador. Si el valor de la función objetivo optimizada es mayor a dicho umbral, entonces se asume que la reconstrucción es capaz de producir dicho compuesto.

#### 2.4.4. Simulación de crecimiento en cocultivo y producción de SCFA

Para realizar la simulación del cocultivo se utilizó el método SteadyCom (Chan et al., 2017). La formulación de este método está diseñada para inferir la distribución de flujos promedio de una comunidad microbiana en estado estacionario y su composición en términos de la abundancia relativa. SteadyCom corresponde a un algoritmo con un enfoque compartimentalizado (Altamirano et al., 2020), en donde cada especie de la comunidad es asignada a un compartimento conectado a las demás por medio de un espacio común.

La nueva formulación del problema de optimización extiende las propiedades del FBA descrito en la sección anterior a una comunidad bacteriana en lugar de un solo organismo o célula, mediante la adición de una variable que representa a la biomasa total aportada por cada especie  $k$ ,  $X^k$ , y a la reescritura del balance de masa en términos del flujo agregado  $V_j^k$ , el cual corresponde al flujo total de la reacción  $j$  producido por los miembros de la especie  $k$ . La relación entre el flujo agregado y el flujo definido en FBA es la siguiente:

$$V_j^k = v_j^k X^k \quad \forall j \in J, \forall k \in K \quad (2.4)$$

Gracias a esta relación los flujos ya no están normalizados respecto a la biomasa de su organismo respectivo y pueden ser considerados en el balance de masa global de la

comunidad. En resumen, el método de SteadyCom resuelve el siguiente problema de optimización:

$$\begin{aligned}
& \text{maximize} && \mu \\
& \text{subject to} && \sum_{j \in J^k} S_{ij}^k V_j^k = 0, \quad \forall i \in I^k, \forall k \in K, \\
& && lb_j^k X^k \leq V_j^k, \quad \forall j \in J^k, \forall k \in K, \\
& && V_j^k \leq ub_j^k X^k, \quad \forall j \in J^k, \forall k \in K, \\
& && V_{biomass}^k = X^k \mu, \quad \forall k \in K, \\
& && u_i^c - e_i^c + \sum_{k \in K} V_{ex(i)}^k = 0, \quad \forall i \in I^{\text{com}}, \\
& && \sum_{k \in K} X^k = X_0, \\
& && \mu, e_i^c \geq 0, \quad \forall i \in I^{\text{com}}
\end{aligned} \tag{2.5}$$

Análogamente, se realizó análisis de variabilidad de flujos sobre la producción de butirato y las abundancias relativas, mediante la siguiente extensión del algoritmo:

$$\begin{aligned}
& \text{max|min} && \sum_{\substack{j \in J^k \\ k \in K}} w_{kj}^V V_j^k + \sum_{k \in K} w_k^X X^k \\
& \text{subject to} && \mu = \mu_0, \\
& && w_{kj}^V, w_k^X \in [0, 1], \\
& && \text{SteadyCom constraints}
\end{aligned} \tag{2.6}$$

Donde  $w_k^X$  corresponde al ponderador de la variable de abundancia relativa de la especie  $k$ , mientras que  $w_{kj}^V$  corresponde al ponderador de la variable del flujo de la reacción  $j$  en la especie  $k$ .

Para la simulación se emplearon las funciones *SteadyComFBA*, *SteadyComFVA* y *CreateMultipleSpeciesModel*, una función diseñada para crear la matriz estequiométrica conjunta del modelo compartimentalizado del cocultivo. Todas estas funciones están implementadas en el paquete COBRA.

#### **2.4.5. Identificación de interacciones por muestreo aleatorio**

Para estudiar las diferencias entre las fibras prebióticas y su efecto en el metabolismo de las bacterias se realizó el muestreo aleatorio del espacio de soluciones factibles de la matriz estequiométrica conjunta empleada para realizar las simulaciones de cocultivo, para xilano e inulina en el medio mZMB. Esta matriz es una unión de las matrices estequiométricas de ambos organismos y es la base para ejecutar el algoritmo de SteadyCom. Debido a que la formulación de este método agrega variables relacionadas con las abundancias relativas, el problema de optimización se transforma en un problema no lineal y dificulta el muestreo aleatorio. Por esta razón, el muestreo fue realizado sobre la matriz conjunta sin considerar las abundancias relativas como variables.

Para el análisis se empleó la implementación del algoritmo CHRR (Haraldsdóttir, Cousins, Thiele, Fleming, & Vempala, 2017) en la función *sampleCbModel* del paquete COBRA. El algoritmo se ejecutó para obtener 10.000 puntos con un paso de 200 por punto generado. Se empleó una versión modificada de la implementación de COBRA (función *custom\_sampleCBModel*) que emplea la función *SteadyComFVA* para reducir el espacio de búsqueda y eliminar las reacciones bloqueadas en el óptimo.

### 3. RESULTADOS

#### 3.1. Reconstrucción de *Phocaeicola dorei* 5\_1\_36/D4

##### 3.1.1. Elección de las especies de referencia por genómica comparativa

Tal como se describió en la metodología, las especies de referencia se definieron bajo tres criterios: ANI, ortología y filogenia. Un mapa de calor con los resultados de ANI entre todos los pares de genomas estudiados se encuentra en la figura A.1 del Anexo 1. En general los valores obtenidos se mueven entre un 65 y un 80%, con la excepción de 6 especies con valores sobre el 94% de identidad de secuencia, tal como se observa en la tabla 3.1.

Tabla 3.1. Identidad media nucleotídica de las especies más cercanas a *Phocaeicola dorei* 5\_1\_36/D4 modeladas en AGORA

Especie	Porcentaje de identidad
<i>Bacteroides</i> sp. 3_1_33FAA	98.91
<i>Bacteroides</i> sp. 9_1_42FAA	98.90
<i>Phocaeicola dorei</i> DSM 17855	98.80
<i>Phocaeicola vulgatus</i> ATCC 8482	94.90
<i>Bacteroides</i> sp. 4_3_47FAA	94.90
<i>Bacteroides</i> sp. 3_1_40A	94.70

Los resultados indican que estas 6 cepas están cerca de alcanzar el umbral del 95% de ANI usado convencionalmente para definir a organismos de una misma especie (M. Kim, Oh, Park, & Chun, 2014). Adicionalmente, estas mismas bacterias fueron ubicadas en posiciones cercanas en el árbol filogenético generado por máxima verosimilitud (ver figura 3.1) y comparten un clado común con la cepa que se desea reconstruir, con un valor de soporte (*bootstrap*) del 100%. Estos antecedentes sugieren que las 6 reconstrucciones son buenas candidatas para ser usadas como referencias para generar el GENRE de la cepa de interés. Sin embargo, el porcentaje de genes ortólogos compartidos entre cada uno de estos genomas respecto al total de secuencias codificantes en la cepa *Phocaeicola dorei* 5\_1\_36/D4 no pasa del 78% (ver figura suplementaria A.2 del Anexo 1). Ya que

un porcentaje importante de genes no es compartido entre cada par de cepas, existe la posibilidad de que los modelos metabólicos no sean equivalentes entre ellos.

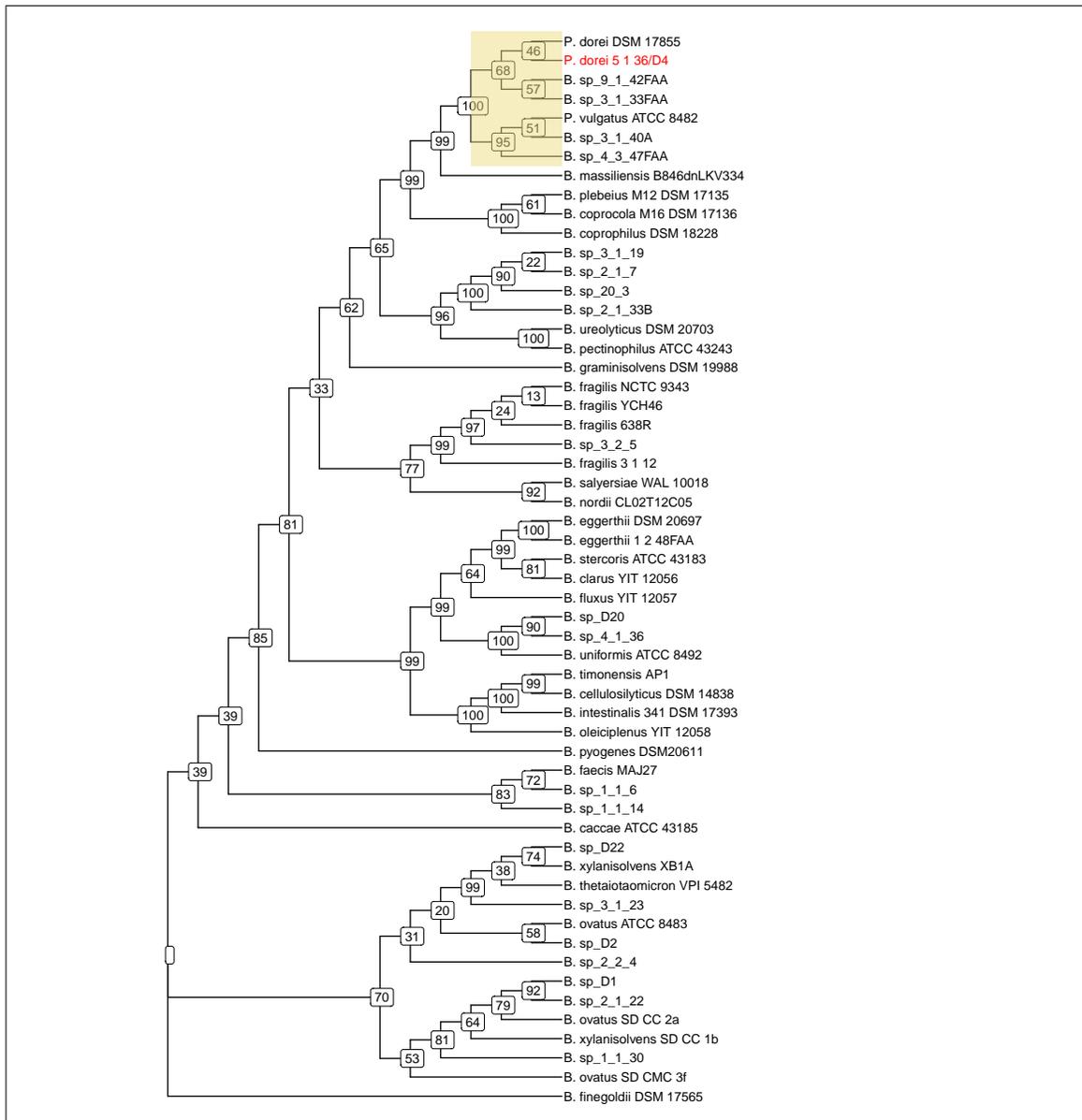


Figura 3.1. Árbol filogenético de las especies de *Bacteroides* con una reconstrucción disponible en AGORA. En rojo se destaca la cepa modelada en el desarrollo de este trabajo. El clado que agrupa a los modelos usados como referencia se encuentra destacado en amarillo.

### 3.1.2. Resultados del proceso de reconstrucción

En vista de estos antecedentes, se utilizó la reconstrucción *Phocaeicola dorei* DSM 17855 como punto de partida para 5\_1\_36/D4, pues pertenece a la misma especie de la cepa de interés, y luego se añadieron las reacciones del resto de las referencias en base a la ortología y anotación funcional de los genomas. Prácticamente todas las reacciones con genes asociados en la reconstrucción de *Phocaeicola dorei* DSM 17855 poseen genes ortólogos en el genoma de *Phocaeicola dorei* 5\_1\_36/D4, motivo por el que la primera versión de la reconstrucción contó con 1186 reacciones, 751 genes y 1061 metabolitos. Por otro lado, se revisaron manualmente todos los genes de la cepa de interés que no poseían ortólogos en la reconstrucción de referencia. No se encontraron genes que codificaran enzimas que pudieran ser añadidas a la red.

Para la segunda versión del modelo se añadieron las reacciones obtenidas por homología en el resto de los modelos de referencia. Con este ejercicio se complementó el metabolismo de folato y se añadieron las reacciones DHFOR2, DHFR, FTHFCL, la enzima D-manonato deshidratasa y su reacción MNNH y las reacciones AMAA y ARGSS de los subsistemas de transporte celular y metabolismo de alanina, entre otras.

Adicionalmente, se agregó la reacción de transporte de alanina por un transportador del tipo protón simporte (ALAt2r) en base a la información obtenida en Biocyc y TransportDB. En la anotación existen transportadores genéricos para aminoácidos, carbohidratos, entre otras moléculas, pero fueron descartados del análisis debido a que su anotación no es concluyente. La versión 2 de la reconstrucción contiene 755 genes, 1065 metabolitos y 1193 reacciones.

### 3.1.3. Consumo de las fibras xilano e inulina

Los resultados de la búsqueda de glicosil hidrolasas (GHs) y de loci de utilización de polisacáridos (PUL) se filtraron para obtener aquellos relacionados con el consumo y la

degradación de inulina y xilano. La tabla 3.2 contiene un resumen de las enzimas GHs buscadas en los genomas analizados.

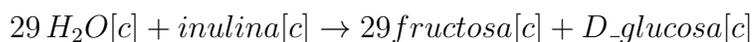
Tabla 3.2. Lista de familias de enzimas glicosil hidrolasas relacionadas al consumo de inulina y xilano.

Familia	Enzimas	Número EC
GH1	$\beta$ -xilosidasa	3.2.1.37
GH5	endo- $\beta$ -1,4-xilanasa	3.2.1.8
GH10	endo- $\beta$ -1,4-xilanasa	3.2.1.8
GH32	endo-inulinasa	3.2.1.7
	endo-levanasa	3.2.1.65
	exo-inulinasa	3.2.1.80
GH43	$\beta$ -xilosidasa	3.2.1.37
GH67	xilano $\alpha$ -1,2-glucuronidasa	3.2.1.131
GH91	inulina liasa	4.2.2.17

Concretamente, se observó la presencia de GH32 en el *locus tag* BSEG\_04149 de la cepa de interés. Según la anotación funcional esta enzima corresponde a una levanasa, enzima que hidroliza los enlaces  $\beta$ -2,6 entre las fructosas del polisacárido levan, otro tipo de fructano similar a inulina. Los resultados de la búsqueda de PUL indican que la enzima pertenece al PUL0425 (Martens et al., 2011), un loci cuya activación está verificada en experimentos de consumo de fructanos y reportado originalmente en *Bacteroides ovatus*. El PUL posee transportadores y enzimas asociadas al metabolismo de fructosa que tienen ortólogos en la cepa de interés. Los resultados de la búsqueda de BLAST y el detalle de esos ortólogos se encuentra en el Anexo 1 en la tabla A.2.

Resultados experimentales indica que *Phocaeicola dorei* hidroliza inulina y utiliza sus derivados como fuente de carbono (Thomson et al., 2018; Hirmas, 2021). Bajo esta evidencia se añadieron las siguientes reacciones en la reconstrucción para simular el consumo de inulina, obtenidas de VMH:

- INULINASE: Degradación de inulina por  $\beta$ -2,1-fructanasa.



- INULINabc: Importación de inulina por sistema ABC de transporte.



- EX\_inulin(e): Reacción de intercambio de inulina.



Las letras encerradas entre corchetes, *e* y *c*, indican el compartimento en el que se encuentran dichos metabolitos y corresponden al espacio extracelular y al citosol, respectivamente. Tal como se observa en las reacciones descritas, el consumo de inulina ocurre intracelularmente, a pesar de que las enzimas que degradan esta fibra, como por ejemplo endo-inulasas y exo-inulasas, son preferentemente extracelulares. En el caso de *Phocaeicola dorei* 5\_1\_36/D4 no se encontró evidencia génica de la existencia de transportadores del principal componente de la inulina, fructosa (no existen entradas en Biocyc, KEGG ni TransportDB asociadas al transporte de este monosacárido a nivel de especie ni de cepa). Bajo esta perspectiva, se modeló el consumo de la fibra como intracelular debido a la evidencia de sistemas de transporte ABC para inulina (TsujiKawa, Ishikawa, Sakane, Yoshida, & Osawa, 2021) que permiten incorporar oligosacáridos de alto grado de polimerización (sobre 8 monómeros) derivados de esta fibra. De esta manera, la fructosa obtenida de la degradación de inulina se consume intracelularmente. Este es uno de los supuestos más fuertes en la modelación.

Respecto al consumo de xilano, se verificó la presencia del PUL0346 (Tauzin et al., 2016), un loci conservado en *Bacteroides*. Este PUL tiene genes de transporte tipo SuS, encargados de adherir el polisacárido a la pared externa de la membrana de la bacteria, lo que da paso a la acción de las enzimas GH43, GH10 y GH67, que tienen actividad xilanasas y son capaces de hidrolizar los enlaces  $\beta$ -1,4 que unen a los polímeros de xilosa. El resumen de los genes ortólogos de este PUL se encuentra en la tabla A.3 del Anexo 1.

El PUL0346 fue clonado originalmente en *Escherichia coli* y se demostró su capacidad para degradar xilano y sus oligosacáridos con grados de polimerización entre 1 y 6 (Tauzin

et al., 2016). En base a esta evidencia se agregaron la siguientes reacciones extracelulares para el consumo de xilano:

- EX\_xylan(e): Reacción de intercambio de xilano.



- XYLAN\_DEGe: Degradación de xilano extracelular.



Con esta información se generó la versión 3 de la reconstrucción de *Phocaeicola dorei*, compuesta por 1197 reacciones, 1068 metabolitos y 755 genes.

#### 3.1.4. Validación de la reconstrucción de *Phocaeicola dorei* 5\_1\_36/D4

La reconstrucción obtenida fue revisada iterativamente con la información experimental de crecimiento en 28 fuentes de carbono y 6 productos de fermentación. El resumen de las métricas calculadas con los experimentos *in silico* se encuentra en la tabla 3.3.

Tabla 3.3. Resultados de la validación para la reconstrucción de *Phocaeicola dorei*. Las categorías en la columna Tipo corresponden a fuentes de carbono (FC) y productos de fermentación (PF).

Tipo	TP	TN	FP	FN	Total exp.	Sensibilidad	Especificidad	N.P.V	P.P.V	Precisión	F-Score
FC	12	9	1	6	28	0.67	0.9	0.6	0.75	0.92	0.77
PF	5	1	0	0	6	1.00	1.0	1.0	1.00	1.00	1.00

Los resultados de la validación indican que el modelo posee una buena precisión y especificidad, con un valor de F-score de un 77%. Respecto a los resultados de las predicciones se puede observar que existen 6 falsos negativos correspondientes a fibras cuyo consumo no ha sido modelado en la reconstrucción.

### 3.2. Genómica comparativa y reconstrucción de *Lachnoclostridium symbiosum*

Para esta reconstrucción se utilizó como punto de partida el modelo de *Lachnoclostridium symbiosum* WAL-14673 disponible en la versión 3 de AGORA, modelo que cuenta con 1178 reacciones, 1078 metabolitos y 935 genes.

Respecto a la genómica comparativa, se observó que existe poca diversidad de consumo de azúcares en las bacterias de la familia *Lachnospiraceae*. Como se aprecia en la figura 3.2, la categoría que posee la mayor cantidad de enzimas anotadas corresponde a GH1. Esta categoría incluye a las enzimas con estructura tipo barril ( $\beta/\alpha$ )<sub>8</sub>, la más común entre todas las enzimas catalíticas, e incluye a  $\beta$ -glucosidasas, galactosidasas, xilosidasas, manosidasas, entre otras enzimas. Llama la atención que en el caso de la cepa WAL-14673 prácticamente no existen enzimas GH1.

Por otro lado, se puede observar que a pesar de no existir una gran diversidad de glicosil hidrolasas en el conjunto de los miembros del cluster, sí existen marcadas diferencias entre cada cepa. Algunas categorías son exclusivas de una sola especie, como es el caso de las enzimas del tipo GH115, GH133, GH24, GH50, GH53, GH73, entre otras. Otras categorías son más abundantes en la familia *Lachnospiraceae*, pero en general cada categoría se encuentra presente en menos de la mitad de las especies estudiadas.



### 3.2.1. Consumo de las fibras xilano e inulina

Respecto a las enzimas que son de interés para el consumo de xilano e inulina, es importante destacar que la categoría GH5 subcategoría 2 solo se encuentra presente en la cepa de este estudio, *Lachnospirillum symbiosum* WAL-14673. Esta categoría incluye a la enzima endo- $\beta$ -1,4-xilanasas (EC 3.2.1.8), por lo que el gen es de potencial interés para el consumo de xilano. El gen corresponde al *locus tag* HMPREF9475\_04133 y está anotado como una proteína hipotética. La búsqueda de PULs indica que el gen es similar a la GH5\_2 del PUL0488, relacionado con el consumo de polímeros de manosa (McBride et al., 2009), pero la identidad de secuencia es menor al 35%, por lo que no es posible hacer hipótesis sobre la función del gen. No se encontraron otros genes que pudieran estar asociados a la degradación de xilano. En vista de estos antecedentes no se agregaron reacciones para el metabolismo de esta fibra.

Por otro lado, la cepa posee una enzima GH32 correspondiente al *locus tag* HMPREF9475\_00225. Esta enzima se encuentra anotada como sucrosa-6-fosfato hidrolasa en PATRIC, pero una búsqueda de BLAST en la base de datos UNIPROT indica que puede corresponder a una  $\beta$ -fructofuranosidasa, una enzima ligada al metabolismo de fructosa. A nivel experimental no existe evidencia de que *Lachnospirillum symbiosum* WAL-14673 consuma inulina, pero sí posee documentado consumo de fructosa, lo que concuerda con la presencia de transportadores de un sistema PTS codificado en el *locus tag* HMPREF9475\_02798. Ya que no existe evidencia suficiente para determinar el consumo de inulina en la bacteria, se mantuvo el transporte de fructosa y no se agregaron otras reacciones.

Por último, no se encontraron transportadores en la base de datos de TransportDB ni Biocyc que fuera necesario añadir a la reconstrucción. Basados en estos antecedentes, no hubo modificaciones relevantes para la reconstrucción de *Lachnospirillum symbiosum* WAL-14673.

### 3.2.2. Validación de la reconstrucción de *Lachnoclostridium symbiosum* WAL-14673

La reconstrucción obtenida fue revisada iterativamente con la información experimental de crecimiento en 18 fuentes de carbono y 7 productos de fermentación. El resumen de las métricas calculadas con los experimentos *in silico* se encuentra en la tabla 3.4.

Tabla 3.4. Resultados de la validación para la reconstrucción de *Lachnoclostridium symbiosum*. Las categorías en la columna Tipo corresponden a fuentes de carbono (FC) y productos de fermentación (PF).

Tipo	TP	TN	FP	FN	Total exp.	Sensibilidad	Especificidad	N.P.V	P.P.V	Precisión	F-Score
FC	3	11	1	3	18	0.5	0.92	0.79	0.78	0.75	0.6
PF	4	1	2	0	7	1.0	0.33	1.00	0.71	0.67	0.8

Tal como se observa en la tabla, la sensibilidad del modelo es baja en el caso de las fuentes de carbono, mientras que la especificidad es de un 33% en el caso de los productos de fermentación. Esto impacta negativamente en el *F-Score*, el cual apenas alcanza un 60%. Respecto a los falsos positivos, el modelo predice crecimiento en ribosa y producción de lactato y succinato, mientras que en el caso de los falsos negativos, el modelo predice que no existe crecimiento en lactosa, ribosa y galactosa, aunque los experimentos de Kaneuchi indican que sí debería existir consumo (el detalle se encuentra en la tabla A.5).

### 3.3. Simulaciones de monocultivo

Se simuló el crecimiento de ambas bacterias con los modelos desarrollados, en el medio mZMB suplementado con inulina o xilano como fuente de carbono. Como resultado se calcularon las siguientes tasas de crecimiento:

Tabla 3.5. Tasas de crecimiento ( $h^{-1}$ ) predichas por FBA en medio mZMB con inulina o xilano.

Cepa	Inulina	Xilano
<i>Phocaeicola dorei</i> 5_1_36/D4	0.562	0.562
<i>Lachnoclostridium symbiosum</i> WAL-14673	0.061	0.061

Tal como se observa en la tabla 3.5, el crecimiento de *Phocaeicola dorei* es mayor al crecimiento de *Lachnoclostridium symbiosum* en las simulaciones de ambas fibras. Por otro lado, no se observan cambios en la tasa de crecimiento de ninguna de las dos bacterias cuando se cambia la fuente de carbono.

Respecto al análisis de variabilidad de flujos, se calcularon los valores mínimos y máximos para los flujos de los ácidos grasos de cadena corta (acetato, propionato, butirato, succinato, lactato) y las reacciones de intercambio, sujetos al valor máximo de crecimiento obtenido para cada reconstrucción. No se observaron diferencias entre los flujos mínimos y máximos de cada uno de los SCFA, lo que indica que en el óptimo la producción de estos compuestos está muy acotada.

### 3.4. Simulaciones en cocultivo

La tasa de crecimiento, las abundancias relativas de *Phocaeicola dorei* y *Lachnoclostridium symbiosum* y los valores mínimos y máximos para los flujos de cada reacción de interés fueron calculados con SteadyCom. Como resultado se obtuvo que las fuentes de carbono inulina y xilano permiten alcanzar una tasa de crecimiento de comunidad de 0.845 y 0.860  $h^{-1}$ , respectivamente. Este valor corresponde al final del eje horizontal en las figuras 3.3 y 3.4 y representa al crecimiento promedio de la comunidad, independientemente de la especie. La tasa de crecimiento predicha en SteadyCom es mayor a la encontrada en las simulaciones en monocultivo (ver tabla 3.5), lo que sugiere un beneficio del cocultivo al crecimiento general de esta pareja de bacterias.

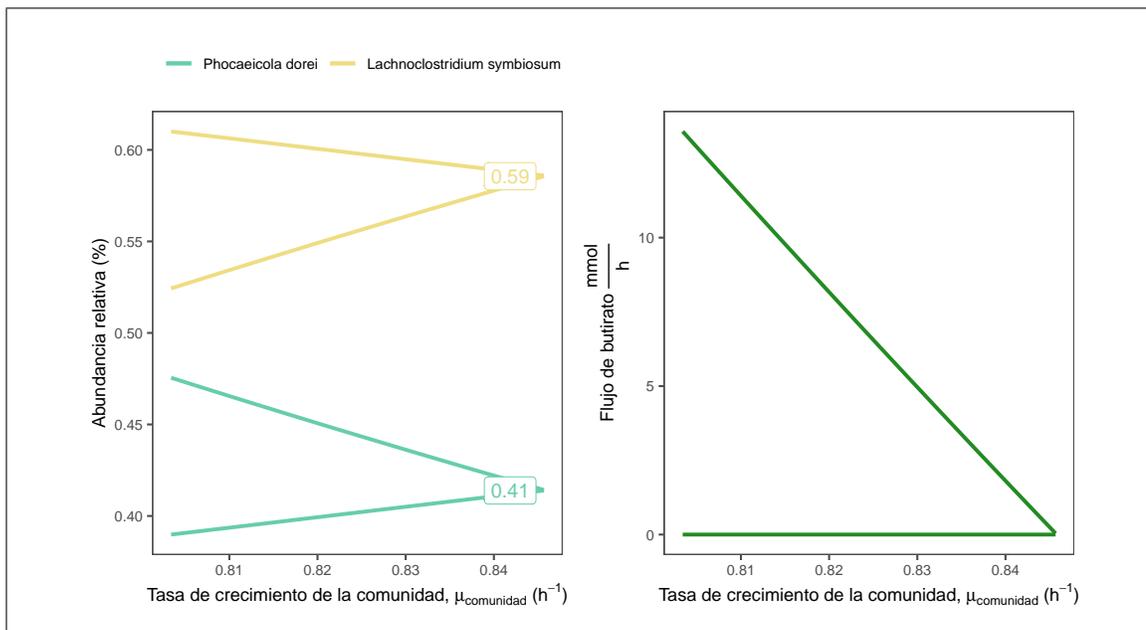


Figura 3.3. Abundancia relativa de las dos especies de interés respecto a la tasa de crecimiento ( $\mu$ ) de la comunidad (panel izquierdo) y producción de butirato total (panel derecho), desde el 95% al 100% de la tasa de crecimiento óptima encontrada por FBA, en medio mZMB suplementado con inulina. El gráfico muestra los valores mínimos y máximos predichos para cada variable. Los valores destacados corresponden a la abundancia encontrada en el punto óptimo para cada especie.

El panel izquierdo de la figura 3.3 muestra la envolvente inferior y superior de las abundancias relativas para los valores de tasa de crecimiento de la comunidad entre un 95% y 100% de la tasa de crecimiento óptima, en cocultivo con medio mZMB suplementado con inulina. Tal como se aprecia en la figura, a medida que nos acercamos al crecimiento óptimo la abundancia relativa de las especies converge en un valor único. Además se observa que la abundancia de *Lachnoclostridium symbiosum* es siempre mayor a la de *Phocaeicola dorei* en estas simulaciones.

Por otro lado, el panel derecho muestra los valores mínimos y máximos del flujo total de butirato de la comunidad respecto a la tasa de crecimiento. Se puede apreciar que la producción de butirato posee una tendencia negativa en relación al crecimiento, hasta

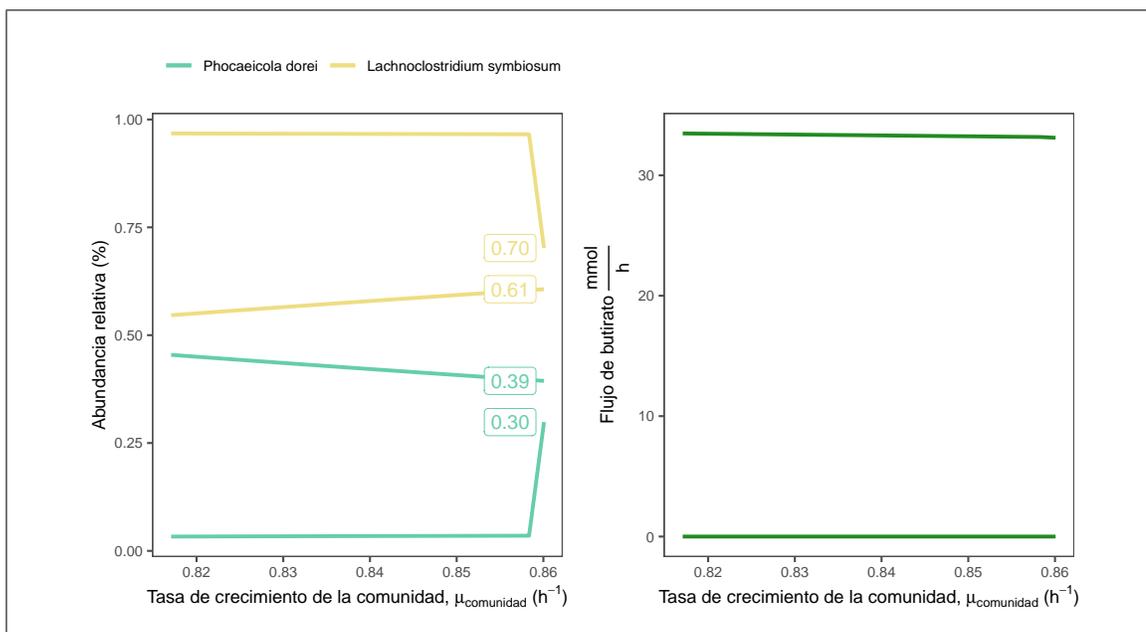


Figura 3.4. Abundancia relativa de las dos especies de interés respecto a la tasa de crecimiento de la comunidad (panel izquierdo) y producción de butirato total (panel derecho), desde el 95% al 100% de la tasa de crecimiento óptima encontrada por FBA, en medio mZMB suplementado con xilano. El gráfico muestra los valores mínimos y máximos predichos para cada variable. Los valores destacados corresponden al mínimo y máximo de abundancia relativa en el punto óptimo de crecimiento.

el punto en el que la producción se hace nula en el punto de máximo crecimiento de la comunidad.

Los resultados de las simulaciones de cocultivo en medio mZMB suplementado con xilano se muestran en la figura 3.4. En el panel izquierdo se puede apreciar que las abundancias relativas no logran converger en el punto óptimo y que existe un rango que sostiene a la tasa de crecimiento, correspondiente al intervalo [61%, 70%] para *Lachnoclostridium symbiosum* y [30%, 39%] para *Phocaeicola dorei*. Esto indica que existen vías metabólicas complementarias activadas por el consumo de xilano que permiten que exista flexibilidad en las proporciones de cada bacteria. Adicionalmente, al igual que en el caso de inulina, *Lachnoclostridium symbiosum* siempre es la especie más abundante.

En lo referente a la producción de butirato, se puede observar en el panel derecho que la envolvente inferior de producción se mantiene en 0 a lo largo del rango de tasas de crecimiento estudiadas. Por otro lado, no existen variaciones en la envolvente superior de producción y esta se mantiene fija mientras aumentamos la tasa de crecimiento. Esto indica que la producción de butirato es independiente del crecimiento cerca del óptimo, y por lo tanto, existen condiciones en la simulación que permiten una producción continua de butirato acotada solamente por la estequiometría de la red.

#### **3.4.1. Búsqueda de interacciones metabólicas**

Las figuras 3.5 y 3.7 contienen las distribuciones de flujo para metabolitos identificados como candidatos a interacciones metabólicas de alimentación cruzada. Las distribuciones calculadas por muestreo aleatorio indican cuáles son los valores más probables de los flujos de intercambio en la red estequiométrica conjunta que sirve como base para las simulaciones realizadas con SteadyCom.

Para definir una interacción se calcularon las medianas de la distribución y se comparó el sentido de su flujo (negativo para consumo, positivo para producción) en cada bacteria. Si las reacciones de intercambio de un mismo metabolito tienen una tendencia complementaria, es decir, consumo en una bacteria y producción en la otra, entonces dicho metabolito es marcado como candidato para ser una interacción metabólica. Este mismo análisis fue realizado en una de las soluciones óptimas encontradas en el análisis de flujos (FBA) en las simulaciones realizadas empleando el método SteadyCom, con la diferencia de que se consideró el valor específico de cada flujo en la solución del problema de optimización en lugar de la mediana de una distribución. Si el metabolito es parte de una interacción en ambos casos, entonces es identificado como interacción metabólica de alimentación cruzada.

### 3.4.2. Interacciones metabólicas en inulina

La figura 3.5 muestra las distribuciones de los flujos de intercambio en las simulaciones de cocultivo para los metabolitos identificados como interacciones metabólicas en medio de cultivo mZMB con inulina. Como se puede observar en la figura, existen metabolitos con distribuciones complementarias que entregan pistas sobre el rol de cada bacteria en esta comunidad. Por ejemplo, en el caso de acetato, dióxido de carbono, prolina, alanina y leucina, la distribución de los flujos de *Phocaeicola dorei* está centrada en los valores negativos del eje horizontal, lo que indica que esta bacteria preferentemente consume estos metabolitos. A su vez, el perfil de la distribución de flujos de *Lachnospirillum symbiosum* es similar al de la bacteria vecina, con la excepción de que sus distribuciones están centradas en el lado positivo del eje horizontal e indican producción.

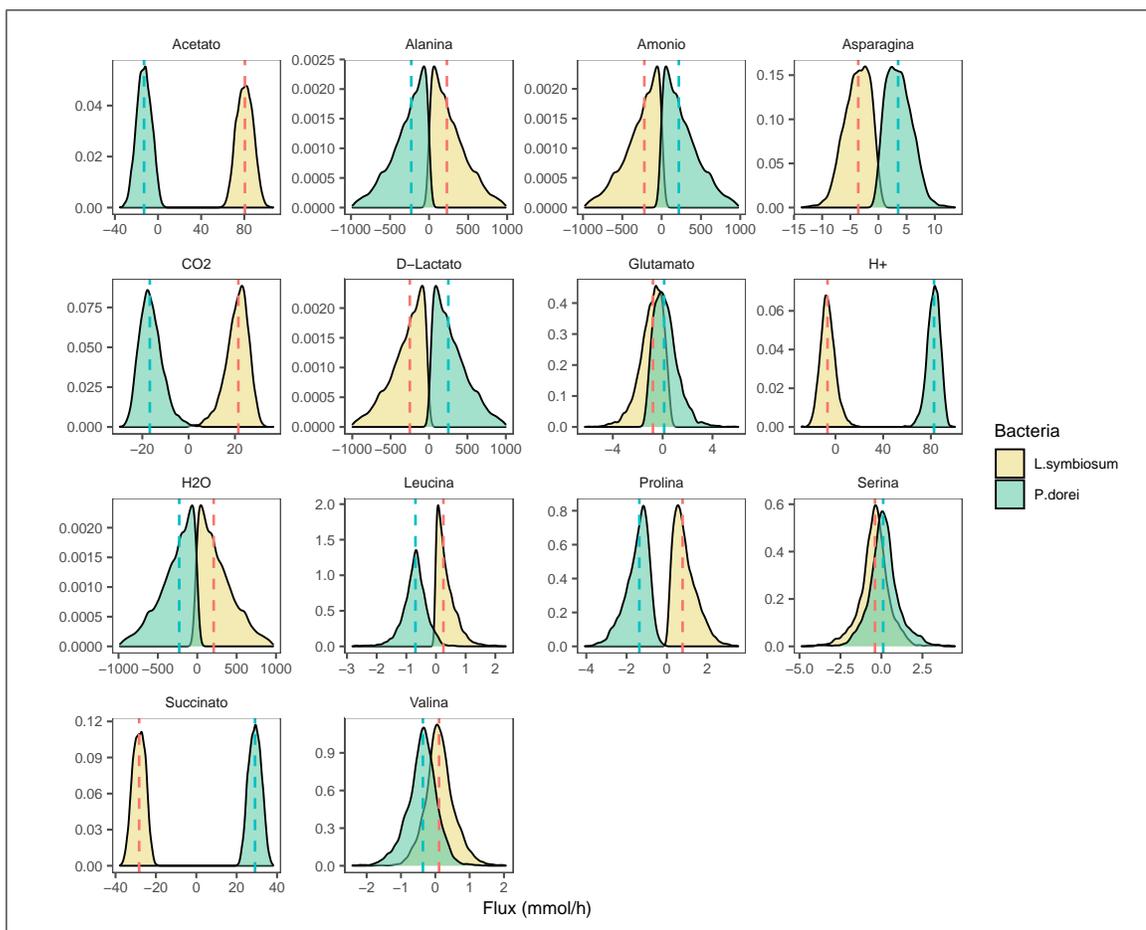


Figura 3.5. Distribuciones de flujo de las reacciones de intercambio identificadas como interacciones en las simulaciones de cocultivo en medio mZMB suplementado con inulina. El eje horizontal indica el valor del flujo medido en  $\frac{mmol}{h}$ , mientras que el eje vertical indica el valor de la probabilidad, ajustado para que el área bajo la curva de la distribución sea equivalente a 1. La línea vertical punteada indica el valor de la mediana de cada distribución.

Existen otros metabolitos que pese a ser identificados como interacciones, poseen distribuciones que indican producción y consumo en una misma bacteria. Por ejemplo, la distribución de los flujos de valina, serina y glutamato muestra que en el espacio de soluciones del problema de optimización existen flujos positivos y negativos para una misma reacción de intercambio, en ambas bacterias. Esto sugiere que en determinados momentos

los roles de productor y consumidor podrían verse intercambiados entre las bacterias del cocultivo, aunque esto solo se desprende a partir de las capacidades de los flujos y del balance de masa inherente a la matriz estequiométrica, y no implican necesariamente una característica de la solución óptima del problema.

Tabla 3.6. Interacciones metabólicas en cocultivo con mZMB e inulina como fuente de carbono predichas por SteadyCom.

Metabolito	<i>Phocaeicola dorei</i>	<i>Lachnoclostridium symbiosum</i>
Acetato	-11.458464	44.600456
Alanina	-402.515526	402.263765
Asparagina	2.383521	-2.608521
CO <sub>2</sub>	-12.650912	14.098107
Etanol	-1.889854	1.889854
Formiato	0.282031	-0.282031
Glutamato	-1.495829	0.675630
H <sup>+</sup>	38.564832	-6.588937
H <sub>2</sub> O	-396.970169	396.260239
D-Lactato	414.031360	-414.031360
Amonio	397.295012	-397.814499
Succinato	17.311622	-17.311622
Treonina	0.135963	-0.478956
Timidina	-0.004113	0.523739

La tabla 3.6 muestra las interacciones encontradas en la solución óptima encontrada con SteadyCom. Se puede observar que ni serina ni valina aparecen como metabolitos de intercambio, a diferencia de lo observado en el análisis anterior. Por otro lado, nuevos candidatos como treonina, etanol, formato o el nucleósido timidina parecen ser parte de una interacción metabólica, aunque no fueron identificados en el análisis de las distribuciones de flujo. Respecto a los metabolitos que coinciden con el análisis de las distribuciones de flujo, se observa que acetato, alanina, asparagina, dióxido de carbono, glutamato, lactato, amonio y succinato son interacciones en la solución óptima del problema.

Es importante destacar que existen valores de flujos muy altos y biológicamente invi-ables para ciertos metabolitos, como lo son lactato y alanina. Estos flujos podrían estar relacionados con un ciclo estequiométrico en la matriz conjunta, detalles que serán analizados más adelante.

La figura 3.6 incluye un diagrama con una simplificación de las interacciones encontradas.

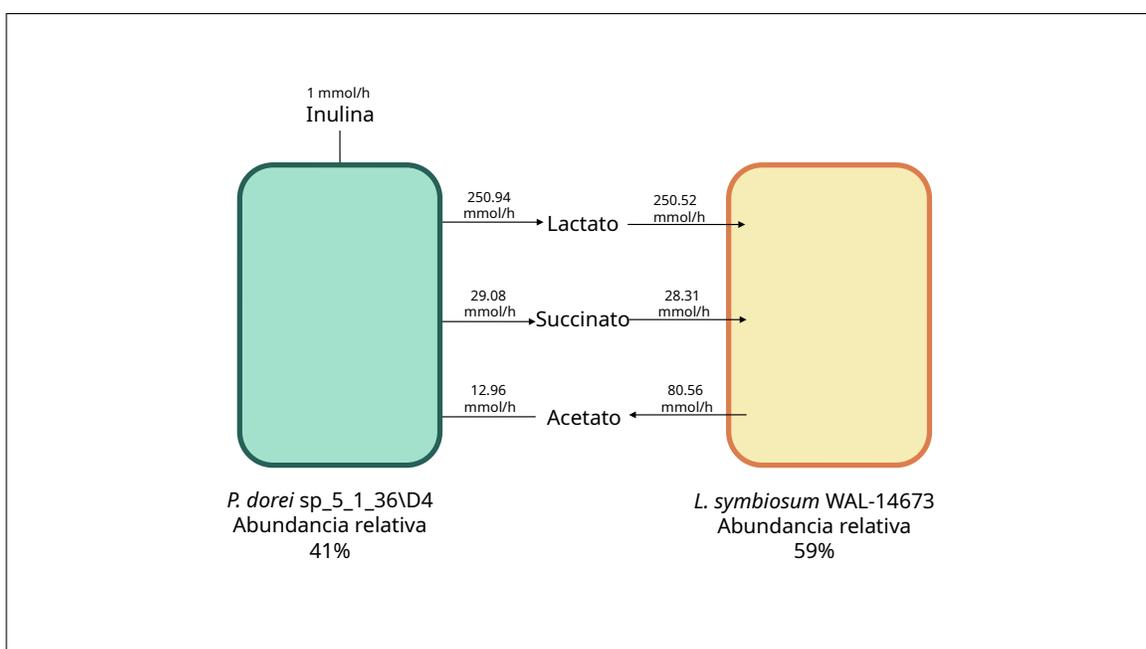


Figura 3.6. Diagrama simplificado de las interacciones en cocultivo en medio mZBM suplementado con inulina. Las números indicados en las interacciones de SCFA corresponden a las medianas de las distribuciones de flujo. No se incluyen las interacciones de aminoácidos debido a la cercanía de sus medianas respecto a las tasas definidas en el medio de cultivo simulado.

### 3.4.3. Interacciones metabólicas en xilano

La figura 3.7 y la tabla 3.7 muestran la distribución de flujos y el valor de las reacciones de intercambio en la solución óptima, respectivamente, para las simulaciones de cocultivo en xilano.

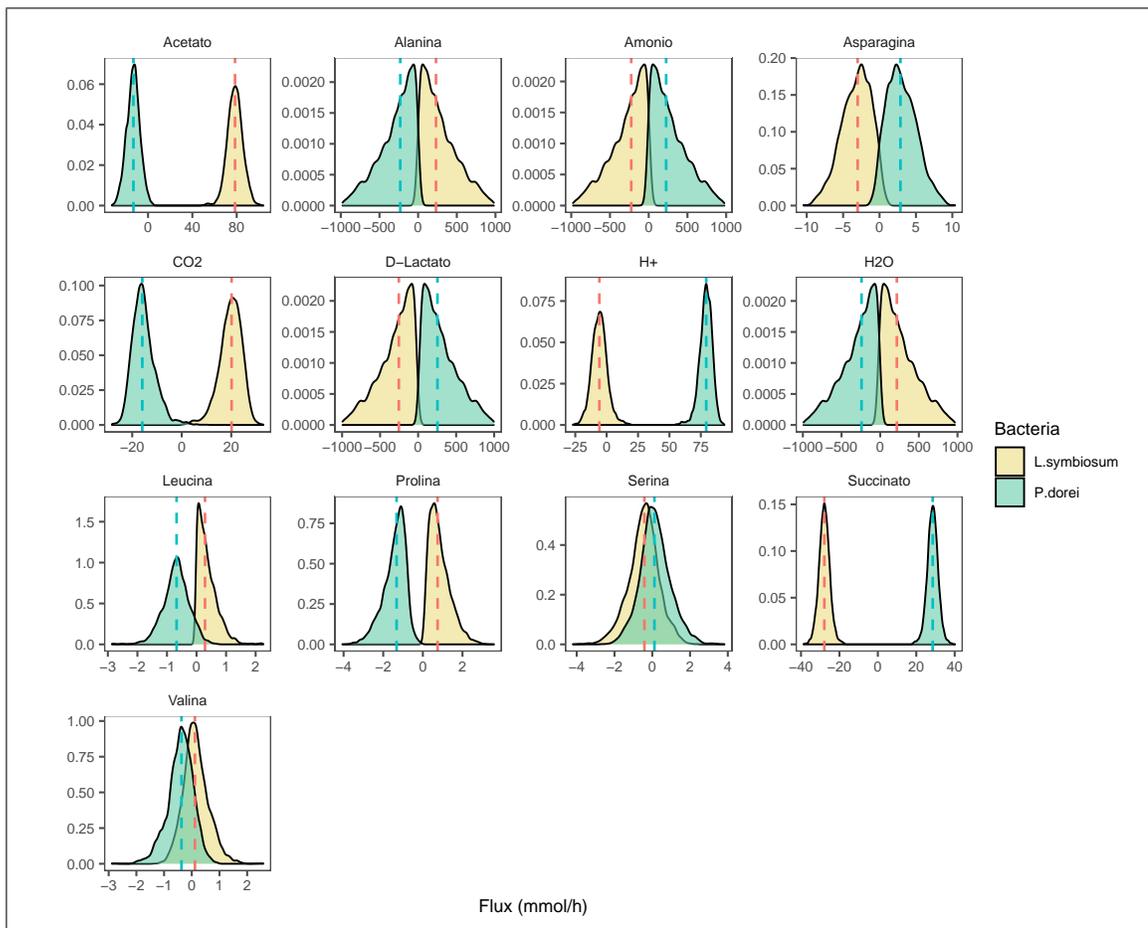


Figura 3.7. Distribuciones de flujo de las reacciones de intercambio identificadas como interacciones en las simulaciones de cocultivo en medio mZBM suplementado con xilano. El eje horizontal indica el valor del flujo medido en  $\frac{mmol}{h}$ , mientras que el eje vertical indica el valor de la probabilidad, ajustado para que el área bajo la curva de la distribución sea equivalente a 1. La línea vertical punteada indica el valor de la mediana de cada distribución.

Tabla 3.7. Interacciones metabólicas en cocultivo con mZMB y xilano como fuente de carbono predichas por SteadyCom.

VMH ID	<i>Phocaeicola dorei</i>	<i>Lachnoclostridium symbiosum</i>
Acetato	-46.968427	50.737645
CO2	-19.851166	19.851166
Formiato	0.070782	-0.070782
Glutamato	-1.477394	0.657195
H+	61.225333	-11.084427
H2O	0.110091	-1.842099
D-Lactato	23.098944	-23.098944
Serina	-0.472736	0.027109
Succinato	43.069424	-19.607489
Treonina	0.169903	-0.512895
Timidina	-0.003980	0.551689

En relación a la distribución de flujos, se puede apreciar que no existen diferencias respecto al resultado obtenido para inulina en la figura 3.5, pues los candidatos a interacciones metabólicas se mantienen y el perfil de las distribuciones es similar entre una fuente de carbono y otra. Sin embargo, los resultados del FBA indican que es posible que existan menos interacciones en esta fuente de carbono. Tal como se observa en la tabla 3.7, acetato, dióxido de carbono, glutamato, lactato y succinato coinciden con los metabolitos sugeridos por la distribución de flujos en la red estequiométrica conjunta. No obstante, no existe intercambio de alanina, asparagina y etanol entre las bacterias simuladas. Por otro lado, nuevamente se sugiere que treonina, timidina y formiato pueden ser un caso de alimentación cruzada. Cabe destacar que en el caso de xilano todos los flujos calculados son inferiores a  $100 \frac{mmol}{h}$ , al contrario de lo ocurrido en el caso de inulina.

La figura 3.8 incluye un diagrama con una simplificación de las interacciones encontradas.

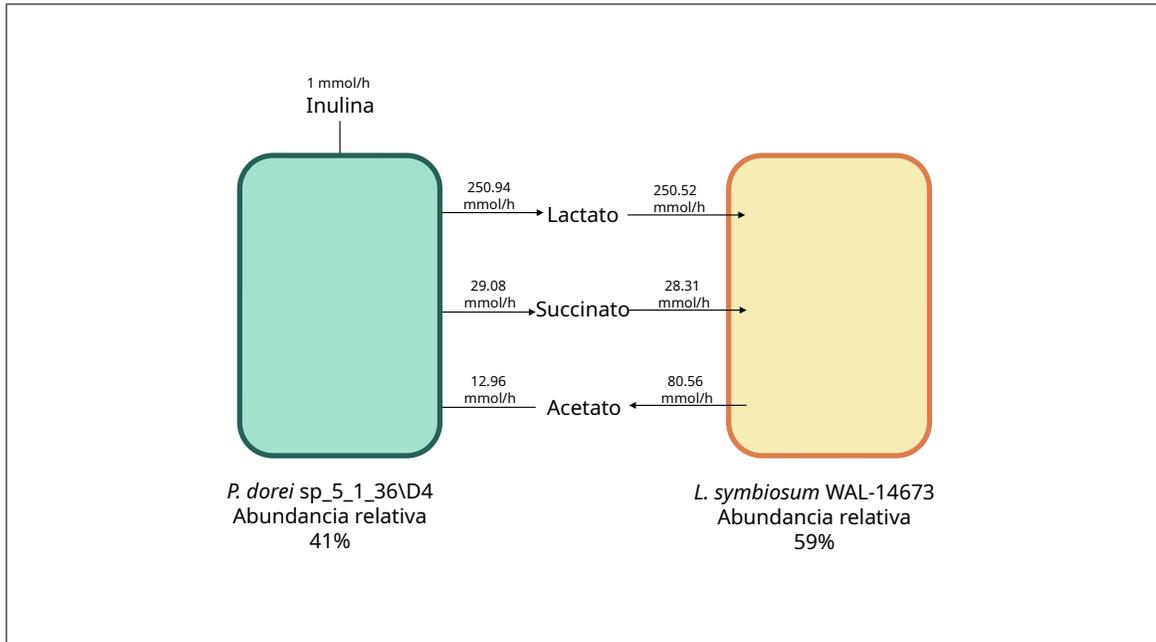


Figura 3.8. Diagrama simplificado de las interacciones en cocultivo en medio mZMB suplementado con xilano. Las números indicados en las interacciones de SCFA corresponden a las medianas de las distribuciones de flujo. En el caso de butirato, los números encerrados en corchetes indican los valores mínimos y máximos estimados por SteadyCom. No se incluyen las interacciones de aminoácidos, salvo el metabolito prolina, debido a la cercanía de sus medianas respecto a las tasas definidas en el medio de cultivo simulado

## 4. DISCUSIÓN

En este trabajo se estudiaron las interacciones metabólicas entre las bacterias *Phocaeicola dorei* y *Lachnoclostridium symbiosum*, dos representantes de la microbiota intestinal, a través de un enfoque de modelamiento a escala genómica. El trabajo de esta tesis fue realizado en paralelo a una investigación experimental para probar la existencia de interacciones metabólicas cruzadas con los resultados de las simulaciones (Hirmas, 2021), trabajo que incluye monocultivos y cocultivos de las dos bacterias de interés en las fibras prebióticas simuladas.

### 4.1. Reconstrucciones metabólicas

El proceso de reconstrucción de un modelo a escala genómica es extenso y requiere una gran cantidad de información experimental y conocimiento exhaustivo del organismo a modelar. Las bacterias modeladas en este trabajo han sido poco estudiadas por la comunidad científica y existen escasos registros experimentales que incluyan información sobre su actividad metabólica, enzimas, expresión de genes, consumo de fuentes de carbono, auxotrofías, entre otros, información esencial para mejorar las capacidades predictivas de las reconstrucciones (Palsson, 2015).

En particular, en este trabajo se realizó la reconstrucción de la cepa *Phocaeicola dorei* 5\_1\_36/D4, de la cual no existen antecedentes previos de modelamiento a escala genómica. A través del análisis de genómica comparativa se logró reconstruir la red bioquímica de la cepa con sus respectivas reglas de asociación gen-proteína-reacción. El análisis comparativo con las bacterias modeladas en el proyecto AGORA (Magnúsdóttir et al., 2017) permitió completar rutas metabólicas presentes en la cepa de interés y que le diferencian de *Phocaeicola dorei* DSM-17855, la única bacteria de la misma especie que ha sido modelada a la fecha. Por otro lado, se realizó una revisión que permitió complementar la información de los transportadores representados en la red estequiométrica.

En el caso de *Lachnoclostridium symbiosum* WAL-14673, no hubo modificaciones respecto a la reconstrucción obtenida desde AGORA, pues no se encontró evidencia experimental que pudiera complementar la información contenida en la reconstrucción ni información experimental sobre sus especies cercanas. La búsqueda de glicosil hidrolasas mostró alta variación en el perfil de las enzimas de los miembros del cluster XIVA, lo que dificulta la extrapolación de los datos experimentales de una cepa a otra debido a las inherentes diferencias en el metabolismo de carbohidratos. En particular, la anotación de los genes de este microorganismo no es muy precisa debido a la baja representatividad de las secuencias codificantes de la familia *Lachnospiraceae* en las bases de datos de secuencias biológicas. Las secuencias de referencia usadas por CAZy y DBCAN-PUL provienen en su mayoría de *Bacteroides* (Ausland et al., 2021), lo que disminuye la precisión de la anotación y la búsqueda de enzimas que puedan degradar fibras alimenticias. No obstante, se logró modelar el consumo de las fibras que son objetivo de este trabajo, inulina y xilano.

#### **4.1.1. Validación de las reconstrucciones**

La principal métrica para estudiar la capacidad predictiva de los modelos, desde una arista cualitativa, es el *F-score*, el cual corresponde a un promedio ponderado de la precisión y la sensibilidad. Los resultados de la validación sugieren que las reconstrucciones requieren más información y refinamiento, pues se obtuvieron valores de 77% y 60% para *Phocaeicola dorei* y *Lachnoclostridium symbiosum*, respectivamente. Otros modelos publicados en literatura, como por ejemplo los modelos de *L. plantarum* y *Oenococcus oeni* poseen valores de *F-score* superiores al 85% (Teusink et al., 2006; Mendoza et al., 2017).

#### **4.2. Modelamiento de fibras alimenticias y su consumo en bacterias**

Una dificultad adicional con la modelación a escala genómica de fibras complejas es la estructura de la fibra y los mecanismos de degradación utilizados por las bacterias estudiadas. En general las fibras alimenticias corresponden a cadenas con distintos grados de polimerización. La hidrólisis de estas cadenas por la acción enzimática genera

una mezcla de oligosacáridos pequeños con distintos largos de cadena, lo que dificulta la representación de la estequiometría de la hidrólisis. Esto agrega dificultades al momento de modelar dietas del microbioma intestinal, además de la dificultad inherente por la complejidad del medio de cultivo (Marinos, Kaleta, & Waschina, 2020). Por otro lado, existen factores regulatorios y cinéticos que no son capturados por los GSMMs ni por las reconstrucciones, tales como la tasa de consumo de la fibra, las proporciones de los oligosacáridos derivados de su hidrólisis según el largo de cadena, el efecto de los reguladores transcripcionales y la represión catabólica, entre otros.

En el caso particular de los modelos, se hizo una simplificación extraordinaria al resumir la degradación de una fibra compleja a cierto número de monómeros en una misma ecuación estequiométrica. Por otro lado, los mecanismos para incorporar oligosacáridos al citosol en las bacterias de interés aún son desconocidos, al igual que la acción de algunas enzimas recuperadas desde la anotación de los genomas. Esto hace necesario un trabajo de modelación más exhaustivo para incorporar el metabolismo de otras fibras alimenticias a las reconstrucciones.

### **4.3. Simulaciones en monocultivo**

Respecto a las simulaciones de monocultivo, se observó que *Phocaicola dorei* posee las mismas tasas de crecimiento en ambas fuentes de carbono. Esto no concuerda con los resultados experimentales (Hirmas, 2021), en donde se observó un mejor crecimiento en inulina. Múltiples hipótesis pueden explicar esta diferencia, entre ellas la velocidad de la cinética enzimática, la concentración de las enzimas responsables de la hidrólisis de cada fibra, la capacidad de los flujos de consumo empleados en la simulación, entre otras variables. No obstante, las simulaciones y la evidencia experimental coinciden en que este microorganismo consume ambas fibras.

Las tasas de crecimiento obtenidas para *Lachnoclostridium symbiosum* indican que en medio mZMB la bacteria no se ve favorecida por la presencia de inulina y xilano y su

tasa de crecimiento alcanza niveles bajos. Del análisis de FVA se desprende que dicho crecimiento se deriva de la fermentación de aminoácidos, abundantes en el medio complejo mZMB, y se relaciona con el modelado del consumo de las fibras, para el cual no se agregaron reacciones en este microorganismo. Este resultado es confirmado por los monocultivos realizados en el trabajo de Hirmas (Hirmas, 2021), en donde no se observan mejoras en el crecimiento de esta bacteria en xilano respecto a su cultivo en mZMB sin fuente de carbono. En el caso de inulina, los experimentos del trabajo citado indican que existe un pequeño crecimiento de *Lachnoclostridium symbiosum*, el cual no se ve reflejado en las simulaciones. Esto debido a que no se encontraron evidencias suficientes para describir el mecanismo de consumo de fructanos, oligosacáridos derivados de inulina, en el genoma de la cepa de interés.

#### **4.4. Simulaciones en cocultivo**

En lo que concierne a las interacciones metabólicas encontradas en las simulaciones de cocultivo, primero se realizará la comparación con el trabajo de Hirmas (Hirmas, 2021) donde se cuantificó la producción de SCFA por HPLC. En el estudio se demostró la existencia de una posible interacción de alimentación cruzada para lactato y succinato, tanto en inulina como en xilano. Esta interacción se obtuvo tanto en las simulaciones maximizando el crecimiento de la comunidad como en el análisis de las distribuciones de los flujos.

En el caso de acetato, en el mismo trabajo experimental se observó la producción del compuesto por parte de las dos bacterias. En contraste, las simulaciones predijeron una interacción metabólica de acetato tanto en el análisis de distribución de flujos como en la solución óptima de SteadyCom. Se ha documentado que miembros del género *Clostridium* consumen lactato y acetato para producir butirato (Louis & Flint, 2017), motivo por el cual esta interacción aún puede ser relevante en el cocultivo de estas dos bacterias.

Respecto a las interacciones que involucran aminoácidos, su validación requiere un estudio detallado para entender si existen condiciones de auxotrofia en las especies de estudio y así profundizar en la esencialidad de cada uno de ellos. Las reconstrucciones que sirvieron como punto de partida para este trabajo fueron refinadas por un proceso de *gap-filling* para completar las rutas metabólicas que son indispensables para las simulaciones de crecimiento de los microorganismos. Este proceso añadió reacciones para la biosíntesis de aminoácidos que no están respaldadas por evidencia genómica (por ejemplo, la reacción CSTS de la ruta de biosíntesis de metionina) y que no pudieron ser descartadas por análisis bioinformático ni evidencia experimental. Por otro lado, el medio utilizado en la simulación es un medio complejo que incluye triptona, un compuesto obtenido por la degradación pancreática de la caseína, y contiene una gran carga de aminoácidos esenciales libres. Es posible que algunas de las interacciones tengan confundido el efecto del consumo de los aminoácidos del medio con el producido por la bacteria que las acompaña en el cocultivo.

Adicionalmente, se predijeron algunas interacciones que son incongruentes con la biología de los microorganismos estudiados. Por ejemplo, el intercambio de protones y agua no debería estar condicionado a una interacción pues estos se encuentran abundantes en el medio de cultivo, por lo que se sugiere que estos corresponden a artificios del muestreo aleatorio y de las restricciones impuestas sobre cada bacteria en la matriz estequiométrica conjunta. De igual manera, en el caso del intercambio de dióxido de carbono, aunque las bacterias del microbioma intestinal suelen ser productoras de este gas no tienen la capacidad para fijarlo.

Por otro lado, las simulaciones indican que la producción de butirato no es independiente de la fuente de carbono, y en particular, que solo existe producción en el caso del cocultivo en xilano. Estos resultados no concuerdan con los experimentos en cocultivo de Hirmas (Hirmas, 2021), donde no se observó un cambio significativo en la cantidad de butirato producida en xilano e inulina, pero sí se detectó producción de butirato en ambas fibras.

Respecto a los valores predichos para el intercambio de lactato en el cocultivo en inulina (ver tabla 3.6), se puede ver que existe un efecto del modelamiento del consumo de la fibra sobre el crecimiento de la bacteria. En el caso de inulina, solo se modeló en *Phocaeicola dorei* la capacidad para degradar la fuente de carbono intracelularmente. Ya que *Lachnospirillum symbiosum* no tiene acceso a la fructosa derivada de la hidrólisis de la fibra, la optimización favorece el consumo de fuentes de carbono alternativas que ayuden a aumentar el crecimiento global de la comunidad (esto corresponde a la función objetivo de SteadyCom). En particular, la reconstrucción de *Lachnospirillum symbiosum* posee la reacción LDH\_D correspondiente a la enzima D-lactato deshidrogenasa, que transforma D-lactato en piruvato y utiliza el cofactor NAD. La solución de SteadyCom indica que existe un sobreflujo en la reacción LDH\_D en inulina comparada con la solución de cocultivo en xilano. La misma situación se observa en la enzima L-alanina deshidrogenasa (ALAD\_L), que produce piruvato a partir de L-alanina. De esta manera, se observa que la solución posee grandes flujos para alanina y lactato con el fin de producir piruvato, un metabolito que es utilizado como precursor para la síntesis de aminoácidos y otros compuestos que son requeridos por la ecuación de biomasa del modelo. La sustitución de inulina por fructosa en el medio de cultivo comprueba que en presencia de fuentes de carbono más eficientes, los flujos se normalizan a valores menores a  $100 \frac{mmol}{h}$  (ver interacciones para fructosa en el Anexo 2, tabla B.3).

A través de esta evidencia se observa que las reconstrucciones tienen capacidad predictiva sobre las interacciones metabólicas de alimentación cruzada en cocultivos de bacterias de la microbiota intestinal. En particular, se logra predecir qué tipo de interacciones favorecen al productor de butirato *Lachnospirillum symbiosum*, y como la presencia de *Phocaeicola dorei* favorece el crecimiento de dicha bacteria. Estimular el crecimiento de los productores de butirato es clave para mejorar la salud del huésped a través de la acción de este SCFA sobre la respuesta antiinflamatoria y la fortaleza del epitelio intestinal.

## 5. CONCLUSIONES

*Lachnoclostridium symbiosum* es un productor de butirato, uno de los SCFA más importantes para la salud humana debido a su potencial antiinflamatorio y su capacidad para restaurar el epitelio intestinal. Esta bacteria se ve beneficiada por la presencia de fermentadores primarios, capaces de degradar las fibras alimenticias y producir metabolitos que sirven como fuente de energía para el crecimiento y producción de SCFA con fines terapéuticos.

En este estudio se identificaron *in silico* las relaciones metabólicas de *Lachnoclostridium symbiosum* WAL-14673 junto a *Phocaeicola dorei*, una de las especies más abundantes del microbioma intestinal. A través de un enfoque de genómica comparativa y modelamiento metabólico se generaron reconstrucciones a escala genómica y simulaciones que permitieron determinar la existencia de una interacción de alimentación cruzada de lactato y succinato al cultivar las dos bacterias en medios suplementados con inulina y xilano. Dicha interacción fue validada a nivel experimental con el trabajo realizado por Hirmas (Hirmas, 2021) y se obtuvo una serie de interacciones de aminoácidos que podrían ser candidatas para próximos estudios experimentales. Por otro lado, se sugiere que existe una interacción de acetato que no ha sido comprobada experimentalmente cuando las bacterias se cultivan en las fibras de interés.

Adicionalmente, se generó una reconstrucción específica para la cepa *Phocaeicola dorei* 5\_1\_36/D4, que puede ser usada en simulaciones posteriores de comunidad. Junto a ello se generó código para buscar interacciones metabólicas en simulaciones de comunidad con SteadyCom y con el muestreo aleatorio del espacio de soluciones factibles de la matriz estequiométrica de la comunidad simulada.

A pesar de que las reconstrucciones generadas requieren más información para mejorar la precisión de las predicciones y representar a cabalidad el metabolismo de los organismos modelados, son un buen punto de partida para continuar el trabajo de modelamiento de las especies del microbioma intestinal y para avanzar en la comprensión de sus interacciones.

## REFERENCIAS

- Aite, M., Chevallier, M., Frioux, C., Trottier, C., Got, J., Cortés, M. P., ... Siegel, A. (2018, May). Traceability, reproducibility and wiki-exploration for “à-la-carte” reconstructions of genome-scale metabolic models. *PLOS Computational Biology*, *14*(5), e1006146. Retrieved 2021-10-13, from <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1006146> (Publisher: Public Library of Science) doi: 10.1371/journal.pcbi.1006146
- Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., ... McArthur, A. G. (2020, January). CARD 2020: antibiotic resistance surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Research*, *48*(D1), D517–D525. doi: 10.1093/nar/gkz935
- Altamirano, A., Saa, P. A., & Garrido, D. (2020, January). Inferring composition and function of the human gut microbiome in time and space: A review of genome-scale metabolic modelling tools. *Computational and Structural Biotechnology Journal*, *18*, 3897–3904. Retrieved 2021-11-08, from <https://www.sciencedirect.com/science/article/pii/S200103702030502X> doi: 10.1016/j.csbj.2020.11.035
- Ausland, C., Zheng, J., Yi, H., Yang, B., Li, T., Feng, X., ... Yin, Y. (2021, January). dbCAN-PUL: a database of experimentally characterized CAZyme gene clusters and their substrates. *Nucleic Acids Research*, *49*(D1), D523–D528. Retrieved 2021-11-05, from <https://doi.org/10.1093/nar/gkaa742> doi: 10.1093/nar/gkaa742
- Bakir, M. A., Sakamoto, M., Kitahara, M., Matsumoto, M., & Benno, Y. . (2006). *Bacteroides dorei* sp. nov., isolated from human faeces. *International Journal of Systematic and Evolutionary Microbiology*, *56*(7), 1639–1643. Retrieved 2021-11-12, from <https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijms.0.64257-0> (Publisher: Microbiology Society,) doi:

10.1099/ijrs.0.64257-0

Bosscher, D., Loo, J. V., & Franck, A. (2006, December). Inulin and oligofructose as prebiotics in the prevention of intestinal infections and diseases. *Nutrition Research Reviews*, *19*(2), 216–226. Retrieved 2021-11-08, from <https://www.cambridge.org/core/journals/nutrition-research-reviews/article/inulin-and-oligofructose-as-prebiotics-in-the-prevention-of-intestinal-infections-and-diseases/640DDDED310BBB4545CAFD37B8F09563> (Publisher: Cambridge University Press) doi: 10.1017/S0954422407249686

Burger-van Paassen, N., Vincent, A., Puiman, P., van der Sluis, M., Bouma, J., Boehm, G., ... Renes, I. (2009, May). The regulation of intestinal mucin MUC2 expression by short-chain fatty acids: implications for epithelial protection. *Biochemical Journal*, *420*(2), 211–219. Retrieved 2021-11-10, from <https://doi.org/10.1042/BJ20082222> doi: 10.1042/BJ20082222

Chan, S. H. J., Simons, M. N., & Maranas, C. D. (2017, May). SteadyCom: Predicting microbial abundances while ensuring community stability. *PLOS Computational Biology*, *13*(5), e1005539. Retrieved 2021-11-05, from <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005539> (Publisher: Public Library of Science) doi: 10.1371/journal.pcbi.1005539

Claesson, M. J., Cusack, S., O'Sullivan, O., Greene-Diniz, R., Weerd, H. d., Flannery, E., ... O'Toole, P. W. (2011, March). Composition, variability, and temporal stability of the intestinal microbiota of the elderly. *Proceedings of the National Academy of Sciences*, *108*(Supplement 1), 4586–4591. Retrieved 2021-11-08, from [https://www.pnas.org/content/108/Supplement\\_1/4586](https://www.pnas.org/content/108/Supplement_1/4586) (Publisher: National Academy of Sciences Section: Colloquium Paper) doi: 10.1073/pnas.1000097107

Darriba, D., Posada, D., Kozlov, A. M., Stamatakis, A., Morel, B., & Flouri, T. (2020,

January). ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. *Molecular Biology and Evolution*, 37(1), 291–294. Retrieved 2021-11-11, from <https://doi.org/10.1093/molbev/msz189> doi: 10.1093/molbev/msz189

Davis, J. J., Wattam, A. R., Aziz, R. K., Brettin, T., Butler, R., Butler, R. M., ... Stevens, R. (2020, January). The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. *Nucleic Acids Research*, 48(D1), D606–D612. Retrieved 2021-11-04, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7145515/> doi: 10.1093/nar/gkz943

Davis-Richardson, A. G., & Triplett, E. W. (2015, July). A model for the role of gut bacteria in the development of autoimmunity for type 1 diabetes. *Diabetologia*, 58(7), 1386–1393. Retrieved 2021-11-10, from <https://doi.org/10.1007/s00125-015-3614-8> doi: 10.1007/s00125-015-3614-8

Diener, C., Gibbons, S. M., & Resendis-Antonio, O. (2020). MICOM: Metagenome-Scale Modeling To Infer Metabolic Interactions in the Gut Microbiota. *mSystems*, 5(1), e00606–19. Retrieved 2021-11-08, from <https://journals.asm.org/doi/full/10.1128/mSystems.00606-19> (Publisher: American Society for Microbiology) doi: 10.1128/mSystems.00606-19

Ebrahim, A., Lerman, J. A., Palsson, B. ., & Hyduke, D. R. (2013, August). COBRAPy: COntstraints-Based Reconstruction and Analysis for Python. *BMC Systems Biology*, 7(1), 74. Retrieved 2021-11-05, from <https://doi.org/10.1186/1752-0509-7-74> doi: 10.1186/1752-0509-7-74

Elbourne, L., Tetu, S. G., Hassan, K. A., & Paulsen, I. T. (2017, January). TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Research*, 45(D1), D320–D324. Retrieved 2021-11-15, from <https://doi.org/10.1093/nar/gkw1068> doi: 10.1093/nar/gkw1068

Emms, D. M., & Kelly, S. (2019, November). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology*, 20(1), 238. Retrieved 2021-10-13, from <https://doi.org/10.1186/s13059-019-1832-y> doi: 10.1186/s13059-019-1832-y

Folsom, J. P., & Carlson, R. P. (2015, August). Physiological, biomass elemental composition and proteomic analyses of *Escherichia coli* ammonium-limited chemostat growth, and comparison with iron- and glucose-limited chemostat growth. *Microbiology*, 161(Pt 8), 1659–1670. Retrieved 2021-11-09, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4681042/> doi: 10.1099/mic.0.000118

Gao, G., Cao, J., Mi, L., Feng, D., Deng, Q., Sun, X., ... Wang, J. (2021, September). BdPUL12 depolymerizes beta-mannan-like glycans into mannoooligosaccharides and mannose, which serve as carbon sources for *Bacteroides dorei* and gut probiotics. *International Journal of Biological Macromolecules*, 187, 664–674. Retrieved 2021-11-06, from <https://linkinghub.elsevier.com/retrieve/pii/S0141813021016330> doi: 10.1016/j.ijbiomac.2021.07.172

Graf, D., Di Cagno, R., Fåk, F., Flint, H. J., Nyman, M., Saarela, M., & Watzl, B. (2015, February). Contribution of diet to the composition of the human gut microbiota. *Microbial Ecology in Health & Disease*, 26(0). Retrieved 2021-11-08, from <http://www.microbecolhealthdis.net/index.php/mehd/article/view/26164> doi: 10.3402/mehd.v26.26164

Gutiérrez, N., & Garrido, D. (2019). Species Deletions from Microbiome Consortia Reveal Key Metabolic Interactions between Gut Microbes. *mSystems*, 4(4), e00185–19. Retrieved 2021-11-08, from <https://journals.asm.org/doi/full/10.1128/mSystems.00185-19> (Publisher: American Society for Microbiology) doi: 10.1128/mSystems.00185-19

Haas, K. N., & Blanchard, J. L. (2020, January). Reclassification of the *Clostridium clostridioforme* and *Clostridium sphenoides* clades as *Enterocloster* gen. nov. and *Lacrimispora* gen. nov., including reclassification of 15 taxa. *International Journal of Systematic and Evolutionary Microbiology*, 70(1), 23–34. Retrieved 2021-10-10, from <https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/ijsem.0.003698> doi: 10.1099/ijsem.0.003698

Haraldsdóttir, H. S., Cousins, B., Thiele, I., Fleming, R. M., & Vempala, S. (2017, June). CHRR: coordinate hit-and-run with rounding for uniform sampling of constraint-based models. *Bioinformatics*, 33(11), 1741–1743. Retrieved 2021-11-09, from <https://academic.oup.com/bioinformatics/article/33/11/1741/2964731> doi: 10.1093/bioinformatics/btx052

Heintz-Buschart, A., & Wilmes, P. (2018, July). Human Gut Microbiome: Function Matters. *Trends in Microbiology*, 26(7), 563–574. Retrieved 2021-11-08, from <https://www.sciencedirect.com/science/article/pii/S0966842X17302512> doi: 10.1016/j.tim.2017.11.002

Hirmas, B. (2021). *Bi-directional cultures reveal metabolic cross-feeding between gut commensals* (Master's thesis). Retrieved 2021-11-22, from <https://repositorio.uc.cl/handle/11534/62472> (Accepted: 2021-09-15T15:50:14Z)

Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., ... Banfield, J. F. (2016, May). A new view of the tree of life. *Nature Microbiology*, 1(5), 16048. Retrieved 2021-11-04, from <http://www.nature.com/articles/nmicrobiol201648> doi: 10.1038/nmicrobiol.2016.48

Kaneuchi, C., Watanabe, K., Terada, A., Benno, Y., & Mitsuoka, T. . (1976). Taxonomic Study of *Bacteroides clostridiiformis* subsp. *clostridiiformis* (Burri and Ankersmit) Holdeman and Moore and of Related Organisms: Proposal of *Clostridium clostridiiformis* (Burri and Ankersmit) comb. nov. and *Clostridium symbiosum* (Stevens) comb.

nov. *International Journal of Systematic and Evolutionary Microbiology*, 26(2), 195–204. Retrieved 2021-11-16, from <https://www.microbiologyresearch.org/content/journal/ijsem/10.1099/00207713-26-2-195> (Publisher: Microbiology Society,) doi: 10.1099/00207713-26-2-195

Karp, P. D., Billington, R., Caspi, R., Fulcher, C. A., Latendresse, M., Kothari, A., ... Subhraveti, P. (2017, August). The BioCyc collection of microbial genomes and metabolic pathways. *Briefings in Bioinformatics*, 20(4), 1085–1093. Retrieved 2021-11-09, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6781571/> doi: 10.1093/bib/bbx085

Katoh, K., & Standley, D. M. (2013, April). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, 30(4), 772–780. Retrieved 2021-11-04, from <https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/mst010> doi: 10.1093/molbev/mst010

Kaur, A., Rose, D. J., Rumpagaporn, P., Patterson, J. A., & Hamaker, B. R. (2011). In Vitro Batch Fecal Fermentation Comparison of Gas and Short-Chain Fatty Acid Production Using “Slowly Fermentable” Dietary Fibers. *Journal of Food Science*, 76(5), H137–H142. Retrieved 2021-11-08, from <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1750-3841.2011.02172.x> (eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1750-3841.2011.02172.x>) doi: 10.1111/j.1750-3841.2011.02172.x

Kim, K., Choe, D., Song, Y., Kang, M., Lee, S.-G., Lee, D.-H., & Cho, B.-K. (2021, November). Engineering *Bacteroides thetaiotaomicron* to produce non-native butyrate based on a genome-scale metabolic model-guided design. *Metabolic Engineering*, 68, 174–186. Retrieved 2021-11-06, from <https://linkinghub.elsevier.com/retrieve/pii/S1096717621001567> doi: 10.1016/j.ymben.2021.10.005

Kim, M., Oh, H.-S., Park, S.-C., & Chun, J. (2014, February). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*, *64*(Pt 2), 346–351. doi: 10.1099/ijs.0.059774-0

Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019, November). RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, *35*(21), 4453–4455. Retrieved 2021-11-04, from <https://doi.org/10.1093/bioinformatics/btz305> doi: 10.1093/bioinformatics/btz305

Kumar, M., Ji, B., Babaei, P., Das, P., Lappa, D., Ramakrishnan, G., ... Nielsen, J. (2018, September). Gut microbiota dysbiosis is associated with malnutrition and reduced plasma amino acid levels: Lessons from genome-scale metabolic modeling. *Metabolic Engineering*, *49*, 128–142. Retrieved 2021-11-03, from <https://www.sciencedirect.com/science/article/pii/S1096717618302222> doi: 10.1016/j.ymben.2018.07.018

Lecerf, J.-M., Dépeint, F., Clerc, E., Dugenet, Y., Niamba, C. N., Rhazi, L., ... Pouillart, P. R. (2012, November). Xylo-oligosaccharide (XOS) in combination with inulin modulates both the intestinal environment and immune status in healthy subjects, while XOS alone only shows prebiotic properties. *British Journal of Nutrition*, *108*(10), 1847–1858. Retrieved 2021-11-08, from <https://www.cambridge.org/core/journals/british-journal-of-nutrition/article/xylooligosaccharide-xos-in-combination-with-inulin-modulates-both-the-intestinal-environment-and-immune-status-in-healthy-subjects-while-xos-alone-only-shows-prebiotic-properties/8F2D78EDB8C1B65D7935F19E4A85FB69> (Publisher: Cambridge University Press) doi: 10.1017/S0007114511007252

Li, Z., Summanen, P. H., Komoriya, T., & Finegold, S. M. (2015, November). In vitro study of the prebiotic xylooligosaccharide (XOS) on the growth of *Bifidobacterium* spp and *Lactobacillus* spp. *International Journal of Food Sciences and Nutrition*, *66*(8), 919–922. Retrieved 2021-11-08, from <https://doi.org/10.3109/09637486.2015.1064869> (Publisher: Taylor & Francis eprint: <https://doi.org/10.3109/09637486.2015.1064869>) doi: 10.3109/09637486.2015.1064869

Liu, B., Zheng, D., Jin, Q., Chen, L., & Yang, J. (2019, January). VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Research*, *47*(D1), D687–D692. doi: 10.1093/nar/gky1080

Louis, P., & Flint, H. J. (2017, January). Formation of propionate and butyrate by the human colonic microbiota. *Environmental Microbiology*, *19*(1), 29–41. Retrieved 2021-11-17, from <https://onlinelibrary.wiley.com/doi/10.1111/1462-2920.13589> doi: 10.1111/1462-2920.13589

Machado, D., Andrejev, S., Tramontano, M., & Patil, K. R. (2018, September). Fast automated reconstruction of genome-scale metabolic models for microbial species and communities. *Nucleic Acids Research*, *46*(15), 7542–7553. Retrieved 2021-11-29, from <https://doi.org/10.1093/nar/gky537> doi: 10.1093/nar/gky537

Magnúsdóttir, S., Heinken, A., Kutt, L., Ravcheev, D. A., Bauer, E., Noronha, A., ... Thiele, I. (2017, January). Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nature Biotechnology*, *35*(1), 81–89. Retrieved 2021-11-04, from <http://www.nature.com/articles/nbt.3703> doi: 10.1038/nbt.3703

Marinos, G., Kaleta, C., & Waschina, S. (2020, August). Defining the nutritional input for genome-scale metabolic models: A roadmap. *PLOS ONE*, *15*(8), e0236890. Retrieved 2021-11-26, from <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0236890> (Publisher: Public Library of Science) doi: 10.1371/journal.pone.0236890

10.1371/journal.pone.0236890

Martens, E. C., Lowe, E. C., Chiang, H., Pudlo, N. A., Wu, M., McNulty, N. P., ... Gordon, J. I. (2011, December). Recognition and Degradation of Plant Cell Wall Polysaccharides by Two Human Gut Symbionts. *PLOS Biology*, *9*(12), e1001221. Retrieved 2021-11-11, from <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1001221> (Publisher: Public Library of Science) doi: 10.1371/journal.pbio.1001221

McBride, M. J., Xie, G., Martens, E. C., Lapidus, A., Henrissat, B., Rhodes, R. G., ... Stein, J. L. (2009, November). Novel Features of the Polysaccharide-Digesting Gliding Bacterium *Flavobacterium johnsoniae* as Revealed by Genome Sequence Analysis. *Applied and Environmental Microbiology*, *75*(21), 6864–6875. Retrieved 2021-11-16, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2772454/> doi: 10.1128/AEM.01495-09

Medina, D. A., Pinto, F., Ovalle, A., Thomson, P., & Garrido, D. (2017, October). Prebiotics Mediate Microbial Interactions in a Consortium of the Infant Gut Microbiome. *International Journal of Molecular Sciences*, *18*(10), E2095. doi: 10.3390/ijms18102095

Mendoza, S. N., Cañón, P. M., Contreras, A., Ribbeck, M., & Agosín, E. (2017, March). Genome-Scale Reconstruction of the Metabolic Network in *Oenococcus oeni* to Assess Wine Malolactic Fermentation. *Frontiers in Microbiology*, *8*, 534. Retrieved 2021-11-15, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5372704/> doi: 10.3389/fmicb.2017.00534

Mendoza, S. N., Olivier, B. G., Molenaar, D., & Teusink, B. (2019, August). A systematic assessment of current genome-scale metabolic reconstruction tools. *Genome Biology*, *20*(1), 158. Retrieved 2021-11-29, from <https://doi.org/10.1186/s13059-019-1769-1> doi: 10.1186/s13059-019-1769-1

Moretti, S., Martin, O., Van Du Tran, T., Bridge, A., Morgat, A., & Pagni, M. (2016,

January). MetaNetX/MNXref – reconciliation of metabolites and biochemical reactions to bring together genome-scale metabolic networks. *Nucleic Acids Research*, 44(D1), D523–D526. Retrieved 2021-11-29, from <https://doi.org/10.1093/nar/gkv1117> doi: 10.1093/nar/gkv1117

Moya, A., & Ferrer, M. (2016, May). Functional Redundancy-Induced Stability of Gut Microbiota Subjected to Disturbance. *Trends in Microbiology*, 24(5), 402–413. Retrieved 2021-11-08, from <https://www.sciencedirect.com/science/article/pii/S0966842X16000263> doi: 10.1016/j.tim.2016.02.002

Niness, K. R. (1999, July). Inulin and Oligofructose: What Are They? *The Journal of Nutrition*, 129(7), 1402S–1406S. Retrieved 2021-11-08, from <https://doi.org/10.1093/jn/129.7.1402S> doi: 10.1093/jn/129.7.1402S

Noronha, A., Modamio, J., Jarosz, Y., Guerard, E., Sompairac, N., Preciat, G., ... Thiele, I. (2019, January). The Virtual Metabolic Human database: integrating human and gut microbiome metabolism with nutrition and disease. *Nucleic Acids Research*, 47(D1), D614–D624. Retrieved 2021-11-04, from <https://doi.org/10.1093/nar/gky992> doi: 10.1093/nar/gky992

Norsigian, C. J., Pusarla, N., McConn, J. L., Yurkovich, J. T., Dräger, A., Palsson, B. O., & King, Z. (2020, January). BiGG Models 2020: multi-strain genome-scale models and expansion across the phylogenetic tree. *Nucleic Acids Research*, 48(D1), D402–D406. Retrieved 2021-11-29, from <https://doi.org/10.1093/nar/gkz1054> doi: 10.1093/nar/gkz1054

Orth, J. D., Thiele, I., & Palsson, B. . (2010, March). What is flux balance analysis? *Nature Biotechnology*, 28(3), 245–248. Retrieved 2021-10-13, from <https://www.nature.com/articles/nbt.1614> doi: 10.1038/nbt.1614

Palsson, B. . (2015). *Systems Biology: Constraint-based Reconstruction and Analysis*. Cambridge University Press. doi: 10.1017/CBO9781139854610

Pritchard, L., Glover, R. H., Humphris, S., Elphinstone, J. G., & Toth, I. K. (2015, December). Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Analytical Methods*, 8(1), 12–24. Retrieved 2021-11-04, from <https://pubs.rsc.org/en/content/articlelanding/2016/ay/c5ay02550h> (Publisher: The Royal Society of Chemistry) doi: 10.1039/C5AY02550H

R Core Team. (2021). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from <https://www.R-project.org/>

Rosario, D., Benfeitas, R., Bidkhor, G., Zhang, C., Uhlen, M., Shoaie, S., & Mardinoglu, A. (2018). Understanding the Representative Gut Microbiota Dysbiosis in Metformin-Treated Type 2 Diabetes Patients Using Genome-Scale Metabolic Modeling. *Frontiers in Physiology*, 9, 775. Retrieved 2021-11-04, from <https://www.frontiersin.org/article/10.3389/fphys.2018.00775> doi: 10.3389/fphys.2018.00775

Schmidt, T. S. B., Raes, J., & Bork, P. (2018, March). The Human Gut Microbiome: From Association to Modulation. *Cell*, 172(6), 1198–1215. Retrieved 2021-11-08, from <https://www.sciencedirect.com/science/article/pii/S0092867418302216> doi: 10.1016/j.cell.2018.02.044

Seaver, S. M. D., Liu, F., Zhang, Q., Jeffryes, J., Faria, J. P., Edirisinghe, J. N., ... Henry, C. S. (2021, January). The ModelSEED Biochemistry Database for the integration of metabolic annotations and the reconstruction, comparison and analysis of metabolic models for plants, fungi and microbes. *Nucleic Acids Research*, 49(D1), D575–D588. Retrieved 2021-11-09, from <https://doi.org/10.1093/nar/gkaa746> doi: 10.1093/nar/gkaa746

Sen, P., & Orešič, M. (2019, February). Metabolic Modeling of Human Gut Microbiota on a Genome Scale: An Overview. *Metabolites*, 9(2), 22. Retrieved 2021-11-02, from

<https://www.mdpi.com/2218-1989/9/2/22> (Number: 2 Publisher: Multidisciplinary Digital Publishing Institute) doi: 10.3390/metabo9020022

Singh, R. D., Banerjee, J., & Arora, A. (2015, January). Prebiotic potential of oligosaccharides: A focus on xylan derived oligosaccharides. *Bioactive Carbohydrates and Dietary Fibre*, 5(1), 19–30. Retrieved 2021-11-08, from <https://www.sciencedirect.com/science/article/pii/S2212619814000552> doi: 10.1016/j.bcdf.2014.11.003

Swainston, N., Smallbone, K., Hefzi, H., Dobson, P. D., Brewer, J., Hanscho, M., ... Mendes, P. (2016, June). Recon 2.2: from reconstruction to model of human metabolism. *Metabolomics*, 12(7), 109. Retrieved 2021-11-29, from <https://doi.org/10.1007/s11306-016-1051-4> doi: 10.1007/s11306-016-1051-4

Tauzin, A. S., Laville, E., Xiao, Y., Nouaille, S., Le Bourgeois, P., Heux, S., ... Bordes, F. (2016). Functional characterization of a gene locus from an uncultured gut *Bacteroides* conferring xylo-oligosaccharides utilization to *Escherichia coli*. *Molecular Microbiology*, 102(4), 579–592. Retrieved 2021-11-10, from <https://onlinelibrary.wiley.com/doi/abs/10.1111/mmi.13480> (eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/mmi.13480>) doi: 10.1111/mmi.13480

Teusink, B., Wiersma, A., Molenaar, D., Francke, C., Vos, W. M. d., Siezen, R. J., & Smid, E. J. (2006, December). Analysis of Growth of *Lactobacillus plantarum* WCFS1 on a Complex Medium Using a Genome-scale Metabolic Model \*. *Journal of Biological Chemistry*, 281(52), 40041–40048. Retrieved 2021-11-26, from [https://www.jbc.org/article/S0021-9258\(20\)76839-0/abstract](https://www.jbc.org/article/S0021-9258(20)76839-0/abstract) (Publisher: Elsevier) doi: 10.1074/jbc.M606263200

Thiele, I., & Palsson, B. . (2010, January). A protocol for generating a high-quality

genome-scale metabolic reconstruction. *Nature Protocols*, 5(1), 93–121. Retrieved 2021-11-09, from <http://www.nature.com/articles/nprot.2009.203> doi: 10.1038/nprot.2009.203

Thomson, P., Medina, D. A., Ortúzar, V., Gotteland, M., & Garrido, D. (2018, July). Anti-inflammatory effect of microbial consortia during the utilization of dietary polysaccharides. *Food Research International*, 109, 14–23. Retrieved 2021-11-11, from <https://www.sciencedirect.com/science/article/pii/S0963996918302795> doi: 10.1016/j.foodres.2018.04.008

Thursby, E., & Juge, N. (2017, May). Introduction to the human gut microbiota. *Biochemical Journal*, 474(11), 1823–1836. Retrieved 2021-11-08, from <https://doi.org/10.1042/BCJ20160510> doi: 10.1042/BCJ20160510

Tsujikawa, Y., Ishikawa, S., Sakane, I., Yoshida, K.-i., & Osawa, R. (2021, August). Identification of genes encoding a novel ABC transporter in *Lactobacillus delbrueckii* for inulin polymers uptake. *Scientific Reports*, 11(1), 16007. Retrieved 2021-11-17, from <https://www.nature.com/articles/s41598-021-95356-1> (Bandiera\_abtest: a Cc\_license\_type: cc\_by Cg\_type: Nature Research Journals Number: 1 Primary\_atype: Research Publisher: Nature Publishing Group Subject\_term: Bacteria;Bacterial evolution;Bacterial systems biology Subject\_term.id: bacteria;bacterial-evolution;bacterial-systems-biology) doi: 10.1038/s41598-021-95356-1

Vandeputte, D., Falony, G., Vieira-Silva, S., Wang, J., Sailer, M., Theis, S., ... Raes, J. (2017, November). Prebiotic inulin-type fructans induce specific changes in the human gut microbiota. *Gut*, 66(11), 1968–1974. Retrieved 2021-11-08, from <https://gut.bmj.com/lookup/doi/10.1136/gutjnl-2016-313271> doi: 10.1136/gutjnl-2016-313271

van der Beek, C. M., Canfora, E. E., Kip, A. M., Gorissen, S. H. M., Olde Damink,

S. W. M., van Eijk, H. M., ... Lenaerts, K. (2018, October). The prebiotic inulin improves substrate metabolism and promotes short-chain fatty acid production in overweight to obese men. *Metabolism*, 87, 25–35. Retrieved 2021-11-08, from <https://www.sciencedirect.com/science/article/pii/S0026049518301513> doi: 10.1016/j.metabol.2018.06.009

Vital, M., Howe, A. C., & Tiedje, J. M. (2014). Revealing the Bacterial Butyrate Synthesis Pathways by Analyzing (Meta)genomic Data. *mBio*, 5(2), e00889–14. Retrieved 2021-11-16, from <https://journals.asm.org/doi/full/10.1128/mBio.00889-14> (Publisher: American Society for Microbiology) doi: 10.1128/mBio.00889-14

Xu, Z., & Knight, R. (2015, January). Dietary effects on human gut microbiome diversity. *British Journal of Nutrition*, 113(S1), S1–S5. Retrieved 2021-11-08, from <https://www.cambridge.org/core/journals/british-journal-of-nutrition/article/dietary-effects-on-human-gut-microbiome-diversity/E9B592DB4E5C570C6D5142B99CE74D3E> (Publisher: Cambridge University Press) doi: 10.1017/S0007114514004127

Zeybek, N., Rastall, R. A., & Buyukkileci, A. O. (2020, May). Utilization of xylan-type polysaccharides in co-culture fermentations of *Bifidobacterium* and *Bacteroides* species. *Carbohydrate Polymers*, 236, 116076. Retrieved 2021-11-08, from <https://www.sciencedirect.com/science/article/pii/S0144861720302502> doi: 10.1016/j.carbpol.2020.116076

Zhang, H., Yohe, T., Huang, L., Entwistle, S., Wu, P., Yang, Z., ... Yin, Y. (2018, July). dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Research*, 46(W1), W95–W101. Retrieved 2021-11-05, from <https://doi.org/10.1093/nar/gky418> doi: 10.1093/nar/gky418

**ANEXO**



## A. INFORMACIÓN COMPLEMENTARIA SOBRE GENÓMICA COMPARATIVA Y RECONSTRUCCIONES

### A.1. *Phocaeicola dorei*

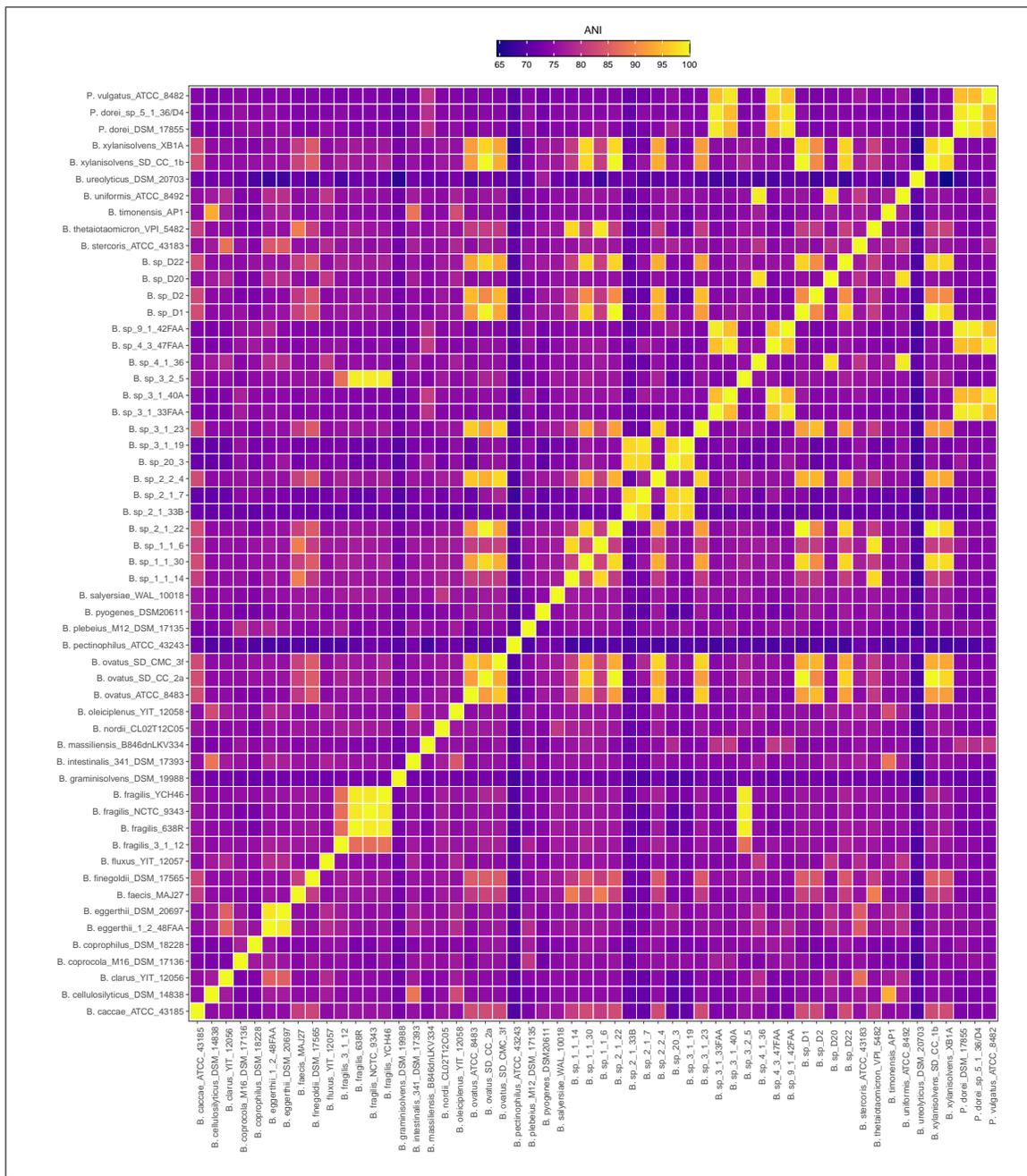


Figura A.1. Mapa de calor del porcentaje de identidad media nucleotídica (ANI) entre las especies de *Bacteroides* modeladas en AGORA.

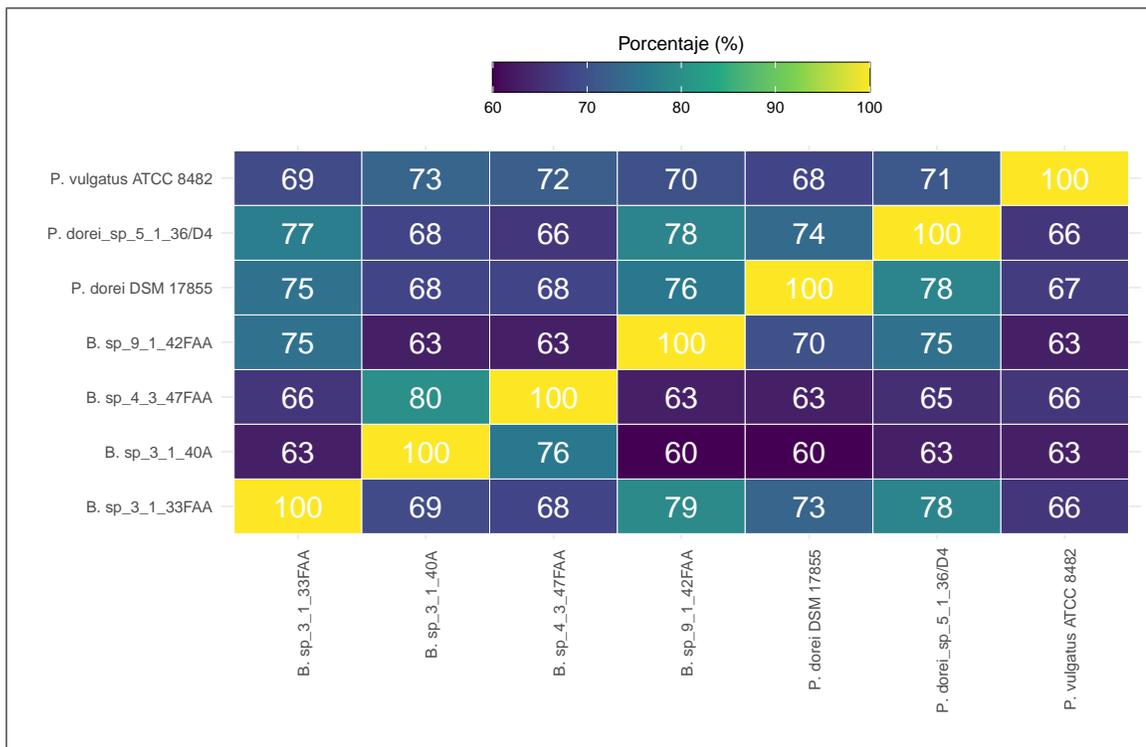


Figura A.2. Mapa de calor del porcentaje de ortólogos respecto al total de regiones codificantes entre los genomas de referencia para la reconstrucción de *Phocaeicola dorei* 5\_1\_36/D4. El eje vertical indica a la especie de referencia de la cual se obtuvo el número de genes totales, mientras que el eje horizontal indica a la especie con la cual se comparó el número de ortólogos.

Tabla A.1. Fuentes de carbono con evidencia experimental para *Phocaeicola dorei* (Bakir et al., 2006; Hirmas, 2021; Gao et al., 2021). El símbolo “+” indica que la bacteria crece en el metabolito y/o produce acidificación del medio suplementado con dicha fuente de carbono, mientras que el símbolo “-” indica lo contrario.

VMH ID	Fuente de carbono	Resultado	Referencias
EX_inulin(e)	Inulina	+	(Hirmas, 2021)
EX_arab_L(e)	L-Arabinosa	+	(Hirmas, 2021)
EX_glc_D(e)	D-Glucosa	+	(Bakir et al., 2006)
EX_lcts(e)	Lactosa	+	(Bakir et al., 2006)
EX_malt(e)	D-Maltosa	+	(Bakir et al., 2006)
EX_man(e)	D-Manosa	+	(Bakir et al., 2006)
EX_raffin(e)	Raffinosa	+	(Bakir et al., 2006)
EX_rmn(e)	L-Ramnosa	+	(Bakir et al., 2006; Gao et al., 2021)
EX_sucr(e)	Sucrosa	+	(Bakir et al., 2006)
EX_xyl_D(e)	D-Xilosa	+	(Bakir et al., 2006; Hirmas, 2021)
EX_xylan(e)	Xilano	+	(Hirmas, 2021)
EX_cellb(e)	Celobiosa	-	(Bakir et al., 2006)
EX_salcn(e)	Salicina	-	(Bakir et al., 2006)
EX_tre(e)	Trehalosa	-	(Bakir et al., 2006)
EX_fru(e)	D-Fructosa	+	(Gao et al., 2021)
EX_acgal(e)	N-Acetilgalactosamina	+	(Gao et al., 2021)
EX_acgam(e)	N-Acetil-D-glucosamina	+	(Gao et al., 2021)
EX_acnam(e)	Ácido N-acetilneuramínico	+	(Gao et al., 2021)
EX_arabinogal(e)	Arabinogalactano	+	(Gao et al., 2021)
EX_arabinoxyl(e)	Arabinoxilano	+	(Gao et al., 2021)
EX_amylopect900(e)	Amilopectina	-	(Gao et al., 2021)
EX_lmn30(e)	Laminarina	-	(Gao et al., 2021)
EX_levan1000(e)	Levan	-	(Gao et al., 2021)
EX_galur(e)	Ácido D-galacturónico	-	(Gao et al., 2021)
EX_pullulan1200(e)	Pululano	+	(Gao et al., 2021)
EX_mnl(e)	D-manitol	-	(Bakir et al., 2006)
EX_glyc(e)	Glicerol	-	(Bakir et al., 2006)
EX_sbt_D(e)	Sorbitol	-	(Bakir et al., 2006)

Tabla A.2. Ortólogos del PUL0425 para el consumo de fructanos encontrados en *Phocaeicola dorei* 5\_1\_36/D4

Locus tag	Tipo	Identidad (%)	e-valor	Anotación
BSEG_03202	CDS	52.921	1.220000e-106	-
BSEG_04149	CAZyme	63.333	0.000000e+00	GH32
BSEG_04150	TC	81.649	0.000000e+00	gnl—TC-DB—Q8A6W8—2.A.1.7.17
BSEG_04151	STP	79.661	4.480000e-177	STP—PfkB

Tabla A.3. Ortólogos del PUL0346 para el consumo de inulina encontrados en *Phocaeicola dorei* 5\_1\_36/D4

Locus tag	Tipo	Identidad (%)	e-valor	Referencia ID
BSEG_01391	-	96.000	0.000000e+00	CCG34979.1
fig—556260.3.peg.1663	-	97.059	3.730000e-19	CCG34969.1
BSEG_01393	-	97.561	0.000000e+00	CCG34983.1
BSEG_01385	-	98.160	0.000000e+00	CCG34970.1
BSEG_01390	-	98.452	0.000000e+00	CCG34978.1
BSEG_01388	-	99.296	0.000000e+00	CCG34976.1
BSEG_01389	-	99.308	0.000000e+00	CCG34977.1
BSEG_01392	-	99.590	0.000000e+00	CCG34981.1
BSEG_01387	-	99.689	0.000000e+00	CCG34975.1

## A.2. *Lachnospiraceae*

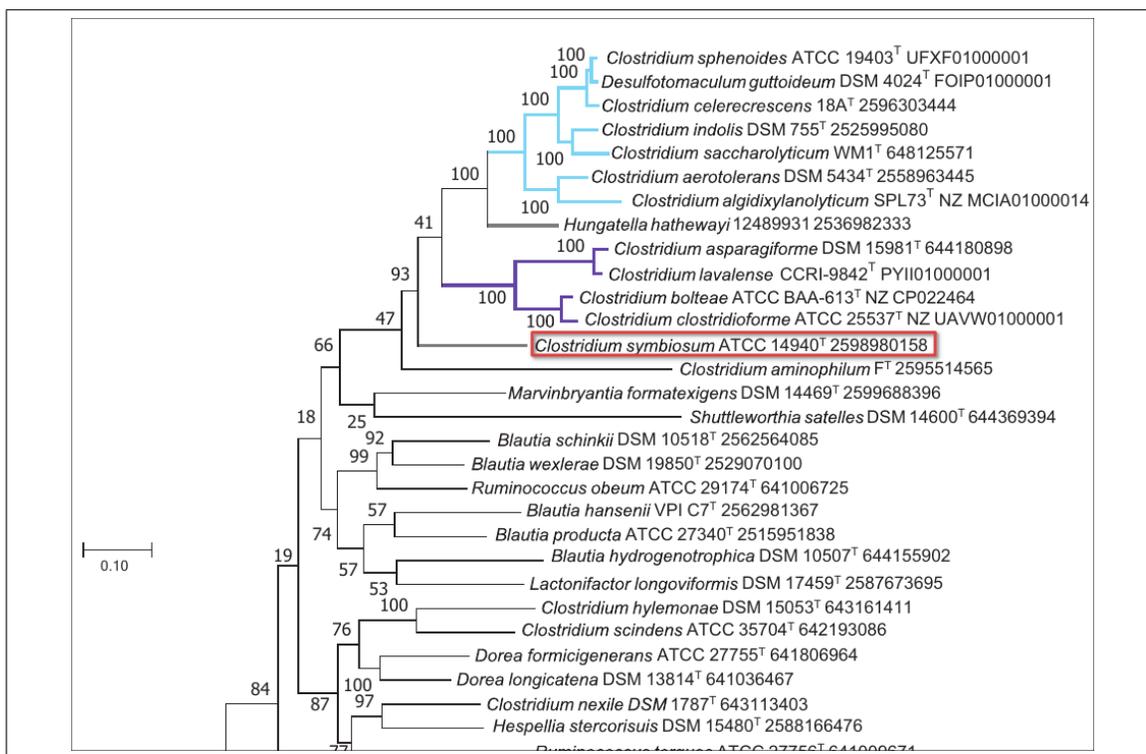


Figura A.3. Árbol filogenético con los miembros de la familia *Lachnospiraceae* basado en el gen *rpoB* (Haas & Blanchard, 2020). En rojo se destaca a la especie *Lachnospiraceae symbiosum*. Las ramas celeste y morada corresponden a los nuevos géneros propuestos en la investigación citada.

Tabla A.4. Genomas seleccionados para el análisis de genómica comparativa en las especies de la familia *Lachnospiraceae*. En total se seleccionaron 12 genomas que incluyen a la cepa de interés *Lachnoclostridium symbiosum* WAL-14673. En la tabla se indica el tamaño de los genomas y el número de contigs por ensamble. La columna CDS contiene el número de regiones codificantes de cada genoma (CDS, por sus siglas en inglés).

PATRIC ID	Genoma	Ensamble	Contigs	Tamaño (Mb)	CDS
1286362.3	<i>Clostridium celerecrescens</i> 18A	GCA_002797975.1	1	5.27	5042
1297793.6	<i>Clostridium sphenoides</i> JCM 1415	GCA_900105615.1	1	5.30	5042
1304881.3	<i>Clostridium indolis</i> DSM 755	GCA_000526995.1	1	6.38	6123
1410651.3	<i>Clostridium aerotolerans</i> DSM 5434	GCA_000687555.1	50	4.73	4373
1531.123	<i>Clostridium clostridioforme</i> NBRC 113352	GCA_006538465.1	2	5.69	6240
208479.8	<i>Clostridium bolteae</i> ATCC BAA-613	GCA_002234575.2	2	6.61	6583
460384.30	<i>Clostridium lavalense</i> MGYG-HGUT-00172	GCA_902364025.1	53	6.50	6228
518636.5	<i>Clostridium asparagiforme</i> DSM 15981	GCA_000158075.1	92	6.42	5652
58134.4	<i>Desulfotomaculum guttoideum</i> DSM 4024	GCA_900105215.1	2	5.93	5804
610130.3	<i>Clostridium saccharolyticum</i> WM1	GCA_000144625.1	1	4.66	4634
742741.3	<i>Clostridium symbiosum</i> WAL-14673	GCA_000189615.1	55	4.92	4704
94868.12	<i>Clostridium algidixylanolyticum</i> SPL73	GCA_003609635.1	35	4.64	4415

Tabla A.5. Fuentes de carbono con evidencia experimental para *Lachnoclostridium symbiosum* (Hirmas, 2021; Kaneuchi et al., 1976). El símbolo “+” indica que la bacteria crece en el metabolito y/o produce acidificación del medio suplementado con dicha fuente de carbono, mientras que el símbolo “-” indica lo contrario.

VMH ID	Fuente de carbono	Resultado	Referencias
EX_fru(e)	D-Fructosa	+	(Hirmas, 2021)
EX_lcts(e)	Lactosa	+	(Hirmas, 2021)
EX_glc_D(e)	D-Glucosa	+	(Kaneuchi et al., 1976)
EX_xyl_D(e)	D-Xilosa	-	(Hirmas, 2021)
EX_arab_L(e)	L-Arabinosa	+	(Kaneuchi et al., 1976)
EX_rib_D(e)	D-ribosa	-	(Kaneuchi et al., 1976)
EX_man(e)	D-Manosa	+	(Kaneuchi et al., 1976)
EX_gal(e)	D-Galactosa	+	(Kaneuchi et al., 1976)
EX_rmn(e)	L-Ramnosa	-	(Kaneuchi et al., 1976)
EX_sucr(e)	Sucrosa	-	(Kaneuchi et al., 1976)
EX_malt(e)	D-Maltosa	-	(Kaneuchi et al., 1976)
EX_cellb(e)	Celobiosa	-	(Kaneuchi et al., 1976)
EX_melib(e)	Melibiosa	-	(Kaneuchi et al., 1976)
EX_raffin(e)	Rafinosa	-	(Kaneuchi et al., 1976)
EX_dextrin(e)	Dextrina	-	(Kaneuchi et al., 1976)
EX_glyc(e)	Glicerol	-	(Kaneuchi et al., 1976)
EX_sbt_D(e)	Sorbitol	-	(Kaneuchi et al., 1976)
EX_salcn(e)	Salicina	-	(Kaneuchi et al., 1976)

## B. INFORMACIÓN COMPLEMENTARIA SOBRE LAS SIMULACIONES

### B.1. Validación de los modelos

Tabla B.1. Composición del medio mZMB (Medina et al., 2017) usado en las simulaciones. Los flujos de minerales, iones, vitaminas y compuestos esenciales para el crecimiento fueron fijados en 1, de acuerdo a los valores observados en las dietas de la microbiota intestinal usadas en las simulaciones del proyecto AGORA (Magnúsdóttir et al., 2017). En el caso de los aminoácidos, los flujos fueron definidos en base al requerimiento nutricional de los modelos.

Nombre reacción	VMH ID	Flujo
Exchange of L-Alanine	EX_ala.L(e)	0.251762
Exchange of L-Arginine	EX_arg.L(e)	0.290359
Exchange of L-Asparagine	EX_asn.L(e)	0.225000
Exchange of L-Aspartate	EX_asp.L(e)	0.571946
Exchange of L-Cysteine	EX_cys.L(e)	0.035089
Exchange of L-Glutamine	EX_gln.L(e)	0.820199
Exchange of L-Glutamate	EX_glu.L(e)	0.820199
Exchange of Glycine	EX_gly(e)	0.157022
Exchange of L-Histidine	EX_his.L(e)	0.200883
Exchange of L-Isoleucine	EX_ile.L(e)	0.392994
Exchange of L-Leucine	EX_leu.L(e)	0.669318
Exchange of L-Lysine	EX_lys.L(e)	0.571069
Exchange of L-Methionine	EX_met.L(e)	0.206146
Exchange of L-Phenylalanine	EX_phe.L(e)	0.358782
Exchange of L-Proline	EX_pro.L(e)	0.758794
Exchange of L-Serine	EX_ser.L(e)	0.445627
Exchange of L-Threonine	EX_thr.L(e)	0.342992
Exchange of L-Tryptophan	EX_trp.L(e)	0.092108
Exchange of L-Tyrosine	EX_tyr.L(e)	0.163163
Exchange of L-Valine	EX_val.L(e)	0.483347
Exchange of Phosphate	EX_pi(e)	1.000000
Exchange of Kalium	EX_k(e)	1.000000
Mg exchange	EX_mg2(e)	1.000000
Exchange of Iron (Fe3+)	EX_fe3(e)	1.000000
Exchange of Iron (Fe2+)	EX_fe2(e)	1.000000
Exchange of Sodium	EX_na1(e)	1.000000
Zinc exchange	EX_zn2(e)	1.000000
Mn2+ exchange	EX_mn2(e)	1.000000
Molybdate exchange	EX_mobd(e)	1.000000
Exchange of Proton	EX_h(e)	1000.000000
Exchange of Ammonia	EX_nh4(e)	1.000000
Exchange of Calcium	EX_ca2(e)	1.000000
Exchange of Sulfate	EX_so4(e)	1.000000
Co2+ exchange	EX_cobalt2(e)	1.000000
Cu2+ exchange	EX_cu2(e)	1.000000
Exchange of Chloride	EX_cl(e)	1.000000
Exchange of (R)-Pantothenate	EX_pnto.R(e)	1.000000
Exchange of Folate	EX_fol(e)	1.000000
Exchange of Nicotinate	EX_nac(e)	1.000000
Exchange of Pyridoxal	EX_pydx(e)	1.000000
4 Aminobenzoate exchange	EX_4abz(e)	1.000000
Exchange of Xanthine	EX_xan(e)	1.000000
Exchange of Myo-Inositol	EX_inost(e)	1.000000
Exchange of Oxidized Glutathione	EX_gthox(e)	1.000000

Tabla B.2. Composición del medio mZMB (Medina et al., 2017) usado en las simulaciones. Los flujos de minerales, iones, vitaminas y compuestos esenciales para el crecimiento fueron fijados en 1, de acuerdo a los valores observados en las dietas del microbioma intestinal usadas en las simulaciones del proyecto AGORA (Magnúsdóttir et al., 2017). En el caso de los aminoácidos, los flujos fueron definidos en base al requerimiento nutricional de los modelos.

Nombre reacción	VMH ID	Flujo
Exchange of Reduced Glutathione	EX_gthrd(e)	1.000000
Exchange of Biotin	EX_btn(e)	1.000000
Exchange of Riboflavin	EX_ribflv(e)	1.000000
Exchange of Thiamin	EX_thm(e)	1.000000
Exchange of Heme	EX_pheme(e)	1.000000
Siroheme exchange	EX_sheme(e)	1.000000
Exchange of Adenine	EX_ade(e)	1.000000
Exchange of Guanine	EX_gua(e)	1.000000
Exchange of Uracil	EX_ura(e)	1.000000
Hydrogen sulfide exchange	EX_h2s(e)	1.000000
Exchange of Hypoxanthine	EX_hxan(e)	1.000000
Ni <sup>2+</sup> exchange	EX_ni2(e)	1.000000
Exchange of Thymidine	EX_thymd(e)	1.000000
Exchange of Water	EX_h2o(e)	1000.000000

Tabla B.3. Interacciones metabólicas en cocultivo con mZMB y fructosa como fuente de carbono predichas por SteadyCom.

VMH ID	Phocaeicola dorei	Lachnoclostridium symbiosum
EX_ac[u]	-1.763135	11.525866
EX_co2[u]	-0.571299	0.824911
EX_etoh[u]	1.671006	-1.671006
EX_glu_L[u]	-1.682400	0.862201
EX_lac_D[u]	-0.081385	0.081385
EX_succ[u]	3.415191	-3.415191
EX_thymd[u]	-0.000010	0.160028