

UNIVERSIDAD CATÓLICA DEL ECUADOR

FACULTAD DE INGENIERÍA

**“DESARROLLO DE ALGORITMOS PARA OPTIMIZAR LA PRODUCCIÓN DE
LICOPENO EN *YARROWIA LIPOLYTICA* USANDO HERRAMIENTAS
COMPUTACIONALES DE INGENIERÍA METABÓLICA EN PYTHON”**

Previa la obtención del Título de:
MAGÍSTER EN BIOLOGÍA COMPUTACIONAL

**ELABORADO POR:
MILTON GIOVANNI QUINGA SOCASI**

**DIRECTOR:
DR. JESÚS ABRAHAM AVELAR RIVAS**

QUITO, 2023

DEDICATORIA

A mis padres Alfonso y Blanca, por su apoyo incondicional, amor y consejos a lo largo de mi vida.

.

A mis familiares, por los momentos compartidos, ayuda y la confianza depositada en mí.

A mi esposa Elena y mis hijos Alberto y Raúl. Que son la inspiración en mi vida.

Milton Quinga

A Dios, por mostrarme que pese a los inconvenientes que se presentan en la vida, siempre hay esperanza para las personas que tienen fe.

A la Pontificia Universidad Católica del Ecuador (PUCE), que hizo posible el programa de maestría y por los excelentes profesores vinculados a este posgrado

Al Doctor Abraham Avelar, por su tiempo, paciencia y apoyo científico en este proyecto de titulación.

A la Doctora María Elena Lienqueo, por su apoyo, vinculación a la Ingeniería Metabólica y guía académica, cuando estuve en Chile.

ÍNDICE GENERAL

RESUMEN	1
ABSTRACT	3
CAPÍTULO 1: INTRODUCCIÓN	5
1.1. Planteamiento del problema	5
1.2. Objetivos	7
1.2.1. Objetivo general.....	7
1.2.2. Objetivos específicos	7
1.3. Justificación	7
CAPÍTULO 2: MARCO TEÓRICO Y CONCEPTUAL	9
2.1. Antecedentes o marco referencial	9
2.2. Marco teórico	11
2.2.1. <i>Yarrowia lipolytica</i>	11
2.2.2. Síntesis de carotenoides	12
2.2.2.1. Producción de carotenoides en levaduras	12
2.2.2.2. Biosíntesis de licopeno en <i>Yarrowia lipolytica</i>	12
2.2.2.3. Ruta del mevalonato.....	13
2.2.2.4. Biosíntesis de los productos de MVA a licopeno	13
Estrategias de bioingeniería para la producción de terpenoides	15
2.2.3. Los modelos metabólicos a escala genómica permiten simular las condiciones óptimas para la producción de metabolitos	15
2.2.4. Herramientas computacionales.....	19
CAPÍTULO 3: METODOLOGÍA	19
3.1. Resumen del algoritmo empleado	20
3.2. Importación del modelo metabólico <i>iYali4</i> de <i>Yarrowia lipolytica</i>	21
3.3. Identificación de las reacciones de la ruta del mevalonato	21
3.4. Adición de rutas heterólogas	23
3.5. Optimización de la producción de licopeno (Evaluación de aceptores/ aceptores reducidos).....	26
3.6. Identificación de las reacciones involucradas en el medio de cultivo	27
3.7. Sobreexpresión de genes y optimización de la producción de licopeno	28
3.8. Eliminación de genes y optimización de la producción de licopeno	29
3.9. Gráficos de planos de fase de fenotipo (<i>Production envelopes</i>)	29
3.10. Visualización de flujos obtenidos	31
3.11. Tipo de investigación	32

	CAPÍTULO 4:
RESULTADOS Y DISCUSIÓN	33
4.1. Análisis de la optimización de la producción de licopeno usando FBA (diferentes factores)	33
4.2. Análisis de los gráficos de planos de fase de fenotipo (Production envelopes)	36
4.3. Análisis de la visualización de flujos metabólicos en Fluxer	42
CAPÍTULO 5: CONCLUSIONES Y RECOMENDACIONES	47
5.1. Conclusiones	47
5.2. Recomendaciones	48
CAPÍTULO 6: REFERENCIAS	49
ANEXOS	56

ÍNDICE DE FIGURAS

	Figura	1.
Diagrama causa efecto, espina de pescado, del problema de investigación.....		6
Figura 2. Esquema de la vía biosintética del licopeno a partir de acetil-CoA en <i>Y. lipolytica</i>		14
Figura 3. El modelado basado en restricciones.....		17
Figura 4. Esquema del algoritmo empleado para la optimización de la producción de licopeno en <i>Yarrowia lipolytica</i> en el modelo iYali4.		20
Figura 5. Resultados obtenidos en la optimización de la tasa de producción de licopeno en <i>Yarrowia lipolytica</i> usando FBA en el modelo iYali4 para diferentes factores.		35
Figura 6. Producción de licopeno y biomasa en función del oxígeno (medio de cultivo) y como fuente de carbono la glucosa (Evaluación del par aceptores/ aceptores reducidos: FAD/FADH ₂).		36
Figura 7. Producción de licopeno y biomasa en función del nitrógeno (medio de cultivo) y como fuente de carbono la glucosa (Evaluación del par aceptores/ aceptores reducidos: FAD/FADH ₂).		37
Figura 8. Producción de licopeno vs biomasa tomando un límite inferior determinado.....		38
Figura 9. Producción de licopeno vs biomasa tomando un límite inferior determinado.....		38
Figura 10. Producción de licopeno y biomasa en función del oxígeno (medio de cultivo) y como fuente de carbono la glucosa (Evaluación de la sobreexpresión del gen HMG1).		39
Figura 11. Producción de licopeno y biomasa en función del nitrógeno (medio de cultivo) y como fuente de carbono la glucosa (Evaluación de la sobreexpresión del gen HMG1).		40
Figura 12. Producción de licopeno vs biomasa tomando un límite inferior determinado.....		41
Figura 13. Producción de licopeno vs biomasa tomando un límite inferior determinado.....		41
Figura 14. Principales k rutas más cortas entre acetyl-CoA C-acetyltransferase y Lycopene demand. Flujos metabólicos obtenidos de la optimización de la producción de licopeno (Evaluación del par aceptores/ aceptores reducidos: FAD/FADH ₂).		44
Figura 15. Principales k rutas más cortas entre acetyl-CoA C-acetyltransferase y Lycopene demand. Flujos metabólicos obtenidos de la sobreexpresión de genes (gen HMG1) y optimización de la producción de licopeno.		45
Figura 16. Principales k rutas más cortas entre acetyl-CoA C-acetyltransferase y Lycopene demand. Flujos metabólicos obtenidos de la eliminación de genes (POX1 y GUT2) y optimización de la producción de licopeno		46

ÍNDICE DE TABLAS

	Tabla	1.
Reacciones de la vía biosintética del licopeno a partir de acetil-CoA hasta farnesil-PP. Identificadas en el modelo metabólico iYali4.....		22
Tabla 2. Reacciones heterólogas adicionadas mediante los comandos de COBRAPy.....		26
Tabla 3. información sobre los nutrientes del medio de crecimiento en el modelo iYali4.....		27

ÍNDICE DE ANEXOS

Anexo 1. Script desarrollado en un entorno Jupyter Notebook de Google Colab	56
Anexo 2. Resultados obtenidos en la optimización de la producción de licopeno en Yarrowia lipolytica.....	56

Aceptores (A):

También llamados aceptores de hidrógeno (aceptan electrones), son coenzimas involucradas en reacciones de bioquímicas de oxido reducción. Comprenden: el dinucleótido de nicotinamida y adenina en su estado oxidado se denomina NAD^+ , el dinucleótido de flavina y adenina en su estado oxidado se llama FAD y la nicotinamida adenina dinucleótido fosfato que es su forma oxida se llama $NADP^+$ (Xiao *et al.*, 2018).

Aceptores reducidos (AH₂): También llamados donadores de hidrógeno (ceden electrones), son las coenzimas involucradas en reacciones de bioquímicas de oxido reducción, son las formas reducidas de los aceptores: NADH, NADPH FADH₂. Estas parejas redox, son esenciales para mantener la homeostasis redox celular y para modular numerosos eventos biológicos, incluido el metabolismo celular (Xiao *et al.*, 2018).

BRENDA: BRAunschweig ENzyme DAtabase. Es una base de datos en línea, que integra información funcional y molecular de enzimas (propiedades, reacción catalizada, cinética, sustratos/productos, inhibidores, cofactores, activadores, estructura y estabilidad), basada en literatura primaria (Schomburg *et al.*, 2002).

Computational Strain Optimization, CSO: Este enfoque consiste en identificar el conjunto de modificaciones genéticas, a ser introducidas en un organismo, que optimicen un objetivo de ingeniería deseado. Por lo general, el objetivo es maximizar la producción de un compuesto de interés, mientras se asegura que el organismo permanezca viable (Pereira *et al.*, 2021).

Fluxer: Es una aplicación web gratuita y de acceso abierto para el cálculo y la visualización de redes de flujo metabólico a escala genómica. Cualquier modelo a escala del genoma basado en el lenguaje de marcado de biología de sistemas (SBML) se puede cargar en esta herramienta (Hari & Lobo, 2020).

Gráficos de planos de fase de fenotipo (*Production envelopes*): Son gráficos desarrollados en COBRApy y Cameo, que sirven para mostrar distintas fases de crecimiento óptimo con un uso diferente de dos sustratos diferente (COBRApy, 2022).

Encyclopedia of Genes and Genomes. Es una base de datos en línea, para el análisis sistemático de las funciones de los genes, que vincula la información genómica con información funcional. Contiene representaciones gráficas de procesos celulares, como el metabolismo, el transporte de membrana, la transducción de señales y el ciclo celular (Kanehisa & Goto, 2000).

Matplotlib: Es una librería de Python especializada en la creación de gráficos en dos dimensiones. Las visualizaciones pueden ser estáticas, animadas e interactivas (Matplotlib, 2022).

Minimización del ajuste metabólico (MOMA): El FBA tiene el objetivo de maximizar el crecimiento y obtener una solución razonable para para células de tipo salvaje, pero no lo hace para mutantes de inactivación de genes (knockout). Dado que la homeostasis gobierna la reprogramación metabólica no se puede suponer que la célula seguirá un objetivo común, como maximizar su crecimiento. Para este tipo de análisis se utiliza MOMA. Y se basa en que, para mantener la homeostasis, la diferencia de flujos antes y después de la perturbación debe ser mínima. Entonces MOMA predice los flujos de una cepa knockout asumiendo que la célula tendrá una redistribución mínima de flujos en comparación con el tipo salvaje (Segrè *et al.*, 2002; Volkova *et al.*, 2020).

RESUMEN

Recientemente la ingeniería metabólica a tomado mayor importancia para mejorar los bioprocesos en los que están implicados microorganismo, su principal objetivo es optimizar la producción de sustancias químicas valiosas y que poseen un potencial de producción industrial. Llevando a cabo modificaciones puntuales de las vías metabólicas que se encuentran en un organismo para comprender y utilizar mejor las vías celulares para la transformación química.

Debido a los diferentes parámetros que controlan el metabolismo tales como: alostería, abundancia de enzimas y modificaciones postraduccionales; el modelado metabólico se ha convertido en una herramienta valiosa para entender los mecanismos que gobiernan el fenotipo celular. Dentro de estos modelos, el modelo basado en restricciones es uno de los más utilizados y permite conocer el metabolismo celular basándose en conocimientos previos como las reacciones que ocurren, sus sustratos y productos (metabolitos), su estequiometría y reversibilidad. El Análisis de Balance de Flujos Metabólicos (FBA) es uno de los métodos basados en restricciones más utilizados para simular el metabolismo a escala del genoma.

La producción a gran escala del licopeno, resulta importante debido a sus funciones terapéuticas, profilácticas y nutracéuticas. Y es que actualmente se ha demostrado su capacidad como agente quimiopreventivo del cáncer y que además posee efecto cardioprotector, antioxidante y antiinflamatorio. La fabricación de licopeno a partir de fuentes vegetales enfrenta problemas y es por ello que actualmente se han realizado grandes esfuerzos para diseñar microorganismos no carotenogénicos para la producción eficiente de licopeno.

En esta investigación, se utilizaron algoritmos computacionales en Python para optimizar la producción de licopeno en la levadura *Yarrowia lipolytica* usando su modelo metabólica a escala del genoma (*iYali4*). Para la producción del carotenoide, se analizó la ruta del mevalonato a fin de convertir los productos de la ruta (IPP y DMAPP) en licopeno, usando para ello rutas nativas y heterólogas (genes *CrtB* y *CrtI*, de *Pantoea ananatis*). Los paquetes para análisis de ingeniería metabólica de Python

utilizados fueron

Cameo y *COBRApy*. Para la optimización se usó los modelos basados en restricciones metabólicas: *FBA* y *FVA*. Se obtuvieron gráficos de planos de fase de fenotipo (*Production envelopes*) para visualizar distintas fases de crecimiento óptimo con un uso diferente de dos sustratos; oxígeno y nitrógeno. Además de visualizaciones en la herramienta web *Fluxer* de los principales flujos metabólicos obtenidos de la optimización. Los mejores resultados en la tasa de crecimiento, tasa de producción de licopeno y rendimiento de licopeno, se obtuvieron usando como par de aceptor/aceptor reducido a *FAD/FADH2* y sobreexpresando los genes de la ruta del mevalonato; *HMG1*, *MVD1* y *EGR8*. Obteniéndose una tasa de producción de licopeno de 0.0562 (mmol/gdcw/h).

El modelo desarrollado en este trabajo se puede usar para guiar la ingeniería metabólica de *Yarrowia lipolytica*. Específicamente, sus predicciones computacionales facilitarán la producción sostenible de metabolitos de interés comercial como se demuestra aquí para el caso del licopeno.

Palabras clave: *Yarrowia lipolytica*, *FBA*, *FVA*, Ingeniería metabólica, *Licopeno*, *COBRApy*, *Cameo*, *Production envelopes*, Sobreexpresión.

ABSTRACT

Recently, metabolic engineering has become more important to improve bioprocesses in which microorganisms are involved. Its main objective is to optimize the production of valuable chemical substances that have industrial production potential. Carrying out specific modifications of the metabolic pathways found in an organism to better understand and use cellular pathways for chemical transformation.

Due to the different parameters that control metabolism such as: allostery, abundance of enzymes and post-translational modifications; metabolic modeling has become a valuable tool to understand the mechanisms that govern the cellular phenotype. Within these models, the constraint-based model is one of the most widely used and allows knowing cell metabolism based on prior knowledge such as the reactions that occur, their substrates and products (metabolites), their stoichiometry and reversibility. Metabolic Flux Balance Analysis (FBA) is one of the most widely used constraint-based methods to simulate metabolism at the genome scale.

The large-scale production of lycopene is important due to its therapeutic, prophylactic and nutraceutical functions. And it is that currently its capacity as a cancer chemopreventive agent has been demonstrated and that it also has a cardioprotective, antioxidant and anti-inflammatory effect. The manufacture of lycopene from plant sources faces problems and that is why currently great efforts have been made to design non-carotenogenic microorganisms for the efficient production of lycopene.

In this research, computational algorithms in Python were used to optimize lycopene production in the yeast *Yarrowia lipolytica* using their genome-wide metabolic model (*iYali4*). For the production of the carotenoid, the mevalonate pathway was analyzed in order to convert the pathway products (IPP and DMAPP) into lycopene, using both native and heterologous pathways (*CrtB* and *CrtI* genes, from *Pantoea ananatis*). The Python metabolic engineering analysis packages used were Cameo and COBRApy. For the optimization, the models based on metabolic restrictions were used: FBA and FVA. Phenotype phase plane plots (*Production envelopes*) were obtained to visualize different phases of optimal growth with a different use of two substrates; oxygen and

nitrogen. In

addition to visualizations in the Fluxer web tool of the main metabolic flows obtained from the optimization. The best results in growth rate, lycopene production rate, and lycopene yield were obtained using FAD/FADH₂ as a reduced acceptor/acceptor pair and overexpressing the mevalonate pathway genes; *HMG1*, *MVD1* and *EGR8*. Obtaining a lycopene production rate of 0.0562 (mmol/gdcw/h).

The model developed in this work can be used to guide the metabolic engineering of *Yarrowia lipolytica*. Specifically, their computational predictions will facilitate the sustainable production of metabolites of commercial interest as demonstrated here for the case of lycopene.

Keywords: *Yarrowia lipolytica*, FBA, FVA, Metabolic engineering, Lycopene, COBRAPy, Cameo, Production envelopes, Overexpression.

CAPÍTULO 1:

INTRODUCCIÓN

1.1. Planteamiento del problema

La ingeniería de células para la producción de químicos y proteínas de interés industrial usados en: la industria alimenticia, detergentes, farmacología, biorremediación, entre otros, utiliza la biotecnología para diseñar las células a fin de brindar soluciones a estas necesidades de la humanidad, a partir de materias primas renovables (Cardoso *et al.*, 2018).

Recientemente, la disciplina de la ingeniería metabólica ha conseguido importancia en la biotecnología, siendo utilizada en gran medida para la mejora de bioprocesos en los que están implicados microorganismos. El objetivo general de la ingeniería metabólica es sobre producir sustancias químicas valiosas para su potencial producción industrial (Kim *et al.*, 2016; Tafur *et al.*, 2021). En esta disciplina se realiza la alteración dirigida e intencionada de las vías metabólicas que se encuentran en un organismo para comprender y utilizar mejor las vías celulares para la transformación química, la transducción de energía y el ensamblaje supramolecular (Kumar & Prasad, 2011).

Debido a los avances en biología de sistemas y biología sintética, un conjunto de herramientas moleculares y computacionales de alto rendimiento se implementan cada vez más en la práctica de la ingeniería metabólica, de modo que el campo ahora se denomina a menudo "ingeniería metabólica de sistema (Kim *et al.*, 2016). Los métodos de biología de sistemas computacionales permiten el diseño racional de fábricas de células a escala del genoma y, por lo tanto, aceleran la ingeniería de células para la producción de sustancias químicas y proteínas valiosas (Cardoso *et al.*, 2018).

Manipular células para realizar tareas para las que no evolucionaron es un desafío y requiere grandes inversiones en tiempo, costo y personal para alcanzar una producción económicamente viable de moléculas objetivo. En este contexto, los métodos de modelado se utilizan cada vez más para acelerar la ingeniería de fábricas de células, lo que ayuda a reducir el tiempo y el costo de desarrollo (Cardoso *et al.*, 2018). Los modelos metabólicos a escala del genoma (GEM) representan una ventaja significativa de la biología de sistemas, ya que sus aplicaciones como modelos de

simulación de flujo

metabólico permiten realizar predicciones para la producción de metabolitos de interés industrial (Nambou *et al.*, 2015).

Aunque se han publicado un gran repertorio de algoritmos que utilizan GEM para calcular estrategias de ingeniería de fábricas de células compuestas por sobreexpresión, regulación negativa, eliminación y adición de genes, el problema general radica en que la mayoría de estos algoritmos no son de fácil acceso para los usuarios, ya que se han publicado sin implementación o la implementación proporcionada por los autores no está documentada o es difícil de instalar. Estos problemas (Figura 1) limitan significativamente la capacidad de los ingenieros metabólicos para utilizar herramientas de diseño computacional como parte de su flujo de trabajo (Cardoso *et al.*, 2018; Maia *et al.*, 2016).

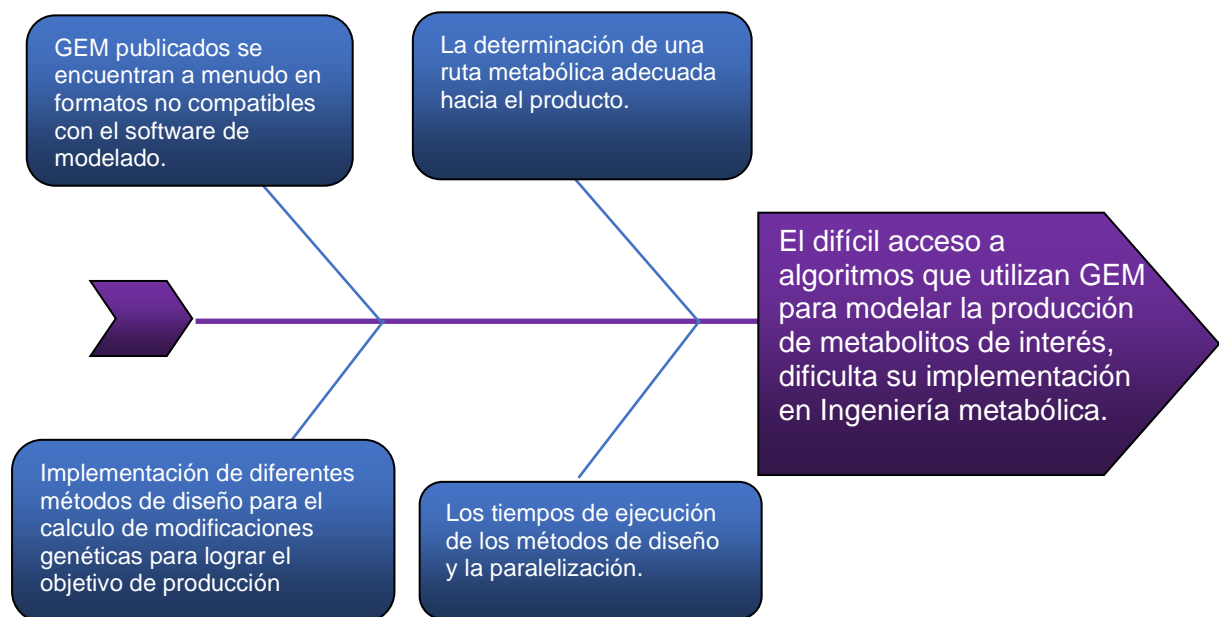


Figura 1. Diagrama causa efecto, espina de pescado, del problema de investigación.

1.2. Objetivos

1.2.1. Objetivo general

Desarrollar algoritmos para optimizar la producción de licopeno en *Yarrowia lipolytica* usando herramientas computacionales de ingeniería metabólica en Python.

1.2.2. Objetivos específicos

- Implementar algoritmos en Python para el modelo metabólico a escala del genoma (GEM) *iYali4* de *Yarrowia lipolytica*.
- Agregar los genes, reacciones y metabolitos necesarios para que el modelo pueda producir licopeno usando la vía del mevalonato.
- Optimizar la producción de licopeno con efectos menores en la tasa de crecimiento en el modelo metabólico.

1.3. Justificación

El licopeno, pertenece a los compuestos tetraterpenoides y a la subcategoría carotenoides, es sintetizado exclusivamente por plantas como el tomate, la sandía, la guayaba, la papaya, el albaricoque, el pomelo rosa, la naranja roja, y por otro lado por microorganismos carotenogénicos. Este carotenoide, es esencial para el cuerpo humano ya que solo puede ser aportado a través de su ingestión bajo múltiples formas como alimentos o drogas. El licopeno promueve la buena salud debido a sus funciones terapéuticas, profilácticas y nutraceuticas. Actualmente se ha demostrado sus aplicaciones en la apoptosis de células cancerígenas y la detención del ciclo celular de diversos tipos de cánceres. También actúa como un fuerte agente quimiopreventivo del cáncer y posee efecto cardioprotector, antioxidante y antiinflamatorio (Nambou *et al.*, 2015; Takeshima *et al.*, 2014; Zhang *et al.*, 2022).

La producción a gran escala de licopeno es fundamental. Sin embargo, la fabricación actual de licopeno a partir de fuentes vegetales enfrenta problemas de competencia alimentaria, medio ambiente, alto costo de producción, largo tiempo de procesamiento y bajo rendimiento. Además, la fermentación por microorganismos carotenogénicos es escasa y no produce títulos considerables de licopeno. Se han realizado grandes

esfuerzos para

diseñar microorganismos no carotenogénicos para la producción eficiente de licopeno (Nambou *et al.*, 2015).

En este contexto, la levadura oleaginoso *Yarrowia lipolytica*, categorizado como GRAS (generalmente reconocido como seguro), se ha establecido como una plataforma de fermentación prometedora, para obtener metabolitos aptos en aplicaciones farmacéuticas y en alimentos. Debido a su sistema de expresión atractivo para la producción de proteínas homólogas y heterólogas, su capacidad de utilizar sustratos hidrófobos de bajo costo y excelente capacidad secretora (Larroude *et al.*, 2018; Yan *et al.*, 2018).

Por lo tanto, se deben desarrollar metodologías *in silico*; que ayuden a la implementación de bioprocesos completos, innovadores, rentables y competitivos para la producción de licopeno a partir de cepas de *Y. lipolytica*. Usado para ello la información de los modelos metabólicos a escala del genoma (GEM), analizando las vías metabólicas, y simulando la eliminación o sobreexpresión de genes específicos junto con enfoques computacionales a fin de optimizar la producción de licopeno en esta levadura.

¿Por qué lo va hacer?

A fin de sugerir una alternativa que busque optimizar el costo, tiempo y trabajo en la Ingeniería metabólica para el diseño de fábricas de células microbianas utilizando las herramientas computacionales. Ofreciendo así una alternativa interesante para el modelamiento de la producción sostenible de metabolitos de interés comercial, en este caso del licopeno. Y en la levadura *Y. lipolytica*, que es un modelo prometedor para ser aplicado de forma industrial en biorrefinerías por su versatilidad de utilizar como fuente de carbono, sustratos de bajo costo tales como margarina rancia, suelo contaminado con aceite, agua de mar, hidrocarburos, ácidos grasos, lípidos y que a cambio es capaz de sintetizar una variedad de metabolitos valiosos (Zieniuk & Fabiszewska, 2018).

MARCO TEÓRICO Y CONCEPTUAL

2.1. Antecedentes o marco referencial

Yarrowia. lipolytica se ha convertido en una levadura atractiva para la producción de carotenoides debido a su capacidad para producir altos niveles de acetil-CoA, un precursor tanto de lípidos como de carotenoides (Liu *et al.*, 2013).

Recientemente, se han desarrollado métodos que utilizan la ingeniería metabólica en *Y. lipolytica* para la producción de licopeno y carotenoides:

- Matthäus y colaboradores (2014), sobreexpresaron los genes cuello de botella *HMG1* y *GGs1* para mejorar en gran medida la formación de licopeno y bloquearon la beta-oxidación provocada por la delección de *POX1* a *POX6* y la pérdida de glicerol-3-fosfato por la delección de *GUT2*.
- Nambu y colaboradores (2015), realizaron el análisis de balance de flujo (FBA) y diseño de Plackett-Burman para seleccionar compuestos adecuados para la producción de licopeno en *Y. lipolytica* modificada.
- Schwartz y colaboradores (2017), descubrieron que la mejor estrategia para mejorar la producción de licopeno en *Y. lipolytica* fue la sobreexpresión de la vía del mevalonato en la cepa PO1f clonando y expresando clonando y expresando los genes *crtE* que codifican GGPP sintasa, *crtB* , *crtI* hasta la producción de licopeno.
- Xie y colaboradores (2021), aplicaron Ingeniería metabólica en *Y. lipolytica* para la biosíntesis de zeaxantina. Para lograrlo, primero permitieron que la cepa PO1f produzca licopeno y β -caroteno, clonando y expresando los genes *crtE* que codifican GGPP sintasa, *crtB* , *crtI* y *CarRP* en la levadura.

En cuanto a las herramientas computacionales que han sido desarrollados en Python para el análisis de Ingeniería metabólica y aplicables a la simulación de producción de metabolitos, podemos indicar los siguientes:

- **COBRApy**,

es un paquete de Python que brinda soporte para los métodos básicos de COBRA. Está diseñado de forma orientada a objetos que facilita la representación de los complejos procesos biológicos del metabolismo y la expresión génica. El modelamiento se basa en restricciones y permite el acceso a métodos COBRA, como el análisis de balance de flujo (FBA), el análisis de variabilidad de flujo (FVA) y análisis de eliminación de genes (Ebrahim *et al.*, 2013).

- **jQMM**, una librería de Python para modelar el metabolismo microbiano, incluye algoritmos para medir y predecir los flujos metabólicos internos utilizando tres técnicas diferentes: análisis de flujo metabólico de ^{13}C (^{13}C MFA), análisis de equilibrio de flujo (FBA) y análisis de flujo metabólico de ^{13}C de dos escalas (2S- ^{13}C MFA) (Birkel *et al.*, 2017)
- **Cameo**, es una librería desarrollada en Python para ingeniería metabólica asistida por computadora y optimización de fábricas de células. Implementa métodos de última generación, utilizando modelos de metabolismo a escala del genoma (GEM) para enumerar y priorizar estrategias de knockout, knock-in, sobreexpresión y down-regulation y combinaciones de las mismas (Cardoso *et al.*, 2018).
- **MEWpy**, es un banco de trabajo desarrollado en Python para ingeniería metabólica, que implementa una gama de enfoques de modelos basados en restricción (restricciones metabólicas, enzimáticas y regulatorias), así como simulación de fenotipos y algoritmos para la optimización computacional de cepas (Computational Strain Optimization, CSO) (Pereira *et al.*, 2021).
- **CNApy**, es una aplicación de escritorio multiplataforma escrita en Python que proporciona una interfaz gráfica de última generación para el análisis intuitivo de modelos metabólicos con métodos COBRA (Thiele *et al.*, 2022).

En el Ecuador la Ingeniería Metabólica es un campo nuevo, cuya aplicación podría facilitarse, al implementar algoritmos computacionales como los descritos en esta propuesta.

2.2. Marco teórico

2.2.1. *Yarrowia lipolytica*

Yarrowia lipolytica es una de las levaduras oleaginosas no convencional más utilizadas en biorrefinerías, debido a su extraordinaria capacidad para utilizar sustratos complejos y sin refinar de bajo costo como fuentes de carbono y nitrógeno. Permitiéndole producir compuestos de valor agregado de manera rentable (Madzak, 2021). Estos sustratos hidrofílicos e hidrofóbicos se los clasifica en convencionales (almidón, melaza, residuos de frutas y verduras) y no convencionales (diferentes fracciones de petróleo, gas natural, etanol, metanol, biomasa lignocelulósica, residuos de aceites de cocina o de motor, grasas de desecho animal o corrientes residuales de diversas industrias). Debido a estas materias primas económicas, la masa microbiana y los lípidos microbianos producidos por *Y. lipolytica* (proteína unicelular “single cell protein” y aceite unicelular) se consideran procesos ecológicos (Madzak, 2021; Zieniuk & Fabiszewska, 2018; Jach *et al.*, 2020).

Debido a que esta levadura contiene una variedad de nutrientes en su biomasa (péptidos, aminoácidos, vitamina B12, minerales traza, ácidos grasos monoinsaturados y lípidos saturados alto valor agregado). Se considera una levadura nutricional. Además, su estatus como GRAS (generalmente reconocido como seguro), la han permitido ser utilizada en la alimentación humana y más aplicada en la alimentación de animales, después de matar las células de esta levadura por secado para garantizar la inocuidad (Jach *et al.*, 2020; Larroude *et al.*, 2018)

También se ha emplea como huésped de producción heterólogo para proteínas y enzimas farmacéuticas e industriales, así como con fines de biorremediación (Groenewald *et al.*, 2013). En el campo de la biomasa de levadura, se ha dirigido mucho esfuerzo hacia la optimización de la producción (Jach *et al.*, 2020).

2.2.2. Síntesis de

carotenoides

2.2.2.1. Producción de carotenoides en levaduras

Los carotenoides son pigmentos tetraterpenoides, producidos naturalmente por algunas plantas, algas, hongos y bacterias. Entre los más comercializados se encuentran el β -caroteno, licopeno, astaxantina, zeaxantina, cantaxantina, luteína y otros. Siendo un mercado competitivo, pues se estimó en \$2000 millones en 2022 y se prevé que crezca a una tasa de crecimiento anual compuesta (CAGR) del 5,7% (BCC Research, 2022).

Recientemente, se han desarrollado nuevas cepas de levadura productoras de carotenos y licopeno. Entre las levaduras que han sido objeto de investigación se encuentran *Yarrowia lipolytica* y *Saccharomyces cerevisiae*. A diferencia de *Y. lipolytica*, *S. cerevisiae* no acumula naturalmente grandes cantidades de lípidos y también tiene una baja formación de acetil-CoA citosólico, que es un precursor tanto de los lípidos como del β -caroteno. Sin embargo, *S. cerevisiae* es un huésped industrial ampliamente utilizado para la producción de proteínas recombinantes, y productos químicos especiales (Fathi *et al.*, 2021)

2.2.2.2. Biosíntesis de licopeno en *Yarrowia lipolytica*

La capacidad de esta levadura de producir altos niveles de acetil-CoA, un precursor tanto de lípidos como de carotenoides, su capacidad para acumular lípidos (produce grandes cantidades de lípidos gotas, que secuestran carotenoides y que evitan la desestabilización de la membrana) y la manejabilidad del metabolismo de los lípidos, convierte a *Y. lipolytica* en un organismo plataforma prometedor para la biosíntesis de licopeno (Liu *et al.*, 2013; Zhang *et al.*, 2020).

2.2.2.3. Ruta del mevalonato

La biosíntesis del licopeno en *Y. lipolytica* depende de los productos de la vía del mevalonato (MVA) como precursores, por lo que se ha demostrado que el aumento del flujo a través de la vía del mevalonato aumente la producción de terpenoides (Zhang *et al.*, 2022).

La vía del mevalonato, desde la acetil-CoA hasta el difosfato de isopentenilo (IPP), consta de 6 reacciones enzimáticas (Figura 2A). Primero, 2 moléculas de acetil-CoA se condensan para formar acetoacetil-CoA mediante la acetoacetil-CoA tiolasa (*ERG10*). Luego se agrega una tercera molécula de acetil-CoA para formar 3-hidroxi-3-metilglutaril-CoA (HMG-CoA) mediante la HMG-CoA sintasa (*ERG13*). El siguiente paso, implica la conversión de HMG-CoA en mevalonato (MVA) por la HMG-CoA reductasa (*HMG1*) en la cual se consume 2 moléculas de NADPH. Después, el mevalonato se fosforila dos veces, primero por la mevalonato quinasa (*ERG12*) y segundo por la fosfomevalonato quinasa (*ERG8*), cada fosforilación consume 1 molécula de ATP y da como resultado mevalonato-5-difosfato (MV5PP). Luego, MV5PP se convierte en IPP mediante la mevalonato pirofosfato descarboxilasa (*MVD1*), que consume 1 molécula de ATP y elimina 1 carbono de la molécula en forma de CO₂. Finalmente, IPP se puede isomerizar a dimetilalil-pirofosfato (DMAPP) mediante isopentenil difosfato:dimetilalil difosfato isomerasa (*IDI1*) (Schwartz *et al.*, 2017)

2.2.2.4. Biosíntesis de los productos de MVA a licopeno

La siguiente fase en la biosíntesis del licopeno, conlleva la ruta biosintética para convertir los productos de la ruta del mevalonato (IPP y DMAPP) en licopeno (Figura 2B). Esta vía metabólica se realiza en combinación de enzimas nativas y enzimas heterólogas ya que *Y. lipolytica* no sintetiza licopeno de forma nativa (Zhang *et al.*, 2022).

Primero, DMAPP e IPP se condensan para formar geranyl-difosfato (geranyl-PP) por farnesil pirofosfato sintetasa (*ERG20*). La misma enzima, *ERG20*, cataliza

posteriormente la adición de una segunda molécula de IPP para producir farnesil-difosfato (farnesil-PP). El farnesil-PP se puede convertir luego en geranylgeranyl-difosfato (GGPP) mediante la geranylgeranyl difosfato sintasa (*BTS1*). La producción de GGPP es el final de la ruta biosintética de la levadura nativa. Los pasos siguientes son heterólogos:

- Se ha encontrado una geranylgeranyl difosfato sintasa heteróloga (*CrtE*, de *Pantoea ananatis*) más eficaz y que se podría adicionar a la ruta.
- La fitoeno sintasa (*CrtB*, de *P. ananatis*) condensa dos moléculas de GGPP para formar fitoeno.
- Finalmente, el fitoeno sufre 4 reacciones de desaturación consecutivas catalizadas por la licopeno sintasa (*CrtI*, de *P. ananatis*) para producir licopeno. (Schwartz et al., 2017; Zhang et al., 2020; Zhang et al., 2022).

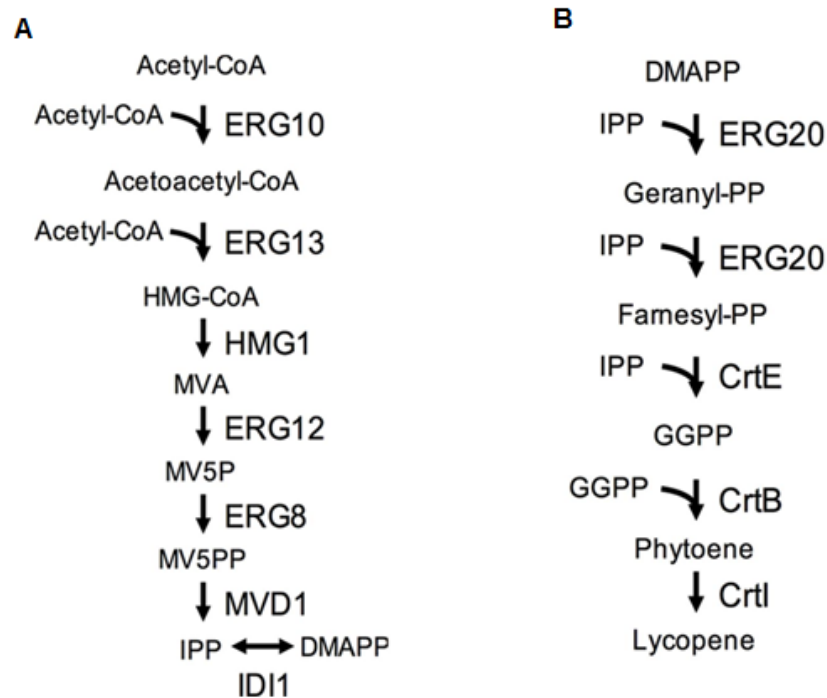


Figura 2. Esquema de la vía biosintética del licopeno a partir de acetil-CoA en *Y. lipolytica*.

(A) Vía de biosíntesis del mevalonato.

(B) Vía biosintética de DMAPP a licopeno

Fuente: (Modificado de Schwartz et al., 2017)

Estrategias de bioingeniería para la producción de terpenoides

- Aumento del suministro de acetil-CoA citosólico mediante la introducción de vías heterólogas o la redirección del metabolismo del carbono central nativo.
- Utilización de nuevos sustratos para minimizar costos de producción.
- Sobreexpresión de los genes en la vía MVA y suministro de cofactores (Zhang *et al.*, 2022).

2.2.3. Los modelos metabólicos a escala genómica permiten simular las condiciones óptimas para la producción de metabolitos

Al encontrarnos actualmente con una gran cantidad de datos biológicos, los modelos computacionales son fundamentales para convertir diferentes fuentes de datos en conocimiento valioso. Como una herramienta de biología de sistemas, los modelos metabólicos a escala del genoma (GEM) conectan genes, proteínas y reacciones, lo que permite predecir las respuestas metabólicas y fenotípicas basadas en restricciones específicas. Estos modelos necesitan además de buenas anotaciones del genoma y datos experimentales para simular comportamientos celulares (Jingbo *et al.*, 2020; Lu *et al.*, 2019).

Modelo metabólico de *Yarrowia lipolytica* (*iYali4*): Este modelo se construyó a partir de la red de consenso *Yeast 7.11* de *S. cerevisiae* y los modelos metabólicos de *Y. lipolytica* *iNL895* (que fue la primera red metabólica de todo el genoma de *Y. lipolytica* desarrollado en 2012) e *iYL619_PCP* (el modelo se lanzó el mismo año y se validó experimentalmente estudiando el crecimiento en glucosa y medio mineral). Este modelo publicado en 2016, y que sido actualizado en base a la secuencia del genoma de la cepa *Y. lipolytica* *W2*. Se ha aplicado al estudio del metabolismo de los aminoácidos y control del flujo metabólico hacia la acumulación de lípidos (Kerkhoven *et al.*, 2016). Se considera al modelo *iYali4*, como el más completo en cuanto al metabolismo de los lípidos, en comparación con sus predecesores, y proporciona a la

científica una herramienta muy útil en ingeniería metabólica para desentrañar la regulación del metabolismo (Da Veiga *et al.*, 2021; Jingbo *et al.*, 2020).

Modelado metabólico: El metabolismo está controlado por mecanismos reguladores dinámicos que involucran alostería, abundancia de enzimas y modificaciones postraduccionales. Debido a la estrecha relación entre el metabolismo y el fenotipo celular, el modelado metabólico se ha convertido en una herramienta valiosa para entender los mecanismos subyacentes que gobiernan el fenotipo celular. El modelado basado en restricciones y el modelado cinético se encuentran entre los métodos más utilizados para estudiar el metabolismo celular a diferentes escalas. Estos enfoques permiten integrar datos metabolómicos, entre otros, para mejorar las capacidades predictivas del modelo (Volkova *et al.*, 2020)

Modelos basados en restricciones (*Constraint-based Methods*): Este modelo es el más utilizado para estudiar metabolismo. Permite conocer el metabolismo celular basándose en conocimientos previos como las reacciones que ocurren, sus sustratos y productos (metabolitos), su estequiometría y reversibilidad. Con base en esta información, se construye una red metabólica y se determina como una matriz estequiométrica (S). En esta matriz, cada fila representa un metabolito y cada columna corresponde a una reacción, mientras que las entradas representan el coeficiente estequiométrico de un metabolito en una reacción específica (Figura 3). (Volkova *et al.*, 2020;)

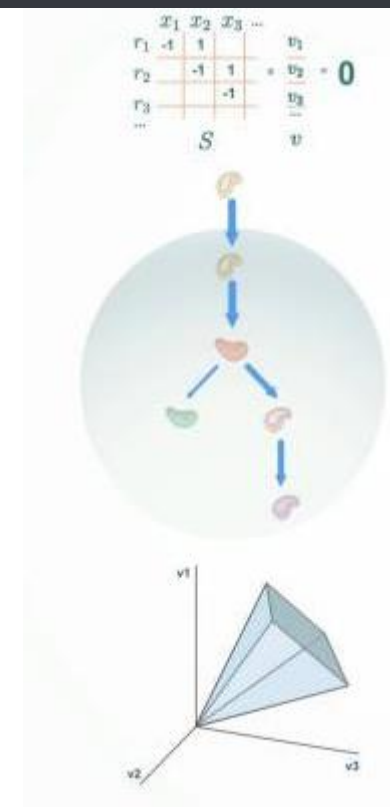


Figura 3. El modelado basado en restricciones.

Permite equilibrar los flujos en el sistema, pero no puede trabajar directamente con la concentración de metabolitos. Los metabolitos consumidos en una reacción tienen un coeficiente estequiométrico negativo, mientras que los metabolitos producidos por la reacción tienen un coeficiente estequiométrico positivo para esa reacción específica

Fuente: (Volkova *et al.*, 2020)

Dado un vector que describe el flujo de cada reacción (v) y la matriz estequiométrica (S), se puede determinar cómo varía la concentración del metabolito a lo largo del tiempo (dX/dt) aplicando (ECU. 1):

$$Sv = \frac{dX}{dt}$$

(ECU. 1)

Donde los cambios en las concentraciones de metabolitos a lo largo del tiempo (dX/dt) son representados por los flujos de reacción y la matriz estequiométrica del sistema. Cuando se asume que el sistema está en estado estacionario, es decir, las concentraciones de metabolitos no cambian con el tiempo, se aplica (ECU. 2)

$$Sv = 0$$

(ECU. 2)

Esto asegura que la suma de los flujos de las reacciones que producen un metabolito dado sea igual a la suma de los flujos que lo consumen, Por lo que no existe acumulación ni agotamiento de metabolitos intracelulares en función del tiempo. Cuando se resuelve este sistema de ecuaciones (reacciones químicas), es posible determinar la distribución de flujo (Bordbar *et al.*, 2014; Volkova *et al.*, 2020)

- **El análisis de balance de flujo (*Flux Balance Analysis, FBA*)**

Se trata de un enfoque matemático para analizar el flujo de metabolitos a través de las reconstrucciones de redes metabólicas a escala del genoma. FBA, es uno de los métodos basados en restricciones más utilizados para simular el metabolismo a escala del genoma. Estas redes contienen todas las reacciones metabólicas conocidas en un organismo y los genes que codifican cada enzima. FBA calcula el flujo de metabolitos a través de esta red metabólica, lo que permite predecir la tasa de crecimiento de un organismo o la tasa de producción de un metabolito (Orth *et al.*, 2010).

FBA usa programación lineal para resolver la ecuación $Sv = 0$ dado un conjunto de límites superior e inferior en v y una combinación lineal de flujos como función objetivo. El resultado de FBA es una solución específica, una distribución de flujo v , que maximiza o minimiza la función objetivo (como la reacción de biomasa al simular el crecimiento máximo). Garantizando que la función objetivo es óptima (Volkova *et al.*, 2020).

- **El análisis de variabilidad de flujo (*Flux Variability Analysis, FVA*)**

Para explorar soluciones alternativas y subóptimas, así como estudiar la flexibilidad de la red, se puede utilizar el análisis de variabilidad de flujo (FVA).

Su aplicación,

consiste en encontrar el flujo mínimo y máximo para las reacciones en la red mientras se mantiene algún estado de la red, por ejemplo, soportando el 90% de la tasa de producción de biomasa máxima posible. FVA se aplica, además, al metabolismo de subproductos y metabolitos secundarios, donde las soluciones FBA no son confiables (Gudmundsson S. & Thiele, 2010; Volkova *et al.*, 2020).

2.2.4. Herramientas computacionales

Entre las principales herramientas computacionales basadas en restricciones tenemos:

- **COBRApy**: Es un paquete de Python que brinda soporte para los métodos básicos de COBRA. Está diseñado de forma orientada a objetos que facilita la representación de los complejos procesos biológicos del metabolismo y la expresión génica. El modelamiento se basa en restricciones y permite el acceso a métodos COBRA, como el análisis de balance de flujo (FBA), el análisis de variabilidad de flujo (FVA) y análisis de eliminación de genes, entre otros (Ebrahim *et al.*, 2013).
- **Cameo**
Es una librería desarrollada en Python para ingeniería metabólica asistida por computadora y optimización de fábricas de células, utiliza los métodos basados en restricciones, siendo además 100 % compatible con COBRApy. Sin embargo, por eficiencia, cameo implementa sus propios métodos de simulación que aprovechan una interfaz de resolución más avanzada. Cameo implementa métodos de última generación, utilizando modelos de metabolismo a escala del genoma (GEM) para enumerar y priorizar estrategias de *knockout*, *knock-in*, sobreexpresión y *down-regulation* y combinaciones de las mismas (Cardoso *et al.*, 2018).

CAPÍTULO 3: METODOLOGÍA

El código fue desarrollado en un entorno **Jupyter Notebook** de Google Colab.

Python utilizada fue: **3.8.16**.

Los paquetes de Python para Ingeniería metabólica que utilizan los modelos basados en restricciones empleados fueron: **COBRAPy y Cameo**.

El script desarrollado se encuentra en el ANEXO 1.

3.1. Resumen del algoritmo empleado

En la Figura 4. Se muestra un resumen del algoritmo empleado para la optimización de la producción de licopeno en *Yarrowia lipolytica* usando el modelo metabólico *iYali4* con los paquetes de Python COBRAPy, Cameo, Matplotlib y el visualizador web Fluxer.

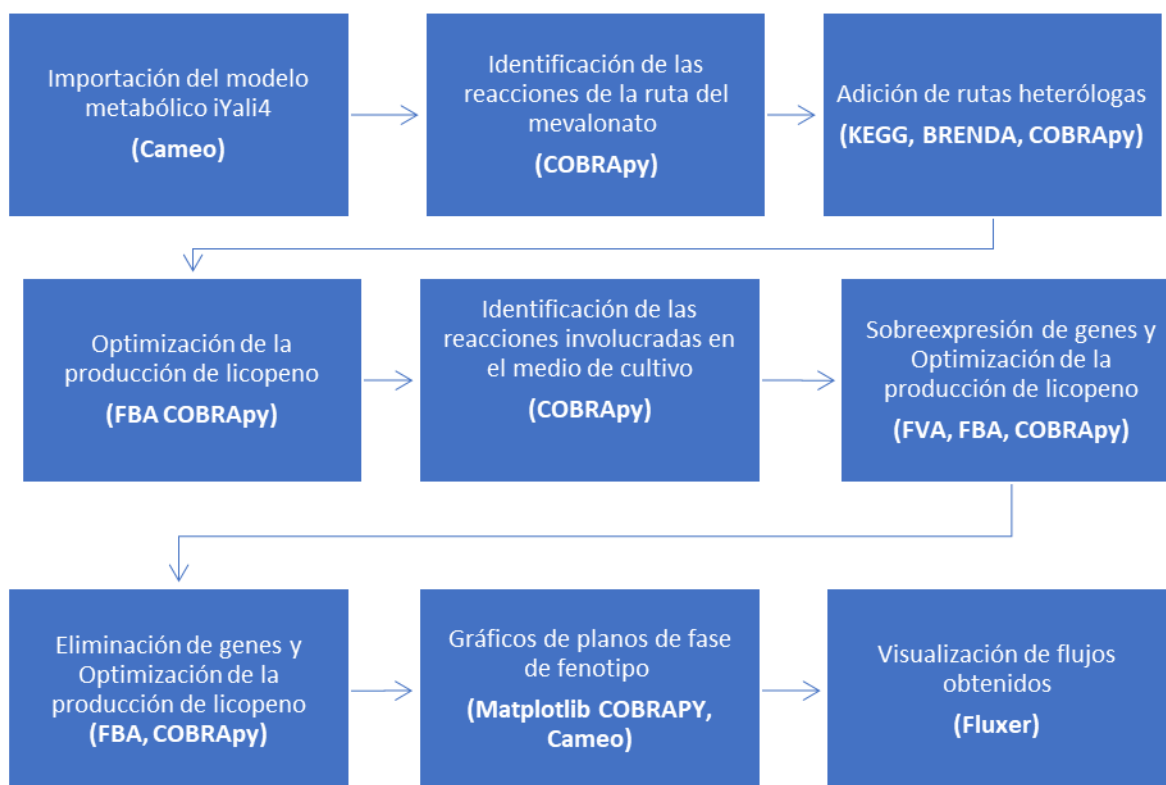


Figura 4. Esquema del algoritmo empleado para la optimización de la producción de licopeno en *Yarrowia lipolytica* en el modelo *iYali4*.

3.2. Importación

del modelo metabólico *iYali4* de *Yarrowia lipolytica*

Se utilizó el comando *models* de *Cameo*, para importar el modelo metabólico a escala del genoma *iYali4* de *Y. lipolytica* de la base de datos de **University of Minho** (<http://darwin.di.uminho.pt/model>).

El modelo cargado, tiene el formato SBML (*Systems Biology Markup Language*). Que es un formato de archivo para representar modelos computacionales compatible con diferentes sistemas de software. SBML está orientado además a describir procesos biológicos tales como vías metabólicas, las vías de señalización celular entre otros (Hucka *et al.*, 2018). El modelo *iYali4* se compone de **1683 metabolitos**, **1985 reacciones**, **901 genes**, la función objetivo con **id: 2111**, pertenece a la reacción de crecimiento de la levadura que tiene como producto la biomasa.

3.3. Identificación de las reacciones de la ruta del mevalonato

Debido que este modelo tiene como id identificadores numéricos para las reacciones bioquímicas, se procedió a ubicar las reacciones presentes de la vía biosintética del licopeno a partir de acetil-CoA en el modelo metabólico

Se utilizo el comando *reactions.get_by_id()* de COBRApy.

Por ejemplo, para ubicar la reacción en la que interviene la enzima *Acetoacetil-CoA tiolasa* con el gen *ERG10* se usó el siguiente código:

```
modelyall.reactions.get_by_id("103")
```

La Tabla 1, presenta la información resumida sobre las reacciones de la vía del mevalonato (MVA) como precursores para la producción del licopeno, presentes en el modelo metabólico *iYali4*. Desde de acetil-CoA hasta la formación de mevalonato-5-difosfato (MV5PP). Y desde este producto, hasta la producción de geranylgeranyl-difosfato (GGPP). Se ha incluido los **id de las reacciones** dentro de este modelo y las reglas de reacción gen-proteína GPR (**Gene-protein-reaction**).

Tabla 1. Reacciones de la vía biosintética del licopeno a partir de acetil-CoA hasta farnesil-PP. Identificadas en el modelo metabólico *iYali4*.

Nombre de la enzima	Genes	Reactivos	Productos	Gene-protein-reaction (GPR) en el modelo	id
Acetoacetil-CoA tiolasa	<i>ERG10</i>	2 moléculas de acetil-CoA	acetoacetil-CoA	YALIOB08536g YALIOE11099g	103(silenciado) 104
HMG-CoA sintasa	<i>ERG13</i>	acetoacetil-CoA + acetil-CoA	3-hidroxi-3-metilglutaril-CoA (HMG-CoA)	YALIOF30481g	560
HMG-CoA reductasa	<i>HMG1</i>	(HMG-CoA) + 2 moléculas de NADPH	mevalonato (MVA)	YALIOE04807g	558
Mevalonato quinasa	<i>ERG12</i>	mevalonato (MVA + ATP)	mevalonato-5-fosfato MV5P	YALIOB16038g	735
Fosfomevalonato quinasa	<i>ERG8</i>	mevalonato-5-fosfato MV5P + ATP	mevalonato-5-difosfato (MV5PP).	YALIOE06193g	904
mevalonato pirofosfato descarboxilasa	<i>MVD1</i>	mevalonato-5-difosfato (MV5PP) * ATP	difosfato de isopentenilo (IPP) + CO ₂	YALIOF05632g	739
isopentenil difosfato:dimetilalil difosfato isomerasa	<i>IDI1</i>	IPP	dimetilalil-pirofosfato (DMAPP)	YALIOF04015g	667
farnesil pirofosfato sintetasa	<i>ERG20</i>	DMAPP e IPP	geranil-difosfato (geranil-PP)	YALIOE05753g	355
farnesil pirofosfato sintetasa	<i>ERG20</i>	(geranil-PP) + IPP	farnesil-difosfato (farnesil-PP).	YALIOE05753g	462
geranilgeranil difosfato sintetasa	<i>BTS1</i>	(farnesil-PP).	geranilgeranil-difosfato (GGPP)	YALIOD17050g	373

Elaborado por: Milton Quinga.

3.4. Adición de

rutas heterólogas

Debido a que *Yarrowia lipolytica* no produce licopeno de forma natural, la ruta biosintética de la levadura nativa termina con la producción de GGPP.

Las vías heterólogas añadidas pertenecen a las vías de producción de licopeno en *Pantoea ananatis*, salvo la reacción de demanda de licopeno (Schwartz *et al.*, 2017; Zhang *et al.*, 2020; Zhang *et al.*, 2022).

Para añadir las reacciones en el modelo fue necesario conocer su estequiometría, metabolitos y genes, que fue consultada en la base de datos KEGG.

Las reacciones heterólogas fueron las siguientes:

- La **fitoeno sintasa** (CrtB, de *P. ananatis*) que condensa dos moléculas de GGPP para formar fitoeno (https://www.genome.jp/entry/pam:PANA_4162):
15-cis-phytoene synthase [EC:2.5.1.32] (<https://www.genome.jp/entry/2.5.1.32>)
15-cis-phytoene forming REACTION: R10177
(<https://www.genome.jp/entry/R10177>):

2 Geranylgeranyl diphosphate \Leftrightarrow 15-cis-Phytoene + 2 Diphosphate

- El fitoeno sufre 4 reacciones de desaturación consecutivas catalizadas por la **licopeno sintasa** (Crtl, de *P. ananatis*) para producir licopeno (https://www.genome.jp/entry/pam:PANA_4161):

phytoene desaturase [EC:1.3.99.31] (<https://www.genome.jp/entry/1.3.99.31>)

15-cis-phytoene:acceptor oxidoreductase (lycopene-forming) REACTION:
R09716 (<https://www.genome.jp/entry/R09716>):

15-cis-Phytoene + 4 Acceptor \Leftrightarrow Lycopene + 4 Reduced acceptor

El par aceptor/aceptor reducido, empleados fueron los siguientes:

NAD/NADH; NADP/NADPH y FAD/FADH₂. Consultados en base a estudios previos (Li *et al.*, 2019) y para el caso del FAD se consultó en la base de datos BRENDA (<https://www.brenda->

[ure.php?e=1.3.99.31&r=712337](http://enzymes.org/literature.php?e=1.3.99.31&r=712337)) y estudios previos (Fraser *et al.*, 1992).

La reacción de demanda de licopeno: Utilizada para garantizar la producción de licopeno, demanda (Erickson, 2018).

Lycopene -->

Para añadir estas reacciones al modelo *iYali4*, se utilizó comandos de COBRAPy. Por ejemplo, para añadir la reacción en la que actúa la **fitoeno sintasa (CrtB)**. Los pasos a seguir fueron los siguientes:

- Crear reacción:

```
#Para crear modelos metabólicos
from cobra import Model, Reaction, Metabolite

#Creación de la reacción
reactionp = Reaction('R_phytoene')
reactionp.name = '15 cis phytoene-forming'
reactionp.subsystem = ' Carotenoid biosynthesis'
reactionp.lower_bound = -1000. # Por default
reactionp.upper_bound = 1000. # Por default
```

- Creación de metabolitos:

```
#Asignación del metabolito PP
PP_C = modelyall1.metabolites.get_by_id('m203')

#Asignación del metabolito GGPP
GGPP_C = modelyall1.metabolites.get_by_id('m513')
# Creación del metabolito 15-cis-Phytoene
phytoene_c = Metabolite(
    'mphytoene_c',
    formula='C40H64',
    name='15-cis-phytoene',
    compartment='c')
```

- Agregación de metabolitos a la reacción:

```
#Adición de los metabolitos a la reacción
```

```
reactionp.add_metabolites({  
    GGPP_C: -2.0,  
    PP_C: 2.0,  
    phytoene_c: 1.0  
})
```

- Asignación de la cadena de reglas de reacción de genes (GPR):

```
#Adición de los genes  
reactionp.gene_reaction_rule = '( CrtB )'  
reactionp.genes
```

- Agregación de la reacción al modelo:

```
#Agregación de la reacción al modelo, que también agregará todos lo  
s metabolitos y genes asociados.  
modelyall.add_reactions([reactionp])
```

También se modificó la función objetivo para que sea en este análisis la reacción de crecimiento (producción de biomasa **id:2111**) y la reacción de demanda de licopeno (**id: LYCO-dem**). Usando los comandos de COBRApy:

```
#Modificación de la función objetivo  
  
modelyall.reactions.get_by_id('2111').objective_coefficient = 1.0  
modelyall.reactions.get_by_id('LYCO-dem').objective_coefficient = 1.0
```

Se cambio la función objetivo a la que maximice la producción de licopeno con efectos menores en la función de crecimiento.

La Tabla 2, presenta la información resumida sobre las reacciones heterólogas adicionadas mediante los comandos de COBRApy, para la producción de licopeno. Desde la conversión del farnesil-PP hasta la producción de licopeno y la reacción de demanda. Se ha incluido los id de las reacciones dentro de este modelo y las reglas de reacción gen-proteína GPR (Gene-protein-reaction). Adicionalmente sean adicionado diferentes cofactores aceptores/aceptores reducidos: (NAD/NADH; NADP/NADPH, FAD/FADH₂), dentro de la reacción de formación de licopeno.

Tabla 2. Reacciones heterólogas adicionadas mediante los comandos de COBRApy.

Nombre de la enzima	Genes	Reactivos	Productos	Gene-protein-reaction (GPR) en el modelo	id
Fitoeno sintasa	<i>CrtB</i>	2geranilgeranil-difosfato (GGPP)	15-cis-fitoeno + difosfato	<i>CrtB</i>	R_phytoene
Licopeno sintasa	<i>Crtl</i>	15-cis-fitoeno + 4NAD	licopeno + 4NADH	<i>Crtl</i>	R_LycopeneNAD
Licopeno sintasa	<i>Crtl</i>	15-cis-fitoeno + 4NADP	licopeno + 4NADPH	<i>Crtl</i>	R_LycopeneNADP
Licopeno sintasa	<i>Crtl</i>	15-cis-fitoeno + 4FAD	licopeno + 4FADH2	<i>Crtl</i>	R_LycopeneFAD
Demanda de licopeno		<i>licopeno_c</i>			LYCO-dem

Elaborado por: Milton Quinga.

3.5. Optimización de la producción de licopeno (Evaluación de aceptores/ aceptores reducidos)

Para la optimización de la producción de licopeno se utilizó el análisis de balance de flujo **FBA**, aplicando los comandos de COBRApy:

Se evaluó para los distintos pares de aceptores/aceptores reducidos: NAD/NADH; NADP/NADPH, FAD/FADH2).

Los datos de salida fueron procesados para obtener las unidades correspondientes de los distintos flujos: *Tasa de crecimiento (biomasa) (1/h)*, *Tasa de producción de licopeno (mmol/gdcw/h)* y *Rendimiento de licopeno (mol/mol glucose)*. Usando los siguientes comandos:

```
#Para la tasa de crecimiento reacción de biomasa
print('Growth Rate (1/h): ' + str(solution4.get_primal_by_id('2111')) #2
111 es la reacción objetivo de producción de biomasa
```

```
#Para obtener el valor del flujo de producción de licopeno
```

```

print('Lycopene
Production Rate (mmol/gdcw/h): ' + str(solution4.get_primal_by_id('LYCO-
dem')))

#Para el rendimiento
print('Lycopene Yield (mol/mol glucose): ' +
str(-solution4.get_primal_by_id('LYCO-
dem')/solution4.get_primal_by_id('1714')))#1714 es la reacción de consumo
de glucosa
    
```

3.6. Identificación de las reacciones involucradas en el medio de cultivo

Para identificar el medio de crecimiento actual del modelo se utilizó los comandos de COBRAPy:

Debido a que los id de las reacciones son valores numéricos, se identificó cada reacción en el modelo. Por ejemplo, para identificar la glucosa se usó el comando:

```

#Para ver las reacciones del medio
modelyall.reactions.get_by_id("1714")
    
```

La Tabla 3, presenta la información sobre los nutrientes del medio de crecimiento utilizados para el modelo metabólico de *Y lipolytica*. Un dato importante son los flujos, que tiene la unidad (concentración por gramo de peso seco de células y hora), **mmol / [gDW h]**.

Tabla 3. información sobre los nutrientes del medio de crecimiento en el modelo iYali4.

Nombre del nutriente	Nomenclatura	Id de la reacción	Id de metabolito	Límites de flujo superiores (mmol/gdcw/h)
Ammonium exchange	H ₃ N	1654	m1031	1000
D-glucose exchange	C ₆ H ₁₂ O ₆	1714	m511	0.649
H ⁺ exchange	H ⁺	1832	m342	1000
Oxygen exchange	O ₂	1992	m1339	1000
Phosphate exchange	H ₃ PO ₄	2005	m215	1000
Sulphate	H ₂ O ₄ S	2060	m1113	1000

exchange				
Water exchange	H ₂ O	2100	m214	1000
<i>Elaborado por: Milton Quinga.</i>				

3.7. Sobreexpresión de genes y optimización de la producción de licopeno

Este análisis partió con las reacciones heterólogas añadidas anteriormente y con la que se obtuvo la mejor optimización de producción de licopeno que fue con el par aceptor/receptor **FAD/FADH₂** de la reacción de formación de licopeno.

Para los análisis de sobreexpresión de genes, se tomaron en cuenta los siguientes genes (**HMG1, MVD1, EGR8, CrtB y CrtI**) que, en investigaciones publicadas, se ha demostrado que aumentan la producción de licopeno en *Yarrowia lipolytica* (Schwartz *et al.*, 2017; Zhang *et al.*, 2022)

Primero se realizó el análisis de variabilidad de flujo, (FVA) usando los comandos de COBRAPy. Se aplicó a las reacciones en las que actúan los genes a ser sobreexpresados (id: 558, 739, 904, R_phytoene y R_LycopeneFAD):

```
from cobra.flux_analysis import flux_variability_analysis

reactions_OE = [modelyall.reactions.get_by_id('558'),modelyall.reactions.get_by_id('739'),modelyall.reactions.get_by_id('904'),modelyall.reactions.get_by_id('R_phytoene' ),modelyall.reactions.get_by_id('R_LycopeneFAD')]
fva = flux_variability_analysis(modelyall, reaction_list = reactions_OE, fraction_of_optimum=0.9)
```

El FVA, obtenido encontró el flujo mínimo y máximo para las reacciones a ser sobre expresadas, soportando el 90% de la tasa de producción de biomasa máxima posible:

	minimum	maximum
558	0.015072	0.451429
739	0.015072	0.451429
904	0.015072	0.451429
R_phytoene	0.000981	0.056423
R_LycopeneFAD	0.000981	0.056423

Los valores de flujos máximos sirvieron para modificar los límites inferiores (*lower_bound*) para simular la sobreexpresión.

Por ejemplo, para sobreexpresar el gen *HMG1* se aplicó los siguientes comandos.

```
#Sobre-expresión del gen HMG1  
hmg1= modelyall.reactions.get_by_id('558').lower_bound = 0.45
```

Se aplicó el algoritmo a los demás genes y se optimizó con FBA para determinar la mejor sobreexpresión.

3.8. Eliminación de genes y optimización de la producción de licopeno

Este análisis partió de la optimización de producción de licopeno con el par aceptor/receptor FAD/FADH₂ sin aplicar la sobreexpresión de genes.

Para los análisis de eliminación de genes, se tomaron en cuenta los siguientes genes:

- *POX1* (**acyl-CoA oxidase**) id: YALI0D24750g (GPR).
- *GUT2* (**glycerol-3-phosphate dehydrogenase**) id: YALI0B13970g (GPR)

En investigaciones publicadas, se ha demostrado que favorecen la producción lípidos y licopeno en *Yarrowia lipolytica* (Matthäus *et al.*, 2014; Zhang *et al.*, 2020).

Se usaron los comandos de COBRAPy para realizar *knockout* de estos genes y así analizar qué sucede si se eliminan estos genes.

Por ejemplo, para la eliminación del gen *POX1* se utilizó el siguiente comando:

```
modelyall.genes.YALI0D24750g.knock_out()
```

Se aplicó el algoritmo a los demás genes y se optimizó con FBA.

3.9. Gráficos de planos de fase de fenotipo (*Production envelopes*)

Este tipo de gráficos permite un modelado teórico de lo que la célula puede o no puede hacer en términos de restricciones estequiométricas. Los planos de fase entre el crecimiento y un producto de interés producen: (*Production envelopes*), que son una representación entre el intercambio entre la producción del producto deseado y el crecimiento. Además, son una forma rápida de inspeccionar las limitaciones del sistema a diseñar y cómo se relaciona la producción con el crecimiento usando sustratos diferentes (Edwards *et al.* 2001)

Se realizaron estos gráficos para analizar los resultados de:

- Optimización de la producción de licopeno (Evaluación del par aceptores/ aceptores reducidos: FAD/FADH₂). Usando los comandos de COBRApy y Cameo y,
- Sobreexpresión de genes y optimización de la producción de licopeno (Sobreexpresión del gen *HMG1*). Usando los comandos de COBRApy y Cameo.

Por ejemplo, para obtener los gráficos de planos de fase de fenotipo de la sobreexpresión del gen *HMG1*, los comandos fueron los siguientes:

Comandos de COBRApy:

Para el gráfico de producción de licopeno y biomasa en función del oxígeno del medio de cultivo y como fuente de carbono la glucosa:

```
Para importar la libreria de COBRAY para desarrollar los graficos de plan  
o de fase  
from cobra.flux_analysis import production_envelope  
  
#Para el oxígeno y producción de licopeno  
#id O2:1992  
#id Glucosa: 1714  
prod_env5 = production_envelope(  
    modelyall, ["1992"], objective="LYCO-dem", carbon_sources="1714")  
  
#Para el oxígeno y producción de biomasa  
#id O2:1992  
#id Glucosa: 1714  
#id de biomasa: 2111  
prod_env6 = production_envelope(  
    modelyall, ["1992"], objective="2111", carbon_sources="1714")  
  
#Para graficar  
import matplotlib.pyplot as plt  
  
fig, ax = plt.subplots()  
ax.plot(prod_env5['1992'],prod_env5['flux_maximum'], label = "Licopeno")  
ax.plot(prod_env6['1992'],prod_env6['flux_maximum'], label = "Biomasa")  
ax.set_xlabel("EX_o2_e, [mmol/gdcw/h]")
```

```
ax.set_ylabel("f  
lux_maximum, [mmol/gdcw/h]")  
ax.legend(loc = 'upper right')  
plt.show()
```

De la misma manera, se aplicó el algoritmo para el gráfico de producción de licopeno y biomasa en función del nitrógeno del medio de cultivo y como fuente de carbono la glucosa:

Comandos de Cameo:

Para el gráfico de producción de licopeno vs biomasa tomando un límite inferior determinado para el oxígeno de **-0.8**:

```
Para general del dataframe  
modelyall.reactions.get_by_id("1992").lower_bound = -0.8  
result1 = phenotypic_phase_plane(modelyall,  
                                variables=[modelyall.reactions.get_by_id(  
"2111")],  
                                objective=modelyall.reactions.get_by_id("  
LYCO-dem"),  
                                points=10)  
  
#Para obtener el gráfico  
from cameo.visualization.plotting.with_plotly import PlotlyPlotter  
plotter = PlotlyPlotter()  
result1.plot(plotter)
```

De la misma manera se aplicó el mismo algoritmo para el gráfico de la producción de licopeno vs biomasa tomando con un límite inferior determinado para el nitrógeno de **-0.2**.

3.10. Visualización de flujos obtenidos

Se utilizó la aplicación web **Fluxer** (<https://fluxer.umbc.edu/>) para cargar y analizar los flujos obtenidos de:

- Optimización de la producción de licopeno (Evaluación del par aceptores/ aceptores reducidos: FAD/FADH₂).
- Sobreexpresión de genes y optimización de la producción de licopeno (Sobreexpresión del gen *HMG1*).

- Eliminación

de genes (*POX1* y *GUT2*) y optimización de la producción de licopeno.

Para guardar los cambios en el modelo metabólico obtenidos en los análisis anteriores y llevarlos al formato SBML necesario para ser utilizado en **Fluxer**, se empleó los siguientes comandos:

```
#Para importar las librerías necesarias para guardar nuestro modelo metabólico en SBML
from pathlib import Path
from cobra.io import load_json_model, save_json_model, load_matlab_model,
    save_matlab_model, read_sbml_model, write_sbml_model
import logging

#Para guardar el modelo en formato SBML
write_sbml_model(modelyall, "modelFAD.xml")
write_sbml_model(modelyall, "modeldelection.xml")
```

3.11. Tipo de investigación

Según el nivel o alcance:

- **Descriptiva:** Puesto que se busca explicar las rutas metabólicas y los procesos que llevar a la optimización de la producción de licopeno en *Y. lipolytica*.

Según el diseño:

- **Modalidad:** Enfoque cuantitativo, ya que se realizará el análisis de datos de simulaciones metabólicas, para contestar la pregunta de investigación.
- **Tipo de diseño:** Experimental, se pretende establecer el posible efecto de una causa que se manipula, análisis de las variables: Variables independientes: Algoritmos y herramientas computacionales de ingeniería metabólica en Python. Variable dependiente: Producción de licopeno en *Y. lipolytica*.

Según el propósito:

- **Aplicada:** Porque busca la generación de conocimiento con aplicación directa a un problema, en nuestro caso la optimización de la producción de licopeno en *Y. lipolytica*.

CAPÍTULO 4:

RESULTADOS Y DISCUSIÓN

4.1. Análisis de la optimización de la producción de licopeno usando FBA (diferentes factores)

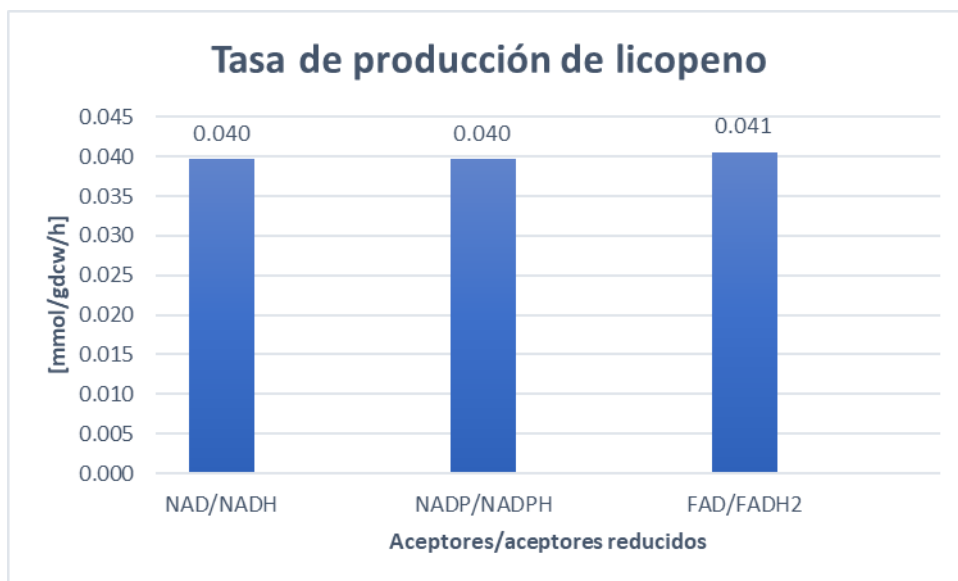
Los resultados de la optimización de la producción de licopeno usando FBA, evaluando diferentes factores (Anexo 2 y Figura 5) en el modelo metabólico *iYali4*. Indicaron lo siguiente:

- **Para el caso de usar diferentes pares de aceptor/aceptores reducidos en la reacción de formación de licopeno (en la que actúa la enzima *licopeno sintasa*).** Se obtuvo resultados similares en la tasa de crecimiento, tasa de producción de licopeno y rendimiento de licopeno (Figura 5A). Observando además un ligero aumento en la tasa de producción y rendimiento de licopeno al usar FAD/FADH₂. Esto pudo darse debido a que se ha reportado que la enzima *licopeno sintasa* o también conocida como *fitoeno desaturasa* [EC 1.3.99.31] requiere de FAD, cuando actúa en la formación de licopeno en el organismo *Pantoea ananatis* (Fraser *et al.*, 1992).
Además, se ha reportado que el cofactor NADPH es clave para la biosíntesis de carotenoides en cepas recombinantes que expresa genes exógenos para sintetizar carotenoides en *Saccharomyces cerevisiae* (Zhao *et al.*, 2015)
- **Para el caso de la sobreexpresión de genes *HMG1*, *MVD1*, *EGR8*, *CrtB* y *Crtl*.** Se obtuvo resultados similares en la tasa de crecimiento, tasa de producción de licopeno y rendimiento de licopeno para los genes *HMG1*, *MVD1*, *EGR8* (Figura 5B). Que resultaron valores mayores a los obtenidos con la sobreexpresión de los genes *CrtB* y *Crtl*. Esto puede darse debido a que resulta crucial la sobreexpresión de los genes de la vía del mevalonato (*HMG1*, *MVD1*, *EGR8*) para aumentar la producción de licopeno, dado que aumentan las reservas de precursores de carotenoides y el flujo al compuesto deseado (Zhang *et al.*, 2022). Estos resultados concuerdan con los obtenidos por Schwartz y colaboradores (2022), que analizaron las sobreexpresiones de estos genes y obtuvieron con éxito un aumento de la producción de licopeno en cepas recombinantes de *Yarrowia lipolytica*.

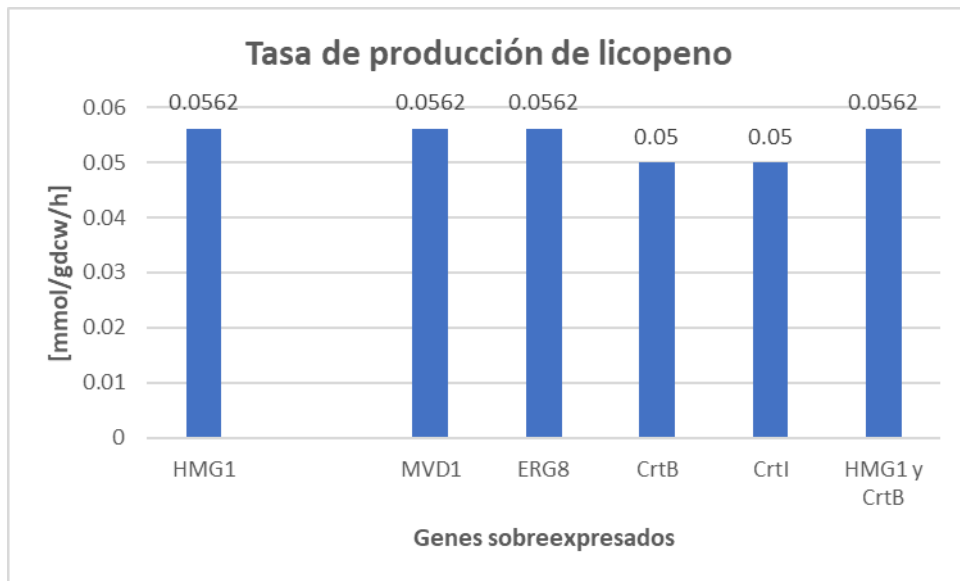
- Para el

caso de la eliminación de los genes *POX1* y *GUT2*. No se observaron cambios en las la tasa de crecimiento, tasa de producción de licopeno y rendimiento de licopeno aplicando FBA (Figura 5C). Pese a que en estudios publicados, se ha visto que la eliminación de genes del primer paso de la β -oxidación (*POX1* a *POX6*) y la delección del gen que codifica la glicerol-3-fosfato deshidrogenasa (*GUT2*, que cataliza la conversión de glicerol-3 -fosfato a fosfato de dihidroxiacetona), aumenta la acumulación de lípidos y licopeno en las cepas mutantes de *Yarrowia lipolytica* (Matthäus *et al.*, 2014; Zhang *et al.*, 2020). Esto se puede explicar debido a que es necesario aplicar más enfoques en la optimización para tratar el caso de cepas *knockout*, como por ejemplo el análisis MOMA (Segrè *et al.*, 2002; Volkova *et al.*, 2020).

A



B



C

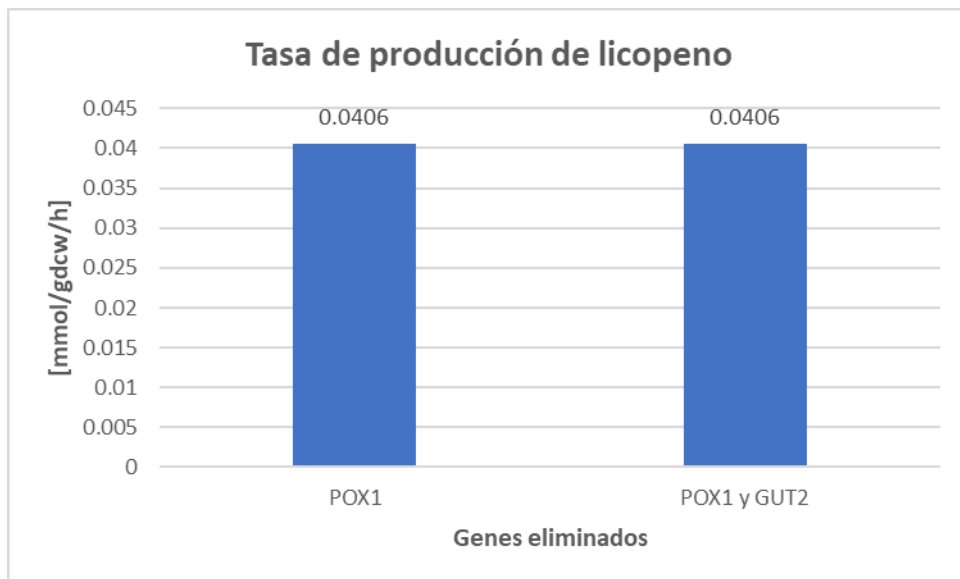


Figura 5. Resultados obtenidos en la optimización de la tasa de producción de licopeno en *Yarrowia lipolytica* usando FBA en el modelo *iYali4* para diferentes factores.

(A) Aceptores/aceptores reducidos, (B) Sobreexpresión de genes y
(C) Eliminación de genes.

4.2. Análisis de los gráficos de planos de fase de fenotipo (Production envelopes)

Los gráficos obtenidos de los planos de fase de fenotipo, muestran distintas fases de crecimiento óptimo con un uso diferente de dos sustratos diferentes. Estos gráficos revelaron lo siguiente:

- **Para la optimización de la producción de licopeno (Evaluación del par aceptores/ aceptores reducidos: FAD/FADH₂):**
 - La Figura 6, simula la producción de licopeno y biomasa, tomando en cuenta el consumo de oxígeno y como fuente de carbono una tasa de captación fija de glucosa de 0.649 [mmol/gdcw/h]. Se pudo comprobar el crecimiento anaerobio estricto de *Yarrowia lipolytica*. Debido a que una vez que el crecimiento de la levadura llega a su flujo máximo decae a media que la tasa de consumo de oxígeno aumenta y de la misma manera la producción de licopeno decae a la par cuando disminuye el flujo de biomasa (Jach & Malm, 2022). Adicionalmente se ha reportado que la limitación de oxígeno, no conduce a la producción de mayor cantidad de licopeno por unidad de DCW (peso seco de células) (Matthäus *et al.*, 2014).

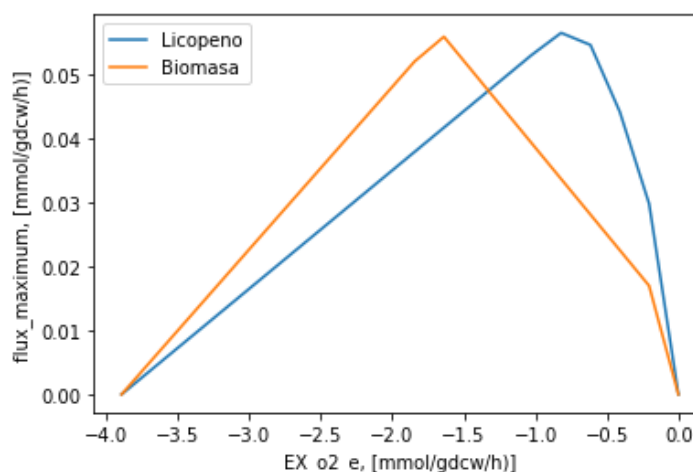


Figura 6. Producción de licopeno y biomasa en función del oxígeno (medio de cultivo) y como fuente de carbono la glucosa (Evaluación del par aceptores/ aceptores reducidos: FAD/FADH₂).

Usando el comando *production_envelope* de COBRApy y la librería *Matplotlib* de Python. El eje de las abscisas representa la tasa de consumo de oxígeno. El eje de las ordenadas representa el flujo máximo de la producción de licopeno y biomasa.

o La

Figura 7, simula la producción de licopeno y biomasa, tomando en cuenta el consumo de nitrógeno y como fuente de carbono una tasa de captación fija de glucosa de 0.649 [mmol/gdcw/h]. Aquí se pudo observar que la tasa de consumo de nitrógeno es un limitante para la producción de biomasa, mientras que la tasa de producción de licopeno permanece en aumento. Esto puede darse debido a que se ha reportado que la limitación de nitrógeno es favorable para la acumulación de carotenoides (Matthäus *et al.*, 2014).

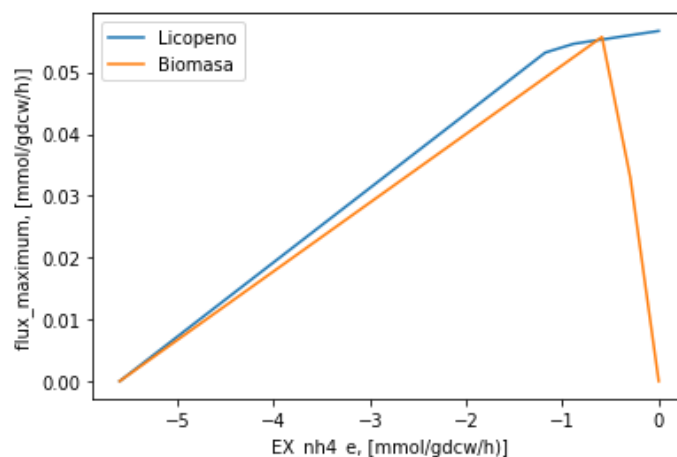


Figura 7. Producción de licopeno y biomasa en función del nitrógeno (medio de cultivo) y como fuente de carbono la glucosa (Evaluación del par aceptores/ aceptores reducidos: FAD/FADH₂).

Usando el comando *production_envelope* de COBRAPy y la librería *Matplotlib* de Python. El eje de las abscisas representa la tasa de consumo de nitrógeno. El eje de las ordenadas representa el flujo máximo de la producción de licopeno y biomasa.

- o Las Figuras 8 y 9, presentan la relación que existe entre la tasa de producción de biomasa y la tasa de producción de licopeno. A una tasa de captación fija de glucosa de 0.649 [mmol/gdcw/h], oxígeno de -1.0 [mmol/gdcw/h] y nitrógeno -0.8 [mmol/gdcw/h] respectivamente. Mostrando que a un rango de entre 0 a 0.0005 [mmol/gdcw/h] de la tasa de crecimiento (biomasa), se alcanza la mayor tasa de producción de licopeno. Cuando empieza a aumentar la tasa de biomasa empieza a decaer la tasa de producción de licopeno. Esto se explicaría debido a

que

el aumento en la tasa de crecimiento requiere los recursos que también son necesarios para la producción de licopeno. Se ha observado en investigaciones que la velocidad de crecimiento reducida parece aumentar la acumulación de licopeno (Matthäus *et al.*, 2014).

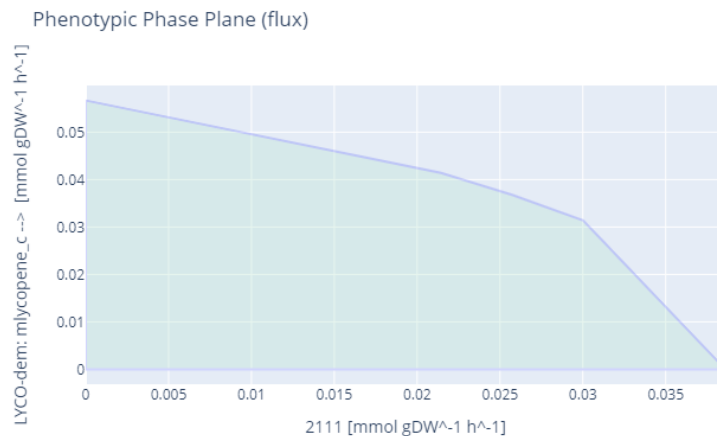


Figura 8. Producción de licopeno vs biomasa tomando un límite inferior determinado (-1.0) para el oxígeno: (Evaluación del par aceptores/ aceptores reducidos: FAD/FADH2).

Usando el comando *phenotypic_phase_plane* y *PlotlyPlotter* de Cameo. El eje de las abscisas representa la tasa de producción de biomasa. El eje de las ordenadas representa la tasa de producción de licopeno.

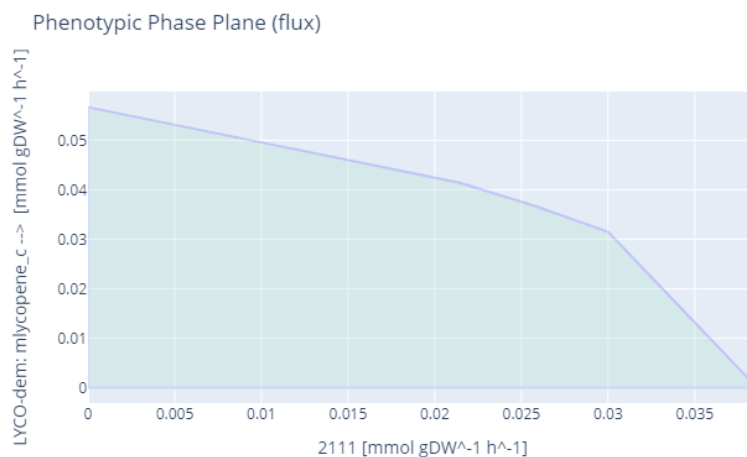


Figura 9. Producción de licopeno vs biomasa tomando un límite inferior determinado (-0.8) para el nitrógeno: (Evaluación del par aceptores/ aceptores reducidos: FAD/FADH2).

Usando el comando *phenotypic_phase_plane* y *PlotlyPlotter* de Cameo. El eje de las abscisas representa la tasa de producción de biomasa. El eje de las ordenadas representa la tasa de producción de licopeno.

- Para la

sobreexpresión de genes (Sobreexpresión del gen *HMG1*):

- La Figura 10, simula la producción de licopeno y biomasa, tomando en cuenta el consumo de oxígeno y como fuente de carbono una tasa de captación fija de glucosa de 0.649 [mmol/gdcw/h]. Adicional a lo ya expuesto en las anteriores figuras de planos de fase de fenotipo. Se pudo apreciar con más detalle el proceso de sobreproducción de licopeno al observar un aumento en la tasa de producción de licopeno y una disminución de la tasa de crecimiento, en condiciones anaerobias.

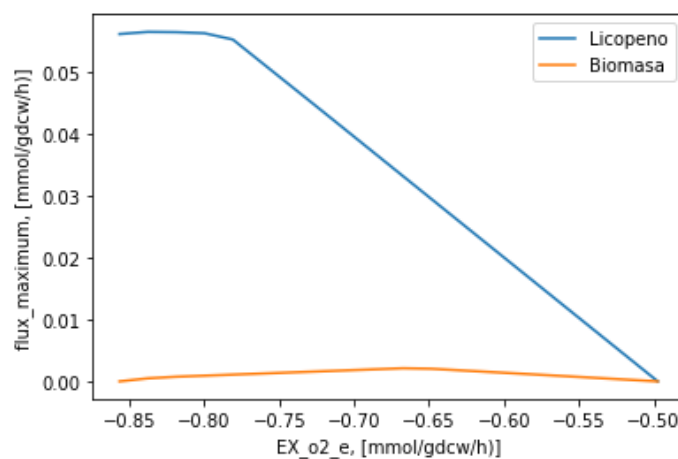


Figura 10. Producción de licopeno y biomasa en función del oxígeno (medio de cultivo) y como fuente de carbono la glucosa (Evaluación de la sobreexpresión del gen *HMG1*).

Usando el comando *production_envelope* de COBRApy y la librería *Matplotlib* de Python. El eje de las abscisas representa la tasa de consumo de oxígeno. El eje de las ordenadas representa el flujo máximo de la producción de licopeno y biomasa.

- La Figura 11, simula la producción de licopeno y biomasa, tomando en cuenta el consumo de nitrógeno y como fuente de carbono una tasa de captación fija de glucosa de 0.649 [mmol/gdcw/h]. Adicional a lo ya expuesto en las anteriores figuras de planos de fase de fenotipo. Se pudo apreciar con más detalle el proceso de sobreproducción de licopeno al observar un aumento en la tasa de producción de licopeno y

una

disminución de la tasa de crecimiento, en condiciones limitantes de nitrógeno.

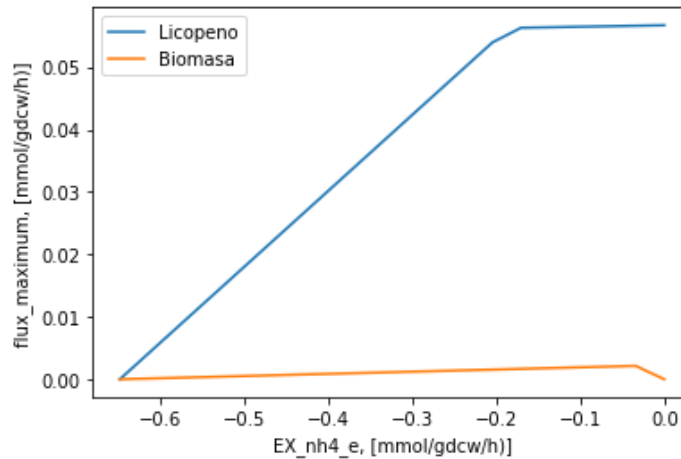


Figura 11. Producción de licopeno y biomasa en función del nitrógeno (medio de cultivo) y como fuente de carbono la glucosa (Evaluación de la sobreexpresión del gen *HMG1*).

Usando el comando *production_envelope* de COBRApy y la librería *Matplotlib* de Python. El eje de las abscisas representa la tasa de consumo de oxígeno. El eje de las ordenadas representa el flujo máximo de la producción de licopeno y biomasa.

- Las Figuras 12 y 13, presentan la relación que existe entre la tasa de producción de biomasa y la tasa de producción de licopeno. A una tasa de captación fija de glucosa de 0.649 [mmol/gdcw/h], oxígeno de -0.8 [mmol/gdcw/h] y nitrógeno -0.2 [mmol/gdcw/h] respectivamente. Mostrando nuevamente que a media que velocidad de crecimiento se reduce, aumentar la acumulación de licopeno.

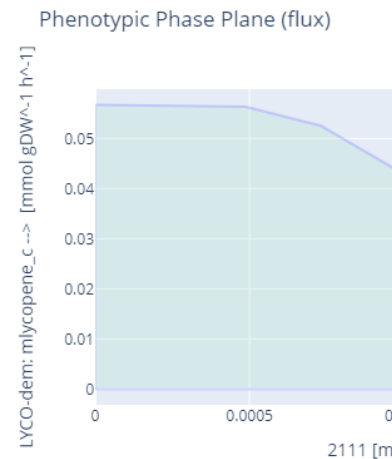


Figura 12. Producción de licopeno vs biomasa tomando un límite inferior determinado

(-0.8) para el oxígeno: (Evaluación de la sobreexpresión del gen *HMG1*)

Usando el comando *phenotypic_phase_plane* y *PlotlyPlotter* de Cameo. El eje de las abscisas representa la tasa de producción de biomasa. El eje de las ordenadas representa la tasa de producción de licopeno.

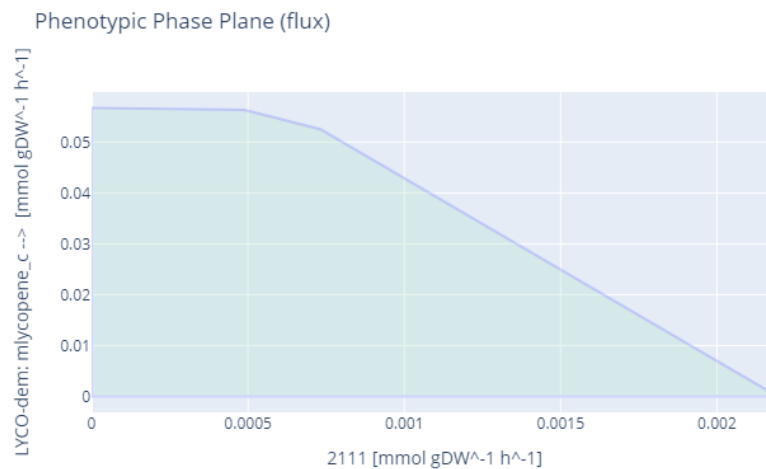


Figura 13. Producción de licopeno vs biomasa tomando un límite inferior determinado

(-0.2) para el nitrógeno: (Evaluación de la sobreexpresión del gen *HMG1*).

Usando el comando *phenotypic_phase_plane* y *PlotlyPlotter* de Cameo. El eje de las abscisas representa la tasa de producción de biomasa. El eje de las ordenadas representa la tasa de producción de licopeno

4.3. Análisis de la

visualización de flujos metabólicos en Fluxer

Los resultados de los flujos metabólicos se visualizaron en *Fluxer*, de la siguiente manera:

- En la Optimización de la producción de licopeno (Evaluación del par aceptores/ aceptores reducidos: FAD/FADH₂) (Figura 14).
- Sobreexpresión de genes y optimización de la producción de licopeno (Sobreexpresión del gen *HMG1*) (Figura 15).
- Eliminación de genes (*POX1* y *GUT2*) y optimización de la producción de licopeno. (Figura 16).

Se uso la opción *Shortest K Paths*. Para obtener las k rutas más cortas entre una fuente dada y un metabolito o reacción de destino. Con la siguiente configuración:

- k: 4, Best Paths/First Paths.
- Source: acetyl-CoA C-acetyltransferase (id: 104).
- Target: Lycopene demand.
- Layout: Dagre.

Los gráficos generados (Figuras 14, 15 y 16), mostraron respectivamente las principales k rutas más cortas entre acetyl-CoA C-acetyltransferase y Lycopene demand para la optimización de la producción de licopeno.

Esta herramienta resulto ser sumamente útil para explorar los flujos metabólicos involucrados en la producción de licopeno en el modelo metabólico *iYali4* de *Yarrowia lipolytica*.

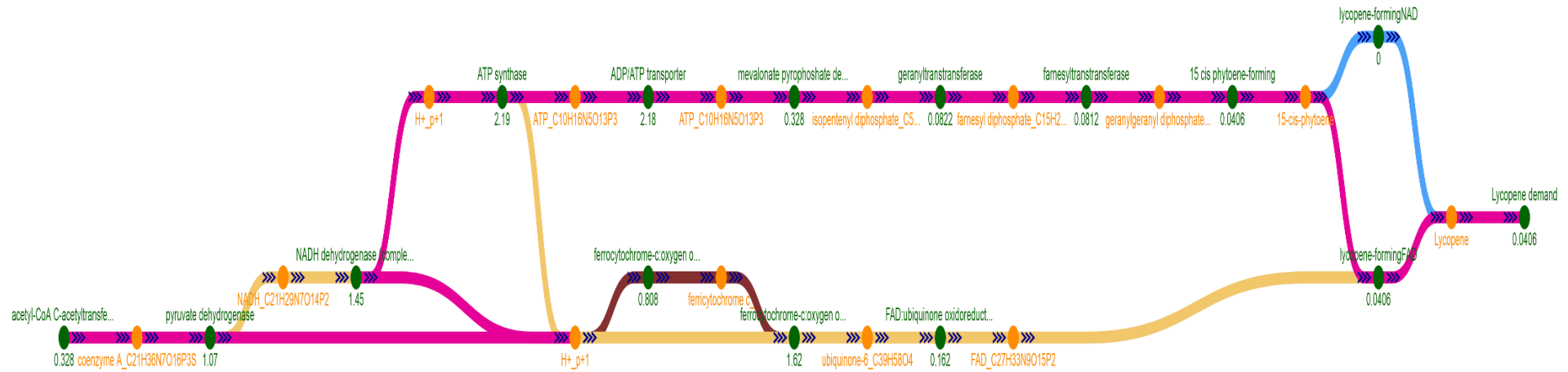
Resulta importante además esta herramienta, pues se observó que presenta más vías importantes en la síntesis de licopeno tales como:

- **Piruvato deshidrogenasa** para mejorar la producción de isoprenoides en levadura (Shibaa *et al.*, 2007).
- **NADH-quinona oxidorreductasa (complejo I)**, convierte NADH en NADPH aumentó la producción de licopeno y β -caroteno (Li *et al.*, 2019).

- ATP

sintasa, los genes relacionados a esta enzima están implicados en la biosíntesis temprana de isoprenoides de la formación de licopeno (Ya *et al.*, 2022).

Que podrían ser consideradas para realizar análisis de sobreexpresión o eliminación de genes.



Leyenda del gráfico

Graph Legend

● : Metabolite/Reaction Node

Reaction Flux, Stoichiometric Coefficient, Edge Weight

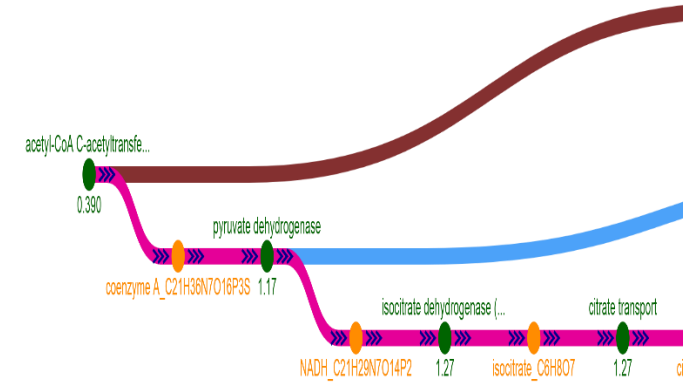
k-path color key



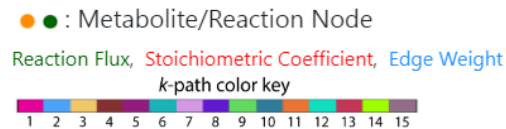
Figura 14. Principales k rutas más cortas entre acetyl-CoA C-acetyltransferase y Lycopene demand. Flujos metabólicos obtenidos de la optimización de la producción de lycopeno (Evaluación del par aceptores/ aceptores reducidos: FAD/FADH2).

Desarrollados en *Fluxer*. Vista Best Paths. Link para visualizar la red metabolica:

<https://fluxer.umbc.edu/model?id=7172c799f386a350e85459a828516debe9a04e0dL>



Graph Legend

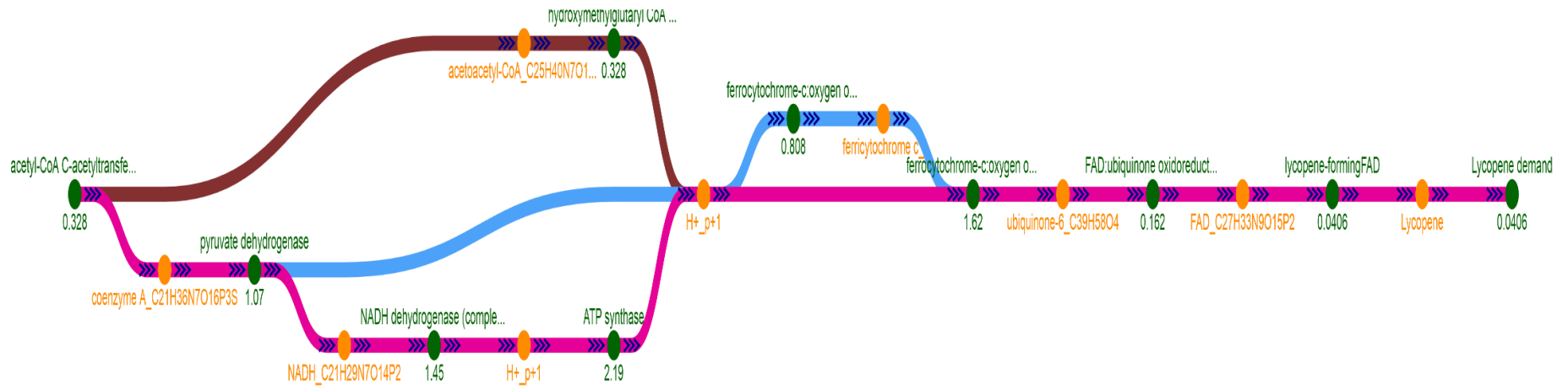


Leyenda del gráfico

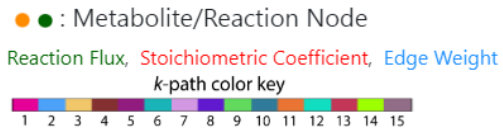
Figura 15. Principales k rutas más cortas entre acetyl-CoA C-acetyltransferase y Lycopene demand. Flujos metabólicos obtenidos de la sobreexpresión de genes (gen HMG1) y optimización de la producción de licopeno.

Desarrollados en *Fluxer*. Vista *First Paths*. Link para visualizar la red metabólica:

<https://fluxer.umbc.edu/model?id=27cf121d83d71cface49d5a5e73a967aba81bf11>



Graph Legend



Leyenda del gráfico

Figura 16. Principales k rutas más cortas entre acetyl-CoA C-acetyltransferase y Lycopene demand. Flujos metabólicos obtenidos de la eliminación de genes (POX1 y GUT2) y optimización de la producción de licopeno

Desarrollados en *Fluxer*. Vista *First Paths*. Link para visualizar la red metabólica:

<https://fluxer.umbc.edu/model?id=5cabd80ad92404e163eb4aa99b2b579907c287ba>

CONCLUSIONES Y RECOMENDACIONES

5.1. Conclusiones

- Se implementó un algoritmo desarrollado en Python para el modelo metabólico a escala del genoma *iYali4* de *Yarrowia lipolytica* usando los paquetes COBRApy y Cameo para simular ingeniería metabólica asistida por computador. Esto permitió agregar los genes, reacciones y metabolitos necesarios para que el modelo pueda producir licopeno usando la vía del mevalonato y vías heterólogas adicionales para análisis metabólicos constreñidos por la estequiometría.
- El modelo con las modificaciones implementadas permitió simular la optimización de la producción de licopeno con efectos menores en la tasa de crecimiento. Los análisis permitieron simular factores tales como: la evaluación del par aceptores/ aceptores reducidos (FAD/FADH₂) en la reacción de formación de licopeno, la sobreexpresión de los genes *HMG1*, *MVD1*, *EGR8*, *CrtB* y *CrtI* y la eliminación de los genes *POX1* y *GUT2*.
- Para obtener gráficos que muestren distintas fases de crecimiento óptimo con un uso diferente de dos sustratos diferentes, se pudo implementar la creación de gráficos de planos de fase de fenotipo (*Production envelopes*) para cada optimización que mostraron como el oxígeno y nitrógeno limitan la producción de biomasa y licopeno. Además, como a medida que aumenta la producción de biomasa, disminuye la producción de licopeno.
- La implementación de la visualización de los principales flujos metabólicos en la herramienta *Fluxer* resultó útil para explorar más vías importantes en la biosíntesis de licopeno tales como: la vía de la Piruvato deshidrogenasa, la vía de la NADH-quinona oxidoreductasa (complejo I), y la vía de la ATP sintasa.
- Se concluye además que la estrategia para optimizar la producción de licopeno en *Yarrowia lipolytica* radicaría en los siguiente: Usar un adecuado par de aceptor/aceptor reducido en la reacción de formación de licopeno, sobre expresar los genes de la ruta del mevalonato y buscar más vías que influyan en la biosíntesis de licopeno en la levadura.

- El algoritmo implementado aquí sería de gran utilidad para los investigadores que se dedican a la ingeniería metabólica, permitiéndoles utilizar herramientas de diseño computacional como parte de su flujo de trabajo. A fin de facilitar el modelamiento de la producción sostenible de metabólicos de interés comercial, en este caso del licopeno en *Yarrowia lipolytica*.

5.2. Recomendaciones

En base a los resultados obtenidos, se recomienda para futuros trabajos:

- Además de usar modelos basados en restricciones metabólicas como FBA y FVA; emplear herramientas basadas en restricción enzimática como los modelos GECKO (modelos a escala del genoma con restricciones enzimáticas utilizando datos cinéticos y ómicos) y restricción regulatoria como el algoritmo *OptRAM* (que puede identificar estrategias de optimización combinatoria que incluyen sobreexpresión o eliminación tanto de genes metabólicos como de factores de transcripción). Para obtener datos más precisos de la predicción del fenotipo y regulación génica (Pereira *et al.*, 2021; Shen *et al.*, 2019).
- Mejorar el suministro de acetil.CoA citosólico, que es crucial para la producción de licopeno. Introduciendo vías heterólogas o redireccionando el metabolismo del carbono central nativo en el modelo metabólico (Zhang *et al.*, 2022).
- Para el caso de eliminación de genes, utilizar algoritmos para predecir los objetivos de la eliminación de genes tales como: algoritmos evolutivos como *OptGene* o programación lineal como *OptKnock* (Burgard *et al.*, 2003; Patil *et al.*, 2005).

CAPÍTULO 6:

REFERENCIAS

BCC Research. (2022). The Global Market for Carotenoids. Disponible en: <https://www.bccresearch.com/market-research/food-and-beverage/the-global-market-for-carotenoids.html>

Birkel, G.W., Ghosh, A., Kumar, V.S. *et al.* (2017). The JBEI quantitative metabolic modeling library (jQMM): a python library for modeling microbial metabolism. *BMC Bioinformatics* 18, 205. Disponible en: <https://doi.org/10.1186/s12859-017-1615-y>

Bordbar, A., Monk, J. M., King, Z. A., & Palsson, B. O. (2014). Constraint-based models predict metabolic and associated cellular functions. *Nature reviews. Genetics*, 15(2), 107–120. Disponible en: <https://doi.org/10.1038/nrg3643>

Burgard, A.P., Pharkya, P., Maranas, C.D. (2003), “*OptKnock: A Bilevel Programming Framework for Identifying Gene Knockout Strategies for Microbial Strain Optimization,*” *Biotechnology and Bioengineering*, 84(6), 647-657. Disponible en: <https://doi.org/10.1002/bit.10803>

Cardoso J, Jensen K., Lieven C., Lærke A., Galkina S., Beber M., Özdemir E., Herrgård M., Redestig H. & Sonnenschein N. (2018). *Cameo: A Python Library for Computer Aided Metabolic Engineering and Optimization of Cell Factories.* *ACS Synthetic Biology* 7 (4), 1163-1166. Disponible en: <https://pubs.acs.org/doi/10.1021/acssynbio.7b00423>

COBRAPy. (2022). Docs 7. Production envelopes. Disponible en: https://cobrapy.readthedocs.io/en/latest/phenotype_phase_plane.html

Da Veiga J., Jolicoeur M., Schwartz L. Peres S. (2021). Fine-tuning mitochondrial activity in *Yarrowia lipolytica* for citrate overproduction. *Sci Rep* 11, 878 (2021). Disponible en: <https://doi.org/10.1038/s41598-020-79577-4>

Ebrahim, A., Lerman, J.A., Palsson, B.O. *et al.* (2013). COBRApy: COntstraints-Based Reconstruction and Analysis for Python. *BMC Syst Biol* 7, 74. Disponible en: <https://doi.org/10.1186/1752-0509-7-74>

Erickson K. E. (2018). FBA with CobraPy. Disponible en: <https://cnls.lanl.gov/external/qbio2018/Slides/FBA%202/qBio-FBA-lab-slides.pdf>

Fathi Z., Larissa Ribeiro Ramos Tramontin, Gholamhossein Ebrahimipour, Irina Borodina, Farshad Darvishi. (2021). Metabolic engineering of *Saccharomyces cerevisiae* for production of β -carotene from hydrophobic substrates, *FEMS Yeast Research*, Volume 21, Issue 1, February 2021, foaa068. Disponible en: <https://doi.org/10.1093/femsyr/foaa068>

Fraser, P. D., Misawa, N., Linden, H., Yamano, S., Kobayashi, K., & Sandmann, G. (1992). Expression in *Escherichia coli*, purification, and reactivation of the recombinant *Erwinia uredovora* phytoene desaturase. *The Journal of biological chemistry*, 267(28), 19891–19895.

Groenewald M, Boekhout T, Neuvéglise C, Gaillardin C, van Dijk PWM, Wyss M (2013) *Yarrowia lipolytica*: safety assessment of an oleaginous yeast with a great industrial potential. *Crit Rev Microbiol* 40(3):187–206

Gudmundsson S. & Thiele I. (2010). Computationally efficient flux variability analysis. *BMC Bioinformatics* 11, 489 . Disponible en: <https://doi.org/10.1186/1471-2105-11-489>

Hari A. & Lobo D.(2020). Fluxer: a web application to compute, analyze and visualize genome-scale metabolic flux networks, *Nucleic Acids Research*, Volume 48, Issue W1, 02 July 2020, Pages W427–W435. Disponible en: <https://doi.org/10.1093/nar/gkaa409>

Hucka, M., Bergmann, F. T., Dräger, A., Hoops, S., Keating, S. M., Le Novère, N., Myers, C. J., Olivier, B. G., Sahle, S., Schaff, J. C., Smith, L. P., Waltemath, D., & Wilkinson, D. J. (2018). The Systems Biology Markup Language (SBML): Language

Level 3 Version 2 Core. *Journal of integrative bioinformatics*, 15(1), 20170081.

Disponibile en <https://doi.org/10.1515/jib-2017-0081>

Jach, M. E., & Malm, A. (2022). *Yarrowia lipolytica* as an Alternative and Valuable Source of Nutritional and Bioactive Compounds for Humans. *Molecules (Basel, Switzerland)*, 27(7), 2300. Disponible en: <https://doi.org/10.3390/molecules27072300>

Jach, M.E., Baj, T. & Juda, M. *et al.* (2020). Statistical evaluation of growth parameters in biofuel waste as a culture medium for improved production of single cell protein and amino acids by *Yarrowia lipolytica*. *AMB Expr* 10, 35 (2020). <https://doi.org/10.1186/s13568-020-00968-x>

Jingbo Ma, Yang Gu, Monireh Marsafari, Peng Xu. (2020). Synthetic biology, systems biology, and metabolic engineering of *Yarrowia lipolytica* toward a sustainable biorefinery platform, *Journal of Industrial Microbiology and Biotechnology*, Volume 47, Issue 9-10, 1, Pages 845–862. Disponible en <https://doi.org/10.1007/s10295-020-02290-8>

Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 28(1), 27–30. Disponible en: <https://doi.org/10.1093/nar/28.1.27>

Kerkhoven, E. J., Pomraning, K. R., Baker, S. E., & Nielsen, J. (2016). Regulation of amino-acid metabolism controls flux to lipid accumulation in *Yarrowia lipolytica*. *NPJ systems biology and applications*, 2, 16005. <https://doi.org/10.1038/npjbsba.2016.5>

Kim H., Charusanti P. b, Yup Lee S. Tilmann Weber T. (2016). Metabolic engineering with systems biology tools to optimize production of prokaryotic secondary metabolites. *Nat. Prod. Rep.* 33, 933-941. Disponible en: <https://pubs.rsc.org/en/content/articlehtml/2016/np/c6np00019c>

Kumar, R. R., &

Prasad, S. (2011). Metabolic engineering of bacteria. *Indian journal of microbiology*, 51(3), 403–409. Disponible en: <https://doi.org/10.1007/s12088-011-0172-8>

Larroude, M., Rossignol, T., Nicaud, J. M., & Ledesma-Amaro, R. (2018). Synthetic biology tools for engineering *Yarrowia lipolytica*. *Biotechnology advances*, 36(8), 2150-2164. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6261845/>

Li, C., Swofford, C. A., & Sinskey, A. J. (2019). Modular engineering for microbial production of carotenoids. *Metabolic engineering communications*, 10, e00118. Disponible en: <https://doi.org/10.1016/j.mec.2019.e00118>

Liu L. Q., Redden H., Alper H. S. (2013). Frontiers of yeast metabolic engineering: diversifying beyond ethanol and *Saccharomyces*. *Curr. Opin. Biotechnol.* 24 1023–1030. 10.1016/j.copbio.2013.03.005

Lu, H., Li, F., Sánchez, B.J. *et al.* (2019). A consensus *S. cerevisiae* metabolic model Yeast8 and its ecosystem for comprehensively probing cellular metabolism. *Nat Commun* 10, 3586 (2019). Disponible en: <https://doi.org/10.1038/s41467-019-11581-3>

Madzak C. (2021). *Yarrowia lipolytica* Strains and Their Biotechnological Applications: How Natural Biodiversity and Metabolic Engineering Could Contribute to Cell Factories Improvement. *Journal of Fungi* 7:7, pages 548. Disponible en: <https://doi.org/10.1080/07388551.2021.1931016>

Maia, P., Rocha, M., and Rocha, I. (2016) In Silico Constraint-Based Strain Optimization Methods: the Quest for Optimal Cell Factories. *Microbiol. Mol. Biol. Rev.* 80, 45– 67, DOI: 10.1128/MMBR.00014-15,

Matplotlib. (2022). Matplotlib: Visualization with Python. Disponible en: <https://matplotlib.org/>

Ketelhot, M., Gatter, M., & Barth, G. (2014). Production of lycopene in the non-carotenoid-producing yeast *Yarrowia lipolytica*. *Applied and environmental microbiology*, 80(5), 1660–1669. Disponible en: <https://doi.org/10.1128/AEM.03167-13>

Nambou K, Jian X, Zhang X, Wei L, Lou J, Madzak C, Hua Q. (2015). Flux Balance Analysis Inspired Bioprocess Upgrading for Lycopene Production by a Metabolically Engineered Strain of *Yarrowia lipolytica*. *Metabolites*. 5(4):794-813. Disponible en: <https://doi.org/10.3390/metabo5040794>

Orth, J. D., Thiele, I., & Palsson, B. Ø. (2010). What is flux balance analysis? *Nature biotechnology*, 28(3), 245–248. Disponible en: <https://doi.org/10.1038/nbt.1614>

Patil, K. R., Rocha, I., Förster, J., & Nielsen, J. (2005). Evolutionary programming as a platform for in silico metabolic engineering. *BMC Bioinformatics*, 6, 308. Disponible en: <https://doi.org/10.1186/1471-2105-6-308>

Pereira V., Cruz F. & Rocha M. (2021). *MEWpy*: a computational strain optimization workbench in Python, *Bioinformatics*, Volume 37, Issue 16, 15, Pages 2494–2496. Disponible en: <https://doi.org/10.1093/bioinformatics/btab013>

Schomburg, I., Chang, A., & Schomburg, D. (2002). BRENDA, enzyme data and metabolic information. *Nucleic acids research*, 30(1), 47–49. Disponible en: <https://doi.org/10.1093/nar/30.1.47>

Schwartz, C., Frogue, K., Misa, J., & Wheeldon, I. (2017). Host and Pathway Engineering for Enhanced Lycopene Biosynthesis in *Yarrowia lipolytica*. *Frontiers in microbiology*, 8, 2233. Disponible en: <https://doi.org/10.3389/fmicb.2017.02233>

Segrè, D., Vitkup, D., & Church, G. M. (2002). Analysis of optimality in natural and perturbed metabolic networks. *Proceedings of the National Academy of Sciences of the*

America, 99(23), 15112–15117. Disponible en: <https://doi.org/10.1073/pnas.232349399>

Shen F, Sun R, Yao J, Li J., Liu Q. & Price N. D., *et al.* (2019) *OptRAM*: In-silico strain design via integrative regulatory-metabolic network modeling. *PLoS Comput Biol* 15(3): e1006835. Disponible en: <https://doi.org/10.1371/journal.pcbi.1006835>

Shibaa Y., Paradisea E., Kirby J., Kyun Ro D. & Keasling J. (2007). Engineering of the pyruvate dehydrogenase bypass in *Saccharomyces cerevisiae* for high-level production of isoprenoids. *Metabolic Engineering*. Disponible en: <https://www.sciencedirect.com/science/article/abs/pii/S1096717606001042>

Tafur A., Ríos W., Mejía D., Ojeda C., Carlson R., Gómez J. & González A. (2021). In silico Design for Systems-Based Metabolic Engineering for the Bioconversion of Valuable Compounds From Industrial By-Products. *Front. Genet.* 12:633073. Disponible en: <https://www.frontiersin.org/articles/10.3389/fgene.2021.633073/full>

Takehima M., Ono M., Higuchi T., Chen C., Hara T. & Nakano S.(2014). Anti-proliferative and apoptosis-inducing activity of lycopene against three subtypes of human breast cancer cell lines. *Cancer Sci.* 105, 252–257. Disponible en: <https://onlinelibrary.wiley.com/doi/10.1111/cas.12349>

Thiele S., Kamp A., Bekiaris P., Schneider P. & Klamt S. (2022). CNApy: a CellNetAnalyzer GUI in Python for analyzing and designing metabolic networks, *Bioinformatics*, Volume 38, Issue 5. Pages 1467–1469. Disponible: <https://doi.org/10.1093/bioinformatics/btab828>

Volkova, S., Matos, M., Mattanovich, M., & Marín de Mas, I. (2020). Metabolic Modelling as a Framework for Metabolomics Data Integration and Analysis. *Metabolites*, 10(8), 303. Disponible en: <https://doi.org/10.3390/metabo10080303>

Xiao, W., Wang, R.

S., Handy, D. E., & Loscalzo, J. (2018). NAD(H) and NADP(H) Redox Couples and Cellular Energy Metabolism. *Antioxidants & redox signaling*, 28(3), 251–272. Disponible en: <https://doi.org/10.1089/ars.2017.7216>

Xie Y., Chen S. & Xiong X (2021). Metabolic Engineering of Non-carotenoid-Producing Yeast *Yarrowia lipolytica* for the Biosynthesis of Zeaxanthin. *Front. Microbiol.* 12:699235. Disponible en: <https://www.frontiersin.org/articles/10.3389/fmicb.2021.699235/full#B39>

Ya-Hui Wang, Rong-Rong Zhang, Y. Yin et al. (2022). Advances in engineering the production of the natural red pigment lycopene: A systematic review from a biotechnology perspective, *Journal of Advanced Research*. Disponible en: <https://doi.org/10.1016/j.jare.2022.06.010>

Yan, J., Han, B., Gui, X. et al. (2018). Engineering *Yarrowia lipolytica* to Simultaneously Produce Lipase and Single Cell Protein from Agro-industrial Wastes for Feed. *Sci Rep* 8, 758. Disponible en: <https://www.nature.com/articles/s41598-018-19238-9#citeas>

Zhang G., Wang H., Zhang Z., Verstrepen K., Wang Q. & Dai Z. (2022) Metabolic engineering of *Yarrowia lipolytica* for terpenoids production: advances and perspectives, *Critical Reviews in Biotechnology*, 42:4, 618-633. Disponible en: <https://www.tandfonline.com/doi/full/10.1080/07388551.2021.1947183>

Zhang, JL., Bai, QY., Peng, YZ. et al. (2020). High production of triterpenoids in *Yarrowia lipolytica* through manipulation of lipid components. *Biotechnol Biofuels* 13, 133 (2020). Disponible en <https://doi.org/10.1186/s13068-020-01773-1>

Zhao X., Shi F. & Zhan W. (2015). Overexpression of ZWF1 and POS5 improves carotenoid biosynthesis in recombinant *Saccharomyces cerevisiae*. *Letters in Applied Microbiology*. Volume 61, Issue 4 p. 354-360. Disponible en: <https://sfamjournals.onlinelibrary.wiley.com/doi/full/10.1111/lam.12463>

Fabiszewska, A. (2018). *Yarrowia lipolytica*: a beneficial yeast in biotechnology as a rare opportunistic fungal pathogen: a minireview. *World journal of microbiology & biotechnology*, 35(1), 10. doi:10.1007/s11274-018-2583-8. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6302869/>

ANEXOS

Anexo 1. Script desarrollado en un entorno *Jupyter Notebook* de Google Colab

Para la optimización de la producción de licopeno en el modelo metabólico *iYali4* de *Yarrowia lipolytica*. Archivo: *Tesis M Quing.ipynb*.

Anexo 2. Resultados obtenidos en la optimización de la producción de licopeno en *Yarrowia lipolytica*.

Usando *FBA* en el modelo *iYali4* para diferentes factores.

Factores	Descripción	Tasa de crecimiento (biomasa) (1/h)	Tasa de producción de licopeno (mmol/gdcw/h)	Rendimiento de licopeno (mol/mol glucose)
Aceptor/aceptores reducidos	NAD/NADH	0.02385	0.0397	0.0613
	NADP/NADPH	0.02385	0.0397	0.0613
	FAD/FADH2	0.0224	0.0406	0.0625
	Combinando todos los aceptores/aceptores reducidos)	0.0939	-1.8777	-2.893e-15
Sobreexpresión de genes	Gen: HMG1	0.000628	0.0562	0.0866
	Gen: <i>MVD1</i>	0.000628	0.0562	0.0866
	Gen: <i>ERG8</i>	0.000628	0.0562	0.0866
	Gen: <i>CrtB</i>	0.00939	0.0500	0.07704
	Gen: <i>CrtI</i>	0.00939	0.0500	0.07704
	Genes: HMG1 y CrtB	0.000628	0.0562	0.0866
Eliminación de genes	Gene: <i>POX1</i>	0.0224	0.0406	0.0625
	Genes: <i>POX1</i> y <i>GUT2</i>	0.0224	0.0406	0.0625
<i>Elaborado por: Milton Quinga.</i>				

