



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

**ANÁLISIS DE FLUJOS METABÓLICOS PARA
ELABORAR ESTRATEGIAS DE INGENIERÍA
METABÓLICA EN UNA CEPA DE
SACCHAROMYCES CEREVISIAE PRODUCTORA
DE MONOTERPENOS.**

FELIPE IGOR PÉREZ ASCENCIO

Tesis para optar al grado de
Magíster en Cs. de la Ingeniería

Profesor Supervisor:
EDUARDO AGOSIN TRUMPER

Santiago de Chile, (Septiembre, 2012)

© 2012, Felipe Igor Pérez Ascencio



PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
ESCUELA DE INGENIERIA

**ANÁLISIS DE FLUJOS METABÓLICOS PARA
ELABORAR ESTRATEGIAS DE INGENIERÍA
METABÓLICA EN UNA CEPA DE
SACCHAROMYCES CEREVISIAE PRODUCTORA
DE MONOTERPENOS.**

FELIPE IGOR PÉREZ ASCENCIO

Tesis presentada a la Comisión integrada por los profesores:

EDUARDO AGOSIN TRUMER

RICARDO PÉREZ CORREA

RAÚL CONEJEROS

RODRIGO CIENFUEGOS

Para completar las exigencias del grado de
Magíster en Cs. de la Ingeniería

Santiago de Chile, (Septiembre, 2012)

(A mis Padres, mi hermano y a los
Feliplonianos, los cuales siempre me
brindaron un apoyo incondicional)

AGRADECIMIENTOS

Agradezco a Jorge y Myriam por su paciencia, comprensión y apoyo incondicional. A Nicolás por no cuestionarme y simplemente apoyarme con las decisiones tomadas. Además, estoy en una deuda infinita con el profesor Ricardo Pérez porque siempre tuvo las palabras precisas en el momento indicado.

Asimismo quiero agradecer a todos quienes de alguna manera me acompañaron en mis tardes de laboratorio, y noches de estudio en el laboratorio, como la Javi, el Marcelo, la Isa, etc. Mención especial a Carlos Galarce cuya voluntad y disposición a enseñar a cualquier persona (las veces que fuese necesario) siempre le dio un espíritu especial al lab. A María Carolina, por hacer propios los desafíos de sus compañeros y por compartir su vasta lista de contactos. A Felipe Aceituno y Felipe Vargas con quienes compartí varias tardes de discusión teórica y alcances del modelo metabólico.

Por último, agradecer a Dios por ayudarme a ponerme de pie.

INDICE GENERAL

Pág.

DEDICATORIA.....	ii
AGRADECIMIENTOS	iii
INDICE DE TABLAS	vi
INDICE DE FIGURAS	vii
RESUMEN.....	viii
ABSTRACT	x
1. Introducción	1
1.1 Terpenos	1
1.1.1 Rutas de síntesis.....	1
1.1.2 Aplicaciones y obtención.....	7
1.2 Ingeniería Metabólica para la Producción de Terpenos	8
1.2.1 Ingeniería metabólica en plantas.....	11
1.2.2 Ingeniería metabólica en microorganismos	13
1.3 FBA y Modelos Metabólicos	18
1.3.1 Representación matemática del metabolismo celular	19
1.3.2 Formulación de un problema FBA tradicional	21
1.3.3 Modificaciones y Aplicaciones del FBA	23
1.3.4 Modelos genómicos de <i>S. cerevisiae</i>	30
2. Métodos	34
2.1. OptGene	34
2.2. FSEOF (flux scanning based on enforced objective flux)	36
2.3. Nodos Principales (NP).....	39
2.4. Análisis Integrado	40

3.	Resultados	44
3.1.	Máximo Teórico	44
3.2.	Análisis Integrado	49
3.2.1.	Glucosa como Única Fuente de Carbono	50
3.2.2.	Uptake de Glucosa y Aminoácidos.....	55
3.3.	Modelación de la Reducción de Flujos en la Vía de las Purinas.....	58
4.	Discusión	63
	BIBLIOGRAFIA.....	69
	A N E X O S.....	80
	Anexo A: Uso del Toolbox COBRA	81
	Anexo B: Máximos Teóricos	83
	Anexo C: Resultados de cada algoritmo de búsqueda por separado	85
	Anexo D: Reacciones de intercambio	86
	Anexo E: Mapas metabólicos según la base de datos Kegg	97
	Anexo F: Ciclo donde participa el glutamato.....	119

INDICE DE TABLAS

	Pág.
Tabla 1: Comparación de los modelos a escala genómica de <i>Saccharomyces cerevisiae</i> . .	32
Tabla 2: Resumen de características de los tres algoritmos de búsqueda utilizados.....	41
Tabla 3: Producción de biomasa y síntesis de limoneno para los modelos.....	46
Tabla 4: Reacciones seleccionadas por el análisis integrado	52
Tabla 5: Reacciones seleccionadas por el análisis integrado (con suplemento de aminoácidos)	56
Tabla 6: Comparación de resultados obtenidos por los distintos modelos en cada una de las optimizaciones.....	61
Tabla 7: Reacciones seleccionadas por el algoritmo de búsqueda FSEOF clasificadas según cada modelo de acuerdo a la ruta a la cual pertenecen.....	85

INDICE DE FIGURAS

	Pág.
Figura 1: Rutas biosintéticas de isoprenoides.	3
Figura 2: Representación matemática del metabolismo celular.....	20
Figura 3: Estructura general del algoritmo OptGene	35
Figura 4: FSEOF y análisis de variabilidad.	38
Figura 5: Segunda etapa del algoritmo NP.....	40
Figura 6: Clasificación y número de reacciones para cada modelo	43
Figura 7: Ruta alternativa para la producción de monoterpenos.....	47
Figura 8: Análisis FSEOF y cálculo de la productividad BPCY	48
Figura 9: Ejemplo de ciclos.....	54
Figura 10: Rutas metabólicas asociadas con la biosíntesis de purinas.....	59

RESUMEN

Los terpenos son una vasta y diversa clase de compuestos orgánicos con más de 50.000 ejemplos. Estos compuestos pueden ser clasificados según el número de átomos de carbono que presenten, siendo los más importantes los monoterpenos (C_5), sesquiterpenos (C_{15}), diterpenos (C_{20}), y los triterpenos (C_{30}). Los terpenos pueden ejercer una amplia variedad de funciones tal que han sido utilizados como anticancerígenos, antivirales, antibióticos, suplementos alimenticios, nutraceuticos, colorantes, cosméticos, o fragancias. Sin embargo, debido a sus bajas concentraciones en el reino vegetal y su compleja purificación se ha justificado el desarrollo de procesos alternativos de producción a gran escala. Entre las alternativas, el enfoque de la ingeniería metabólica para la producción de terpenos en *Saccharomyces cerevisiae* ha sido considerada como una metodología adecuada para la biosíntesis, extracción y purificación de estos compuestos. Entre las ventajas de dicho microorganismo como plataforma de producción se encuentran, la disponibilidad de su secuencia genómica, el conocimiento de su fisiología y de sus técnicas de manipulación génica, además de la aceptación general como un organismo inocuo. Para obtener altos rendimientos con este microorganismo no sólo se requiere la sobreexpresión de los genes involucrados en la ruta de síntesis del terpeno de interés, sino que involucra también la optimización de la red metabólica de la levadura. Para esto se han utilizado diversas estrategias como delección y sobreexpresión de genes aleatorios, la sobreexpresión de rutas biosintéticas paralelas, o la búsqueda de blancos génicos en modelos metabólicos a escala genómica.

La gran mayoría de los estudios que consideran un modelo metabólico en particular, utilizan solo un algoritmo de búsqueda determinado. Sin embargo, ninguno ha combinado las distintas construcciones ni las distintas metodologías de búsqueda para la identificación de blancos génicos lo que tiene la ventaja de aportar en la robustez del análisis debido a enfrentar un mismo objetivo bajo diferentes enfoques. En este trabajo los modelos a escala genómica iND750, iMM904 y Yeast 4.0, fueron utilizados para sugerir un nuevo blanco génico para aumentar la biosíntesis de monoterpenos en *S. cerevisiae*. Para ello, a cada

modelo se le adicionó la reacción de síntesis de limoneno, una molécula utilizada como representante general de los monoterpenos; y luego, cada modelo fue analizado con tres algoritmos de búsqueda para seleccionar las reacciones más importantes que compiten por los precursores de la vía de biosíntesis. Tales algoritmos fueron: el OptGene, que identifica a través de programación evolutiva una combinación de blancos génicos; el FSEOF que clasifica todas las reacciones según su comportamiento ante la redistribución de flujos; y un algoritmo de selección de los nodos principales de la red metabólica. De esta forma, el análisis integrado de sus resultados mostró que varias reacciones vinculadas con la ruta de síntesis de purinas disminuyen sus flujos a medida que aumenta la producción de monoterpenos, por lo que se propone alterar esta vía y se discute la mejor forma de lograrlo.

Palabras Claves: Ingeniería metabólica, iND750, iMM904, Yeast 4.0, OptGene, FSEOF.

ABSTRACT

Terpenes are a large and diverse class of organic compounds with more than 50,000 examples. These compounds can be classified according to the number of carbon atoms that they present, where the most important of them are the monoterpenes (C₅), sesquiterpenes (C₁₅), diterpenes (C₂₀), and triterpenes (C₃₀). Terpenes may have a wide variety of functions such that they have been used as anticancerous, antivirals, antibiotics, food supplements, nutraceuticals, dyes, cosmetics or fragrances. However, due to their low concentrations in the plant kingdom and its complex purification it has been justified the need to develop alternative processes for large scale production. Amongst the alternatives, the approach of metabolic engineering to produce terpenes in *Saccharomyces cerevisiae* has been considered as an appropriate methodology for the biosynthesis, extraction and purification of these compounds. Some of the advantages of this microorganism as a production platform are, the availability of its genomic sequence, the knowledge of their physiology and their genetic manipulation techniques, in addition to general acceptance as an innocuous organism. To obtain high yields with this micro organism it's not only required overexpression of the genes involved in the synthesis route of the terpene of interest, but it also involves the optimization of the metabolic network of yeast. For this, various strategies have been used as deletion and overexpression of random genes, overexpression of parallel biosynthetic pathways, or search for gene targets in metabolic models at a genomic scale.

The majority of the studies that consider a specific metabolic model, use only one determined search algorithm. However, none has combined the different constructions nor the different search methodologies for identification of genetic targets, which has the advantage of providing the robustness of the analysis, due to confronting a same target under different approaches. In this work the models at a genomic scale iND750, iMM904 and Yeast 4.0, were used to suggest a new genetic target to increase the biosynthesis of terpenes in *S. cerevisiae*. For that each model was explored with three search algorithms to select the most important reactions that compete for precursors of the terpene

biosynthesis path. Such algorithms were: the OptGene, which identifies a combination of genetic targets through evolutionary programming; the FSEOF which classifies all the reactions according to their behavior in the redistribution of flows; and a selection algorithm of the main nodes of the metabolic network. This way, the integrated analysis of their results showed that many reactions whose flows decrease at the same time that the monoterpenes production increase, are linked with the purines synthesis path, so that this work proposes the reduction of the flows in this path and the discussion of the best way to achieve it.

Keywords: Metabolic engineering, iND750, iMM904, Yeast 4.0, OptGene, FSEOF.

1. INTRODUCCIÓN

1.1 Terpenos

Los metabolitos secundarios producidos en plantas y microorganismos son de gran interés debido a su variedad estructural y diversidad funcional. Entre ellos, los terpenos representan una de las clases más grande de metabolitos secundarios (Breitmaier 2006), identificándose más de 50.000 compuestos (Withers & Keasling 2007) con una gran diversidad de estructuras químicas (lineales, cíclicas, policíclicas) y con amplias modificaciones, ya sean hidroxílicas, carbonilos, cetonas, aldehídos y/o con grupos peróxidos. Estos compuestos, también llamados isoprenoides, son encontrados en plantas como constituyentes de sus aceites esenciales (McCaskill & Croteau 1997) y cumplen roles como hormonas (giberelinas), pigmentos fotosintéticos (fitol, carotenoides), o transportadores de electrones (ubiquinonas), y en algunos casos pueden actuar como atractores de polinizadores o como repelentes de cierto tipo de insectos (Langenheim 1994; Kessler & Baldwin 2001; Pichersky & Gershenzon 2002; Aharoni *et al.* 2003).

1.1.1 Rutas de síntesis

Los terpenos, junto con los esteroides, son derivados de la unidad de 5 átomos de carbono isopreno (2-metil-1,3-butadieno), y la combinación sucesiva de estas unidades forma los distintos tipos de isoprenos. Es así que la clasificación de ellos

se basa en la cantidad de átomos de carbono que contienen, siendo los más importantes los monoterpenos, sesquiterpenos, diterpenos, y los triterpenos con 10, 15, 20, y 30 átomos de carbono respectivamente (Kirby & Keasling 2009). Los precursores universales de todos ellos son el isopentenil difosfato (IPP) y su isómero dimetilalil difosfato (DMAPP), los cuales pueden ser sintetizados por dos rutas: la vía del mevalonato (MVA) o la del deoxi-xilulosa-fosfato (DXP) (**Figura 1**). La primera es encontrada en hongos, animales, y en el citosol de las plantas, y la segunda está presente en bacterias y en los plastídeos de las plantas (Kirby & Keasling 2009).

a) La ruta del MVA

La ruta del mevalonato utiliza siete reacciones enzimáticas para convertir el precursor acetil-CoA en IPP y DMAPP. Esta vía citosólica inicia con la condensación de tres moléculas de **acetil-CoA** para formar hidroximetilglutaril-CoA (**HMG-CoA**), proceso catalizado por dos enzimas, la tiolasa y la HMG-CoA sintasa. Luego, este producto es reducido con dos moléculas de NADPH, lo que es catalizado por la HMG-CoA reductasa, formando **ácido mevalónico** el cual posteriormente sufre dos fosforilaciones dependientes de ATP y una descarboxilación por la acción consecutiva de la mevalonato kinasa, la fosfomevalonato kinasa, y la mevalonato difosfato descarboxilasa. El producto de esto es el precursor general de 5 átomos de carbono **IPP**, que puede formar **DMAPP** con la acción de la IPP isomerasa.

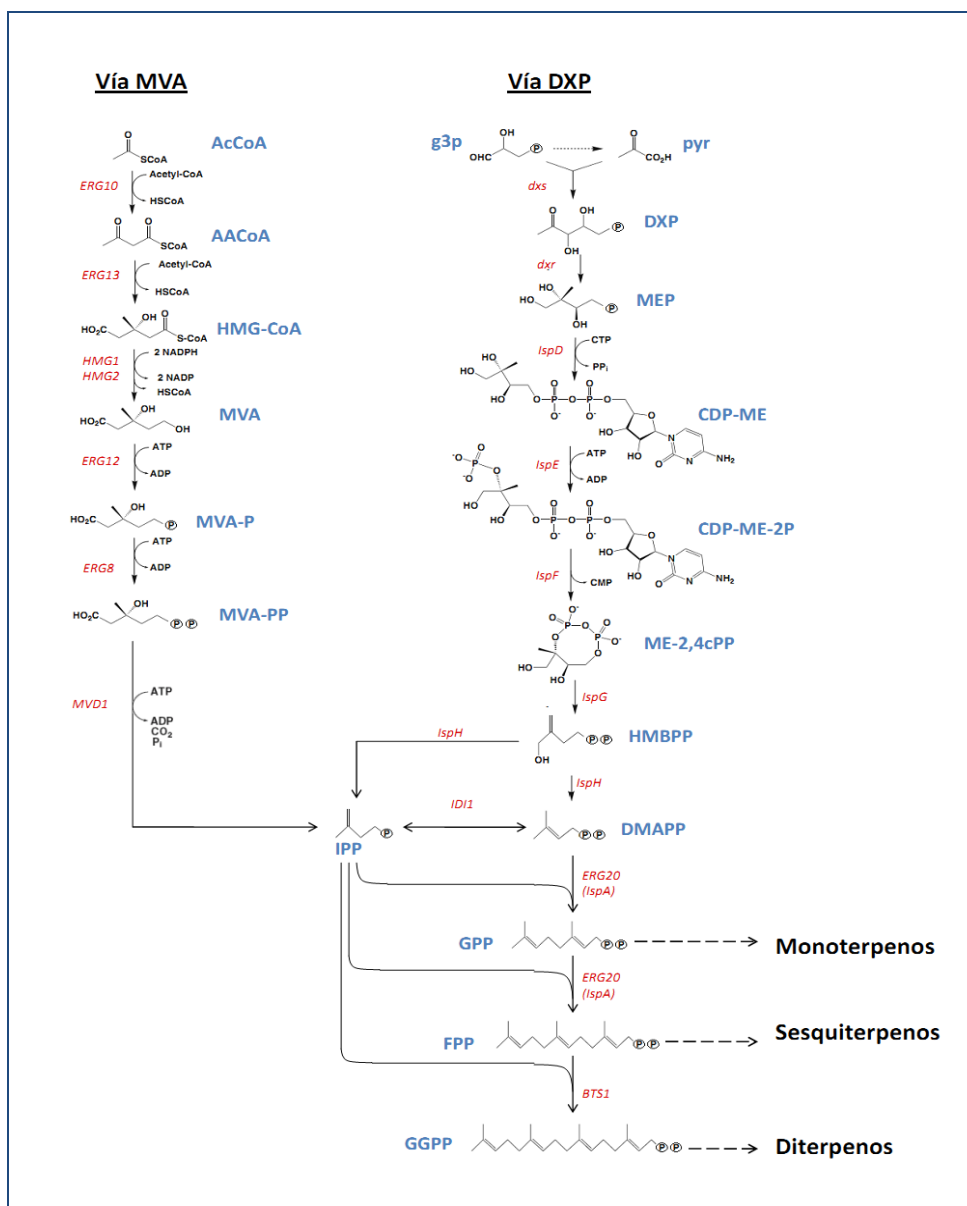


Figura 1: Rutas biosintéticas de isoprenoides. En azul se muestra la abreviación de cada metabolito, y en rojo la abreviación del gen que codifica para la enzima respectiva.

AcCoA: acetil CoA; **AACoA:** acetoacetil CoA; **HMG-CoA:** hidroximetilglutaril-CoA; **MVA:** mevalonato; **MVA-P:** mevalonato-5-fosfato; **MVA-PP:** mevalonato-5-difosfato; **g3p:** gliceraldehido-3-fosfato; **pyr:** piruvato; **DXP:** deoxi-xilulosa-fosfato; **MEP:** 2C-metil-D-eritrol-4-fosfato; **CDP-ME:** 4-difosfocitidil-2C-metil-D-eritrol; **CDP-ME-2P:** 2-fosfo-4-difosfocitidil-2C-metil-D-eritrol; **ME-2,4cPP:** 2C-metil-D-eritrol-2,4-ciclodifosfato; **HMBPP:** 1-hidroxi-2-metil-2-(E)-butenil 4-difosfato; **IPP:** isopentenil difosfato; **DMAPP:** dimetilalil difosfato; **GPP:** geranyl difosfato; **FPP:** farnesil difosfato; **GGPP:** geranylgeranyl difosfato.

La enzima mejor estudiada es la HMG-CoA reductasa (HMGR), la cual cataliza el paso limitante en la vía. La levadura tiene dos isoenzimas codificadas por el gen *HMG1*, que proporciona el 80% de la actividad total, y por el gen *HMG2* (Basson *et al.* 1986). Ambas enzimas son reguladas diferencialmente, la HMGR codificada por la *HMG1* (Hmg1p) es dominante en la fase exponencial del crecimiento aeróbico, mientras que la codificada por la *HMG2* (Hmg2p) es más importante en el crecimiento anaeróbico o en la fase estacionaria. Estas enzimas tienen además un dominio transmembrana de anclaje al retículo endoplasmático que regula la actividad enzimática, y por lo tanto, al sobreexpresar sólo el dominio catalítico se logra aumentar el flujo por la vía (Donald *et al.* 1997).

b) La ruta del DXP

La primera reacción de esta vía condensa al **piruvato** y al **gliceraldehido-3-fosfato** para formar deoxi-xilulosa-fosfato (**DXP**), lo que es catalizado por la DXP sintasa (DXS). Luego, la enzima DXP reductoisomerasa produce 2C-metil-D-eritrol-4-fosfato (**MEP**) con el consumo de una molécula de NADPH. La tercera reacción introduce el nucleótido CTP para producir 4-difosfocitidil-2C-metil-D-eritrol (**CDP-ME**), el cual es subsecuentemente fosforilado gastando una molécula de ATP, formando 2-fosfo-4-difosfocitidil-2C-metil-D-eritrol (**CDP-ME-2P**). Tales reacciones son catalizadas por la MEP citidiltransferasa y por la CDP-ME kinasa. La quinta reacción requiere de los cationes divalentes Zn^{+2} y Mg^{+2} (Kemp *et al.* 2002; Steinbacher *et al.* 2002) y consiste en la ciclación del CDP-ME-2P, lo que libera al nucleótido CMP y produce 2C-metil-D-eritrol-2,4-ciclodifosfato (**ME-2,4cPP**). A continuación, este metabolito es reducido generando a 1-hidroxi-2-metil-2-(E)-butenil 4-diphosphate (**HMBPP**), el que vuelve a ser reducido formando **IPP** y su isómero **DMAPP** en una proporción de 6:1, la cual es

posteriormente ajustada a 7:3 con la isopentenil difosfato isomerasa (IDI) que cataliza la interconversión entre ambos compuestos (Rohdich *et al.* 2003).

La primera reacción corresponde a la etapa limitante de la vía, y la sobreexpresión del gen *Dxs* resulta en el incremento de la síntesis de isoprenosidos, y por otro lado, en la reducción de la tasa de crecimiento (Kim & Keasling 2001). Con respecto al gen *Idi*, éste también mejora la producción de terpenos en *E. coli* y es el único que no es esencial, ya que sólo cataliza la interconversión entre dos metabolitos.

c) La ruta del MVA y del DXP en plantas

Aunque la biosíntesis de los terpenos en las plantas, animales, y microorganismos involucra familias de enzimas similares, existen diferencias importantes entre ellos. La variedad de terpenos producidos por las plantas es mucho mayor, al igual que la complejidad en la organización y regulación de su biosíntesis.

En primer lugar, la biosíntesis puede tener regulación tejido específica (hojas, raíces, frutas, semillas, flores, madera o corteza) (Daviet & Schalk 2010). Por ejemplo, los tricomas glandulares y las cavidades secretorias de las hojas, además de las glándulas epidermales de los pétalos pueden generar, almacenar, o secretar aceites esenciales terpénicos que son importantes para favorecer la polinización por insectos. Otros ejemplos son los ductos de resinas de las coníferas que producen y acumulan compuestos defensivos (trementina y colofonia); o los laticíferos que producen ciertos triterpenos y politerpenos como el caucho (Langenheim 1994).

En segundo lugar, existe una regulación subcelular donde los sesquiterpenos (C_{15}), triterpenos (C_{30}), y politerpenos son producidos en el citosol, lugar donde se

expresa la vía del mevalonato. Por otro lado, los monoterpenos (C_{10}), diterpenos (C_{20}), y los tetraterpenos (C_{40}) se sintetizan en los plastideos, donde se expresa la vía DXP. Un hecho importante es que el IPP puede ser suministrado por el plastideo al citosol, y *vice versa*, lo que hace muy difícil el estudio de las vías por separado (Croteau *et al.* 2000).

d) Rutas biosintéticas río abajo del IPP

La condensación consecutiva del IPP en el DMAPP, catalizada por preniltransferasas, forma los precursores de los distintos tipos de isoprenos (**Figura 1**). Tales precursores corresponden al geranil difosfato (GPP) de 10 átomos de carbono, farnesil difosfato (FPP) de 15, y al geranilgeranil difosfato (GGPP) de 20 átomos de carbono. Las reacciones siguientes ciclan de distintas formas a tales precursores, y son catalizadas por las terpeno sintasas (terpenos ciclasas). Finalmente, son las citocromo P450 monooxigenasas y oxidoreductasas las que están involucradas en las modificaciones subsecuentes del esqueleto básico de los terpenos, dando miles de compuestos distintos (Kirby & Keasling 2009).

Las terpenos sintasas forman un amplio rango de compuestos estructuralmente diversos, de los que unos pocos se producen preferencialmente. Los rasgos estructurales que determinan la especificidad de sustrato y el diverso perfil de productos son desconocidos, y los análisis de comparación de secuencia por sí solos no entregan mucha información para entender tal diversidad. Dichos análisis han mostrado que algunas terpenos sintasas que producen distintos compuestos en un organismo vegetal, son más similares que las mismas terpenos sintasas de especies distintas (Martin *et al.* 2004). Por lo tanto, no se puede predecir con precisión la función bioquímica de una terpeno sintasa basado únicamente en la similaridad de secuencia, incluso para enzimas.

1.1.2 Aplicaciones y obtención

Muchos isoprenos han encontrado interesantes aplicaciones en medicina y agricultura. Por ejemplo, se destaca el diterpeno **Paclitaxel** como un potente agente antimitótico con gran actividad contra el cáncer de mamas y de ovarios (Dieras 1998). Este compuesto es obtenido de la corteza de la conífera *Taxus brevifolia* y se comercializa bajo el nombre comercial de Taxol. Otros anticancerígenos son los monoterpenos **limoneno** y **perilil alcohol**, conocidos por inhibir el desarrollo de algunos carcinomas como el de hígado, pulmón, colon, próstata, y páncreas entre otros (Crowell & Gould 1994; Reddy *et al.* 1997; Shi & Gould 2002). La **artemisina** es un sesquiterpeno extraído de la hierba *Artemisia annua* utilizado para combatir la malaria (Karunajeewa *et al.* 2007). Por otro lado, varios monoterpenos tienen actividades insecticida y acaricida, por ejemplo, el **carvacrol** es efectivo contra algunas plagas de productos agrícolas almacenados (Ahn *et al.* 1998); el α - y β -**pineno** en conjunto son una potente mezcla contra larvas de moscas en champiñones (Choi *et al.* 2006); y el **tímol** tiene actividad insecticida contra gusanos de raíces, arañas y moscas (Lee *et al.* 1997).

La extracción de los terpenos desde tejidos vegetales ha sido la opción tradicional para la producción a gran escala. Sin embargo, en muchos casos esto no es factible económicamente debido a los bajos rendimientos, además de la variabilidad en la composición y concentración de isoprenos en los tejidos vegetales (de Carvalho & da Fonseca 2006). Por ejemplo, para el paclitaxel el porcentaje de recuperación en

peso seco es de 0.01-0.06% p/p (Zu *et al.* 2009), llegando en algunos casos a 0.128% p/p (Kawamura *et al.* 1999). Esto tiene un gran impacto medioambiental, ya que para extraer 1 kg de paclitaxel de *Taxus brevifolia*, se necesitan cortar cerca de 3.000 árboles (Bui-Khac & Dupuis 2002), lo que alcanza para el tratamiento de cerca de 500 pacientes con cáncer. En relación a la artemisina, la cantidad total encontrada en diversas variedades de *A. annua* varía entre 0.01 y 1.4% p/p en peso seco (Lapkin *et al.* 2006).

Por otro lado, la síntesis química tiene poco valor práctico para desarrollar complejos productos naturales debido al bajo rendimiento y a los numerosos pasos involucrados. Usualmente, una ruta sintética para moléculas complejas es impráctica si el proceso total requiere más de 10 pasos, debido a que en cada uno de ellos el rendimiento total disminuye mientras que incrementan los desechos y los recursos necesarios (Chemler & Koffas 2008). En el caso del Taxol se requieren entre 35 a 51 pasos, con un rendimiento máximo de 0.4% (Ajikumar *et al.* 2010).

1.2 Ingeniería Metabólica para la Producción de Terpenos

El enfoque de la ingeniería metabólica presenta una alternativa, en donde se trata de entender y modular el metabolismo celular, con un énfasis en el estado global de la célula (Tyo *et al.* 2007; Alper & Stephanopoulos 2009). Su objetivo es mejorar la conversión de la materia prima en un compuesto de interés al modificar

la dirección de los flujos metabólicos (Rigoutsos & Stephanopoulos 2007). Este enfoque se constituye por el diseño inteligente y sistemático de las redes intracelulares, y se logra al optimizar, introducir (Stephanopoulos & Sinskey 1993; Stephanopoulos *et al.* 2004; Raab *et al.* 2005) y/o eliminar (Causey *et al.* 2003; Causey *et al.* 2004; Alper *et al.* 2005; Chang & Keasling 2006; Jarboe *et al.* 2007; Leonard *et al.* 2008; Trinh *et al.* 2008) una o varias reacciones químicas.

Al trabajar bajo este enfoque se pueden evaluar cuatro estrategias importantes: la sobreexpresión o silenciamiento de genes estructurales en la ruta metabólica, regulación transcripcional, control de flujos metabólicos, y sobreexpresión de transportadores (Wang *et al.* 2011).

En primer lugar, hay que analizar la estructura de la red metabólica y decidir por la introducción de uno o varios genes. Por ejemplo, para incrementar la biosíntesis y acumulación en plantas, la sobreexpresión de enzimas que catalizan el paso limitante, sin los dominios regulatorios, es el primer paso a realizar. Sin embargo, a menudo el efecto de la introducción de un único gen es absorbido por el sistema en un intento de restaurar la homeostasis (Capell & Christou 2004). Es por esto que apuntar a múltiples pasos puede ayudar a controlar el flujo metabólico de forma más predecible. Para esto se requiere de la sobreexpresión de varias enzimas en una misma vía y/o la supresión de otras en rutas competitivas o paralelas a la de interés. En microorganismos ha sido común la introducción de

rutas biosintéticas foráneas, las cuales carecen de todo tipo de regulación al pertenecer a un organismo distinto (Maury *et al.* 2008).

En segundo lugar, con el mismo desafío de regular varios pasos en la red metabólica, es que la maquinaria de transcripción debe ser examinada. Los factores de transcripción son una conveniente forma de regular toda una ruta biosintética ya que están involucrados en la activación y represión de genes que participan en la misma vía de interés y/o en otras (Ro *et al.* 2006; Petersen 2007; Grotewold 2008).

Una tercera estrategia es la identificación de problemas con los flujos metabólicos a través del análisis de datos genómicos, transcriptómicos, o metabolómicos. Los trabajos con modelos a escala genómica han mostrado ser muy útiles para identificar cuellos de botella y vías que compiten por sustratos comunes. Basado en tales análisis, la sobreexpresión o la supresión de algunos genes ha permitido variar la distribución de flujos, incrementando los rendimientos de los productos de interés (Patil *et al.* 2005; Bro *et al.* 2006; Asadollahi *et al.* 2009).

Finalmente, el almacenamiento o exportación de los productos sintetizados deben ser controlados en plantas y microorganismos. Los transportadores desempeñan roles importantes en la ingeniería metabólica (Yazaki 2005) ya que evitan que la producción de los compuestos de interés tenga efectos tóxicos. Además facilitan el

equilibrio metabólico a favor de la biosíntesis ya que remueven los productos del espacio donde se encuentran las enzimas (Wang *et al.* 2011).

1.2.1 Ingeniería metabólica en plantas

La ingeniería metabólica vegetal de terpenos ha permitido mejorar numerosos rasgos propios de las plantas además de generar una plataforma de producción de tales compuestos. Es así que varios trabajos se han realizado con el fin de mejorar directa o indirectamente la defensa vegetal, el nivel nutricional de alimentos, o la calidad aromática en frutas y flores (Aharoni *et al.* 2005).

Por ejemplo, en menta se logró aumentar en cerca del 50% la producción de mentol, compuesto principal del aceite esencial de dicha planta y que se utiliza ampliamente en la industria (Mahmoud & Croteau 2001). Para ello la menta se transformó con dos genes: el que codifica para la deoxyxilulosa fosfato reductasa (segunda reacción de la vía del DXP); y con un RNA antisentido para la mentofurano sintasa, enzima que cataliza una reacción lateral en la síntesis de mentol. En un trabajo desarrollado en petunias y en tomates se logró acumular linalool, un monoterpeno con propiedades defensivas contra plagas de insectos (Lewinsohn *et al.* 2001; Lucker *et al.* 2001). Otro ejemplo es el conocido “arroz dorado”, un cereal al cual se le adicionó la ruta biosintética completa del β -caroteno (provitamina A) para aumentar su valor nutricional (Ye *et al.* 2000). Además, en la planta de tabaco se han expresado tres monoterpenos sintasas de

limón capaces de sintetizar β -pineno, limoneno, y γ -terpineno (Luckner *et al.* 2004).

En particular, cuando el objetivo es desarrollar una plataforma para la producción de compuestos volátiles, hay que tener en cuenta que la acumulación de terpenos en muchos casos está asociada con la reducción del crecimiento de la planta y con clorosis (reducción de clorofila). Esto indica un posible efecto tóxico y/o la reducción del flujo de carbono a metabolitos esenciales (Daviet & Schalk 2010).

Por otro lado, algunas plantas no acumulan productos volátiles comúnmente, por lo que puede haber una gran proporción de terpenos sintetizados y perdidos a la atmósfera (Wu *et al.* 2006). Sin embargo, en otras especies existen ciertas estructuras llamadas tricomas glandulares y que justamente se encargan de la acumulación de tales compuestos. Varios trabajos han caracterizado promotores de tricomas, lo que es útil para la expresión específica, con el fin de tener limitados efectos tóxicos junto con un aumento en la acumulación de terpenos (Daviet & Schalk 2010). Otra ventaja de utilizar estos promotores es que la posterior extracción de terpenos se realizará a partir de las hojas, el tejido más fácil para la recuperación de aceites esenciales.

1.2.2 Ingeniería metabólica en microorganismos

La biosíntesis de compuestos naturales en microorganismos ha surgido como una alternativa a su síntesis química o a su extracción desde plantas, ya sean modificadas genéticamente o no. Además, la ingeniería microbiana ofrece muchas ventajas sobre los sistemas vegetales en cuanto a producción y purificación. Las más importantes son: el conocimiento genético y metabólico es típicamente mucho mayor que en el huésped nativo, lo que permite analizar más y mejores estrategias de ingeniería genética; y en segundo lugar, el conocimiento de procesos biotecnológicos y de los protocolos de escalamiento han sido bien estudiados, lo que reduce el tiempo en el desarrollo de una plataforma de producción.

Debido a las limitaciones en el suministro de los precursores universales IPP y DMAPP, diversas estrategias se han desarrollado para aumentar los flujos metabólicos en las rutas que los sintetizan. El primer enfoque empleado es la sobreexpresión de los genes que codifican para las enzimas de cada vía. Por ejemplo en *E. coli*, la sobreexpresión de la *dxs* (**Figura 1**) de *Bacillus subtilis* resultó en el aumento de 2 veces la producción de licopeno en comparación a la cepa control (de ~450 a ~1000 µg/gDW) (Harker & Bramley 1999). En otro trabajo, se incorporaron 4 genes de la vía (*dxs*, *ispD*, *ispF*, *idi*), además del gen que participa en la síntesis de quinonas (*IspB*), todos bajo el promotor fuerte del bacteriófago T5, lo que permitió producir β-caroteno a concentraciones de 6 mg/L

(Yuan *et al.* 2006). Recientemente, la sobreexpresión combinada de los genes *dxs*, *dxr*, y de *idi1* (**Figura 1**) ha mostrado la mejoría en la biosíntesis del diterpeno abietadieno de acuerdo al siguiente orden: $dxr < dxs \leq idi + dxr < dxs + dxr < control < idi + dxs \leq idi < idi + dxs + dxr$ (Morrone *et al.* 2010). Los autores discuten sus resultados sugiriendo que la isomerización del IPP y del DMAPP, mediada por *idi*, es la etapa limitante de la vía, y que además los genes *dxs* y *dxr* ejercen un control en la distribución de flujos hacia otras vías.

Por otro lado, en *S. cerevisiae* la etapa limitante en la vía del mevalonato es la reacción catalizada por la enzima HMG-CoA reductasa, que es codificada por el gen *HMG1*. En una cepa que produce 1.1 mg/gDW de licopeno, la sobreexpresión de tal gen incrementa al doble dicha producción, y si se expresa sólo el dominio catalítico (*tHMG*) se logran niveles de 4.3 mg/gDW. El aumento puede ser de hasta 7 veces si además se elimina el gen *ERG9* que codifica para la escualeno sintasa, parte de una vía competitiva de los terpenos (Shimada *et al.* 1998). En otro estudio se diseñó un alelo condicional para este gen, el cual es regulado por el promotor MET3, de forma tal que la expresión es reprimida en presencia de altas concentraciones extracelulares de metionina (Gardner & Hampton 1999). El uso de este gen condicional tiene la ventaja de que reduce el flujo de carbono a ergosterol, y además puede aumentar el flujo por la vía MVA ya que reduce la inhibición de la HMG-CoA sintasa producida por el mismo ergosterol. Aunque esto puede reducir el crecimiento al desviar el carbono destinado originalmente a

la síntesis de biomasa, el promotor MET3 soluciona el problema ya que es reprimido en la fase de crecimiento y activado en la fase estacionaria (Maury *et al.* 2005).

Otra estrategia utilizada en *S. cerevisiae* ha sido la expresión del gen *upc2-1* el cual codifica para un factor de transcripción que participa en el *uptake* de esteroides. En una cepa que sobreexpresa dicho gen junto con el de la *tHMG* se han logrado obtener concentraciones de 370 µg/L de epi-cedrol, un sesquiterpeno poco volátil utilizado como reportero (Jackson *et al.* 2003). Este factor de transcripción se eligió debido a que la producción de esteroides es un proceso que interfiere con la biosíntesis de terpenos al competir por los mismos precursores; por lo tanto, al proporcionar una forma alternativa en la cual la célula pueda captar tales compuestos desde el medio extracelular, se intenta inhibir su biosíntesis e incrementar el flujo hacia los terpenos. En otro trabajo, en donde se combina la sobreexpresión de los genes *tHMG* y *upc2-1*, además de la regulación del gen *ERG9* con el promotor MET3, se ha logrado obtener 153 mg/L del sesquiterpeno amorfadieno, precursor de la artemisina.

Si bien el enfoque anterior logró mejorar la producción de isoprenoides, tiene limitaciones debido a los mecanismos de control presentes en el huésped. Así, una segunda estrategia ha sido la expresión heteróloga de una ruta paralela, la cual no es regulada por el organismo huésped (Martin *et al.* 2003). En *E. coli* se incorporó

la ruta del mevalonato al diseñar y transformar dicha bacteria con tres plásmidos: (i) uno con el operón *atoB:ERG13:tHMG1*, donde el gen *atoB* corresponde a la acetoacetyl-CoA tiolasa de *E. coli*; (ii) un segundo plásmido que poseía el operón *ERG12:ERG8:MVD1:idi:ispA* con los dos últimos genes de *E. coli*; (iii) y finalmente, un tercer plásmido tenía el gen de la amorfadieno sintasa de *A. annua*. Con esta cepa y con un sistema de cultivo de dos fases se logró determinar una producción de 0.5 g/L de amorfadieno (Newman *et al.* 2006), lo que mejoró a 25 g/L al cambiar los genes *ERG13* y *tHMG* de *S. cerevisiae* por los homólogos de *Staphylococcus aureus* (Tsuruta *et al.* 2009).

Por otro lado, la incorporación de la ruta DXP en *S. cerevisiae* se ha logrado al diseñar dos plásmidos, donde cada uno porta 4 genes heterólogos. Esto logró concentraciones de 1 mg/L de *valence*, un sesquiterpeno cítrico utilizado como reportero (Maury *et al.* 2008). Además se demostró que esta vía por sí sola puede suministrar los precursores necesarios para el crecimiento, ya que fue viable un cultivo al cual se le adicionó lovastatina, un inhibidor fuerte de la ruta MVA.

Un tercer enfoque ha sido la sobreexpresión de una librería de DNA en una cepa de *E. coli* productora de licopeno (Kang *et al.* 2005). La ventaja de este sistema es que debido a que el licopeno le confiere un fuerte color rojo a las colonias, es fácilmente identificable aquellas que tienen un incremento en la producción de dicho compuesto. De esta forma, luego de un extenso *screening* se lograron

identificar tres genes regulatorios que aumentan la producción de licopeno (*crl*, *rpoS*, y *appY*).

Finalmente, una cuarta estrategia ha sido el estudio holístico del metabolismo celular bajo un enfoque matemático. Dicho enfoque, que es detallado extensamente en el apartado siguiente, ha sido útil para encontrar blancos génicos que aumentan la producción de terpenos. Por ejemplo, en *E. coli* se han identificados tres genes a suprimir para mejorar la producción de licopeno (*gdhA*, *aceE*, *fdhF*) (Alper *et al.* 2005). Esta cepa se desarrolló en una que previamente sobreexpresaba los genes *dxs*, *idi*, e *ispDF*, logrando un rendimiento de 6.6 mg/gDW, que corresponde a aproximadamente un 40% de incremento en comparación con la cepa parental. Posteriormente, la optimización del cultivo (Alper *et al.* 2006), y la sobreexpresión de genes identificados por el *screening* de librerías genómicas (Jin & Stephanopoulos 2007) permitieron aumentar los rendimientos a 16 mg/gDW. Por otro lado, en *S. cerevisiae* el modelo desarrollado por Förster *et al.* (2003) identificó al gen *GDH1*, que codifica la glutamato deshidrogenasa, para mejorar la producción de sesquiterpenos. La delección de dicho gen aumenta la disponibilidad de NADPH citosólico para otras enzimas dependientes de este metabolito, incluyendo la HMG-CoA reductasa. Esta modificación aumenta en un 85% la concentración final de cubebol, un sesquiterpeno utilizado como reportero (Asadollahi *et al.* 2009).

1.3 FBA y Modelos Metabólicos

Para establecer a un microorganismo como plataforma de producción de algún compuesto de interés es necesario aumentar los flujos de carbono hacia las vías de síntesis de dicho compuesto. Para ello, además de sobreexpresar los genes que codifican las enzimas que catalizan los pasos limitantes, se puede redirigir el flujo del metabolismo. Sin embargo, por los complejos mecanismos de interacción y control que existen entre las rutas biosintéticas, es difícil predecir las mínimas modificaciones genéticas para explotar la máxima capacidad celular en función del fenotipo deseado (Asadollahi *et al.* 2009). Este desafío puede ser abordado bajo el enfoque del **análisis de balance de flujos** (*Flux Balance Analysis - FBA*) ya que representa al metabolismo celular de una forma holística a través de un modelo matemático (*in silico*), y que es usado para estimar cuantitativamente el comportamiento del metabolismo celular bajo determinadas condiciones (ambientales o genéticas).

El FBA es una metodología basada en restricciones ampliamente utilizada para estudiar redes metabólicas, en especial las reconstrucciones a escala genómica. Este análisis permite calcular los flujos metabólicos intracelulares basándose en un balance de masa realizado sobre cada metabolito (nodo) y en el supuesto de estado estacionario. Su ventaja es que no requiere de información cinética o regulatoria debido a que se centra en la topología de la red metabólica.

1.3.1 Representación matemática del metabolismo celular

El primer paso en el FBA es la descripción matemática de las reacciones metabólicas para lo cual se requiere de un modelo metabólico que represente las reacciones más importantes del sistema estudiado. Para la construcción de dicho modelo se debe establecer un balance de masa sobre cada metabolito (**Figura 2C**), lo que representa la primera restricción del modelo. Dichos balances pueden ser escritos de forma matricial como: $\frac{dC}{dt} = S \cdot v$; donde C es el vector de concentraciones de cada metabolito al interior de la célula, v es el vector de tasas de reacción, y S es la matriz estequiométrica de tamaño $m \times n$, con m metabolitos y n reacciones. Una reacción muy importante en esta matriz es la síntesis de biomasa, la cual es matemáticamente representada al adicionar una reacción que consuma diversos precursores metabólicos como proteínas, RNA, lípidos, ATP, *etc.*

A continuación se incorpora el supuesto de estado estacionario ($0 = S \cdot v$), implicando que la concentración de cada metabolito al interior de la célula permanece constante. Este supuesto es razonable debido a que el metabolismo opera a escalas mucho más rápidas que la dinámica extracelular (Terzer *et al.* 2009).

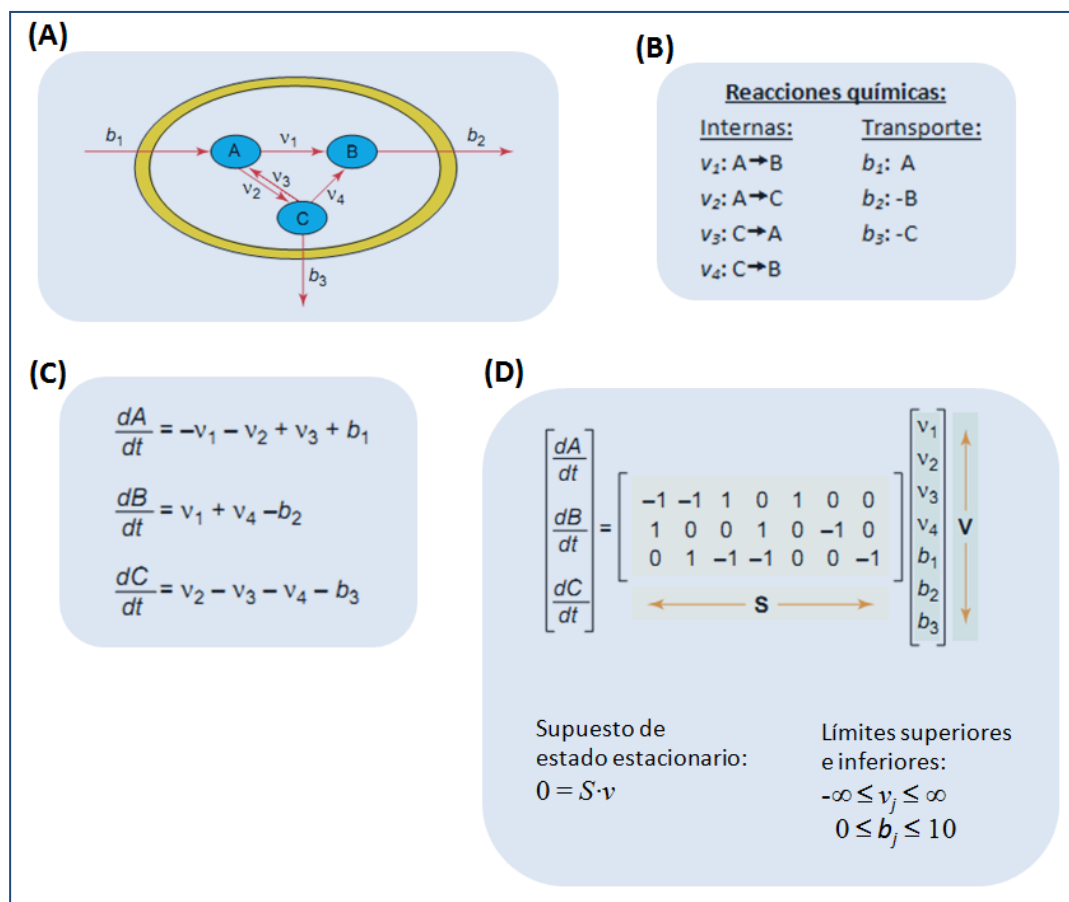


Figura 2: representación matemática del metabolismo celular. (A) una pequeña red de reacciones que presenta tres metabolitos, cuatro reacciones (v_i), y tres transportes (b_j). (B) Resumen de las reacciones. (C) Balance de masa dinámico para cada metabolito. (D) La matriz estequiométrica S es construida a partir de los balances de masa. Se supone estado estacionario, y se establecen los límites para cada reacción. Adaptado de (Kauffman et al. 2003; Lee et al. 2006; Terzer et al. 2009).

De esta forma se imponen restricciones de flujo sobre los metabolitos, tal que se asegura que la cantidad total de cualquier compuesto que se produce debe ser igual a la cantidad total que se consume en estado de equilibrio. Un segundo tipo de restricción impuesta en el sistema corresponde a los límites superiores e inferiores, es decir, un rango de flujo permitido para cada reacción, lo que en su conjunto genera el **espacio solución o espacio nulo**. Estas restricciones diferencian al FBA de otros tipos de modelos que se basan en ecuaciones biofísicas y que requieren de parámetros cinéticos difíciles de obtener (Orth *et al.* 2010). Restricciones adicionales a menudo son consideradas para reducir el espacio nulo, de tal forma que las predicciones del FBA puedan ser mejoradas. Algunas de ellas son por ejemplo, la compartimentalización de metabolitos y de reacciones, la reversibilidad de las mismas, el mínimo y máximo teórico para algunos flujos, la cantidad de ATP requerida para mantención, o la mínima y máxima tasa de crecimiento (Otero & Nielsen 2010).

1.3.2 Formulación de un problema FBA tradicional

El segundo paso en el análisis FBA es la **optimización** de uno o más flujos. En un sistema de ecuaciones subdeterminado, como los que existen en los modelos a escala genómica ($n > m$), existen múltiples soluciones dentro del espacio nulo. Dicho espacio define el rango de todos los fenotipos permisibles, de los cuales sólo un conjunto particular suele ser de interés. Para encontrar tal conjunto hay que establecer una función objetivo que represente al estado biológico buscado,

que posteriormente se debe optimizar. Este enfoque se basa en el supuesto de que las células luchan por un objetivo determinado. Por ejemplo, es razonable pensar que la evolución ha influido para que los microorganismos maximicen el crecimiento o la razón entre la producción de energía y la cantidad de sustrato consumido. Este objetivo se traduce en una función que se expresa matemáticamente como una combinación lineal de los flujos metabólicos (**ec.1. 1**).

$$\text{Max } Z = c^T \cdot v$$

s.a.

$$S \cdot v = 0$$

$$lb_j \leq v_j \leq ub_j$$

ec.1. 1

Donde lb_j y ub_j son los límites de flujo inferiores y superiores para la reacción j , y donde c es un vector de ponderación que indica la contribución de cada reacción a la función objetivo. De esta forma, el sistema de ecuaciones puede ser resuelto usando algoritmos de programación lineal, tal que se encuentre un conjunto de flujos que optimice la función objetivo (Orth *et al.* 2010). Las funciones objetivos pueden ser utilizadas para probar las capacidades de la red metabólica y para representar posibles diseños que la mejoren. Hasta la fecha, la función objetivo más utilizada ha sido la maximización de biomasa, que ha entregado predicciones coincidentes con datos experimentales (Varma & Palsson 1994; Edwards *et al.* 2001). Sin embargo, hay que tener especial cuidado en la definición de dicha ecuación ya que se ha visto un aumento en las discrepancias (del 16% al 80%)

entre los datos experimentales y las predicciones si se utiliza una composición de biomasa incorrecta (Pramanik & Keasling 1997).

1.3.3 Modificaciones y Aplicaciones del FBA

El FBA ha sido utilizado para **predecir genes esenciales o para estudiar la robustez** de una célula bajo una determinada condición ambiental o genética. Para ello se puede observar el cambio en el valor de la función objetivo, en respuesta a variaciones de flujo de una reacción particular (Kauffman *et al.* 2003; Orth *et al.* 2010). Por ejemplo, si el valor de flujo para la síntesis de biomasa se reduce a cero luego de fijar una reacción también a cero, significa que dicha reacción es esencial. Fisiológicamente, el restringir el flujo de una reacción a cero puede interpretarse como una cepa *knockout* del gen que codifica para la enzima que cataliza dicha reacción, o por la adición de un compuesto inhibitorio. Es así que este tipo de análisis ha sido importante para el desarrollo de nuevas drogas contra ciertos microorganismos (Jamshidi & Palsson 2007; Raman & Chandra 2009). Adicionalmente, al restringir las tasas de consumo o de síntesis de los sustratos o productos en el medio externo, el FBA puede ser usado para: cuantificar la tasa de crecimiento celular bajo diferentes condiciones ambientales (Edwards *et al.* 2001); investigar la secreción de subproductos bajo *uptakes* crecientes de oxígeno (Varma *et al.* 1993); evaluar la capacidad de utilización de la fuente de carbono (Orth *et al.* 2010); o para identificar la composición óptima del medio de cultivo (Song *et al.* 2008).

Además, **otras funciones objetivos** se han utilizado aparte de la maximización de biomasa. Entre ellas se incluyen: la maximización de la producción de algún metabolito para determinar las capacidades productivas de una determinada red metabólica (Varma *et al.* 1993); la minimización del *uptake* de nutrientes (Famili *et al.* 2003); o la minimización de la producción de ATP, que se ha establecido para determinar condiciones óptimas de eficiencia energética y ha sido empleada para estudiar las propiedades de la mitocondria (Ramakrishna *et al.* 2001).

Por otro lado, con el fin de explotar al máximo el potencial metabólico e identificar blancos génicos para la ingeniería metabólica, varios **algoritmos basados en el FBA** se han desarrollado para analizar el espacio nulo de los modelos a escala genómica. El FBA se ha considerado sólo como una “primera mirada” para evaluar las estrategias de optimización de la red metabólica (Boghigian *et al.* 2010) ya que se ha sugerido que el comportamiento metabólico sub-óptimo visto luego de la delección de uno o varios genes, es el resultado de un ajuste del organismo como una forma de minimizar los cambios en la distribución de flujos. Para tener en cuenta dicho ajuste, se han desarrollado los algoritmos MOMA (*Minimization Of Metabolic Adjustment*) y ROOM (*Regulatory on/off Minimization*) (Segre *et al.* 2002; Shlomi *et al.* 2005) los cuales utilizan métricas de distancias continuas o discretas para cuantificar la desviación entre la distribución de flujo en un estado original, cepa *wild-type* (WT), y en un estado perturbado, *knockout*. El **MOMA** asume que la cepa mutante “se alejará” de la

cepa *WT* lo menos posible, por lo que encuentra una distribución de flujos cuya diferencia con la cepa *WT* es minimizada. Para esto minimiza la norma Euclidiana del vector de diferencias de flujos a través de programación cuadrática (*Quadratic Programming* – QP). Por otro lado, el **ROOM** utiliza programación lineal entera mixta (*Mixed Integer Linear Programming* – MILP) para buscar una distribución de flujos que minimiza el número de cambios de flujo en el *WT*. Es así que el MOMA puede generar varios cambios pequeños en gran parte de los flujos totales, a diferencia del método ROOM que intenta minimizar dichos cambios.

El MOMA ha sido utilizado tanto en *E. coli* como en *S. cerevisiae* para identificar una o más combinaciones de genes a deletar para aumentar la producción de ciertos compuestos. Por ejemplo, para la producción de L-valina en una cepa de *E. coli* previamente modificada se mejoró la producción al identificar la combinación de tres genes candidatos que aumentaban la producción 3 veces (Park *et al.* 2007). Para esto se utilizó el modelo MBEL979 (Lee *et al.* 2005), una modificación del iJR904 (Reed *et al.* 2003), bajo el enfoque del MOMA y generando varias rondas de deletaciones *in silico* (simples, dobles y triples). En otro estudio, una cepa de levadura productora de vanilina fue modelada con la red metabólica iFF708 (Förster *et al.* 2003) bajo el mismo enfoque anterior de forma tal que se logró identificar la deletación de dos genes que incrementaron 5 veces la producción (Brochado *et al.* 2010).

El MILP ha sido utilizado junto con el FBA para desarrollar otros algoritmos, como el **OptKnock** y **OptStrain** que utilizan un sistema de optimización en dos niveles para diseñar una cepa mejorada. Para esto el análisis se basa en la teoría de la dualidad para optimizar dos funciones objetivos, una dentro de otra; o mejor dicho, el enfoque consiste en un problema externo y uno interno, donde el externo optimiza una función objetivo de diseño, mientras que el problema interno optimiza una función objetivo celular. Así, para el OptKnock se maximiza la producción de un compuesto sujeto a la maximización de biomasa (Burgard *et al.* 2003), o para el OptStrain se maximiza la producción de un compuesto sujeto a la minimización de la incorporación de reacciones foráneas (Pharkya *et al.* 2004).

Basado en el OptKnock, se desarrolló el método **OptGene** para identificar blancos génicos para la optimización de la producción de un compuesto usando un algoritmo genético en vez de MILP para reducir el tiempo computacional (Patil *et al.* 2005). Los algoritmos genéticos comúnmente se han utilizado para optimizar una función objetivo que es discontinua, no diferenciable, estocástica, o altamente no linear, y están basados en la evolución Darwiniana (Boghigian *et al.* 2010). El enfoque consiste en evolucionar una población de posibles soluciones durante varias generaciones, hasta que eventualmente se llegue a un óptimo global. Una población de varios genotipos es inicializada al asignar un estado *on/off* para cada gen, y luego a cada genotipo se le asigna un valor de *fitness* al usar FBA, MOMA, ROOM, u otro algoritmo. Los individuos con mayores *fitness* son seleccionados

para la generación de una nueva población aplicando modificaciones genéticas (*crossover* y mutaciones aleatorias). Este ciclo se repite hasta que se cumpla una de las condiciones de salida, como número máximo de generaciones o tiempo máximo de corrida. El algoritmo OptGene ha sido utilizado para sugerir blancos de delección para mejorar la producción de vanilina, glicerol, ácido succínico, y sesquiterpenos en *Saccharomyces cerevisiae* (Patil et al. 2005; Asadollahi et al. 2009).

Por otro lado, MILP ha sido utilizado para encontrar **óptimos alternativos**, es decir, donde la función objetivo toma un mismo valor pero las distribuciones de flujo son distintas. Un ejemplo de esto lo entrega el trabajo de Lee *et al.* (2000) en donde se identificaron 9 distribuciones de flujos en una red metabólica de *E. coli* con 30 metabolitos y 33 reacciones. Extrapolando tal trabajo, en una red a escala genómica el número de distribuciones distintas puede aumentar considerablemente junto con los costos computacionales requeridos; de hecho, en un trabajo realizado con el modelo a escala genómica de *E. coli* iJR904, se mostraron más de 50.000 distribuciones de flujo alternativos bajo 136 condiciones de crecimiento (Reed & Palsson 2004). Una forma de enfrentar tal problema es con un **análisis de variabilidad de flujos** (*Flux Variability Analysis* - FVA) el cual se enfoca en cada reacción en vez de buscar el número total de distribuciones posibles (Mahadevan & Schilling 2003). Esta estrategia fija el valor de la función objetivo al valor

óptimo, para luego maximizar y minimizar cada una de las reacciones entregando el rango de variación de cada reacción ante un valor óptimo dado.

Otro análisis basado en el FBA, pero con el objetivo de identificar genes a sobreexpresar es el **FSEOF** (*flux scanning based on enforced objective flux*) (Choi *et al.* 2010). El análisis consiste en agrupar los flujos metabólicos según su respuesta ante el aumento (forzado) de la síntesis de un compuesto de interés. De esta forma las distintas reacciones se clasifican según el comportamiento de sus flujos los que pueden ser: crecientes, decrecientes constantes, u oscilantes. Este método se aplicó en *E. coli* para hallar blancos génicos que aumentan la producción de licopeno, identificando 3 genes (*dxs*, *idi* y *mdh*) que al amplificarlos mejoraron en 3 veces la producción.

La metodología del FSEOF se relaciona con el concepto desarrollado por Stephanopoulos y Vallino (1991) para identificar los cuellos de botella en la sobreproducción de compuestos de interés. De acuerdo a su análisis, el rendimiento del producto depende de la partición de flujos de un pequeño subconjunto de todos los nodos que comprenden la red metabólica, los que se denominan **nodos principales** (NP). La identificación de ellos facilita la elaboración de estrategias para la ingeniería metabólica, para lo cual se deben ubicar aquellos nodos donde existe la mayor variación en la partición de flujos a medida que aumenta la síntesis del producto de interés. Esto último es lo que

vincula ambos análisis, ya que en los dos casos se requiere de la evolución de la distribución de flujo desde un estado inicial, a un estado final de producción del compuesto de interés. La ventaja de la búsqueda de los NP es que debido a que se relaciona con la arquitectura de la red, es posible realizarlo con o sin datos experimentales.

Sin embargo, aunque la arquitectura de la red afecta la respuesta metabólica ante modificaciones enzimáticas, es el grado de rigidez del nodo en donde reside en última instancia la facilidad con que la partición de flujo puede ser alterada (Stephanopoulos & Vallino 1991). Tal característica está determinada tanto por las diferencias en los parámetros cinéticos de las enzimas participantes en cada rama del NP, como en la estructura de control (*feedback*, activación/inactivación por fosforilaciones, *etc*). De esta forma se pueden establecer tres tipos de nodos: los flexibles, los débilmente rígidos, y los fuertemente rígidos.

Los **nodos flexibles** son aquellos donde la partición de flujos puede ser cambiada fácilmente al aumentar la demanda por el metabolito en cuestión. En ellos, las enzimas participantes en cada rama tienen afinidades similares por el nodo, al igual que las velocidades de reacción son similares en magnitud. Por otro lado, la partición de flujos en un **nodo rígido** está fuertemente regulada y es dominada por la cinética de una de sus ramas. Dichos nodos pueden formar cuellos de botella

debido a su resistencia a modificar su partición de flujos a favor de la biosíntesis de los productos de interés.

A diferencia de los NP, para estudiar la rigidez de los nodos se requieren de datos experimentales. Los métodos que se han propuesto para dicho fin contemplan la desregulación de ciertos metabolitos para incrementar la carga metabólica, la alimentación con distintas fuentes de carbono, o el aumento y disminución de la actividad enzimática a través de modificaciones genéticas o por la adición de inhibidores específicos.

1.3.4 Modelos genómicos de *S. cerevisiae*

El primer modelo a escala genómica de *Saccharomyces cerevisiae* fue el **iFF708**. Esta red consiste en 708 genes, 1175 reacciones, 584 metabolitos, y además compartamentaliza las reacciones en el citosol, mitocondria y el medio extracelular (Förster *et al.* 2003). Las predicciones de dicho modelo se ajustan a los datos experimentales de distintos cultivos y con diversas delecciones (Famili *et al.* 2003). A partir de dicha reconstrucción, varios modelos se han desarrollado por diversos equipos de investigación, los cuales han ido aportando al conocimiento del metabolismo de *S. cerevisiae*. El modelo **iND750** introdujo 5 compartimentos extras (peroxisoma, núcleo, aparato de golgi, vacuola, y el retículo endoplasmático rugoso) (Duarte *et al.* 2004), y el modelo **iLL672** eliminó algunas reacciones que no estaban conectadas a la red metabólica (Kuepfer *et al.* 2005). Luego, la versión

iIN800 introdujo el metabolismo de lípidos y modificó la ecuación de biomasa (Nookaew *et al.* 2008). Posteriormente, el modelo **iMM904** adicionó genes involucrados en la degradación de esfingolípidos y glicerolípidos, además de incorporar la síntesis y degradación de esteril ésteres y almacenamiento de esteroides (Mo *et al.* 2009). Este modelo también adicionó varias reacciones de transporte para conectar algunos *gaps* entre los compartimentos intracelulares, y agregó ciertas vías de secreción.

Todas estas reconstrucciones se realizaron independientemente, usando distintos datos o interpretando de distinta forma la información de literatura, lo que motivó al desarrollo de un modelo consenso. Este modelo ha sido actualizado regularmente y su última versión, denominada **Yeast 4.0**, contiene 932 genes, 1865 reacciones y 1319 metabolitos (Dobson *et al.* 2010).

Tabla 1: Comparación de los modelos a escala genómica de *Saccharomyces cerevisiae*.

Modelo	Genes	Reacciones	Metabolitos	Templado	Referencia
iFF708	708	1175	584		Förster et al. 2003
iND750	750	1489	1061	iFF708	Duarte et al. 2004
iLL672	672	1038	636	iFF708	Kuepfer et al. 2005
iIN800	800	1446	1013	iFF708	Nookaew et al. 2008
iMM904	904	1412	1228	iND750	Mo et al. 2009
Yeast 4.0	932	1865	1319	consenso	Dobson et al. 2010

En este trabajo se utilizaron los algoritmos de búsqueda OptGene, FSEOF, y NP sobre los modelos iND750, iMM904, y Yeast 4.0 con el fin de **identificar una vía cuyos flujos disminuyen a medida que aumenta la síntesis de limoneno**, un compuesto utilizado como molécula representativa de los monoterpenos en general. El uso de cada método en un mismo modelo le confiere al análisis una mayor robustez debido a que se explora la red metabólica bajo distintos enfoques, pero con un mismo objetivo: la maximización de la biosíntesis de monoterpenos. Asimismo, los distintos modelos también aportan a la robustez del análisis debido

a que están contruidos con distintos supuestos. Por ejemplo, un modelo puede asumir una vía biosintética que en otro esté ausente, o las ecuaciones de biomasa pueden asumirse distintas, *etc.*

2. MÉTODOS

Tres modelos metabólicos a escala genómica (iND750, iMM904, y yeast4) (Duarte *et al.* 2004; Mo *et al.* 2009; Dobson *et al.* 2010) fueron implementados en el *software* Matlab con la ayuda del *toolbox* COBRA (*Constraints Based Reconstruction and Analysis*). En dicha plataforma se modificaron los modelos al incorporar a cada uno la reacción de síntesis de limoneno a partir de GPP (**ANEXO Aa**). Este compuesto es utilizado como representante de los monoterpenos en general. Luego, se ejecutaron los algoritmos OptGene, FSEOF, y uno de búsqueda de los nodos principales, con el fin de destacar las reacciones más importantes en la redistribución de flujos.

2.1. OptGene

Dentro del *Toolbox* COBRA se encuentra la rutina de optimización evolutiva (Patil *et al.* 2005), pero que requiere del paquete de optimización *tomlab_cplex* para el correcto funcionamiento. Luego de instalar tal paquete se pudo correr el algoritmo (**ANEXO Ab**), el cual entrega como resultado una población de 500 individuos, donde cada uno corresponde a un conjunto de genes a ser deletados. Tales individuos están ordenados según un valor de *fitness* que surge del resultado de la optimización realizada bajo el enfoque MOMA, sin embargo, el programa sólo entrega el detalle del mejor individuo, lo que impide analizar al resto de ellos. Por lo tanto, se modificó el programa para acceder a

todas las poblaciones evaluadas (**Figura 3**), tal que se aumenta la posibilidad de encontrar un gen que en la práctica pueda ser mutado.

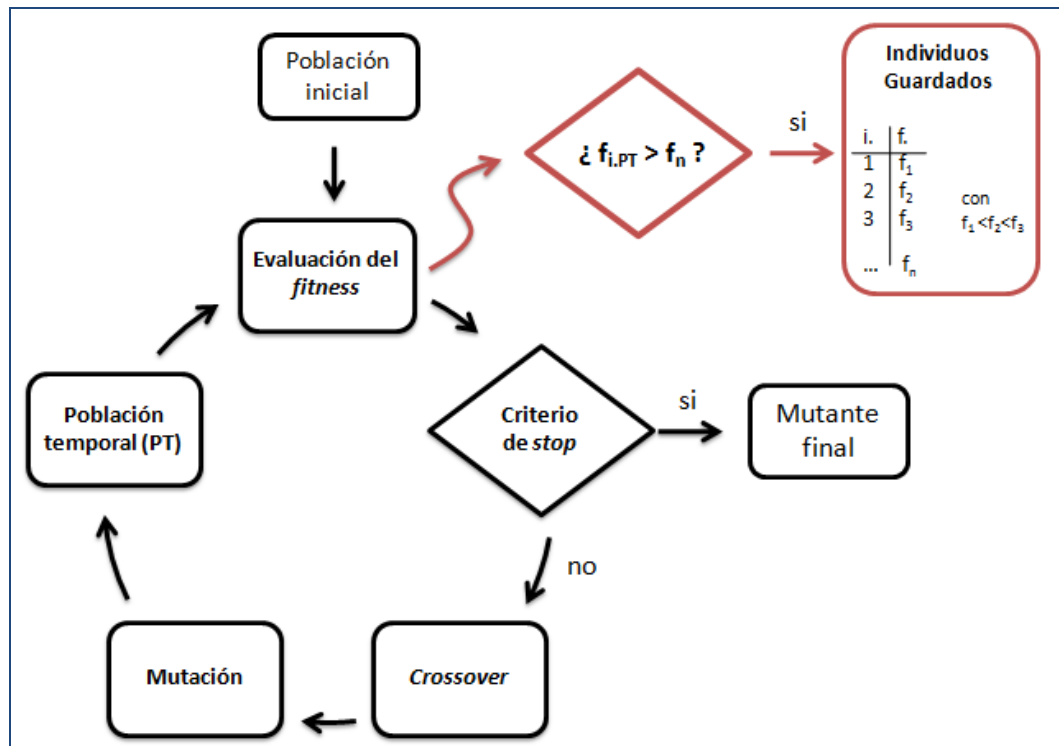


Figura 3: Estructura general del algoritmo OptGene desarrollado por Patil et. al. 2005. En negro se presenta la base del algoritmo OptGene, y en rojo lo adicionado al *script* del Toolbox COBRA para almacenar un número mayor de individuos evaluados. El criterio fue que si el *fitness* del mejor individuo de la nueva población ($f_{i,PT}$) es mayor que el *fitness* del mejor individuo de los que están almacenados (f_n), entonces se guarda dicho individuo.

2.2. FSEOF (flux scanning based on enforced objective flux)

Para aplicar el algoritmo descrito por (Choi et al. 2010) se escribió un *script* que genera 10 distribuciones de flujos (*steps*) (**Figura 4A**), en donde la primera distribución (v_j^{step1}) corresponde a la maximización de biomasa (**ec.2. 1**), y la décima (v_j^{step10}) corresponde a la maximización de Limoneno (**ec.2. 2**), compuesto utilizado como representante de los monoterpenos. Los *steps* intermedios (v_j^{step-k}) corresponden a la maximización de biomasa, pero con flujos crecientes en la síntesis de limoneno (**ec.2. 3**).

Step1

$Max v_{biomasa}$

ec.2. 1

s.a.

$$\sum_{j=1}^n S_{ij} v_j^{step1} = 0; \quad (\forall i \in M, \text{ y } \forall j \in N)$$

$$lb_j \leq v_j \leq ub_j$$

$$0 \leq v_{Limoneno} \leq 1000$$

Step10

$Max v_{Limoneno}$

ec.2. 2

s.a.

$$\sum_{j=1}^n S_{ij} v_j^{step10} = 0; \quad (\forall i \in M, \text{ y } \forall j \in N)$$

$$lb_j \leq v_j \leq ub_j$$

$$0 \leq v_{Limoneno} \leq 1000$$

Step K

Max $v_{biomasa}$

s.a.

$$\sum_{j=1}^n S_{ij} v_j^{step_K} = 0; \quad (\forall i \in M, \text{ y } \forall j \in N)$$

$$lb_j \leq v_j \leq ub_j$$

$$k R \leq v_{Limoneno} \leq 1000; \quad (\forall k \in [1, 2, \dots, 9])$$

$$R = \frac{v_{Limoneno}^{step10}}{10}$$

ec.2. 3

Donde S_{ij} es el coeficiente estequiométrico del metabolito i en la reacción j , M es el set de metabolitos, N es el set de reacciones, lb_j y ub_j son los límites inferiores y superiores para la reacción j ($-1000 \leq v_j \leq 1000$, para reacciones reversibles; $0 \leq v_j \leq ub_j$, para reacciones irreversibles).

Con este método, cuatro tipos de flujos fueron identificados a medida que se avanza por los steps: flujos crecientes, decrecientes, invariables, y flujos oscilantes (**Figura 4 B-E**). Sin embargo, debido a la existencia de soluciones múltiples, no existe la certeza de que las reacciones dentro de un grupo tengan siempre el mismo comportamiento, por lo que fue necesario realizar un análisis de variabilidad (Mahadevan & Schilling 2003) a cada una. Este análisis consiste en fijar la función objetivo (producción de biomasa) en su valor óptimo, y luego maximizar y minimizar el flujo por cada reacción. Para esto se utilizó el comando *fluxVariability* del *Toolbox* COBRA (**ANEXO Ac**) en cada uno de los steps (**Figura 4 F y G**) de los flujos decrecientes. Este método, si bien no entrega el

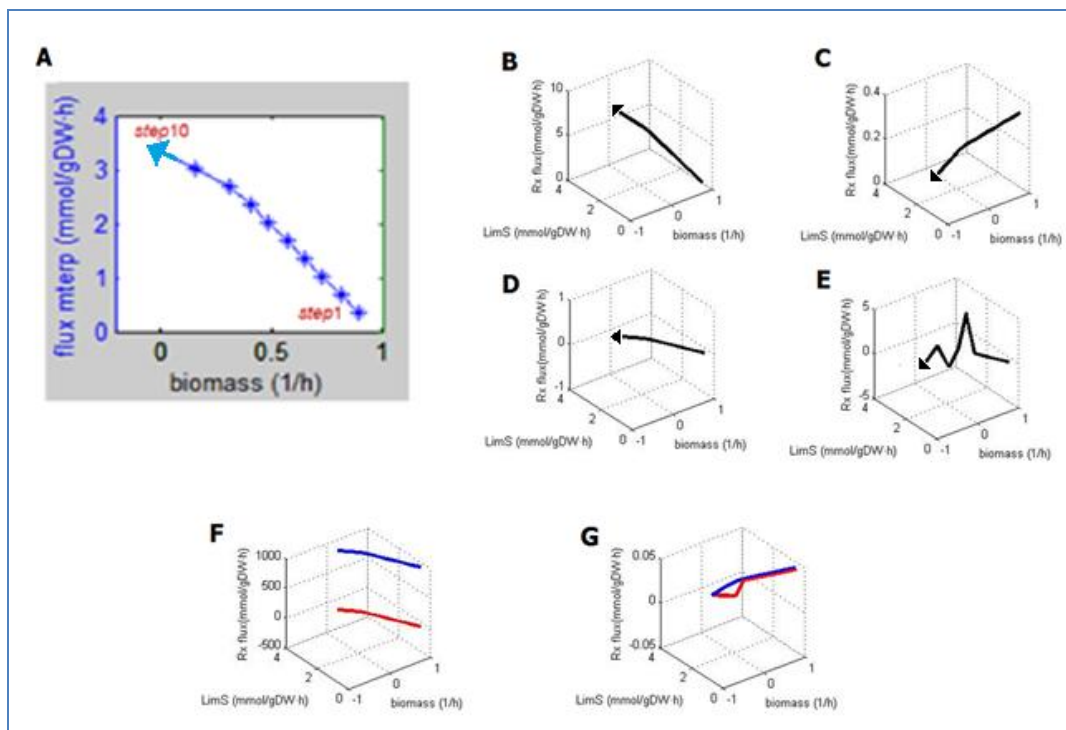


Figura 4: FSEOF y análisis de variabilidad. (A) Base del algoritmo FSEOF donde se maximiza la biomasa (*step1*) o la producción de monoterpenos (*step10*). Los *steps* intermedios corresponden a la maximización de biomasa pero con flujos crecientes para la síntesis de monoterpenos. (B-E) Reacciones clasificadas como (B) crecientes, (C) decrecientes, (D) invariables, u (E) oscilantes según se avanza del *step1* al *step10*. (F y G) Análisis de variabilidad para reacciones (F) muy variables o (G) poco variables. El color azul indica el valor máximo de flujo para una determinada reacción, mientras que el color rojo indica su valor mínimo.

número de cuantas distribuciones de flujo distintas pueden existir, entrega el rango de variación para cada reacción. De esta forma, las reacciones decrecientes y que además son poco variables se clasificaron como FSEOF_down.

Las reacciones clasificadas así pueden ser utilizadas como base de búsqueda de aquellas más plausibles para ser deleccionadas con el fin de aumentar el flujo a la síntesis de monoterpenos. En otras palabras, se pueden descartar del análisis de búsqueda aquellas reacciones que no interesan debido a su comportamiento creciente, oscilante o invariable ante el aumento del flujo hacia los monoterpenos.

2.3. Nodos Principales (NP)

Los NP son aquellos donde su partición de flujos varía en mayor medida al pasar de una condición a otra (Stephanopoulos & Vallino 1991). Para poder identificarlos en un modelo a escala genómica se desarrolló el siguiente algoritmo dividido en 3 etapas: i) FSEOF, ii) diferencia porcentual (*score*), iii) *ranking* y selección de los NP. La primera de ellas consiste en generar y guardar los datos del algoritmo FSEOF, donde la condición inicial es la producción de biomasa, y la final es la producción de limoneno. En la segunda etapa (**Figura 5**) se calcula un *score* para cada nodo (o metabolito), el cual está dado por la norma de la diferencia porcentual entre cada *step*. De esta forma se puede ordenar cada nodo de acuerdo a tal valor para los 10 *steps*. Finalmente, en la tercera etapa se promedia el puesto (*ranking*) de cada nodo, y se elige un número arbitrario de los

nodos con los mejores *rankings* para representar a los NP, que en nuestro caso fueron sólo los 40 primeros. De estos 40 nodos, se descartaron algunos (*e.g.* ATP, NADH, H₂O, CO₂) debido a su participación en varias reacciones lo que dificulta el análisis (ANEXO C).

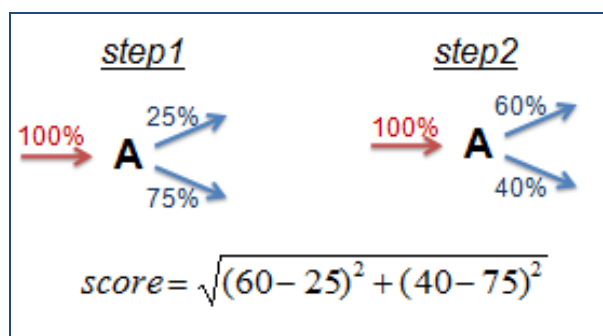


Figura 5: Segunda etapa del algoritmo NP. Cálculo del *score* utilizado para ordenar los nodos. En este ejemplo, todos los flujos que producen el metabolito “A” son estandarizados al 100% y comparados con los flujos que lo consumen.

2.4. Análisis Integrado

Los enfoques de cada algoritmo de búsqueda anteriormente descritos son distintos (Tabla 2), pero los tres apuntan al mismo objetivo: aumentar los flujos metabólicos hacia los monoterpenos. Por ejemplo, el algoritmo OptGene busca un conjunto de **genes** que pueden ser delecionados, el FSEOF encuentra **reacciones** que disminuyen sus flujos continuamente, y los NP busca a los **metabolitos** cuyos destinos metabólicos cambian en mayor medida. Sin embargo, cuando son

ejecutados con la producción de monoterpenos como función objetivo, se deben encontrar coincidencias entre sus resultados si son analizados desde el punto de vista de las reacciones químicas, o mejor dicho, desde el punto de vista de la topología de la red metabólica.

Tabla 2: Resumen de características de los tres algoritmos de búsqueda utilizados.

Algoritmo	Detalles	Función Obj.	Búsqueda de
OptGene	Busca genes que al ser deletados puedan aumentar el rendimiento de un compuesto de interés.	MOMA	genes
FSEOF	Clasifica cada reacción de la red metabólica según su comportamiento al aumentar progresivamente el flujo a una reacción en particular.	Biomasa	reacciones
NP	Selecciona los metabolitos con los mayores cambios en sus particiones de flujo cuando aumenta progresivamente el flujo a una reacción en particular.	Biomasa	metabolitos

De esta forma, los resultados de los NP y los del OptGene fueron analizados según las reacciones a las cuales están asociadas, identificando aquellas que coinciden con las del FSEOF_down (**Figura 6**). Al encontrar las coincidencias resulta un análisis más robusto ya que contiene la combinación de tres enfoques distintos, y permite un estudio con menos reacciones lo que facilita la comprensión de los resultados. Análogamente, se compararon las reacciones seleccionadas para cada modelo metabólico, lo que también aporta en la robustez del análisis debido a que cada modelo tiene sus propias características. Por ejemplo, el modelo iND750

carece de algunas rutas de síntesis de lípidos que sí tiene el iMM904, y el modelo Yeast 4.0 es la versión consenso del metabolismo de *Saccharomyces cerevisiae* y tiene más reacciones y metabolitos que los modelos anteriores.

Cabe destacar que se consideró como base sólo el conjunto de reacciones FSEOF_down, ya que las que están fuera de él tienen un comportamiento que no es de interés (flujos crecientes, invariables u oscilantes). De esta forma las reacciones provenientes del algoritmo OptGene y del NP permitieron únicamente descartar (o filtrar) 59, 126 y 125 reacciones para los modelos iND750, iMM904 y yeast4.0 respectivamente (**Figura 6**).

Una dificultad en esta metodología es la nomenclatura utilizada en el modelo Yeast 4.0 ya que es totalmente distinta a la de los otros dos modelos. Por este motivo, el análisis de integración de datos se realizó de forma manual, lo que fue facilitado por los filtros realizados con los algoritmos de búsqueda. Dichos algoritmos seleccionaron 19, 49, y 23 reacciones para los modelos iND750, iMM904, y Yeast 4.0 respectivamente, cantidades manejables para emparejar de forma visual aquellas que son idénticas en una plantilla de *Microsoft Excel*.

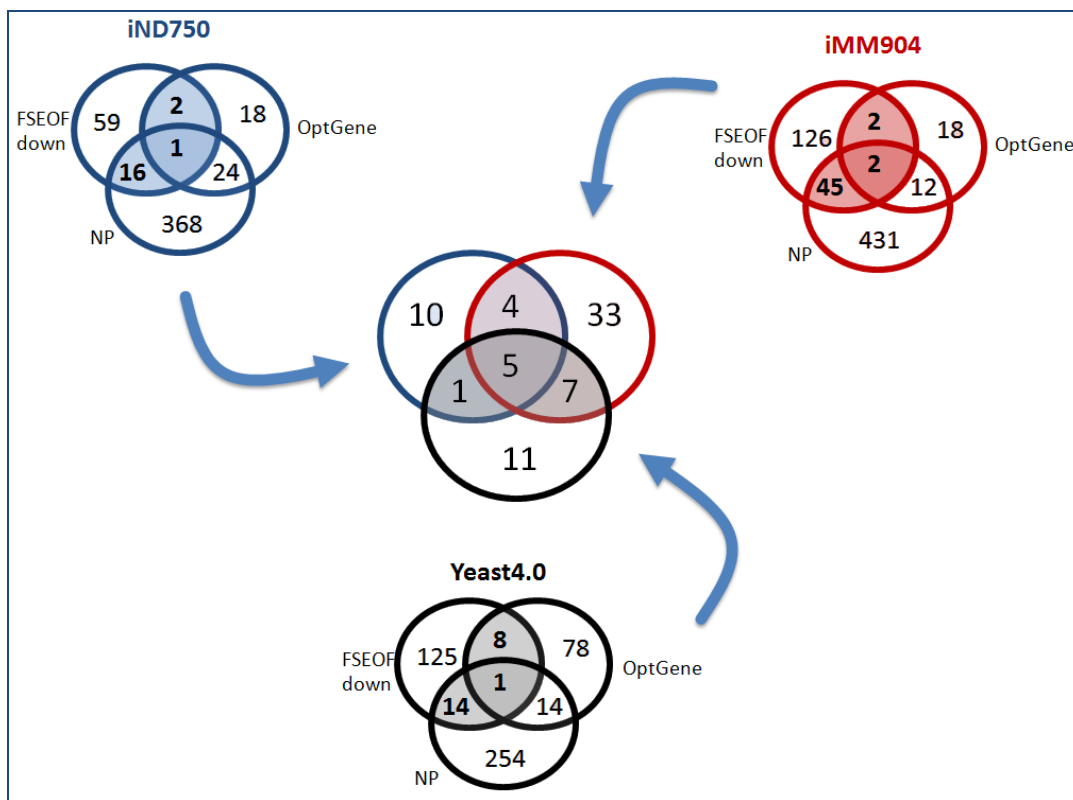


Figura 6: Clasificación y número de reacciones para cada modelo. A cada modelo metabólico se le aplicaron los algoritmos OptGene, FSEOF y NP, para luego agrupar las reacciones que se repiten entre ellos.

3. RESULTADOS

3.1. Máximo Teórico

Para tener una idea general del proceso de producción microbiana de limoneno ($C_{10}H_{16}$), lo primero que se realizó fue calcular el rendimiento teórico máximo en un cultivo alimentado únicamente con glucosa. Para ello se puede plantear la ecuación general:

$$-a \cdot C_6H_{12}O_6 - b \cdot O_2 + c \cdot CO_2 + d \cdot H_2O + C_{10}H_{16} = 0$$

ec.3. 1

En donde tres de los cuatro coeficientes pueden ser calculados a partir del grado de reducción y de los balances de masa de carbono e hidrógeno, quedando la siguiente expresión en función de un solo parámetro.

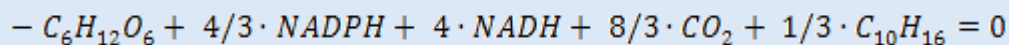
$$-\left[\frac{d+8}{6}\right] \cdot C_6H_{12}O_6 - [d-6] \cdot O_2 + [d-2] \cdot CO_2 + d \cdot H_2O + C_{10}H_{16} = 0$$

ec.3. 2

Pero debido a que en el proceso no se produce O_2 se puede plantear la desigualdad: $d-6 \geq 0$; de la que se obtiene que: $d \geq 6$. Es así que el mayor

rendimiento de limoneno se logra con un $d = 6$; tal que $Y_{PS} = 3/7 = 0,43$ [mol limoneno/mol glucosa].

Sin embargo, esta metodología no considera los requerimientos de cofactores ni supone una ruta metabólica. Para incorporar dichos parámetros la formulación de todas las reacciones desde glucosa hasta limoneno pueden ser planteadas (**ANEXO B**), de tal forma que al considerar la ruta más corta (la menor cantidad de reacciones) se obtiene la ecuación **ec.3. 3**, lo que entrega un rendimiento de $Y_{PS} = 1/3 = 0,33$ [mol limoneno/mol glucosa] (y $d = 10$).



ec.3. 3

Posteriormente se montaron tres modelos metabólicos de *Saccharomyces cerevisiae* y se compararon dichos valores. Los modelos utilizados fueron el iND750, el iMM904, y el modelo consenso Yeast 4.0 (Duarte *et al.* 2004; Mo *et al.* 2009; Dobson *et al.* 2010), los que se utilizaron por su mayor accesibilidad de descarga e implementación. A continuación, con el *software* MATLAB y el *toolbox* COBRA se estableció un *uptake* de glucosa a 10 mmol/gDW·h y se les incorporó la reacción de producción de limoneno a partir de GPP (**ANEXO A; Figura 1**) para luego maximizar su biosíntesis al igual que la de biomasa (**Tabla 3**).

Tabla 3: Valores para la producción de biomasa (en h^{-1}) y para la biosíntesis de limoneno (en $\text{mmol/gDW}\cdot\text{h}$) para los tres modelos utilizados cuando el *uptake* de glucosa es fijado a $10 \text{ mmol/gDW}\cdot\text{h}$.

función objetivo	iND750	iMM904	Yeast 4.0
biomasa	0.97	0.97	1.70
limoneno	3.33	3.47	4.29

Las diferencias en la producción de limoneno se deben a las distintas rutas que se utilizan entre los nodos **piruvato** y **acetil-CoA** (**Figura 7**). El flujo de entrada a piruvato en los tres modelos es de $20 \text{ mmol/gDW}\cdot\text{h}$, pero debido a las distintas rutas utilizadas a partir de dicho nodo los flujos de entrada a la vía del mevalonato son de 20, 20.8, y $25.7 \text{ mmol/gDW}\cdot\text{h}$ para los modelos iND750, iMM904 y Yeast 4.0 respectivamente. El modelo iND750 utiliza casi la misma ruta que la descrita en el **ANEXO B** para llegar a la **ec.3. 3**, la única diferencia es que en vez de utilizar la piruvato deshidrogenasa para formar acetil-CoA, utiliza la piruvato descarboxilasa que produce acetaldehído. Por otro lado, los modelos iMM904 y Yeast 4.0 utilizan además una rama del nodo piruvato que se origina con la piruvato carboxilasa para formar oxalacetato, el cual luego de una serie de reacciones puede formar acetaldehído y glicina (**Figura 7**). El acetaldehído entra a la ruta del mevalonato, mientras que la glicina luego de transformarse en serina puede formar nuevamente piruvato. El mayor rendimiento de esta ruta alternativa

se debe a que no se pierde un átomo de carbono en forma de CO_2 con la piruvato deshidrogenasa o la piruvato descarboxilasa.

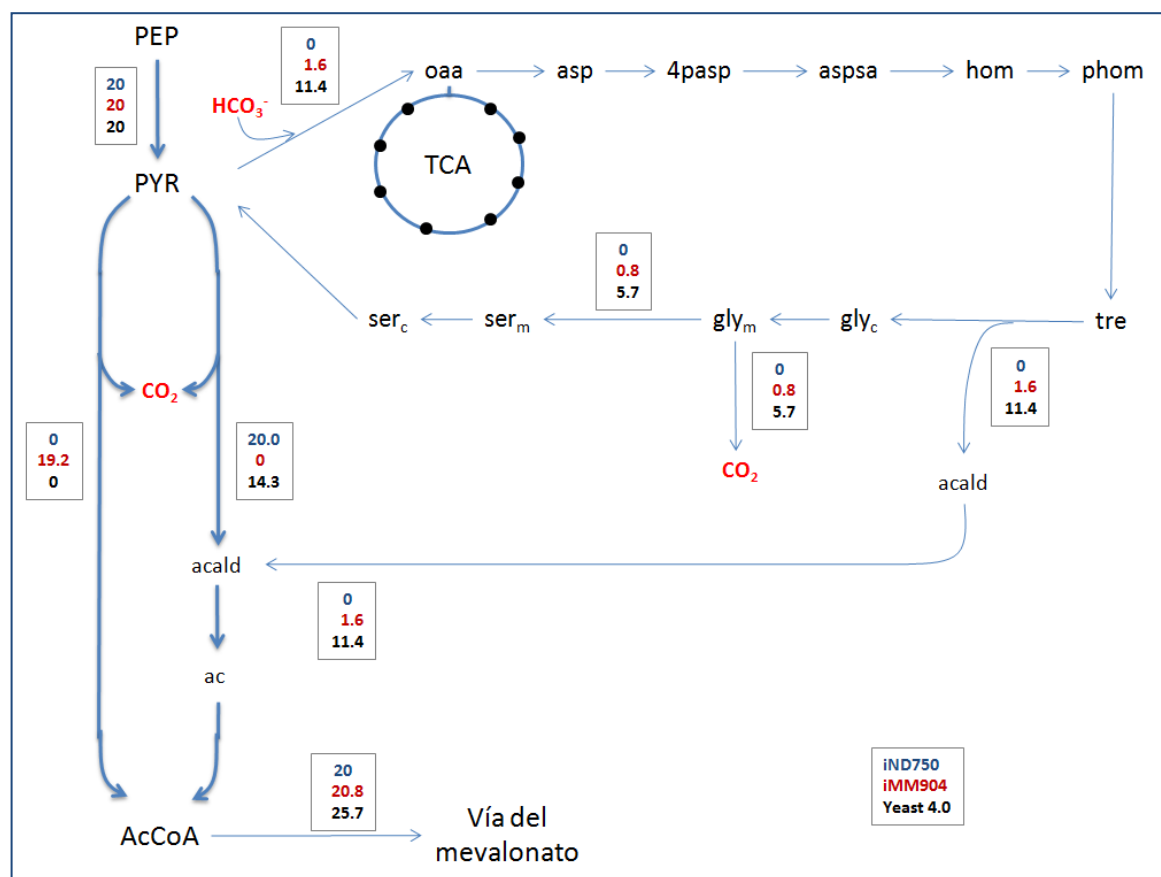


Figura 7: Ruta alternativa para la producción de monoterpénos. Los recuadros muestran el valor de flujo en mmol/gDW·h para cada vía según un uptake de glucosa de 10 mmol/gDW·h. El primer valor en azul corresponde al modelo iND750, el segundo en rojo al modelo iMM904, y el último valor en negro corresponde al modelo Yeast 4.0.

oaa: oxalacetato; asp: aspartato; 4pasp: 4P-aspartato; aspsa: aspartato semialdehído; hom: homoserina; phom: P-homoserina; tre: treonina; gly: glicina; ser: serina; Acald: acetaldehído; PYR: piruvato; PEP: fosfoenolpiruvato.

Por otro lado se analizó la productividad para los modelos estequiométricos según la definición de Patil (2005). La productividad es la cantidad de producto sintetizado por unidad de tiempo y por unidad de sustrato consumido, por lo que la multiplicación del rendimiento con la tasa de crecimiento entrega dicho valor. Sin embargo, debido a que en los modelos estequiométricos la tasa de consumo de sustrato es fijada a un valor determinado, Patil (2005) hace la distinción entre la productividad real y la calculada para los modelos, la cual la denomina como BPCY (*Biomass-Product Coupled Yield*). Este análisis bajo el enfoque del FSEOF para los tres modelos puede encontrar fácilmente el máximo de productividad BPCY según lo muestra la **Figura 8**.

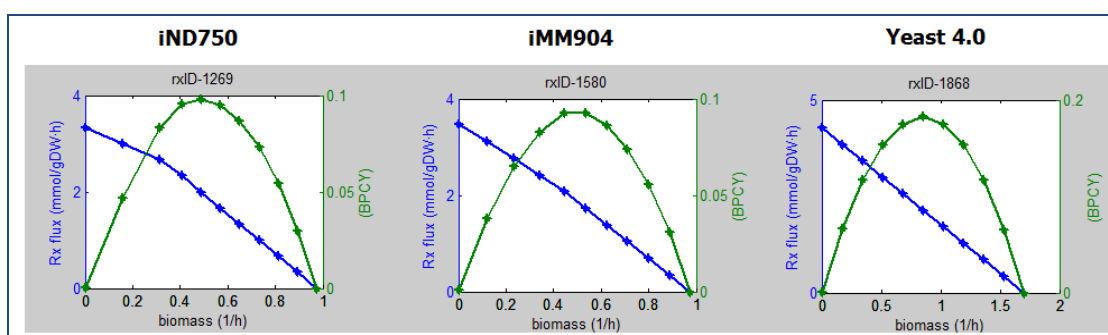


Figura 8: Análisis FSEOF y cálculo de la productividad BPCY para los modelos iND750, iMM904, y Yeast 4.0.

3.2. Análisis Integrado

Los modelos metabólicos a escala genómica iND750, iMM904, y Yeast 4.0 fueron utilizados junto con los algoritmos de búsqueda OptGene, FSEOF, y NP para seleccionar un conjunto de reacciones cuyos flujos disminuyen continuamente a medida que se fuerza la producción de limoneno. Tales reacciones sirven como puntos claves dentro de la red metabólica para poder entender la redistribución de flujos, ya que presentan un metabolito cuya partición de flujos es uno de los que se ve mayormente afectado, llamado nodo principal (NP) (Stephanopoulos & Vallino 1991). La selección consiste en ubicar todas las reacciones que son comunes en los tres algoritmos para un mismo modelo, y luego identificar aquellas que son comunes a los tres (**Figura 6**).

La ventaja de esta metodología es el uso de tres enfoques distintos para la exploración de las redes metabólicas. De esta forma se puede reducir el número de reacciones seleccionadas facilitando el análisis. Por ejemplo, si se toman únicamente los datos entregados por el algoritmo de búsqueda FSEOF_down se obtienen 78, 175, y 148 reacciones para los modelos iND750, iMM904, y Yeast 4.0 respectivamente. El análisis no se clarifica si además dichas reacciones son agrupadas de acuerdo a las rutas en las cuales participan, y ocurre lo mismo con cada algoritmo de búsqueda por separado (**ANEXO C**). Sin embargo, cuando los algoritmos son integrados la cantidad de reacciones seleccionadas disminuyen a

19, 49, y 23 respectivamente, lo que puede ser mejorado a 17 reacciones cuando se combinan los tres modelos (**Figura 6**).

Esta metodología se utilizó bajo dos condiciones: con un *uptake* de glucosa como única fuente de carbono, y con un suplemento de aminoácidos.

3.2.1. Glucosa como Única Fuente de Carbono

Para todas las reacciones de transporte entre el medio extracelular y el citosol el comportamiento y variabilidad fueron similares en los tres modelos (**ANEXO D**). Esto asegura que los modelos pueden ser comparables entre sí, aportando a la robustez del análisis. Por otro lado, bajo el escenario en el cual la glucosa es la única fuente de carbono (con un *uptake* de 10 mmol/gDW h), es claro que el flujo de carbono debe ingresar por la glicólisis y luego ramificarse hacia la síntesis de los precursores de biomasa en el *step1*, o hacia los precursores de limoneno en el *step10*.

Es así que las reacciones que presentan las mayores disminuciones de flujo tienen un NP en la glicólisis (**Tabla 4**). En primer lugar la reacción **CSm** tiene como NP al Acetil-CoA, y corresponde a la entrada al ciclo TCA que proporciona el ATP necesario para la producción de biomasa. Cabe destacar que en la ecuación de biomasa el coeficiente estequiométrico promedio para cada precursor es de 0.17, sin embargo, para el ATP tal valor es de 59, lo que explica que la reacción CSm

tenga la mayor reducción de flujo dentro de la **Tabla 4**. En segundo lugar se encuentra la reacción **MAN6PI**, la que a partir del NP fructosa-6P sintetiza manosa-6P. Éste último metabolito sirve como precursor de *mannan*, un compuesto que está presente en la ecuación de biomasa con un coeficiente estequiométrico de 0.8, lo que justifica, al igual que en el caso anterior, la elevada reducción de flujo. En tercer lugar, la reacción **ACLSm** origina a partir del piruvato una ruta que culmina (sin ramificaciones) con la síntesis de valina y leucina, precursores directos de biomasa.

Por otro lado, al utilizar los mapas metabólicos de la base de datos kegg (www.genome.jp/kegg) (**ANEXO E**), nueve de las reacciones de la **Tabla 4** son agrupadas en vías de síntesis de aminoácidos, de las cuales ocho tienen al glutamato como NP. Este metabolito está asociado a más de 40 reacciones, donde el mayor aporte porcentual de síntesis y de consumo está restringido a sólo 4, las que además presentan la mayor variabilidad. Tales reacciones participan de un gran ciclo, con más de 20 nodos (**ANEXO F**), en el cual existe un consumo neto de $\text{NADH}_{\text{mitocondrial}}$ y una producción de $\text{NADH}_{\text{citósolico}}$ en la misma cantidad, tal que una parte del ciclo funciona como un *shuttle*, similar al *malate-aspartate shuttle* (Lehninger *et al.* 2005).

Tabla 4: Reacciones seleccionadas por el análisis integrado para aumentar el flujo de síntesis de limoneno. Cada reacción seleccionada por el análisis fue buscada en la base de datos kegg (www.genome.jp/kegg), y se clasificaron en una o varias rutas metabólicas según tal página (**ANEXO E**). El nombre sigue el formato de los modelos iND750 e iMM904, y la reducción de flujo presenta el valor absoluto de tal disminución (en mmol/gDW·h) para cada modelo (iND750, iMM904, y yeast4.0 respectivamente).

Nombre	Clasificación según Kegg	OptGene	NP	Reducción de flujo.		
Grupo A (iND750, iMM904, yeast4.0)						
IG3PS	Histidine metabolism		'L_Glutamate'	0.06	0.1	0.2
ANS	Phenylalanine, tyrosine and tryptophan biosynthesis		'L_Glutamate'	0.03	0.03	0.05
HSTPT	Histidine metabolism	✓	'L_Glutamate'	0.06	0.1	0.2
TRPS1	Phenylalanine, tyrosine and tryptophan biosynthesis Glycine, serine and threonine metabolism		Glyceraldehyde_3_phosphate'	0.03	0.03	0.03
CSm	Citrate cycle (TCA cycle)		'Acetyl_CoA'	1	8	1.5
Grupo B (iMM904, yeast4.0)						
PHETA1	Phenylalanine, tyrosine and tryptophan biosynthesis		'L_Glutamate'		0.1	0.2
LEUTA	Valine, leucine and isoleucine biosynthesis		'L_Glutamate'		0.2	0.4
ACLSm	Butanoate metabolism Valine, leucine and isoleucine biosynthesis	✓	'Pyruvate'		0.6	1
DB4PS	Riboflavin metabolism		D_Ribulose_5_phosphate'		0.020	0.003
DRTPPD	Riboflavin metabolism		'Ammonium'		0.001	0.002
MAN6PI	Fructose and mannose metabolism		D_Fructose_6_phosphate'		0.6	1.3
TRDR	Pyrimidine metabolism Selenocompound metabolism	✓	Reduced_thioredoxin'		0.06	0.15
Grupo C (iND750, iMM904)						
ASNS1	Alanine, aspartate and glutamate metabolism Nitrogen metabolism	✓	'L_Glutamate'		0.1	0.1
AATA	Lysine biosynthesis Lysine degradation		'L_Glutamate'		0.3	0.3
SACCD1	Lysine biosynthesis Lysine degradation		'L_Glutamate'		0.3	0.3
FAS80COA_L	Fatty acid biosynthesis		'Coenzyme_A'		0.06	0.05
Grupo D (iND750, yeast4.0)						
ACCOACr	Pyruvate metabolism		'Acetyl-CoA'		0.35	0.16

Los ciclos se pueden detectar por sus elevados flujos y alta variabilidad. En el caso del modelo iND750 el 6% de las reacciones (en valor absoluto) son superiores a 10 mmol/gDW h, llegando incluso a valores cercanos de 1000 mmol/gDW h (**Figura 9A**). De ellas, las que tienen un flujo superior a 40 mmol/gDW h son altamente variables. Para ejemplificar, la **Figura 9B** muestra un pequeño ciclo, cuyo efecto neto es la reducción en la síntesis de NADH a medida que se favorece la producción de limoneno. Esta reducción corresponde a 6 mmol/gDW h, valor comparable al 94% de las reacciones en el modelo. Además, resulta interesante el hecho de que esta reducción puede ser independiente del flujo inicial, es decir, que se podría esperar la misma caída de 6 mmol/gDW h si la restricción de límite superior para todas las reacciones en vez de 1000 mmol/gDW h, fuera de 500 u otro valor. Esta característica, en la cual permiten casi cualquier valor de flujo, explica la alta variabilidad en sus reacciones.

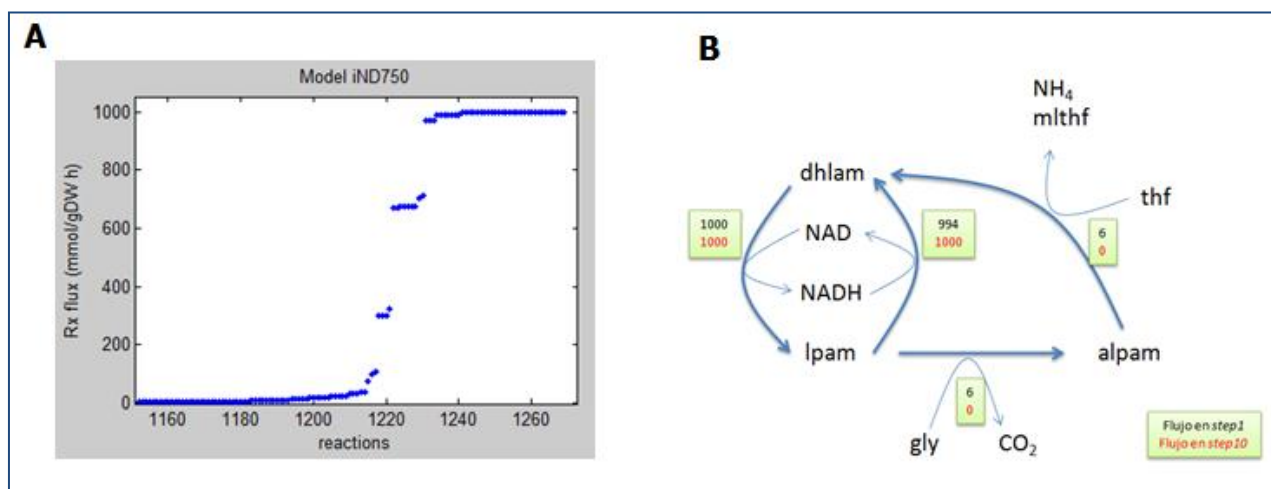


Figura 9: Ejemplo de ciclos en el modelo iND750. A) Para 1191 reacciones (94%) el valor absoluto del promedio de los flujos en los 10 steps es menor a 10 mmol/gDW h, 26 reacciones están entre 10 y 110 mmol/gDW h, 4 entre 300 y 350 mmol/gDW h, 9 entre 650 y 750 mmol/gDW h, y 39 reacciones están entre 950 y 1000 mmol/gDW h. **B)** Un ciclo pequeño de 2 nodos (dhlam, lpam) y dos reacciones laterales (lpam -> alpam -> dhlam), en el cual los valores dentro de los rectángulos representan el flujo (en mmol/gDW h) ya sea en el *step1* (dato superior de color negro) o en el *step10* (dato inferior de color rojo). Con respecto al NADH, el efecto neto es la disminución de su síntesis.

alpam: S-aminomethyldihydrolipoamide; *dhlam*: Dihydrolipoamide; *gly*: Glycine; *lpam*: Lipoamide; *mlthf*: Methylenetetrahydrofolate; *thf*: Tetrahydrofolate.

Por otro lado, las reacciones **DB4PS** y **DRTPPD**, están involucradas en la síntesis de riboflavina, un precursor directo de biomasa. El NP de la reacción DB4PS es la ribulosa-5P, un metabolito de la vía de las pentosas fosfato ubicado a la entrada de la ruta de biosíntesis de purinas; y la reacción DRTPPD pertenece a una ramificación de esta última vía (a partir del GTP), y cuyo NP es el amonio (**ANEXO E**, *Riboflavin metabolism*).

3.2.2. Uptake de Glucosa y Aminoácidos

Debido a que en las condiciones anteriores muchas vías están relacionadas con la biosíntesis de aminoácidos, y a que los cultivos se puedan realizar con estos compuestos en el medio, es que se realizó el mismo análisis pero con un suplemento de tales sustratos. De esta forma 16 reacciones fueron seleccionadas (**Tabla 5**), de las cuales 4 se repiten en las dos condiciones (MAN6PI, ACCOACr, FAS80COA_L, y TRDR). La reacción **MAN6PI** presenta la segunda mayor disminución de flujo en ambas condiciones debido a su participación en la síntesis de *mannan* como se explicó anteriormente. Por otro lado, las reacciones **ACCOACr** y **FAS80COA_L** están en la vía de síntesis de ácidos grasos, donde la primera también tiene una de las mayores disminuciones de flujo ya que sintetiza malonil-CoA, el precursor común a todos los ácidos grasos.

Tabla 5: Reacciones seleccionadas por el análisis integrado para aumentar el flujo de síntesis de limoneno con suplemento de aminoácidos. Cada reacción seleccionada por el análisis fue buscada en la base de datos kegg (www.genome.jp/kegg), y se clasificaron en una o varias rutas metabólicas según tal página (ANEXO E). El nombre sigue el formato de los modelos iND750 e iMM904, y la reducción de flujo presenta el valor absoluto de tal disminución (en mmol/gDW·h) para cada modelo (iND750, iMM904, y yeast4.0 respectivamente).

Nombre	Clasificación según Kegg	OptGene	NP	Reducción de flujo.		
Grupo A (iND750, iMM904, yeast4.0)						
GLUPRT	Purine metabolism Alanine, aspartate and glutamate metabolism	✓	'L_Glutamate'	0.12	0.12	0.26
MAN6PI	Fructose and mannose metabolism		'D_Fructose_6_phosphate'	1	1	2
PRFGS	Purine metabolism		'L_Glutamate'	0.12	0.12	0.26
Grupo B (iMM904, yeast4.0)						
GHMT2r	Glycine, serine and threonine metabolism One carbon pool by folate		'L_Serine'		0.35	0.65
GRTT	Terpenoid backbone biosynthesis		'Geranyl_diphosphate'		0.005	0.03
TRDR	Pyrimidine metabolism Selenocompound metabolism	✓	'Reduced_thioredoxin'		0.08	0.09
Grupo C (iND750, iMM904)						
ACCOACr	Pyruvate metabolism		'Acetyl_CoA'	0.47	0.5	
ATPS3m	Oxidative phosphorylation	✓		65	65	
CYOR_u6m	Oxidative phosphorylation		Ubiquinol_6' ; 'Ubiquinone_6'	58	60	
FAS80COA_L	Pyruvate metabolism		Coenzyme_A' ; 'Acetyl_CoA'	0.07	0.07	
GMPS2	Purine metabolism		'L_Glutamate'	0.06	0.06	
IMPD	Purine metabolism	✓		0.06	0.06	
PRPPS	Purine metabolism Pentose phosphate pathway	✓	'alpha_D_Ribose_5_phosphate'	0.25	0.27	
SLFAT	Sulfur metabolism Purine metabolism	✓	'Sulfate'	0.07	0.06	
Grupo D (iND750, yeast4.0)						
TKT1	Pentose phosphate pathway		glyceraldehyde 3-phosphate'	0.025		0.16
TKT2	Pentose phosphate pathway		glyceraldehyde 3-phosphate' ; 'D-Erythrose 4-phosphate' ; 'D-fructose 6-phosphate'	0.2		0.3

Por otro lado, la reacción **TRDR** participa en la vía de incorporación de sulfato al reducir la tioredoxina (**Figura 10**). En esta ruta participan además las reacciones **SLFAT** y **GHMT2r** en donde la primera fija el sulfato al ATP, y la segunda (con la tercera mayor disminución de flujo) interconvierte la serina en glicina transfiriendo una unidad de carbono al cofactor tetrahidrofolato (THF), metabolito que participa en varias rutas metabólicas (**ANEXO E, folate biosynthesis**).

En cuanto al metabolismo energético las reacciones **CYOR_u6m** y **ATPS3m** son las que presentan las mayores disminuciones de flujo (**Tabla 5**). La primera se ubica en la cadena transportadora de electrones mientras que la segunda corresponde a la ATP sintasa, y al igual que en el caso anterior, presentan las mayores reducciones debido a la menor demanda de ATP para la producción de biomasa.

Con respecto a la reacción **GRTT**, ésta corresponde a la rama de síntesis de ergosterol a partir del NP geranil difosfato (GPP). Tal metabolito es el precursor de limoneno, por lo tanto es evidente el cambio en su partición de flujos, por lo que no aporta información relevante al análisis. Por otro lado, todas las reacciones restantes están involucradas en la reducción de flujo en la vía de síntesis de purinas (**Figura 10**), por un lado participando directamente de dicha ruta (**PRPPS, GLUPRT, PRFGS, IMPD, y GMPS2**) o estando presente en la vía de las pentosas fosfato (**TKT1 y TKT2**).

3.3. Modelación de la Reducción de Flujos en la Vía de las Purinas

Al considerar todas las reacciones de la **Tabla 4** y **Tabla 5** se obtienen 29 reacciones ([17+16]-[4 repetidas]). De ellas 12 están vinculadas a la síntesis de purinas (**Figura 10**), otras 10 pertenecen a vías de síntesis de diversos aminoácidos, 3 se relacionan con la producción de ATP, 2 con los ácidos grasos, y 1 a la síntesis de de FPP. Todas ellas se entienden bajo el argumento de disminución de precursores de biomasa, **destacándose el porcentaje de reacciones vinculadas al metabolismo de purinas**. Por lo tanto, el análisis desarrollado bajo este enfoque *in silico*, propone al trabajo experimental la reducción de los flujos metabólicos en esta vía para aumentar el rendimiento de la producción de monoterpenos.

Para lograr dicha disminución, una alternativa es la reducción de flujos por la reacción PRPPS (**Figura 10**) tal que se bloquee la puerta de entrada a la ruta de biosíntesis de purinas. *Saccharomyces cerevisiae* tiene 5 genes capaces de codificar la enzima que cataliza dicha reacción (*PRS1-PRS5*) (Carter *et al.* 1994; Hernando *et al.* 1998), y con la delección de cualquiera de ellos la actividad enzimática disminuye sobre un 80% y el tiempo de duplicación aumenta de 1.7 h a 2 – 3 h dependiendo de la versión del gen suprimido (Hernando *et al.* 1999), lo que equivale a una reducción de 0.41 [h⁻¹] a 0.35 – 0.23 [h⁻¹], es decir un 14.6 – 43.9% en relación a la tasa de crecimiento.

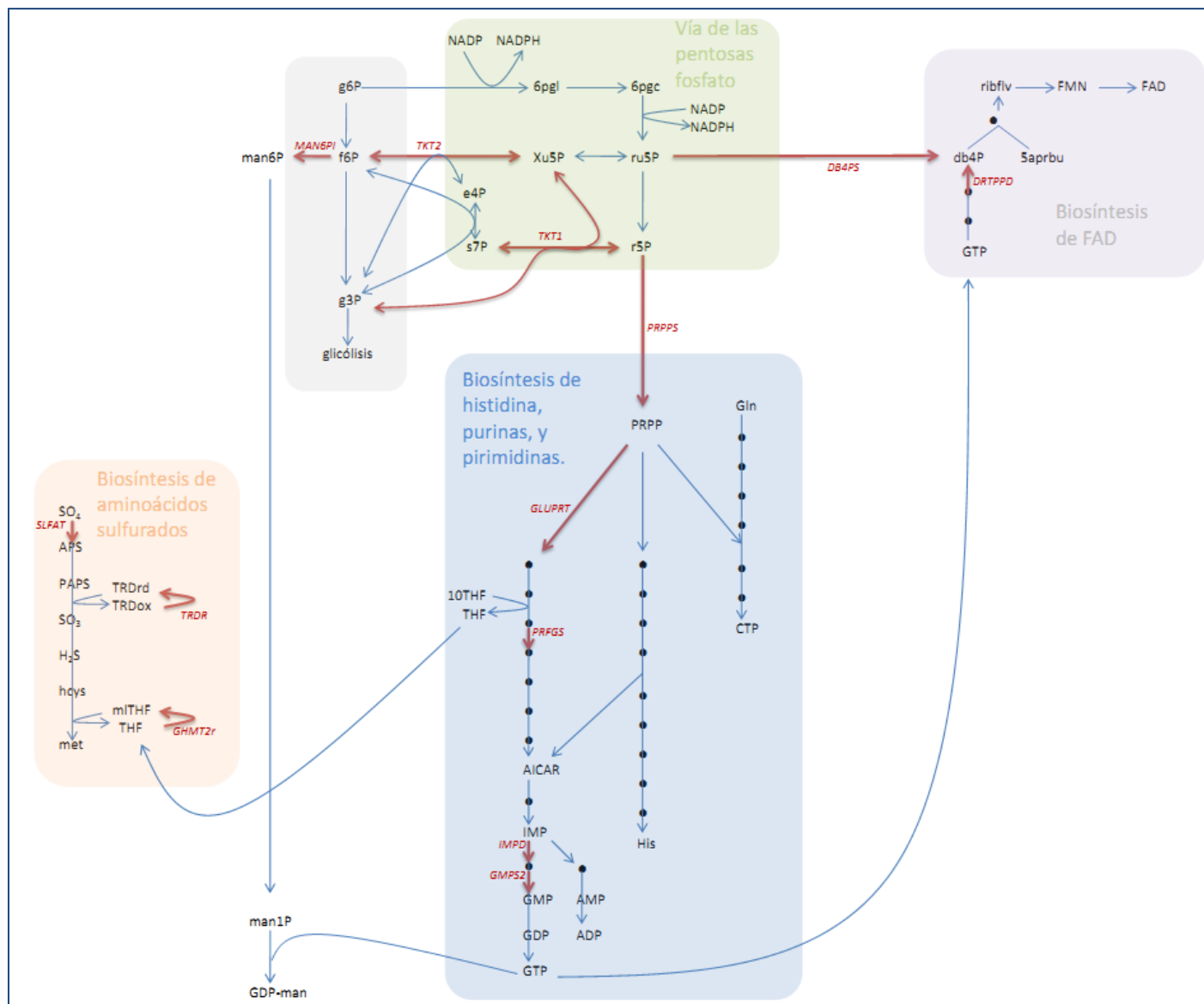


Figura 10: Rutas metabólicas asociadas con la biosíntesis de purinas luego del análisis. Las reacciones que están marcadas en rojo corresponden a las más importantes de aquellas seleccionadas por el análisis de intersección entre los tres algoritmos (Optgene, FSEOF_down, y NP) y por los tres modelos (iND750, iMM904, y yeast4.0).

10THF: '10_Formyltetrahydrofolate'; **5aprbu:** '5_Amino_6__5__phosphoribitylamino_uracil'; **6pgc:** '6_Phospho_D_gluconate'; **6pgl:** '6_phospho_D_glucono_1_5_lactone'; **AICAR:** '5_Amino_1_5_Phospho_D_ribose_1_imidazole_4_carboxamide'; **APS:** 'Adenosine_5_phosphosulfate'; **db4P:** '3_4_dihydroxy_2_butanone_4_phosphate'; **e4P:** 'D_Erythrose_4_phosphate'; **f6p:** 'D_Fructose_6_phosphate'; **FAD:** 'Flavin_adenine_dinucleotide_oxidized'; **FMN:** 'Flavin_mononucleotide'; **g3P:** 'Glyceraldehyde_3_phosphate'; **g6p:** 'D_Glucose_6_phosphate'; **GDP-man:** 'GDP_D_mannose'; **Gln:** 'L_Glutamine'; **H2S:** 'Hydrogen_sulfide'; **hcys:** 'L_Homocysteine'; **His:** 'L_Histidine'; **IMP:** 'inosine-monophosphate'; **man1P:** 'D_Mannose_1_phosphate'; **man6P:** 'D_Mannose_6_phosphate'; **met:** 'L_Methionine'; **mITHF:** '5_10_Methylenetetrahydrofolate'; **PAPS:** '3_Phosphoadenylyl_sulfate'; **PRPP:** '5_Phospho_alpha_D_ribose_1_diphosphate'; **r5P:** 'alpha_D_Ribose_5_phosphate'; **ribflv:** 'Riboflavin'; **ru5P:** 'D_Ribulose_5_phosphate'; **s7P:** 'Sedoheptulose_7_phosphate'; **SO3:** 'Sulfite'; **SO4:** 'Sulfate'; **THF:** '5_6_7_8_Tetrahydrofolate'; **TRDox:** 'Oxidized_thioredoxin'; **TRDRd:** 'Reduced_thioredoxin'; **Xu5P:** 'D_Xylulose_5_phosphate'.

Por lo tanto, para simular dicho escenario se realizaron 3 optimizaciones seguidas (**Tabla 6**). La primera solución (sol1) se obtiene al maximizar biomasa, y con el valor de flujo para la reacción PRPPS se ajusta su límite superior pero con una reducción del 80%. Así se genera una segunda optimización de biomasa, donde su valor de flujo es utilizado como límite inferior en la tercera optimización, para así poder maximizar la biosíntesis de limoneno (LimS). En otras palabras, la tercera solución (sol3) corresponde a la maximización de la biosíntesis de limoneno, pero con el límite inferior de la producción de biomasa ajustado al valor donde el flujo por la PRPPS disminuye un 80% en comparación a su *wild type*.

Con esta metodología se aumentó la síntesis de limoneno, pero la producción de biomasa disminuyó considerablemente. Para los modelos iND750 e iMM904 la reducción en la tasa de crecimiento fue de un 80.4%, mientras que en el modelo Yeast 4.0 fue de 80.0%, lo que es muy distinto de los 14.6 – 43.9% obtenidos experimentalmente.

Tabla 6: Comparación de resultados obtenidos por los distintos modelos en cada una de las optimizaciones. Los flujos para las reacciones PRPPS y LimS están dados en [mmol/gDW·h]; la síntesis de biomasa está en [h^{-1}]; y la productividad BPCY está en [mmol de Lim/mmol de glc·h].

Reacción	iND750				iMM904				Yeast 4.0			
	sol1	sol2	sol3	sol4	sol1	sol2	sol3	sol4	sol1	sol2	sol3	sol4
PRPPS	0.30	0.06	0.06	0.15	0.30	0.06	0.06	0.15	0.57	0.11	0.11	0.29
biomasa	0.97	0.19	0.19	0.49	0.97	0.19	0.19	0.49	1.70	0.34	0.34	0.85
LimS	0	0	2.92	2.01	0	0	2.89	1.91	0	0	3.43	2.14
BPCY	0	0	0.06	0.10	0	0	0.06	0.09	0	0	0.12	0.18

* **sol1**: maximización de biomasa; **sol2**: maximización de biomasa pero con el límite superior de la reacción PRPPS reducido en un 80% del valor obtenido en la sol1; **sol3**: maximización de LimS pero con el límite inferior para la producción de biomasa igual al obtenido en sol2; **sol4**: maximización de biomasa pero con el límite inferior para la reacción LimS ajustado para obtener el máximo BPCY

Por otro lado, como se obtuvieron valores de productividad BPCY menores al máximo, se buscó el porcentaje de reducción necesario para poder alcanzarlo. Es así que se realizó una cuarta optimización, donde se maximizó la producción de biomasa pero con el límite inferior de LimS ajustado a la máxima productividad BPCY. Tal ajuste representa una disminución en la tasa de crecimiento del 50% en los tres modelos, lo que se puede observar gráficamente en la **Figura 8**. De esta

forma se vio que la reducción en el flujo de la PRPPS debe ser cercana al 50% en los tres modelos (**Tabla 6**).

4. DISCUSIÓN

La metodología propuesta en este trabajo combina distintos enfoques y supuestos para identificar una ruta metabólica a modificar para aumentar el rendimiento de la biosíntesis de monoterpenos. Los algoritmos de búsqueda por sí solos no son capaces de identificar un conjunto pequeño de reacciones, pero con la combinación de ellos y junto con los distintos modelos a escala genómica, se puede mejorar el análisis. **De esta forma se ha logrado identificar la ruta del metabolismo de purinas como una vía a reducir sus flujos** para mejorar la producción de monoterpenos, lo que se convierte en la PROPUESTA de este trabajo.

Sin embargo, hay que tener presente que los supuestos detrás de los análisis no son menores. Por ejemplo, se asume que las modificaciones genéticas no afectan el *uptake* de glucosa (que se mantiene constante a 10 mmol/gDW·h). Otro supuesto importante es que todo el flujo de carbono que no ingresa a la vía de las purinas, se redirige a la síntesis de monoterpenos, lo que en condiciones reales no ocurre ya que el excedente de carbono se redistribuye en varias rutas metabólicas. Es por esto que la propuesta planteada debe ser evaluada tomando en cuenta la mayor cantidad de información posible, en donde las bases de datos MetaCyc, brenda-enzymes, y Kegg pueden orientar para tomar la mejor decisión.

Es así que, al estudiar la ruta de las purinas junto a sus vías asociadas (**Figura 10**), se reconoce a la ribosa-5P (r5P) como la única puerta de entrada a dicha vía, y por lo tanto es un NP clave para lograr la propuesta. En tal metabolito se definen tres destinos

posibles al flujo de carbono según las demandas metabólicas. En primer lugar, si existe demanda por ácidos nucleicos, como lo es en un estado de proliferación celular, la enzima **fosforibosilpirofosfato sintetasa** (PRS, EC 2.7.6.1) cataliza la síntesis de PRPP (5-fosforibosil pirofosfato), un precursor de histidina, triptófano, purinas y pirimidinas entre otros compuestos (Hove-Jensen 1988; Murray *et al.* 2003; Lehninger *et al.* 2005). En segundo lugar, si los ácidos nucleicos son suficientes y hay demandas de ATP o de NADPH, la PRS es inhibida y la ribosa-5P es reciclada por la vía no oxidativa de las pentosas fosfato, produciendo gliceraldehído-3P y fructosa-6P los cuales pueden volver a entrar a la vía de las pentosas fosfato o seguir por la glicólisis (Garrett & Grisham 1999; Metzler & Metzler 2003). Las enzimas encargadas de la vía no oxidativa son la **trascetolasa** (EC 2.2.1.1) y la **transaldolasa** (EC 2.2.1.2). En tercer lugar la ribosa-5P puede ser transformada a 3,4-dihidroxi-2-butanona-4P (db4P), un precursor de riboflavina (Richter *et al.* 1992; García-Ramírez *et al.* 1995; Jin *et al.* 2003), gracias a la acción de la **DB4P sintasa** (EC 4.1.99.12).

Sin embargo, a pesar de que la ribosa-5P sea un nodo principal, hay que estudiar las regulaciones que se ejercen sobre él ya que los modelos metabólicos no consideran estos parámetros. Según Stephanopoulos (1991) tales regulaciones definen la rigidez de un nodo, que se puede considerar como las resistencias de la red metabólica a modificar la partición de flujos en un determinado nodo. Por este motivo, dicho parámetro puede ser estudiado únicamente con datos experimentales.

Stephanopoulos (1991) define tres tipos de nodos: los rígidos, los flexibles, y los débilmente rígidos. Un **nodo rígido** es aquel que no modifica su partición de flujos cuando la ruta metabólica es perturbada lo que puede ocurrir debido a que los productos (generalmente finales) de cada rama retroalimentan positivamente la vía paralela, mientras que en la propia retroalimentan negativamente. Por lo tanto, al disminuir el flujo por una rama del nodo que no es de interés, no se activa la rama paralela (la que sí es de interés), por lo que el resultado neto es la disminución de flujo por ambas vías. Por el contrario, los **nodos flexibles** pueden modificar su partición de flujos cuando la vía metabólica es perturbada ya que la retroalimentación negativa de cada rama es inducida sólo por sus propios productos. Finalmente, los **nodos débilmente rígidos** son aquellos donde la partición de flujos es dominada por la cinética de una de las ramas como resultado de la mayor actividad enzimática o afinidad, y por la falta de inhibición en la rama dominante (Stephanopoulos & Vallino 1991).

Bajo tal perspectiva, el control existente en la ribosa-5P para definir su destino metabólico **clasifica a este NP como flexible**. Tal control se basa en la inhibición de la PRS por parte de las purinas y pirimidinas mono-, di-, y trifosfatos, productos finales de la vía que se inicia con esta enzima. Por el contrario, la DB4P sintasa no presenta inhibición, y las enzimas de la vía no oxidativa sólo tienen inhibición por sulfato (base de datos: MetaCyc.com; brenda-enzymes.org), favoreciendo la fijación de tal compuesto en metionina ya que el proceso requiere del NADPH producido en la vía oxidativa.

Por lo tanto, como **primera estrategia** para lograr el objetivo de la propuesta, es la delección de una versión de la fosforibosilpirofosfato sintetasa (*PRS1-PRS5*). El efecto de esto es la reducción de la actividad enzimática sobre un 80% (Hernando *et al.* 1999), y debido a que la ribosa-5P es un nodo flexible, los flujos metabólicos se redistribuirían, generando un excedente de carbono que puede ser utilizado, entre otras cosas, para la biosíntesis de monoterpenos.

Sin embargo existen dos inconvenientes al querer desarrollar una cepa con esta estrategia. La primera es que para lograr el máximo de productividad BPCY según la **Tabla 6**, la reducción del flujo debe ser sólo del 50%, lo que no se puede saber *a priori* con una reducción del 80% de la actividad enzimática. En segundo lugar, se ha visto que la delección de algún gen *PRS* genera daño en la pared celular y mayor sensibilidad térmica (Wang *et al.* 2004; Vavassori *et al.* 2005), lo que puede ser perjudicial para una cepa de levadura productora de un antifúngico como limoneno.

Considerando lo anterior, una **segunda estrategia** para disminuir el flujo por la vía de las purinas es la regulación fina de la expresión del promotor del gen *PRS* (Nevoigt *et al.* 2006), de tal forma que exista un óptimo de producción en función de la expresión. Una tercera estrategia, que puede ser complementaria a la anterior, es la utilización de un promotor inducible (Shibasaki *et al.* 2003; Shin *et al.* 2010; Matsuyama *et al.* 2011) tal que luego de producir la biomasa suficiente se anule la inducción del gen *PRS* y se active la síntesis de limoneno.

Por otro lado, dejando de lado la propuesta y tomando en cuenta el resto de la información entregada por el análisis integrado, se observa que la segunda mayor reducción de flujo se localiza en la reacción MAN6PI (**Tabla 4** y **Tabla 5**). Dicha reacción sintetiza manosa-6P a partir de fructosa-6P y la enzima encargada es la fosfomanosa isomerasa (*PMI40*, EC 5.3.1.8). Su delección resulta en una cepa auxótrofa para manosa (Payton *et al.* 1991), por lo que no resulta ser un buen blanco génico. La auxotrofía se explica por la incapacidad de la levadura para sintetizar GDP-manoza, un precursor en la glicosilación de las proteínas de la pared celular (Foret *et al.* 2009), y que se obtiene luego de dos reacciones a partir de manosa-6P (**Figura 10**) que son catalizadas por la fosfomanomutasa (*SEC53*, EC 5.4.2.8) y la manosa-1P guaniltransferasa (*PSA1*, EC 2.7.7.13). Un sustrato de la última enzima es el GTP, producto final de la vía de biosíntesis de las purinas, y como se mencionó anteriormente, la síntesis de éste metabolito puede ser reducida al deleccionar cualquiera de las 5 versiones del gen *PRS*. Por lo tanto, el daño en la pared celular producido por la delección de tal gen se puede explicar por la menor disponibilidad de GTP para la síntesis de glicoproteínas.

Otra vía cuyos flujos se ven reducidos es la biosíntesis de riboflavina, un precursor del cofactor FAD que es utilizado en una amplia variedad de reacciones redox (Burgess *et al.* 2009). La ruta de síntesis de la riboflavina se origina a partir de GTP y de ribulosa-5P (**Figura 10**), y al igual que en el caso anterior, debido a que el GTP es su precursor, la delección de uno de los genes *PRS* también puede reducir la síntesis de riboflavina

(Jiménez *et al.* 2008). Por lo tanto, esto reafirma la propuesta de deleccionar tal gen para disminuir el flujo en la vía de síntesis de purinas.

Finalmente una segunda propuesta podría considerarse a partir de la **Figura 7**, la cual consiste en forzar los flujos a través del metabolismo de aminoácidos. Sin embargo, debido a la existencia de varias ramificaciones en dicha ruta, resulta poco probable que llegue una cantidad apreciable al acetaldehído. Los detalles completos de la ruta se encuentran en el **ANEXO E**, *Citrate cycle (TCA cycle), Alanine, aspartate and glutamate metabolism*, y *Glycine, serine and threonine metabolism*.

BIBLIOGRAFIA

Aharoni A, AP Giri, S Deuerlein, F Griepink, WJ de Kogel, FWA Verstappen, HA Verhoeven, MA Jongsma, W Schwab & HJ Bouwmeester (2003): Terpenoid metabolism in wild-type and transgenic Arabidopsis plants. *The Plant Cell Online* **15**: 2866.

Aharoni A, MA Jongsma & HJ Bouwmeester (2005): Volatile science? Metabolic engineering of terpenoids in plants. *Trends in plant science* **10**: 594-602.

Ahn YJ, SB Lee, HS Lee & GH Kim (1998): Insecticidal and acaricidal activity of carvacrol and β -thujaplicine derived from *Thujopsis dolabrata* var. *hondai* sawdust. *Journal of Chemical Ecology* **24**: 81-90.

Ajikumar PK, WH Xiao, KEJ Tyo, Y Wang, F Simeon, E Leonard, O Mucha, TH Phon, B Pfeifer & G Stephanopoulos (2010): Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science* **330**: 70.

Alper H, YS Jin, J Moxley & G Stephanopoulos (2005): Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metabolic Engineering* **7**: 155-164.

Alper H, K Miyaoku & G Stephanopoulos (2005): Construction of lycopene-overproducing *E. coli* strains by combining systematic and combinatorial gene knockout targets. *Nature biotechnology* **23**: 612-616.

Alper H, K Miyaoku & G Stephanopoulos (2006): Characterization of lycopene-overproducing *E. coli* strains in high cell density fermentations. *Applied microbiology and biotechnology* **72**: 968-974.

Alper H & G Stephanopoulos (2009): Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? *Nature Reviews Microbiology* **7**: 715-723.

Asadollahi MA, J Maury, KR Patil, M Schalk, A Clark & J Nielsen (2009): Enhancing sesquiterpene production in *Saccharomyces cerevisiae* through in silico driven metabolic engineering. *Metabolic Engineering* **11**: 328-334.

Basson ME, M Thorsness & J Rine (1986): *Saccharomyces cerevisiae* contains two functional genes encoding 3-hydroxy-3-methylglutaryl-coenzyme A reductase. *Proceedings of the National Academy of Sciences* **83**: 5563.

Boghigian B, H Shi, K Lee & B Pfeifer (2010): Utilizing elementary mode analysis, pathway thermodynamics, and a genetic algorithm for metabolic flux determination and optimal metabolic network design. *BMC systems biology* **4**: 49.

Boghigian BA, G Seth, R Kiss & BA Pfeifer (2010): Metabolic flux analysis and pharmaceutical production. *Metabolic Engineering* **12**: 81-95.

Breitmaier E (2006): Terpenes: flavors, fragrances, pharmaca, pheromones. Wiley-VCH.

Bro C, B Regenberg, J Förster & J Nielsen (2006): In silico aided metabolic engineering of *Saccharomyces cerevisiae* for improved bioethanol production. *Metabolic Engineering* **8**: 102-111.

Brochado AR, C Matos, BL Moller, J Hansen, UH Mortensen & KR Patil (2010): Improved vanillin production in baker's yeast through in silico design. *Microbial Cell Factories* **9**: 84.

Bui-Khac T & N Dupuis (2002): Process for extraction and purification of paclitaxel from natural sources Book Process for extraction and purification of paclitaxel from natural sources. City. Google Patents.

Burgard AP, P Pharkya & CD Maranas (2003): Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnology and bioengineering* **84**: 647-657.

Burgess CM, EJ Smid & D Van Sinderen (2009): Bacterial vitamin B2, B11 and B12 overproduction: An overview. *International journal of food microbiology* **133**: 1-7.

Capell T & P Christou (2004): Progress in plant metabolic engineering. *Current Opinion in Biotechnology* **15**: 148-154.

Carter AT, A Narbad, BM Pearson, KF Beck, B Baum, M Logghe, R Contreras & M Schweizer (1994): Phosphoribosylpyrophosphate synthetase (PRS): a new gene family in *Saccharomyces cerevisiae*. *Yeast* **10**: 1031-1044.

Causey T, K Shanmugam, L Yomano & L Ingram (2004): Engineering *Escherichia coli* for efficient conversion of glucose to pyruvate. *Proceedings of the National Academy of Sciences* **101**: 2235.

Causey T, S Zhou, K Shanmugam & L Ingram (2003): Engineering the metabolism of *Escherichia coli* W3110 for the conversion of sugar to redox-neutral and oxidized products: homoacetate production. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 825.

Chang M & J Keasling (2006): Production of isoprenoid pharmaceuticals by engineered microbes. *Nature chemical biology* **2**: 674-681.

Chemler JA & MAG Koffas (2008): Metabolic engineering for plant natural product biosynthesis in microbes. *Current Opinion in Biotechnology* **19**: 597-605.

Choi HS, SY Lee, TY Kim & HM Woo (2010): In silico identification of gene amplification targets for improvement of lycopene production. *Applied and environmental microbiology* **76**: 3097.

Choi WS, BS Park, YH Lee, DY Jang, HY Yoon & SE Lee (2006): Fumigant toxicities of essential oils and monoterpenes against *Lycoriella mali* adults. *Crop Protection* **25**: 398-401.

Croteau R, TM Kutchan & NG Lewis (2000): Natural products (secondary metabolites). *Biochemistry and molecular biology of plants*: 1250-1318.

Crowell PL & MN Gould (1994): Chemoprevention and therapy of cancer by d-limonene. *Critical reviews in oncogenesis* **5**: 1.

Daviet L & M Schalk (2010): Biotechnology in plant essential oil production: progress and perspective in metabolic engineering of the terpene pathway. *Flavour and Fragrance Journal* **25**: 123-127.

de Carvalho CCCR & MMR da Fonseca (2006): Carvone: Why and how should one bother to produce this terpene. *Food Chemistry* **95**: 413-422.

Dieras V (1998): Taxanes in combination with doxorubicin in the treatment of metastatic breast cancer Book Taxanes in combination with doxorubicin in the treatment of metastatic breast cancer. City: 18.

Dobson PD, K Smallbone, D Jameson, E Simeonidis, K Lanthaler, P Pir, CA Lu, N Swainston, WB Dunn, P Fisher, D Hull, M Brown, O Oshota, NJ Stanford, DB Kell, RD King, SG Oliver, RD Stevens & P Mendes (2010): Further developments towards a genome-scale metabolic model of yeast. *Bmc Systems Biology* **4**.

Donald K, RY Hampton & IB Fritz (1997): Effects of overproduction of the catalytic domain of 3-hydroxy-3-methylglutaryl coenzyme A reductase on squalene synthesis in *Saccharomyces cerevisiae*. *Applied and environmental microbiology* **63**: 3341-3344.

Duarte NC, MJ Herrgard & BO Palsson (2004): Reconstruction and validation of *Saccharomyces cerevisiae* iND750, a fully compartmentalized genome-scale metabolic model. *Genome Research* **14**: 1298-1309.

Edwards JS, RU Ibarra & BO Palsson (2001): In silico predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nature biotechnology* **19**.

Famili I, J Förster, J Nielsen & BO Palsson (2003): *Saccharomyces cerevisiae* phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 13134.

Foret J, B de Courcy, N Gresh, JP Piquemal & L Salmon (2009): Synthesis and evaluation of non-hydrolyzable d-mannose 6-phosphate surrogates reveal 6-deoxy-6-dicarboxymethyl-d-mannose as a new strong inhibitor of phosphomannose isomerases. *Bioorganic & medicinal chemistry* **17**: 7100-7107.

Förster J, I Famili, P Fu, BÅ Palsson & J Nielsen (2003): Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome research* **13**: 244.

García-Ramírez JJ, MA Santos & JL Revuelta (1995): The *Saccharomyces cerevisiae* RIB4 gene codes for 6, 7-dimethyl-8-ribityllumazine synthase involved in riboflavin biosynthesis. *Journal of Biological Chemistry* **270**: 23801.

Gardner RG & RY Hampton (1999): A highly conserved signal controls degradation of 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase in eukaryotes. *Journal of Biological Chemistry* **274**: 31671.

Garrett R & C Grisham (1999): *Biochemistry. Second Book Biochemistry*. Second. City. Fort Worth: Saunders College Publishing.

Grotewold E (2008): Transcription factors for predictive plant metabolic engineering: are we there yet? *Current Opinion in Biotechnology* **19**: 138-144.

Harker M & P Bramley (1999): Expression of prokaryotic 1-deoxy-xylulose-5-phosphatases in *Escherichia coli* increases carotenoid and ubiquinone biosynthesis. *FEBS letters* **448**: 115-119.

Hernando Y, AT Carter, A Parr, B Hove-Jensen & M Schweizer (1999): Genetic analysis and enzyme activity suggest the existence of more than one minimal functional unit capable of synthesizing phosphoribosyl pyrophosphate in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* **274**: 12480.

Hernando Y, A Parr & M Schweizer (1998): PRS5, the fifth member of the phosphoribosyl pyrophosphate synthetase gene family in *Saccharomyces cerevisiae*, is essential for cell viability in the absence of either PRS1 or PRS3. *Journal of bacteriology* **180**: 6404.

Hove-Jensen B (1988): Mutation in the phosphoribosylpyrophosphate synthetase gene (*prs*) that results in simultaneous requirements for purine and pyrimidine nucleosides, nicotinamide nucleotide, histidine, and tryptophan in *Escherichia coli*. *Journal of bacteriology* **170**: 1148.

Jackson BE, EA Hart-Wells & SPT Matsuda (2003): Metabolic engineering to produce sesquiterpenes in yeast. *Organic letters* **5**: 1629-1632.

- Jamshidi N & B Palsson (2007): Investigating the metabolic capabilities of *Mycobacterium tuberculosis* H37Rv using the in silico strain iNJ661 and proposing alternative drug targets. *BMC systems biology* **1**: 26.
- Jarboe L, T Grabar, L Yomano, K Shanmugan & L Ingram (2007): Development of ethanologenic bacteria. *Biofuels*: 237-261.
- Jiménez A, M Santos & J Revuelta (2008): Phosphoribosyl pyrophosphate synthetase activity affects growth and riboflavin production in *Ashbya gossypii*. *BMC biotechnology* **8**: 67.
- Jin C, A Barrientos & A Tzagoloff (2003): Yeast dihydroxybutanone phosphate synthase, an enzyme of the riboflavin biosynthetic pathway, has a second unrelated function in expression of mitochondrial respiration. *Journal of Biological Chemistry* **278**: 14698.
- Jin YS & G Stephanopoulos (2007): Multi-dimensional gene target search for improving lycopene biosynthesis in *Escherichia coli*. *Metabolic Engineering* **9**: 337-347.
- Kang MJ, YM Lee, SH Yoon, JH Kim, SW Ock, KH Jung, YC Shin, JD Keasling & SW Kim (2005): Identification of genes affecting lycopene accumulation in *Escherichia coli* using a shot-gun method. *Biotechnology and bioengineering* **91**: 636-642.
- Karunajeewa HA, L Manning, I Mueller, KF Ilett & TME Davis (2007): Rectal administration of artemisinin derivatives for the treatment of malaria. *JAMA: the journal of the American Medical Association* **297**: 2381.
- Kauffman KJ, P Prakash & JS Edwards (2003): Advances in flux balance analysis. *Current Opinion in Biotechnology* **14**: 491-496.
- Kawamura F, Y Kikuchi, T Ohira & M Yatagai (1999): Accelerated solvent extraction of paclitaxel and related compounds from the bark of *Taxus cuspidata*. *Journal of natural products* **62**: 244-247.
- Kemp LE, CS Bond & WN Hunter (2002): Structure of 2C-methyl-D-erythritol 2, 4-cyclodiphosphate synthase: an essential enzyme for isoprenoid biosynthesis and target for antimicrobial drug development. *Proceedings of the National Academy of Sciences* **99**: 6591.
- Kessler A & IT Baldwin (2001): Defensive function of herbivore-induced plant volatile emissions in nature. *Science* **291**: 2141.
- Kim SW & J Keasling (2001): Metabolic engineering of the nonmevalonate isopentenyl diphosphate synthesis pathway in *Escherichia coli* enhances lycopene production. *Biotechnology and bioengineering* **72**: 408-415.

- Kirby J & JD Keasling (2009): Biosynthesis of plant isoprenoids: perspectives for microbial engineering. *Annual review of plant biology* **60**: 335-355.
- Kuepfer L, U Sauer & LM Blank (2005): Metabolic functions of duplicate genes in *Saccharomyces cerevisiae*. *Genome research* **15**: 1421.
- Langenheim JH (1994): Higher plant terpenoids: a phytocentric overview of their ecological roles. *Journal of Chemical Ecology* **20**: 1223-1280.
- Lapkin AA, PK Plucinski & M Cutler (2006): Comparative assessment of technologies for extraction of artemisinin. *Journal of natural products* **69**: 1653-1664.
- Lee JM, EP Gianchandani & JA Papin (2006): Flux balance analysis in the era of metabolomics. *Briefings in bioinformatics* **7**: 140-150.
- Lee S, C Phalakornkule, MM Domach & IE Grossmann (2000): Recursive MILP model for finding all the alternate optima in LP models for metabolic networks. *Computers & Chemical Engineering* **24**: 711-716.
- Lee S, R Tsao, C Peterson & JR Coats (1997): Insecticidal activity of monoterpenoids to western corn rootworm (Coleoptera: Chrysomelidae), twospotted spider mite (Acari: Tetranychidae), and house fly (Diptera: Muscidae). *Journal of economic entomology* **90**: 883-892.
- Lee SY, HM Woo, DY Lee, HS Choi, TY Kim & H Yun (2005): Systems-level analysis of genome-scale in silico metabolic models using MetaFluxNet. *Biotechnology and Bioprocess Engineering* **10**: 425-431.
- Lehninger AL, DL Nelson & MM Cox (2005): *Lehninger principles of biochemistry*. Wh Freeman.
- Leonard E, Y Yan, Z Fowler, Z Li, C Lim, K Lim & M Koffas (2008): Strain improvement of recombinant *Escherichia coli* for efficient production of plant flavonoids. *Molecular Pharmaceutics* **5**: 257-265.
- Lewinsohn E, F Schalechet, J Wilkinson, K Matsui, Y Tadmor, KH Nam, O Amar, E Lastochkin, O Larkov & U Ravid (2001): Enhanced levels of the aroma and flavor compound S-linalool by metabolic engineering of the terpenoid pathway in tomato fruits. *Plant physiology* **127**: 1256-1265.
- Lucker J, HJ Bouwmeester, W Schwab, J Blaas, LHW Van Der Plas & HA Verhoeven (2001): Expression of *Clarkia* S-linalool synthase in transgenic petunia plants results in the accumulation of S-linalyl- β -D-glucopyranoside. *The Plant Journal* **27**: 315-324.

Lucker J, W Schwab, B Van Hautum, J Blaas, L van der Plas, H Bouwmeester & H Verhoeven (2004): Increased and altered fragrance of tobacco plants after metabolic engineering using three monoterpene synthases from lemon. *Plant physiology* **134**: 510.

Mahadevan R & CH Schilling (2003): The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metabolic Engineering* **5**: 264-276.

Mahmoud SS & RB Croteau (2001): Metabolic engineering of essential oil yield and composition in mint by altering expression of deoxyxylulose phosphate reductoisomerase and menthofuran synthase. *Proceedings of the National Academy of Sciences* **98**: 8915.

Martin DM, J F  ldt & J Bohlmann (2004): Functional characterization of nine Norway spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. *Plant physiology* **135**: 1908-1927.

Martin V, D Pitera, S Withers, J Newman & J Keasling (2003): Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nature biotechnology* **21**: 796-802.

Matsuyama T, M Yamanishi & H Takahashi (2011): Improvement of galactose induction system in *Saccharomyces cerevisiae*. *Journal of bioscience and bioengineering*.

Maury J, M Asadollahi, K M  ller, A Clark & J Nielsen (2005): Microbial isoprenoid production: an example of green chemistry through metabolic engineering. *Biotechnology for the Future*: 864-866.

Maury J, MA Asadollahi, K M  ller, M Schalk, A Clark, LR Formenti & J Nielsen (2008): Reconstruction of a bacterial isoprenoid biosynthetic pathway in *Saccharomyces cerevisiae*. *FEBS letters* **582**: 4032-4038.

McCaskill D & R Croteau (1997): Prospects for the bioengineering of isoprenoid biosynthesis. *Biotechnology of Aroma Compounds*: 107-146.

Metzler DE & CM Metzler (2003): *Biochemistry: the chemical reactions of living cells*. Academic Pr.

Mo ML, BO Palsson & MJ Herrgard (2009): Connecting extracellular metabolomic measurements to intracellular flux states in yeast. *Bmc Systems Biology* **3**.

Morrone D, L Lowry, MK Determan, DM Hershey, M Xu & RJ Peters (2010): Increasing diterpene yield with a modular metabolic engineering system in *E. coli*: comparison of MEV and MEP isoprenoid precursor pathway engineering. *Applied microbiology and biotechnology* **85**: 1893-1906.

Murray R, D Granner, P Mayes & V Rodwell (2003): *Harper's Illustrated Biochemistry*. McGraw-Hill.

- Nevoigt E, J Kohnke, CR Fischer, H Alper, U Stahl & G Stephanopoulos (2006): Engineering of promoter replacement cassettes for fine-tuning of gene expression in *Saccharomyces cerevisiae*. *Applied and environmental microbiology* **72**: 5266.
- Newman JD, J Marshall, M Chang, F Nowroozi, E Paradise, D Pitera, KL Newman & JD Keasling (2006): High-level production of amorpha-4,11diene in a two-phase partitioning bioreactor of metabolically engineered *Escherichia coli*. *Biotechnology and bioengineering* **95**: 684-691.
- Nookaew I, M Jewett, A Meechai, C Thammamongtham, K Laoteng, S Cheevadhanarak, J Nielsen & S Bhumiratana (2008): The genome-scale metabolic model iIN800 of *Saccharomyces cerevisiae* and its validation: a scaffold to query lipid metabolism. *BMC systems biology* **2**: 71.
- Orth JD, I Thiele & BÅ Palsson (2010): What is flux balance analysis? *Nature biotechnology* **28**: 245-248.
- Otero JM & J Nielsen (2010): Industrial systems biology. *Biotechnology and bioengineering* **105**: 439-460.
- Park JH, KH Lee, TY Kim & SY Lee (2007): Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. *Proceedings of the National Academy of Sciences* **104**: 7797.
- Patil K, I Rocha, J Förster & J Nielsen (2005): Evolutionary programming as a platform for in silico metabolic engineering. *BMC bioinformatics* **6**: 308.
- Payton M, M Rheinnecker, L Klig, M DeTiani & E Bowden (1991): A novel *Saccharomyces cerevisiae* secretory mutant possesses a thermolabile phosphomannose isomerase. *Journal of bacteriology* **173**: 2006.
- Petersen M (2007): Current status of metabolic phytochemistry. *Phytochemistry* **68**: 2847-2860.
- Pharkya P, AP Burgard & CD Maranas (2004): OptStrain: a computational framework for redesign of microbial production systems. *Genome research* **14**: 2367-2376.
- Pichersky E & J Gershenzon (2002): The formation and function of plant volatiles: perfumes for pollinator attraction and defense. *Current opinion in plant biology* **5**: 237-243.
- Pramanik J & J Keasling (1997): Stoichiometric model of *Escherichia coli* metabolism: Incorporation of growth-rate dependent biomass composition and mechanistic energy requirements. *Biotechnology and bioengineering* **56**: 398-421.

Raab R, K Tyo & G Stephanopoulos (2005): Metabolic engineering. *Biotechnology for the Future*: 1-17.

Ramakrishna R, JS Edwards, A McCulloch & BO Palsson (2001): Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* **280**: R695-R704.

Raman K & N Chandra (2009): Flux balance analysis of biological systems: applications and challenges. *Briefings in bioinformatics* **10**: 435-449.

Reddy BS, C Wang, H Samaha, R Lubet, VE Steele, GJ Kelloff & CV Rao (1997): Chemoprevention of colon carcinogenesis by dietary perillyl alcohol. *Cancer research* **57**: 420.

Reed JL & BÅ Palsson (2004): Genome-scale in silico models of E. coli have multiple equivalent phenotypic states: assessment of correlated reaction subsets that comprise network states. *Genome research* **14**: 1797-1805.

Reed JL, TD Vo, CH Schilling & BO Palsson (2003): An expanded genome-scale model of Escherichia coli K-12 (iJR904 GSM/GPR). *Genome Biol* **4**: R54.

Richter G, R Volk, C Krieger, H Lahm, U Rothlisberger & A Bacher (1992): Biosynthesis of riboflavin: cloning, sequencing, and expression of the gene coding for 3, 4-dihydroxy-2-butanone 4-phosphate synthase of Escherichia coli. *Journal of bacteriology* **174**: 4050.

Rigoutsos I & G Stephanopoulos (2007): *Systems Biology: Genomics*. Oxford University Press, USA.

Ro D, E Paradise, M Ouellet, K Fisher, K Newman, J Ndungu, K Ho, R Eachus, T Ham, J Kirby, M Chang, S Withers, Y Schiba, R Sarpong & J Keasling (2006): Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**: 940-943.

Rohdich F, F Zepeck, P Adam, S Hecht, J Kaiser, R Laupitz, T GrÅwert, S Amslinger, W Eisenreich & A Bacher (2003): The deoxyxylulose phosphate pathway of isoprenoid biosynthesis: studies on the mechanisms of the reactions catalyzed by IspG and IspH protein. *Proceedings of the National Academy of Sciences* **100**: 1586.

Segre D, D Vitkup & GM Church (2002): Analysis of optimality in natural and perturbed metabolic networks. *Proceedings of the National Academy of Sciences* **99**: 15112.

Shi W & MN Gould (2002): Induction of cytostasis in mammary carcinoma cells treated with the anticancer agent perillyl alcohol. *Carcinogenesis* **23**: 131.

Shibasaki S, A Tanaka & M Ueda (2003): Development of combinatorial bioengineering using yeast cell surface display--order-made design of cell and protein for bio-monitoring. *Biosensors and Bioelectronics* **19**: 123-130.

Shimada H, K Kondo, PD Fraser, Y Miura, T Saito & N Misawa (1998): Increased carotenoid production by the food yeast *Candida utilis* through metabolic engineering of the isoprenoid pathway. *Applied and environmental microbiology* **64**: 2676.

Shin SY, NS Han, YC Park, MD Kim & JH Seo (2010): Production of resveratrol from p-coumaric acid in recombinant *Saccharomyces cerevisiae* expressing 4-coumarate: coenzyme A ligase and stilbene synthase genes. *Enzyme and Microbial Technology*.

Shlomi T, O Berkman & E Ruppin (2005): Regulatory on/off minimization of metabolic flux changes after genetic perturbations. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 7695.

Song H, TY Kim, BK Choi, SJ Choi, LK Nielsen, HN Chang & SY Lee (2008): Development of chemically defined medium for *Mannheimia succiniciproducens* based on its genome sequence. *Applied microbiology and biotechnology* **79**: 263-272.

Steinbacher S, J Kaiser, J Wungsintaweeikul, S Hecht, W Eisenreich, S Gerhardt, A Bacher & F Rohdich (2002): Structure of 2C-methyl--erythritol-2, 4-cyclodiphosphate synthase involved in mevalonate-independent biosynthesis of isoprenoids¹. *Journal of molecular biology* **316**: 79-88.

Stephanopoulos G, H Alper & J Moxley (2004): Exploiting biological complexity for strain improvement through systems biology. *Nature biotechnology* **22**: 1261-1267.

Stephanopoulos G & A Sinskey (1993): Metabolic engineering--methodologies and future prospects. *Trends in biotechnology* **11**: 392-396.

Stephanopoulos G & JJ Vallino (1991): Network rigidity and metabolic engineering in metabolite overproduction. *Science* **252**: 1675.

Terzer M, ND Maynard, MW Covert & J Stelling (2009): Genome-scale metabolic networks. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* **1**: 285-297.

Trinh C, P Unrean & F Srienc (2008): Minimal *Escherichia coli* cell for the most efficient production of ethanol from hexoses and pentoses. *Applied and environmental microbiology* **74**: 3634.

Tsuruta H, CJ Paddon, D Eng, JR Lenihan, T Horning, LC Anthony, R Regentin, JD Keasling, NS Renninger & JD Newman (2009): High-level production of amorpha-4, 11-diene, a precursor of the antimalarial agent artemisinin, in *Escherichia coli*. *PLoS One* **4**: e4489.

Tyo KE, HS Alper & GN Stephanopoulos (2007): Expanding the metabolic engineering toolbox: more options to engineer cells. *Trends in biotechnology* **25**: 132-137.

Varma A, BW Boesch & BO Palsson (1993): Biochemical production capabilities of *Escherichia coli*. *Biotechnology and bioengineering* **42**: 59-73.

Varma A & BO Palsson (1994): Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Applied and environmental microbiology* **60**: 3724.

Vavassori S, K Wang, L Schweizer & M Schweizer (2005): Ramifications of impaired PRPP synthesis in *Saccharomyces cerevisiae*. *Biochemical Society Transactions* **33**: 1418-1420.

Wang K, S Vavassori, LM Schweizer & M Schweizer (2004): Impaired PRPP-synthesizing capacity compromises cell integrity signalling in *Saccharomyces cerevisiae*. *Microbiology* **150**: 3327.

Wang Y, S Chen & O Yu (2011): Metabolic engineering of flavonoids in plants and microorganisms. *Applied microbiology and biotechnology*: 1-8.

Withers ST & JD Keasling (2007): Biosynthesis and engineering of isoprenoid small molecules. *Applied microbiology and biotechnology* **73**: 980-990.

Wu S, M Schalk, A Clark, RB Miles, R Coates & J Chappell (2006): Redirection of cytosolic or plastidic isoprenoid precursors elevates terpene production in plants. *Nature biotechnology* **24**: 1441-1447.

Yazaki K (2005): Transporters of secondary metabolites. *Current opinion in plant biology* **8**: 301-307.

Ye X, S Al-Babili, A Klöti, J Zhang, P Lucca, P Beyer & I Potrykus (2000): Engineering the provitamin A (-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science* **287**: 303.

Yuan LZ, PE Rouvière, RA LaRossa & W Suh (2006): Chromosomal promoter replacement of the isoprenoid pathway for enhancing carotenoid production in *E. coli*. *Metabolic Engineering* **8**: 79-90.

Zu Y, Y Wang, Y Fu, S Li, R Sun, W Liu & H Luo (2009): Enzyme-assisted extraction of paclitaxel and related taxanes from needles of *Taxus chinensis*. *Separation and Purification Technology* **68**: 238-243.

ANEXOS

ANEXO A: USO DEL TOOLBOX COBRA

a) Adición de la reacción de síntesis de limoneno a cada modelo:

Luego de descargar cada modelo metabólico se utilizó el comando `addReaction` para adicionar las reacciones de síntesis y de exportación de limoneno. A continuación se muestra un ejemplo del código, utilizando el modelo iND750.

```
load('iND750_Bigg_SBML_export');
newmodel = addReaction(model, 'LimS', 'grdp[c] -> lim[c] + ppi[c]',...
                        [],true,0,1000,0,'Monoterpene synthesis');

newmodel = addReaction(newmodel, 'Lim_trs', 'lim[c] -> lim[e]',...
                        [],true,0,1000,0,'Monoterpene synthesis');

newmodel = addReaction(newmodel, 'Lim_exp', 'lim[e] -> ',...
                        [],true,0,1000,0,'Monoterpene synthesis');
```

b) OptGene:

Para ejecutar el algoritmo se utilizó el comando `optGene`, según se muestra a continuación.

```
[x, population, scores, optGeneSol, bst] = optGene(model, targetRxns, rx,
rxns, del);
```

Donde **x** es el individuo seleccionado por tener el mayor valor de *fitness*, y corresponde a un vector binario de *ceros* y *unos* de $1 \times g$ (con g igual al número de genes), donde cada *uno* representa al gen a ser deletado; **population** es una matriz binaria de *ceros* y *unos* de $500 \times g$, que representa la población a la cual pertenece el individuo x ; **scores** es un vector de 500×1 que almacena el valor de *fitness* para cada individuo de *population*; **optGeneSol** es un “*structure*” con un resumen del individuo x (*uptakes* de sustratos, la reacción a maximizar, una lista de los genes deletados, el número total de genes deletados, una lista de reacciones afectadas, el valor de *fitness*, *population*, y *scores*); y finalmente, **bst** es un *output* añadido al *script* que entrega los mejores individuos evaluados desde el comienzo de la rutina hasta el final (ver texto principal).

En relación a los *inputs* del algoritmo, **model** es el modelo metabólico; **targetRxns** es la reacción a maximizar su producción (producción de limoneno); **rx** son las reacciones de uptakes; rxns son las reacciones que serán evaluadas para delecionar (todas); y **del** es el número máximo de delecciones evaluadas, que en este caso fue de 4.

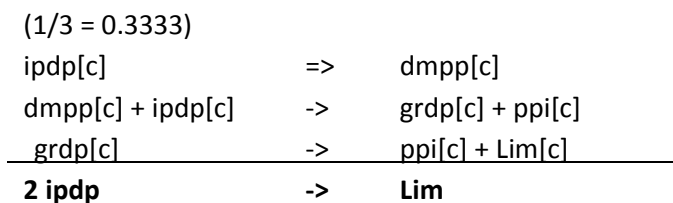
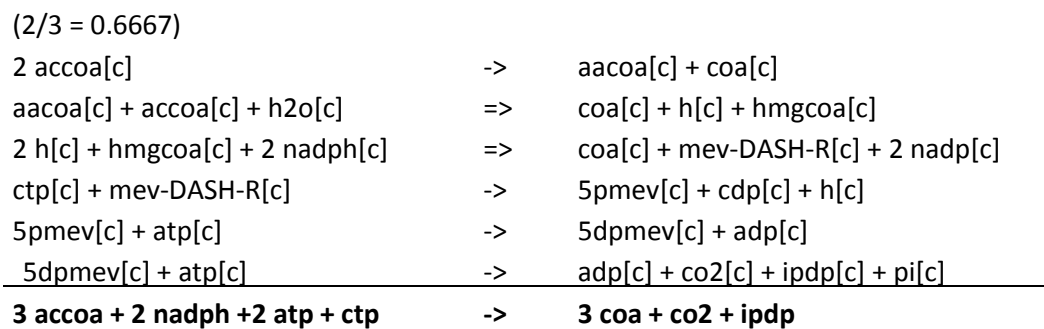
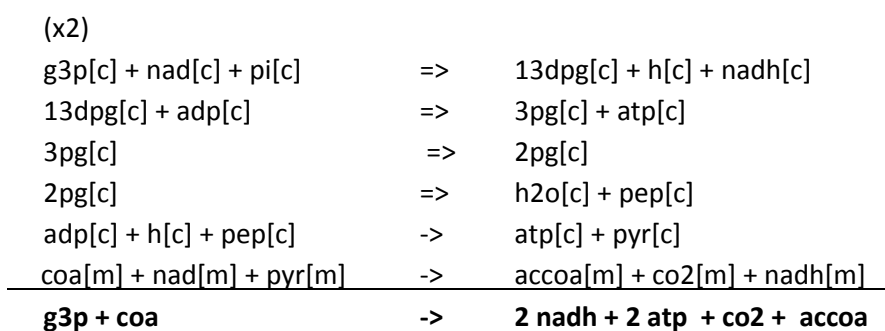
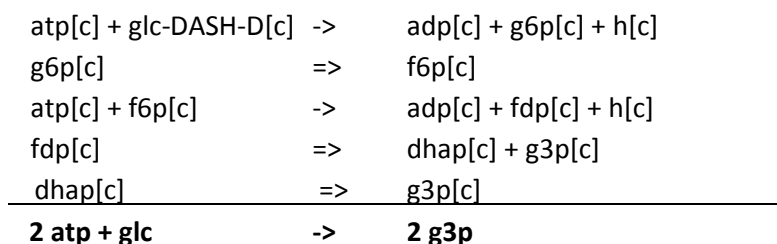
c) Análisis de variabilidad:

El análisis de variabilidad se ejecutó para cada uno de los *steps* con el comando fluxVariability como se muestra a continuación.

```
[minFlux,maxFlux] = fluxVariability(model,percent,'max',rxAbb);
```

Donde minFlux y maxFlux corresponden a vectores con los valores mínimos y máximos para las reacciones ingresadas en rxAbb; percent corresponde a un porcentaje del valor de la función objetivo, el cual se fijó por defecto a 100%; y 'max'corresponde al sentido de la optimización, es decir, maximizar biomasa.

ANEXO B: MÁXIMOS TEÓRICOS



TOTAL

(x1)	2 atp + glc	->	2 adp + 2 g3p
(x2)	g3p + coa	->	2 nadh + 2 atp + co2 + accoa
(x2/3)	3 accoa + 2 nadph + 2 atp + ctp	->	3 coa + co2 + ipdp
(x1/3)	2 ipdp	->	Lim

Suponiendo ctp = atp

	2 atp + glc	->	2 g3p
	2 g3p + 2 coa	->	4 nadh + 4 atp + 2 co2 + 2 accoa
2 accoa + 4/3 nadph + 2 atp		->	2 coa + 2/3 co2 + 2/3 ipdp
	2/3 ipdp	->	1/3 Lim
	glc	->	8/3 co2 + 1/3 Lim + 4/3 nadph + 4 nadh

ANEXO C: RESULTADOS DE CADA ALGORITMO DE BÚSQUEDA POR SEPARADO

Tabla 7: Reacciones seleccionadas por el algoritmo de búsqueda FSEOF clasificadas según cada modelo de acuerdo a la ruta a la cual pertenecen.

Ruta	iND750	iMM904	Yeast 4.0
'Alanine and Aspartate Metabolism'	1	2	2
'Arginine and Proline Metabolism'	9	12	4
'Cysteine Metabolism'	5	1	
'Glutamine Metabolism'	1	1	
'Histidine Metabolism'	9	10	10
'Methionine Metabolism'	1	1	1
'Threonine and Lysine Metabolism'	8	9	
'Tyrosine Tryptophan and Phenylalanine Metabolism'	5	15	16
'Valine Leucine and Isoleucine Metabolism'	4	13	8
'Glutamine Metabolism'			2
'Glycine and Serine Metabolism'			1
'Alternate Carbon Metabolism'		4	5
'Fructose and Mannose Metabolism'		3	3
'Glycoprotein Metabolism'		2	2
'Pentose Phosphate Pathway'	2	2	1
'Purine and Pyrimidine Biosynthesis'	5	20	12
'Nucleotide Salvage Pathway'	1	6	9
'Riboflavin Metabolism'		7	7
'Sterol Metabolism'		14	2
'Fatty Acid Biosynthesis'	11	8	4
'Phospholipid Biosynthesis'	1	9	5
'Glycerolipid Metabolism'		1	1
'Fatty Acid Degradation'			1
'Sphingolipid Metabolism'			3
'Pyruvate Metabolism'	1	2	1
'Oxidative Phosphorylation'		4	
'Citric Acid Cycle'	1	2	1
'Folate Metabolism'	1	4	4
'Anaplerotic reactions'	1		
Transport'	9	23	17
[]	2		26
TOTAL	78	175	148

* Las reacciones clasificadas como [] son aquellas que no están definidas en una ruta en particular

ANEXO D: REACCIONES DE INTERCAMBIO

A continuación se presenta un resumen con las reacciones de intercambio para cada modelo. En primer lugar se define la reacción a evaluar según el formato del modelo dado, y se entregan los valores de flujo para dicha reacción. Luego se muestran dos gráficos, el primero con el análisis FSEOF y el segundo con el análisis de variabilidad para dicha reacción.

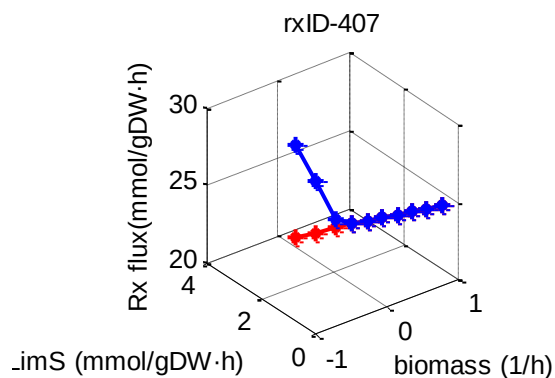
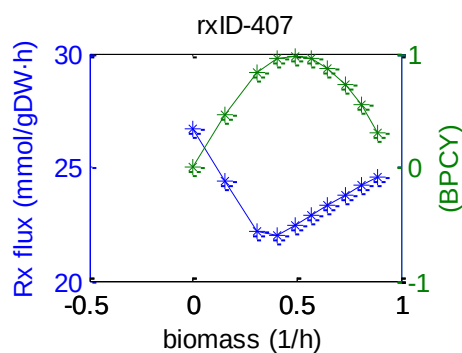
Model: iND750

rxID: 407; rxAbb: EX_co2(e)

CO2 $\rightleftharpoons$

co2[e] $\rightleftharpoons$

Fluxes: 24.6 24.1637 23.7273 23.291 22.8546 22.4183 21.9819
22.1369 24.4018 26.6667



En el primer gráfico con línea azul se muestra el comportamiento del flujo de la reacción 407 del modelo iND750. Tal reacción corresponde a la producción de CO₂ definida como “ co2[e] $\rightleftharpoons$ ”, lo que quiere decir que los flujos positivos liberan CO₂ al medio. Se puede ver que inicialmente (*step1*) el flujo es de 24.6 [mmol/gDW·h] y que la producción de biomasa es cercana a 1[h⁻¹], lo que progresivamente comienza a disminuir a medida que aumenta la producción de limoneno. La disminución llega a su mínimo 21.9

[mmol/gDW·h] para luego aumentar en los últimos 3 *steps* hasta 26.6 [mmol/gDW·h]. Se muestr además, en línea verde, la productividad BPCY de limoneno.

En el segundo gráfico se ocuparon los mismos *steps* para hacer un análisis de variabilidad. El plano XZ corresponde al primer gráfico, al cual se le ha añadido el eje Y que muestra la producción de limoneno. De esta forma, en cada *step* se ha maximizado (puntos azules) y minimizado (puntos rojos) la producción de CO₂ (reacción 407). Como se aprecia, la variabilidad aumenta en los últimos 3 *steps*.

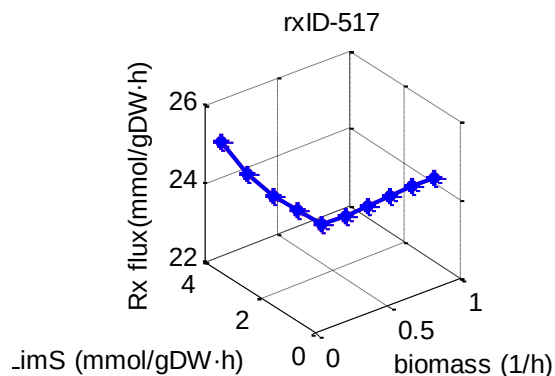
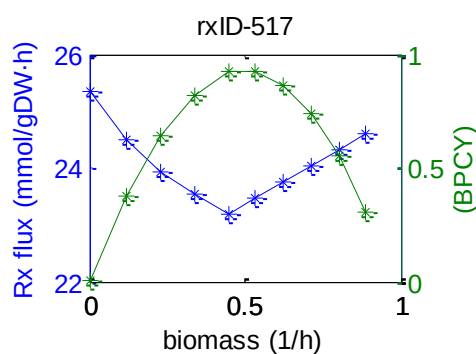
Model: iMM904

rxID: 517; rxAbb: EX_co2(e)

CO₂ <=>

co2[e] <=>

Fluxes: 24.5933 24.309 24.0246 23.7402 23.4558 23.1973 23.5566
23.9236 24.4835 25.3333



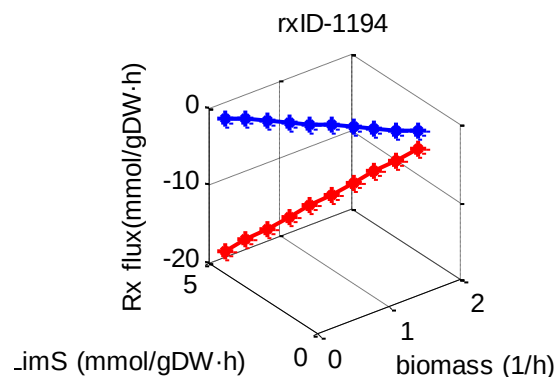
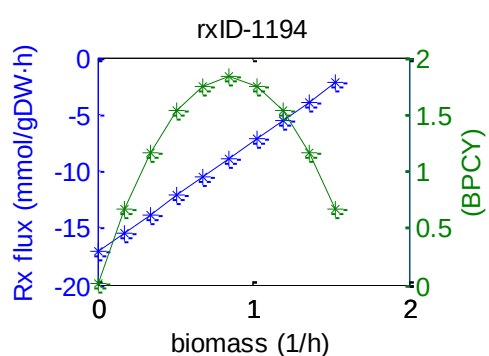
Model: yeast4.0

rxID: 1194; rxAbb: r_1194

<=> carbon dioxide [cytoplasm]

<=> s_0470

Fluxes: -2.26402 -3.91723 -5.57043 -7.22363 -8.87684 -10.53 -12.1832
-13.8364 -15.4897 -17.1429



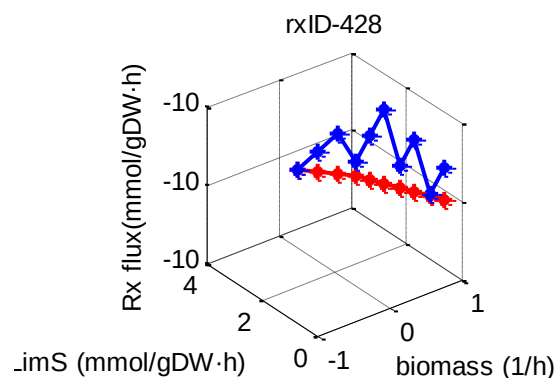
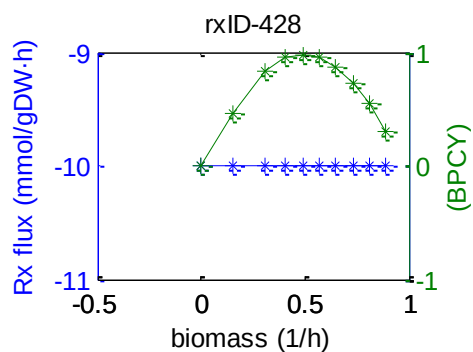
Model: iND750

rxID: 428; rxAbb: EX_glc(e)

D-Glucose $\rightleftharpoons$

glc-DASH-D[e] $\rightleftharpoons$

Fluxes: -10 -10 -10 -10 -10 -10 -10 -10 -10 -10



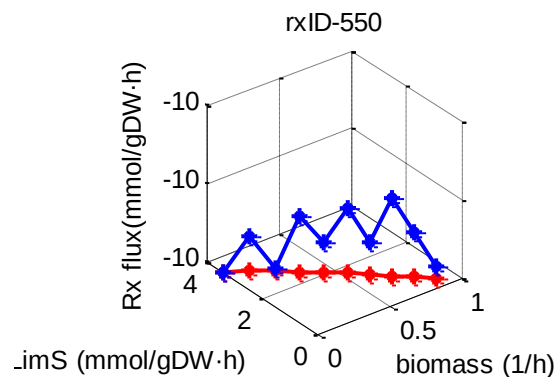
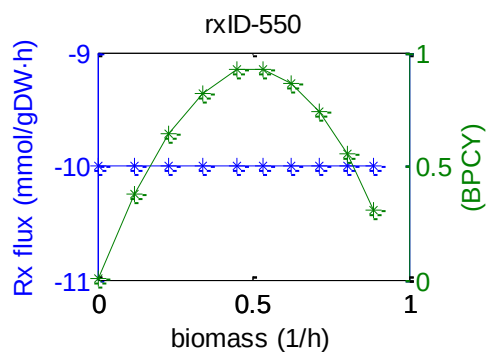
Model: iMM904

rxID: 550; rxAbb: EX_glc(e)

D_Glucose $\rightleftharpoons$

glc_D[e] $\rightleftharpoons$

Fluxes: -10 -10 -10 -10 -10 -10 -10 -10 -10 -10



Model: yeast4.0

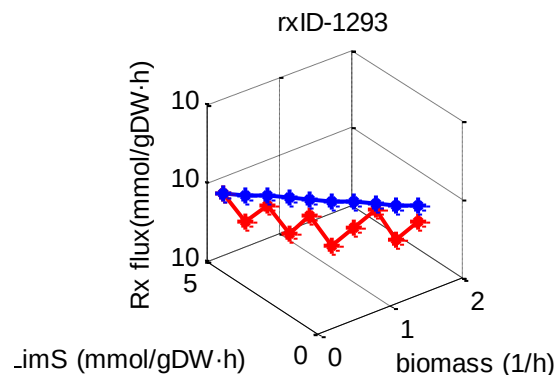
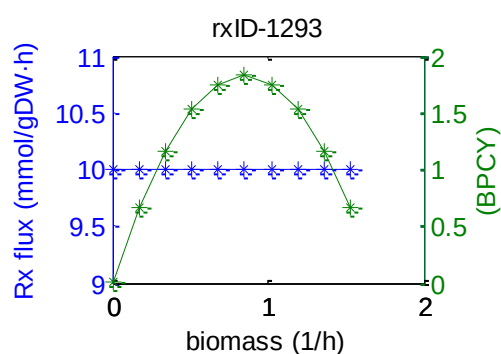
rxID: 1293; rxAbb: r_1293

YDL245C; YDR342C; YDR343C; YDR345C; YDR536W; YEL069C; YFL011W;
YHR092C; YHR094C; YHR096C; YJL214W; YJL219W; YJR158W; YLR081W;
YMR011W; YNR072W; YOL156W;

<=> D-glucose [cytoplasm]

<=> s_0545

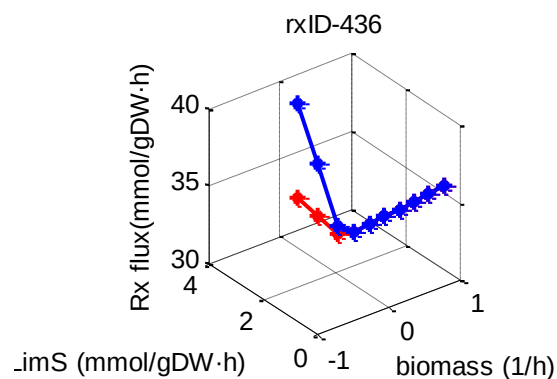
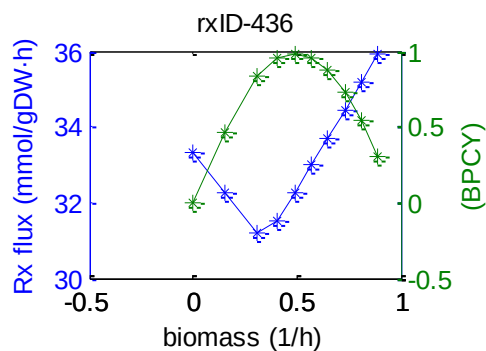
Fluxes: 10 10 10 10 10 10 10 10 10 10

**Model: iND750**

rxID: 436; rxAbb: EX_h2o(e)

H2O <=>

h2o[e] <=>

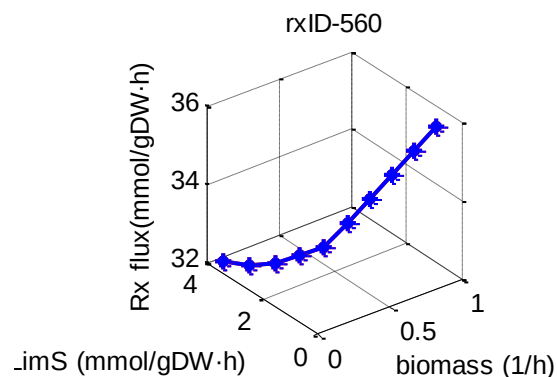
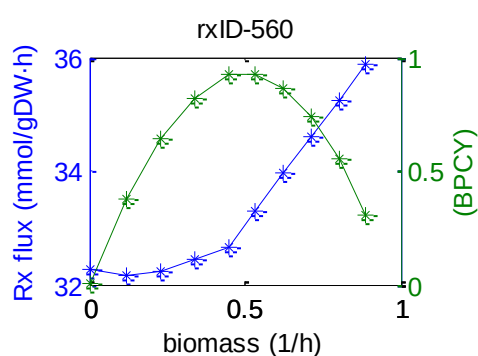
Fluxes: 35.9057 35.1749 34.4441 33.7132 32.9824 32.2515 31.5207
31.185 32.2591 33.3333**Model: iMM904**

rxID: 560; rxAbb: EX_h2o(e)

H2O <=>

h2o[e] <=>

Fluxes: 35.8902 35.2427 34.5952 33.9477 33.3002 32.6699 32.4524
32.24 32.1565 32.2667



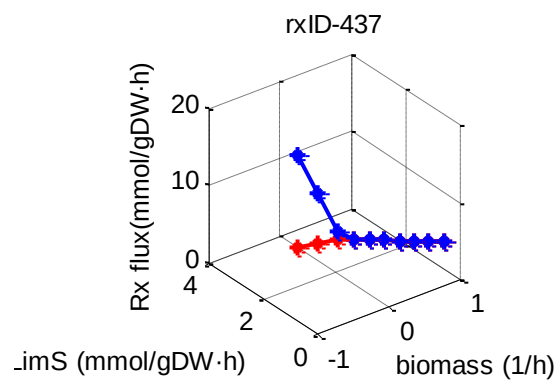
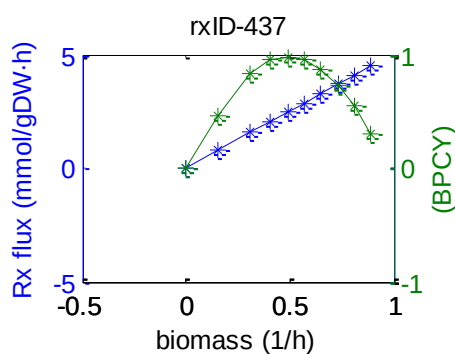
Model: iND750

rxID: 437; rxAbb: EX_h(e)

$H^+ \rightleftharpoons$

$h[e] \rightleftharpoons$

Fluxes: 4.5009 4.0943 3.6877 3.281 2.8744 2.4678 2.0612 1.5715
0.78577 -9.1899e-014



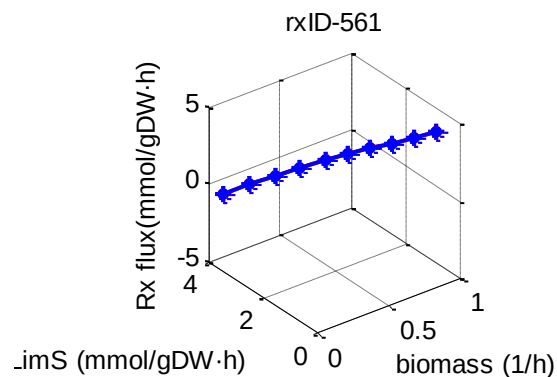
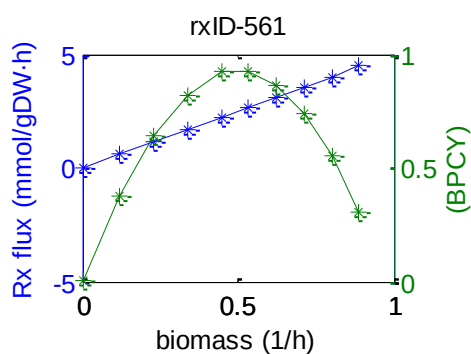
Model: iMM904

rxID: 561; rxAbb: EX_h(e)

$H \rightleftharpoons$

$h[e] \rightleftharpoons$

Fluxes: 4.4702 4.0248 3.5794 3.134 2.6887 2.2397 1.7042 1.1677
0.60412 -8.4721e-014

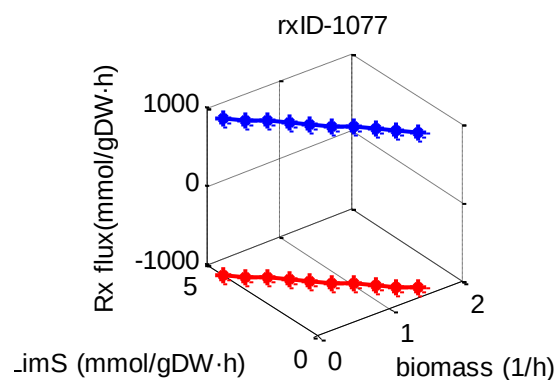
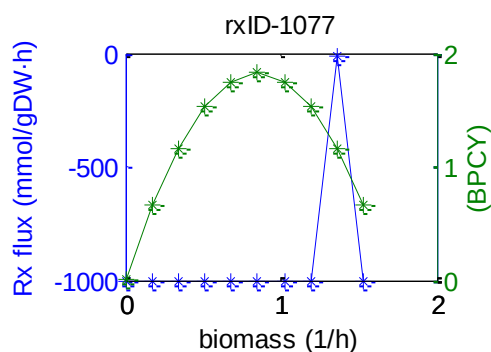


Model: yeast4.0

rxID: 1077; rxAbb: r_1077

hydroxide [cytoplasm] $\rightleftharpoons$ s_0806 $\rightleftharpoons$

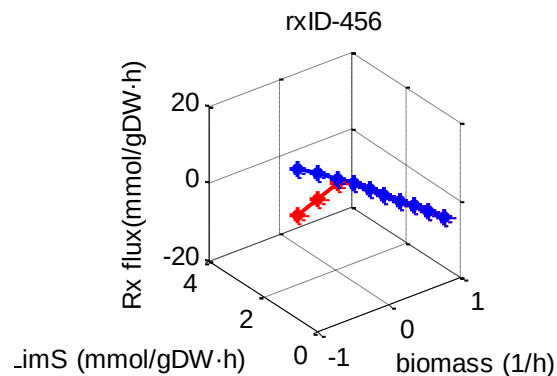
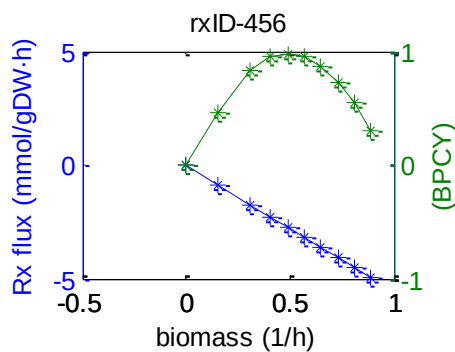
Fluxes: -1000 -10.245785 -1000 -1000 -1000 -1000 -1000
 -1000 -1000 -1000

**Model: iND750**

rxID: 456; rxAbb: EX_nh4(e)

Ammonium $\rightleftharpoons$ nh4[e] $\rightleftharpoons$

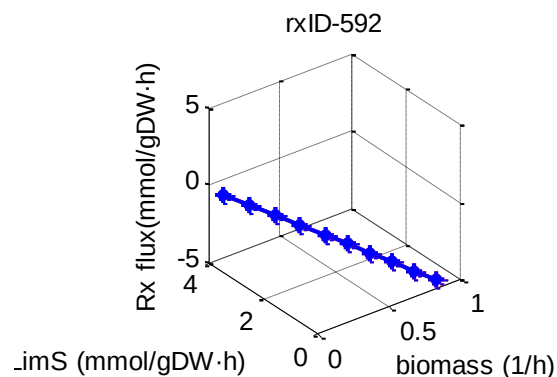
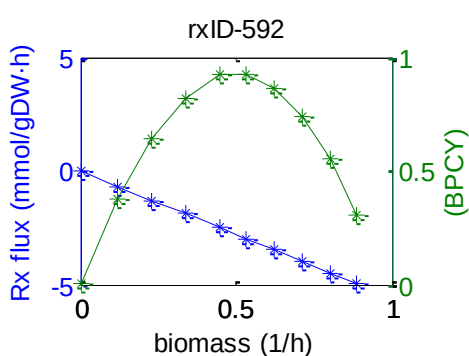
Fluxes: -4.9918 -4.5409 -4.0899 -3.6389 -3.1879 -2.737 -2.286 -1.743
 -0.87148 2.3159e-012

**Model: iMM904**

rxID: 592; rxAbb: EX_nh4(e)

Ammonium $\rightleftharpoons$ nh4[e] $\rightleftharpoons$

Fluxes: -4.9733 -4.4778 -3.9823 -3.4868 -2.9913 -2.4917 -1.896 -1.2991
 -0.67212 4.2942e-013



Model: yeast4.0

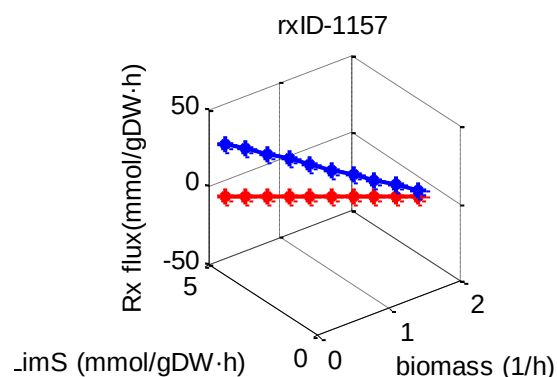
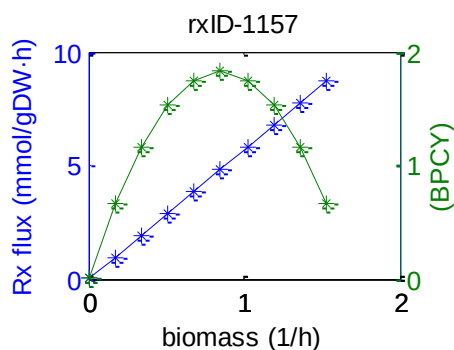
rxID: 1157; rxAbb: r_1157

YDR384C; YGR121C; YNL142W; YPR138C;

<=> ammonium [cytoplasm]

<=> s_0430

Fluxes: 8.7551 7.7823 6.8095 5.8367 4.8639 3.8911 2.9184 1.9456
0.97279 9.851e-013



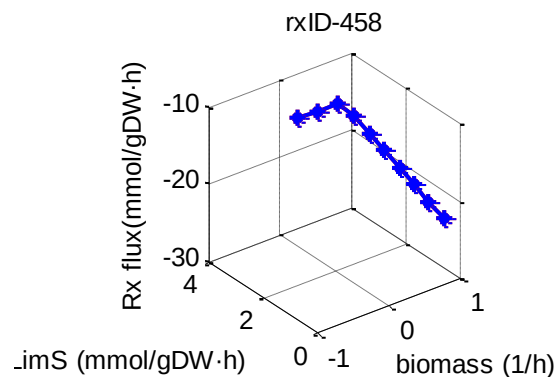
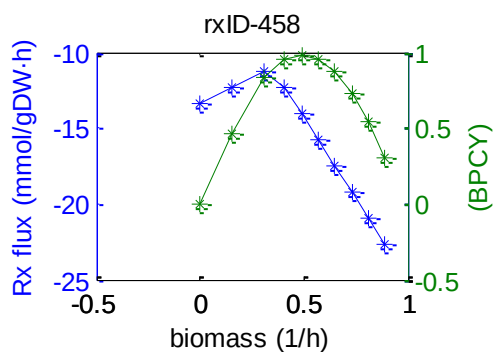
Model: iND750

rxID: 458; rxAbb: EX_o2(e)

O2 <=>

o2[e] <=>

Fluxes: -22.626 -20.9142 -19.2024 -17.4906 -15.7788 -14.067 -12.3552
-11.2465 -12.2899 -13.3333

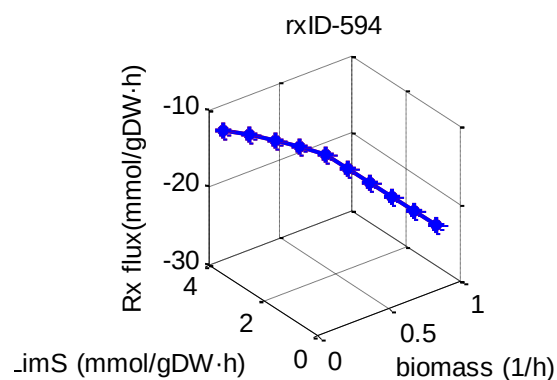
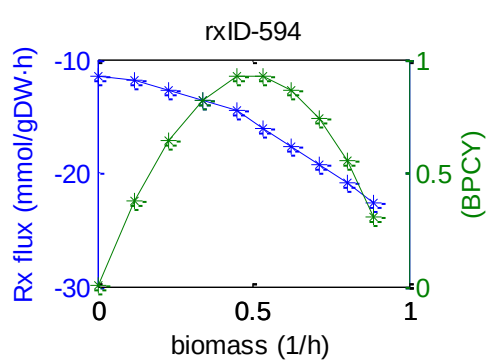


Model: iMM904

rxID: 594; rxAbb: EX_o2(e)

O2 $\rightleftharpoons$ o2[e] $\rightleftharpoons$

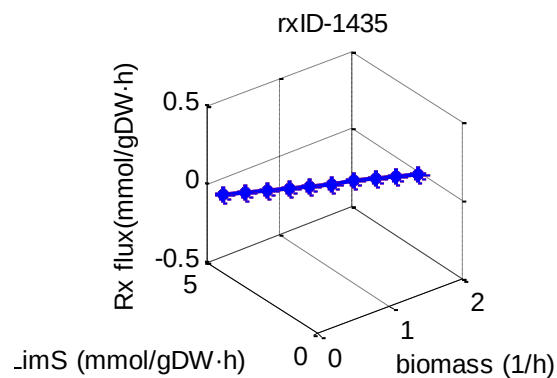
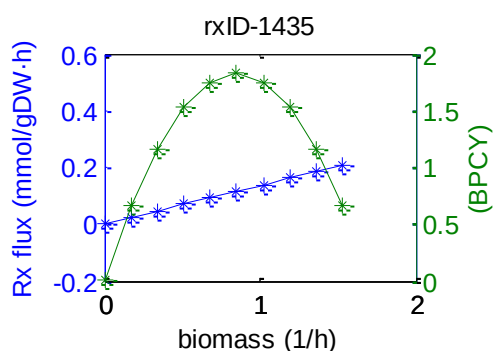
Fluxes: -22.5696 -20.9621 -19.3545 -17.7469 -16.1394 -14.5581 -13.6071
 -12.6639 -11.9174 -11.4667

**Model: yeast4.0**

rxID: 1435; rxAbb: r_1435

 $\rightleftharpoons$ oxygen [cytoplasm] $\rightleftharpoons$ s_1160

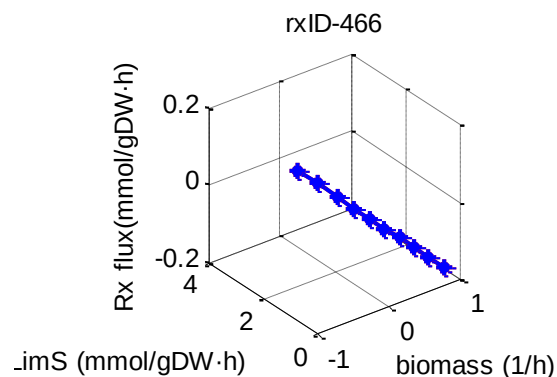
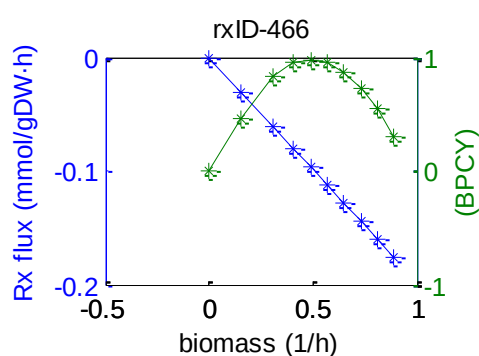
Fluxes: 0.20894 0.18572 0.16251 0.13929 0.11608 0.092861 0.069646
 0.046431 0.023215 -6.8376e-013

**Model: iND750**

rxID: 466; rxAbb: EX_pi(e)

Phosphate $\rightleftharpoons$ pi[e] $\rightleftharpoons$

Fluxes: -0.17647 -0.16052 -0.14458 -0.12864 -0.1127 -0.096754 -0.080811 -
 0.061615 -0.030808 -2.2424e-013



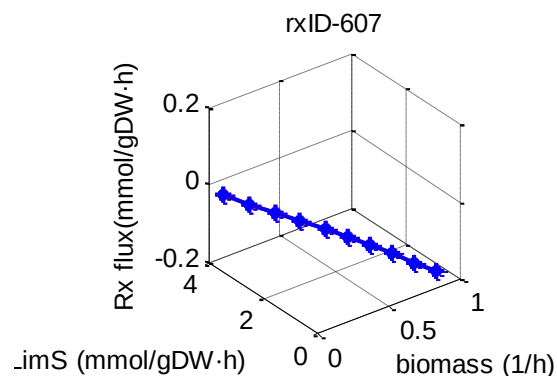
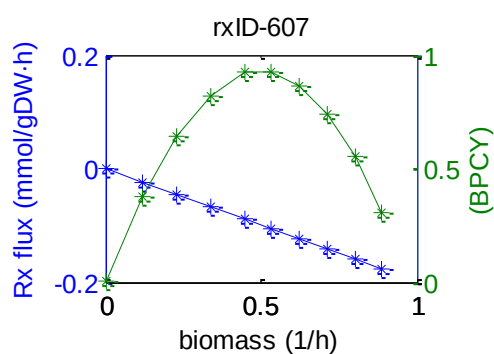
Model: iMM904

rxID: 607; rxAbb: EX_pi(e)

Phosphate $\rightleftharpoons$

pi[e] $\rightleftharpoons$

Fluxes: -0.17569 -0.15818 -0.14068 -0.12317 -0.10567 -0.088023 -0.066978 -
0.045891 -0.023743 2.1944e-013



Model: yeast4.0

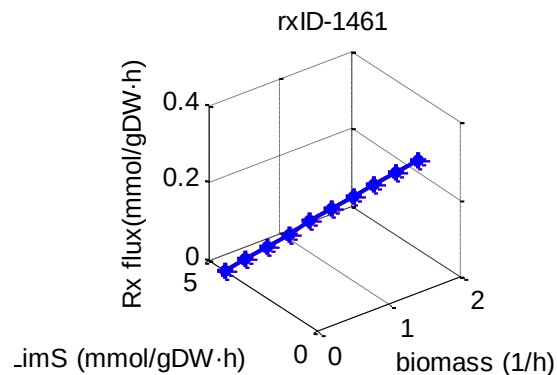
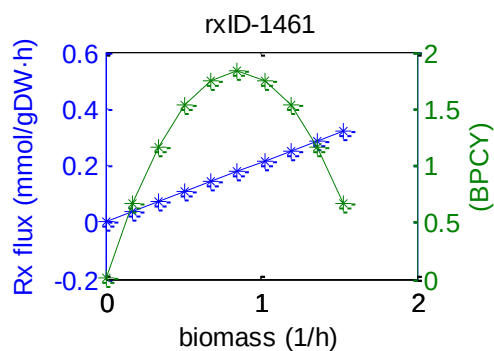
rxID: 1461; rxAbb: r_1461

YBR296C; YCR037C; YJL198W; YML123C; YNR013C;

$\rightleftharpoons$ phosphate [cytoplasm]

$\rightleftharpoons$ s_1207

Fluxes: 0.3205 0.28489 0.24928 0.21367 0.17806 0.14245 0.10683
0.071223 0.035612 -1.5088e-013

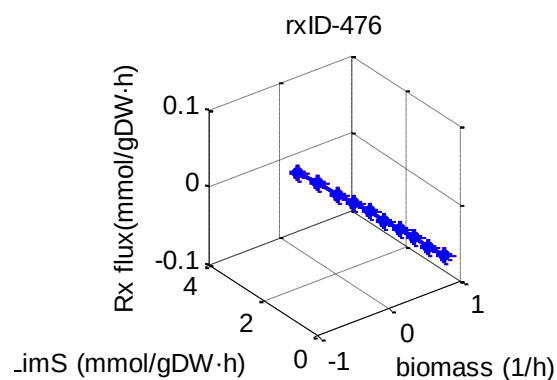
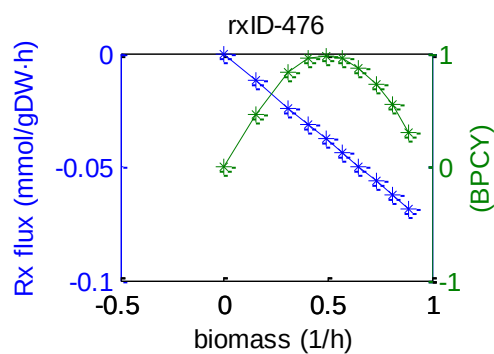


Model: iND750

rxID: 476; rxAbb: EX_so4(e)

Sulfate $\rightleftharpoons$ so4[e] $\rightleftharpoons$

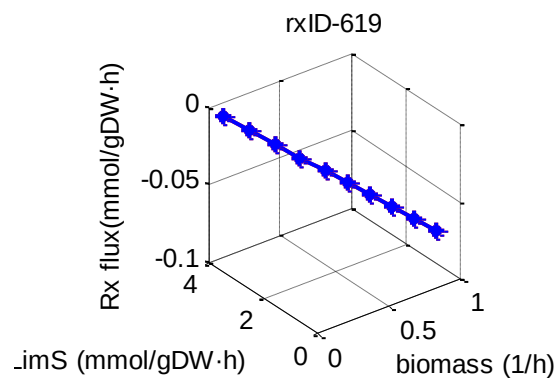
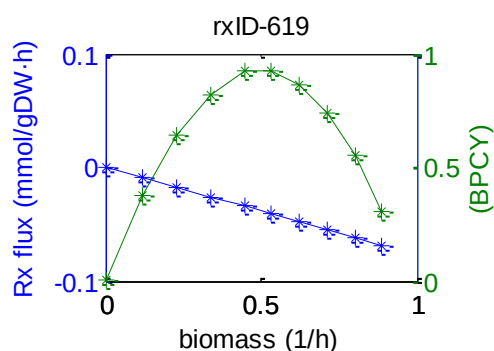
Fluxes: -0.068998 -0.062764 -0.056531 -0.050297 -0.044064 -0.03783 -0.031597
 -0.024091 -0.012046 -0

**Model: iMM904**

rxID: 619; rxAbb: EX_so4(e)

Sulfate $\rightleftharpoons$ so4[e] $\rightleftharpoons$

Fluxes: -0.068693 -0.061849 -0.055005 -0.048161 -0.041316 -0.034417 -0.026188
 -0.017943 -0.0092835 1.4285e-013

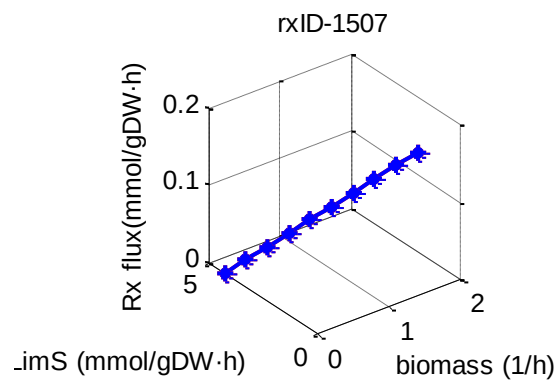
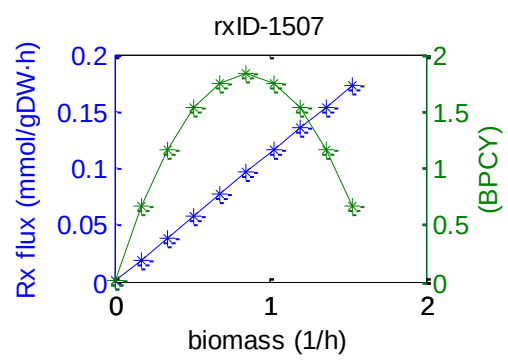
**Model: yeast4.0**

rxID: 1507; rxAbb: r_1507

YBR294W; YLR092W;

 $\rightleftharpoons$ sulphate [cytoplasm] $\rightleftharpoons$ s_1347

Fluxes: 0.17299 0.15377 0.13454 0.11532 0.096103 0.076883 0.057662
 0.038441 0.019221 2.3069e-015

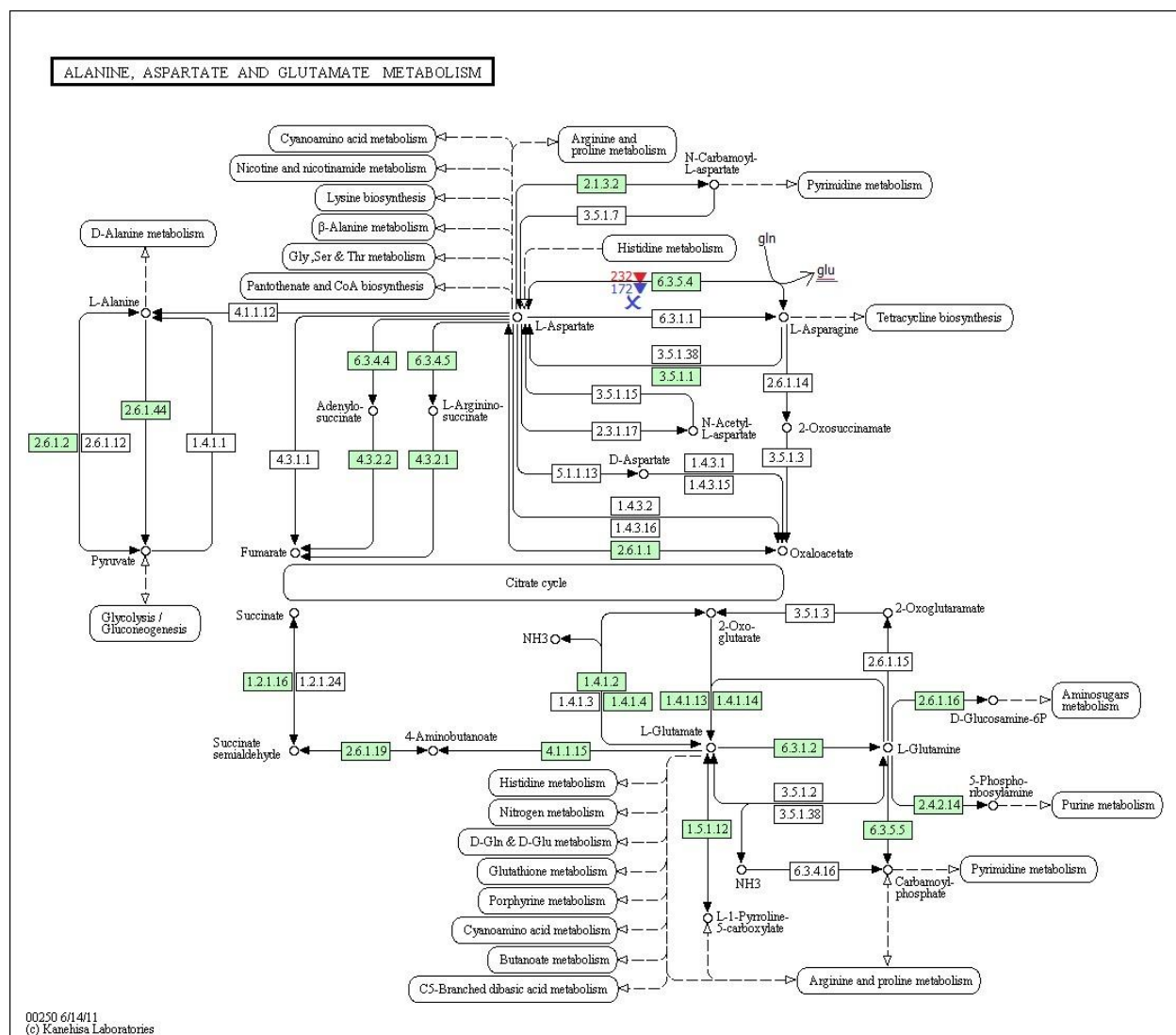


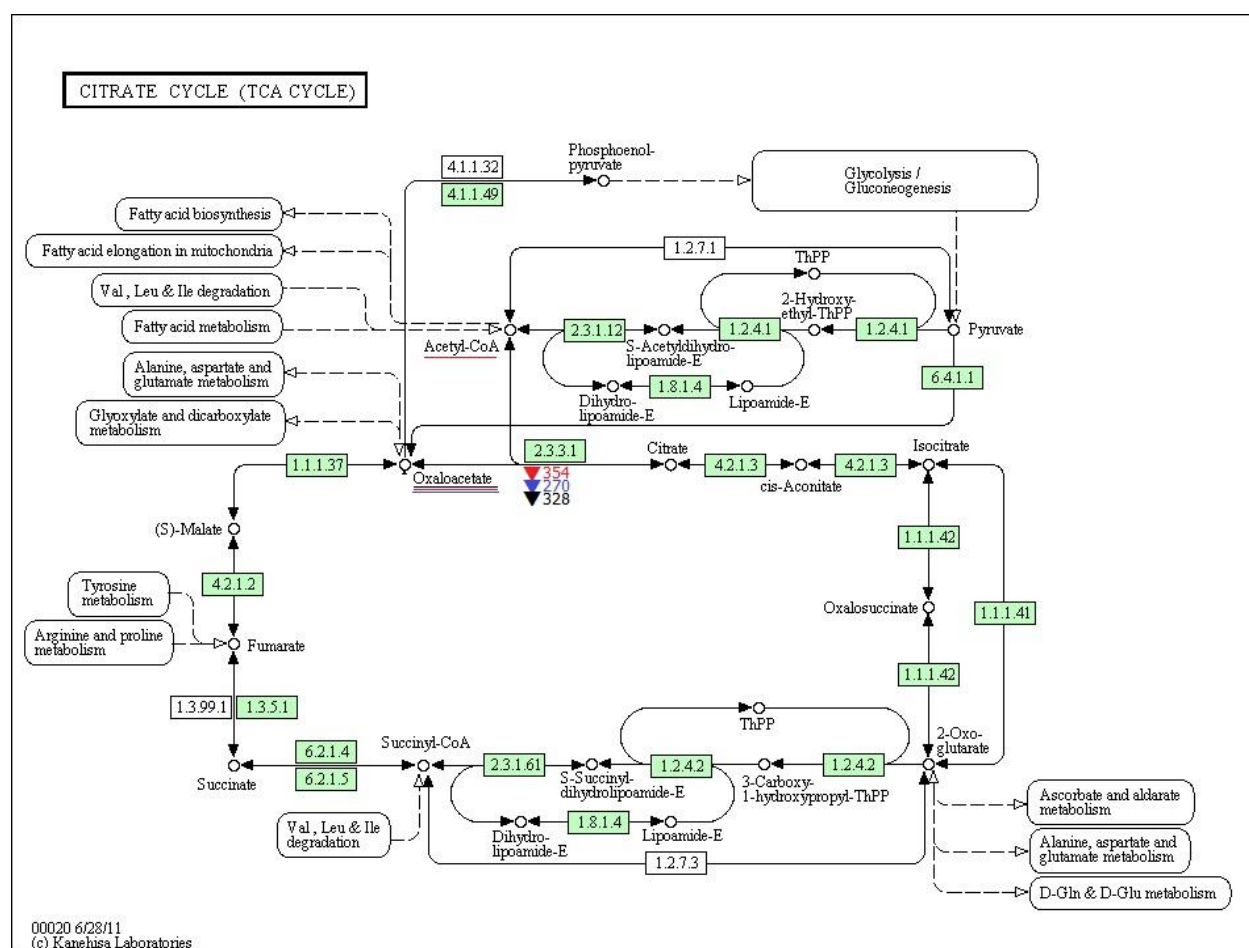
ANEXO E: MAPAS METABÓLICOS SEGÚN LA BASE DE DATOS KEGG

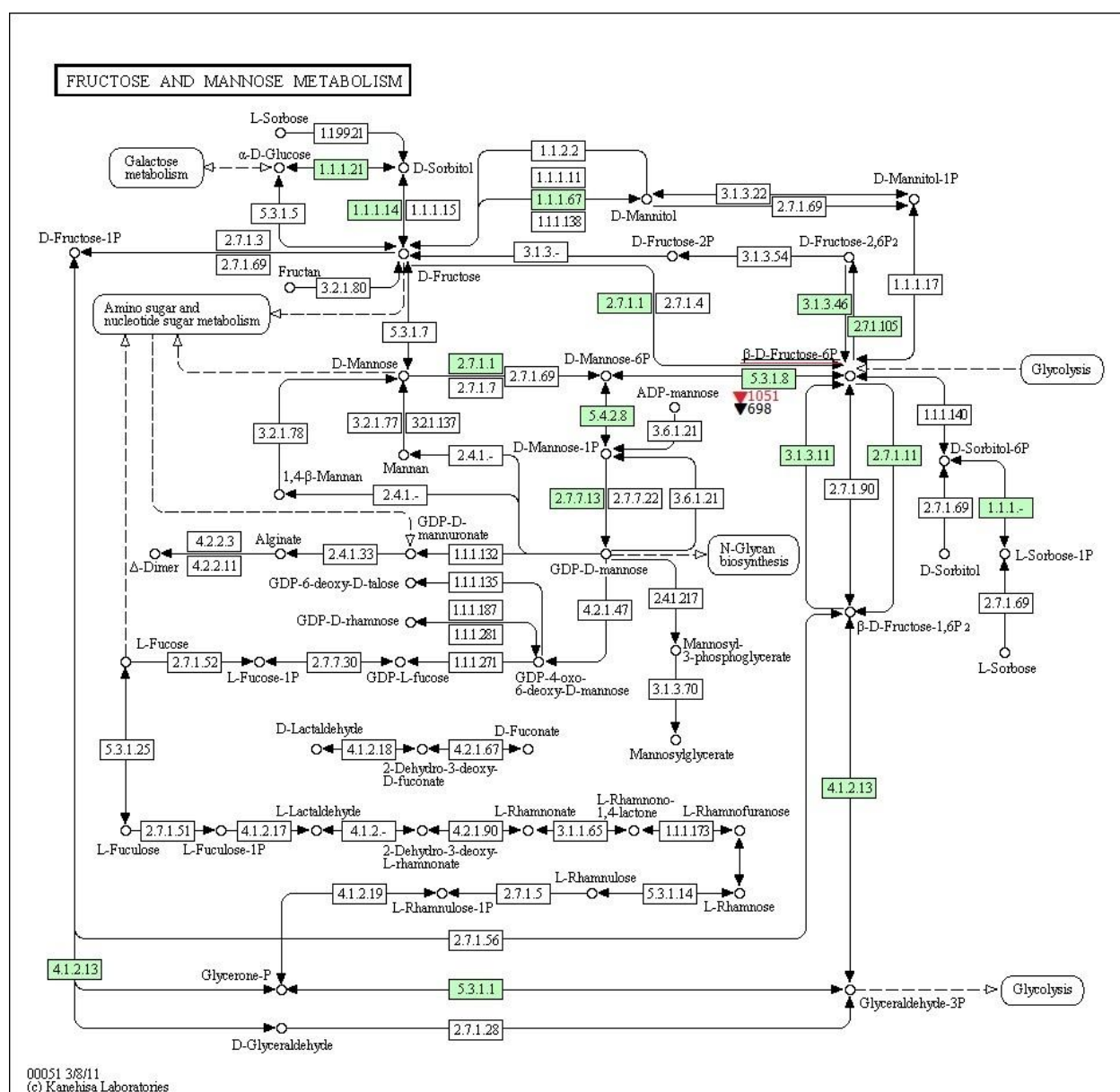
a) Nomenclatura utilizada

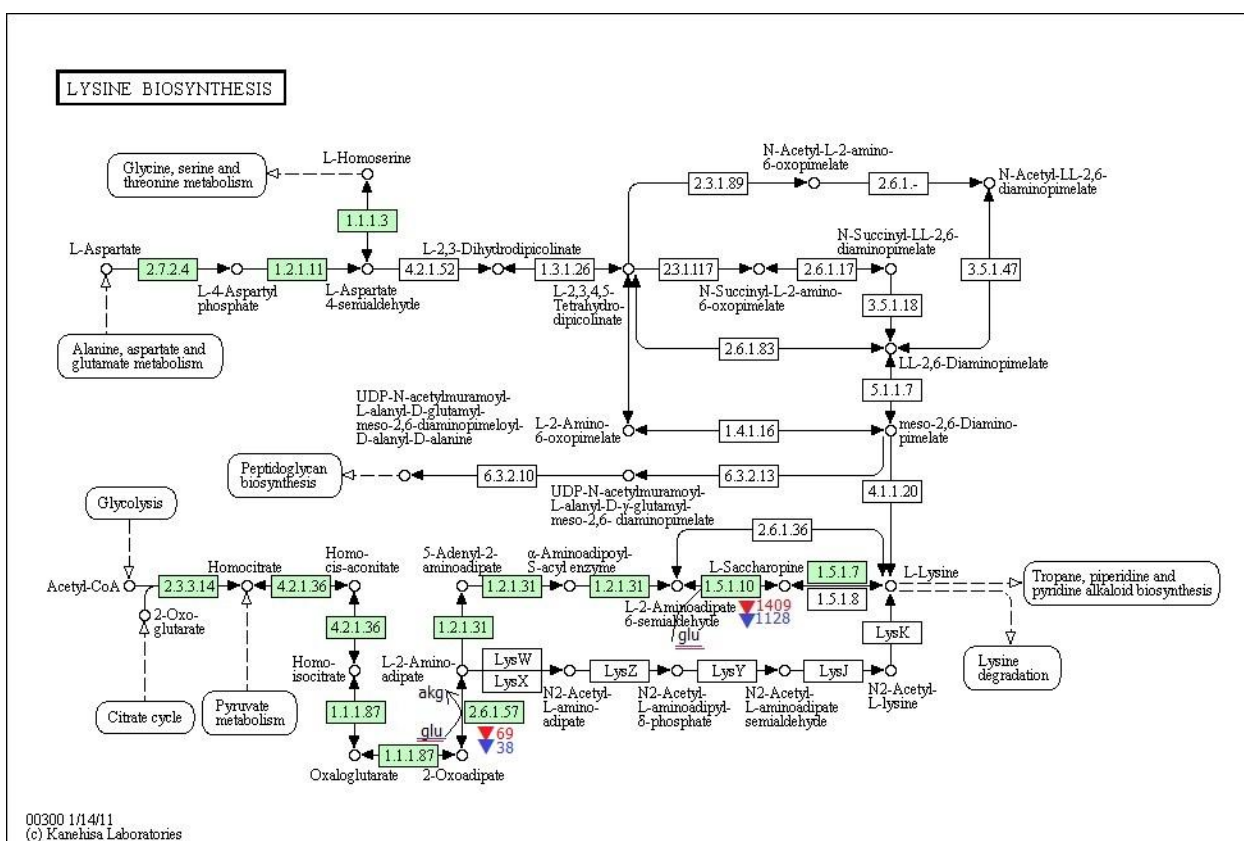
En cada mapa metabólico se incorpora un triángulo o una cruz (▼,✕) en cada reacción seleccionada por el análisis. Cuando el símbolo es de color azul, rojo o negro, se alude al modelo iND750, iMM904 o Yeast 4.0 respectivamente. Además se incluye un número que corresponde al número de reacción del modelo respectivo. Finalmente, cada metabolito que es subrayado representa un nodo principal.

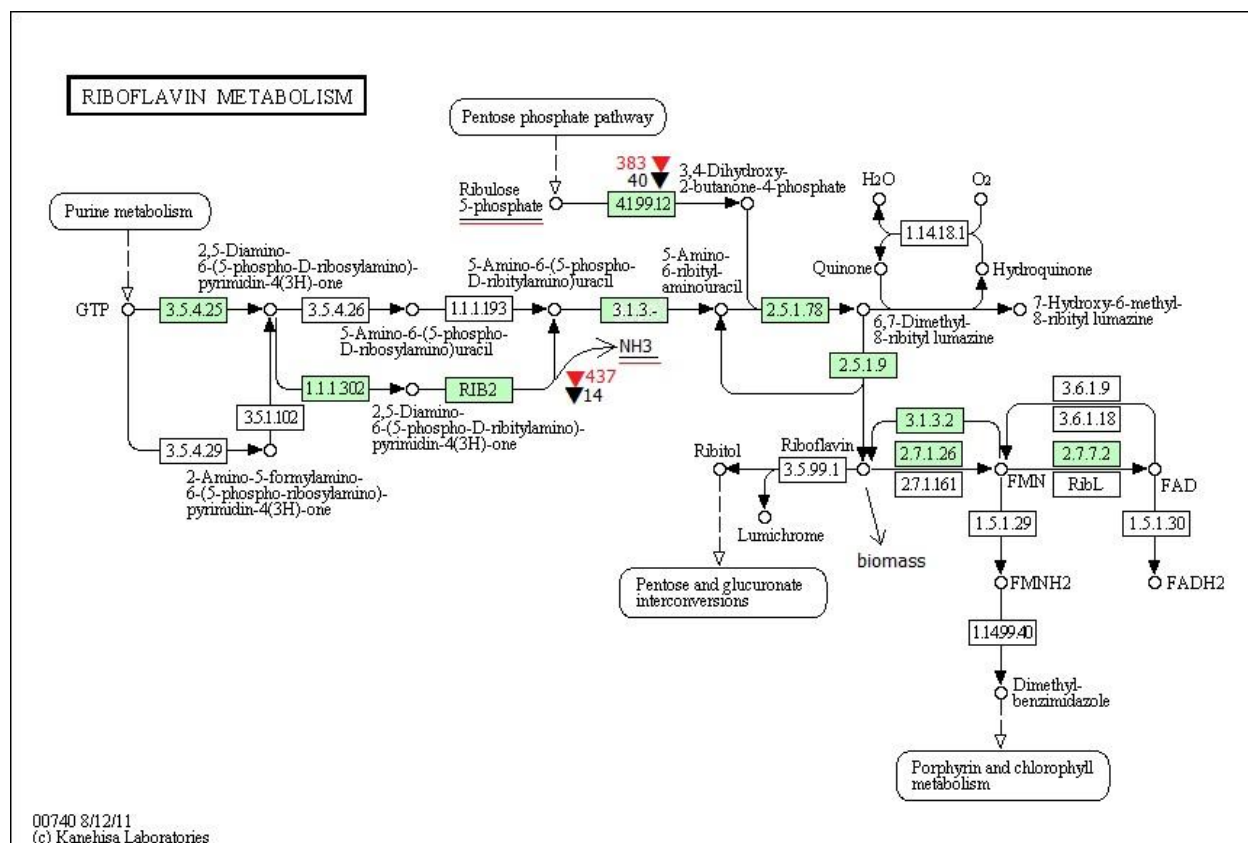
a) Condición sin aminoácidos

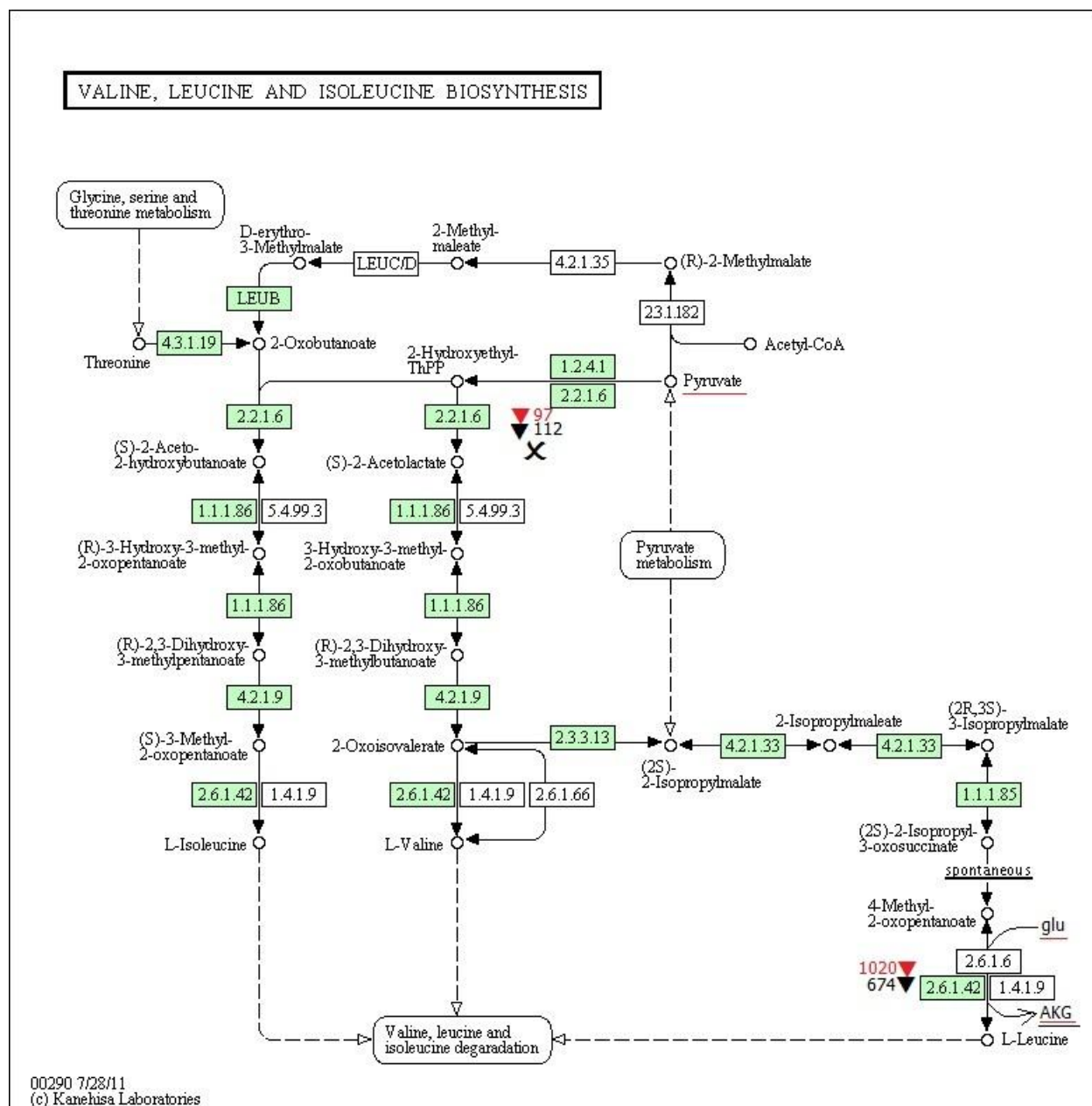


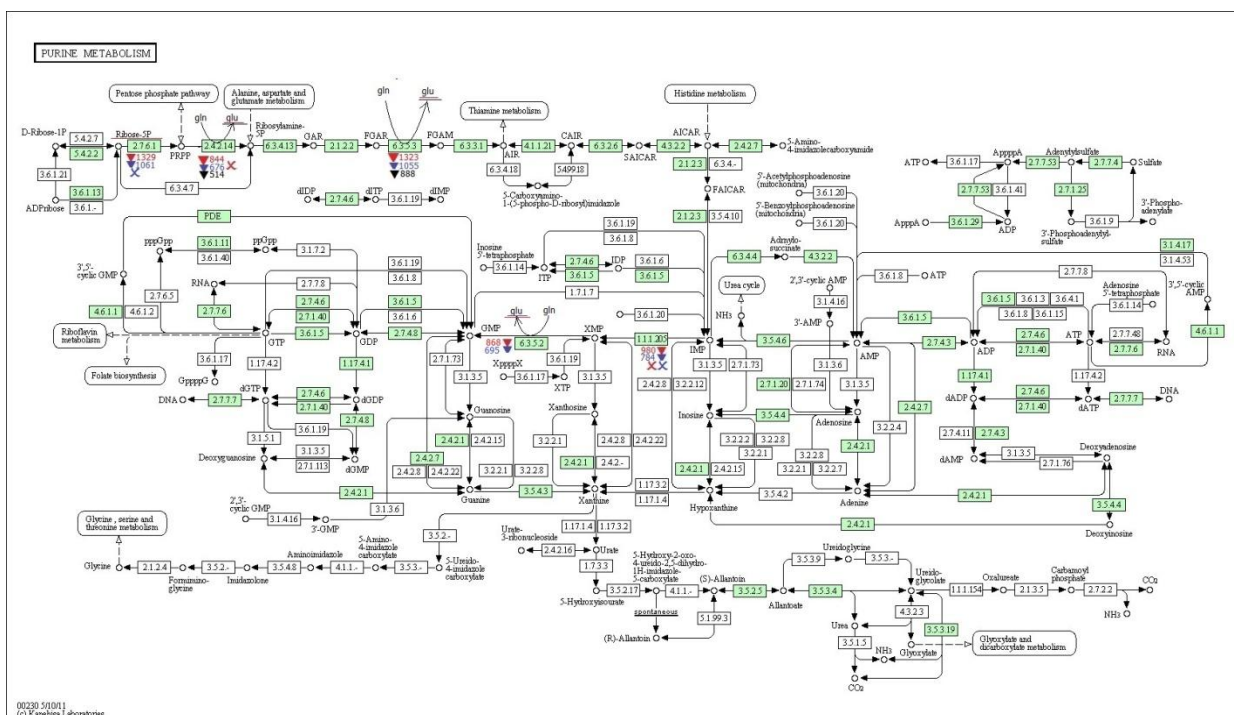


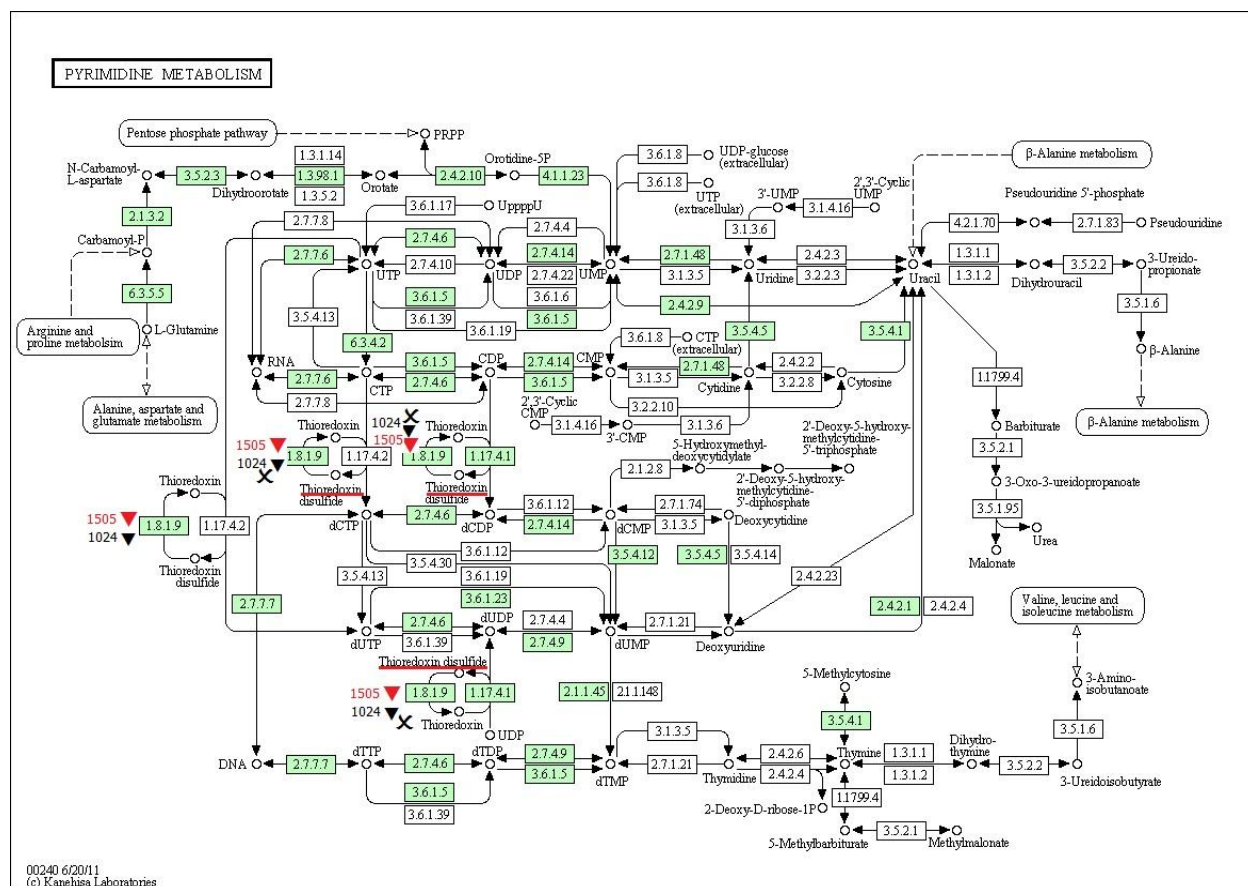


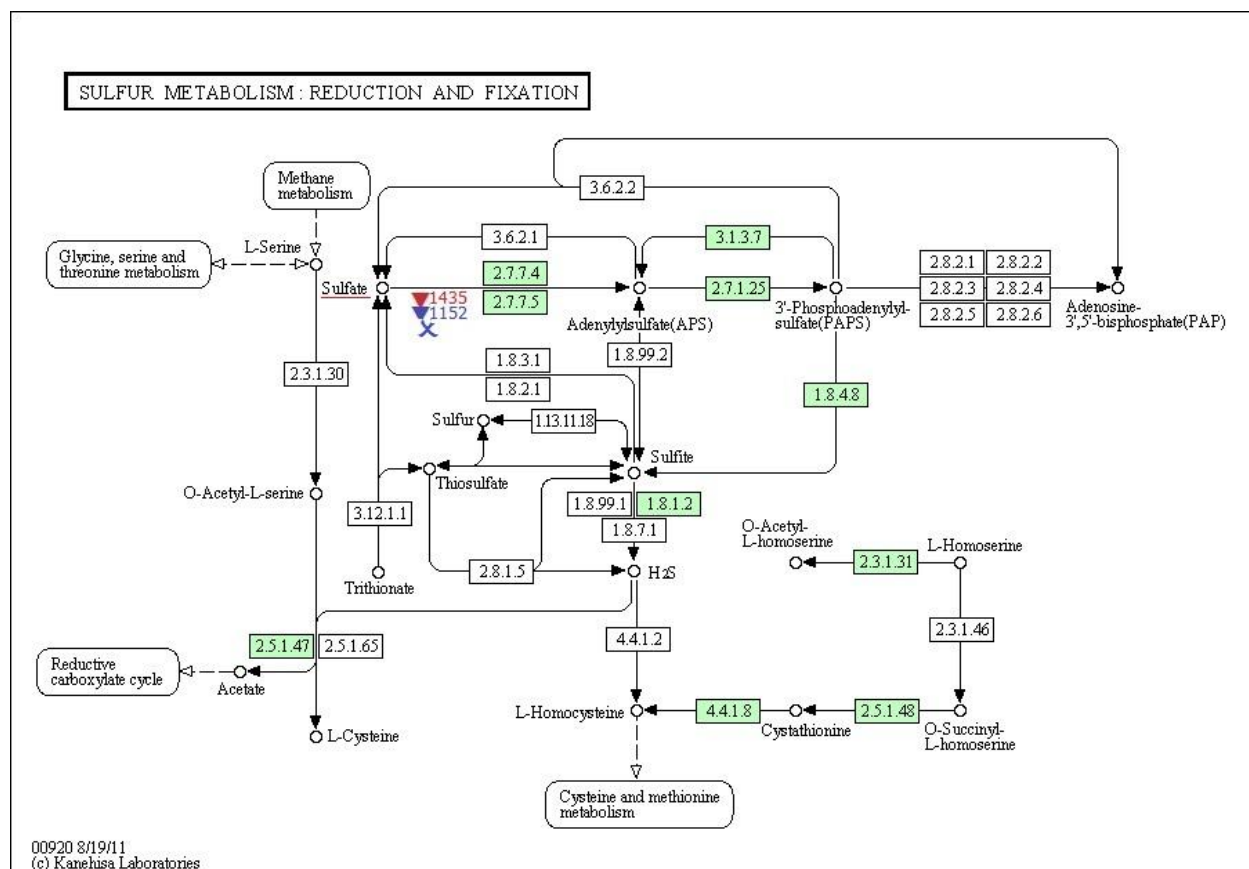




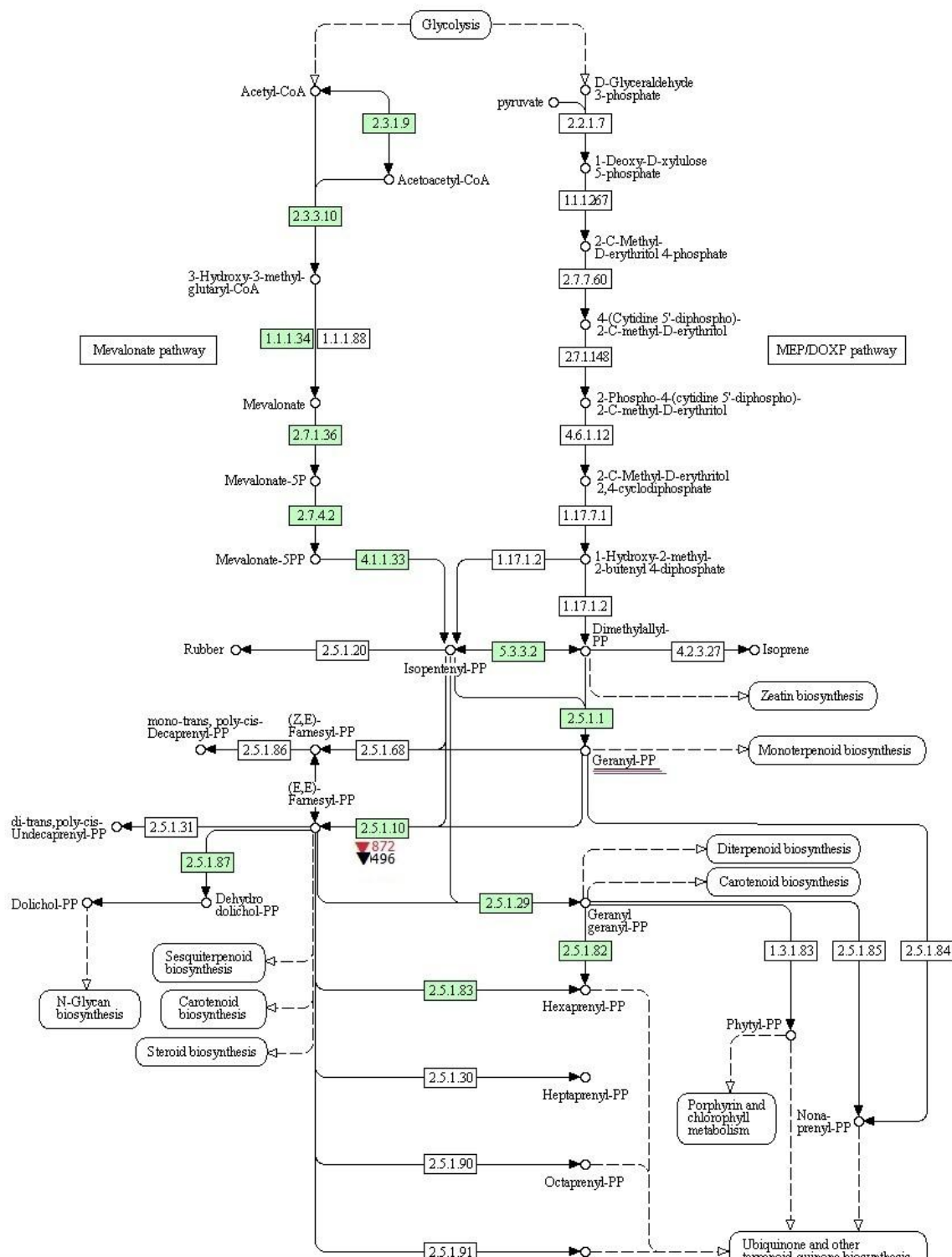




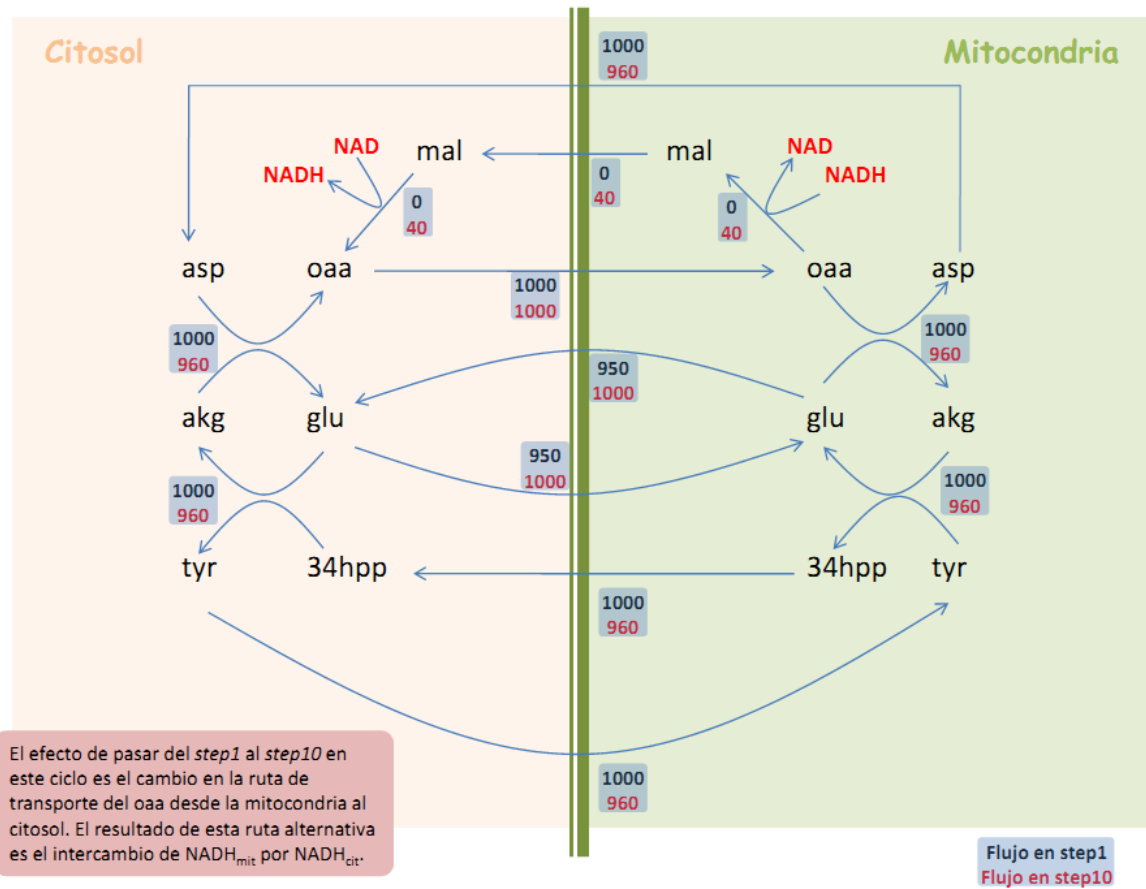




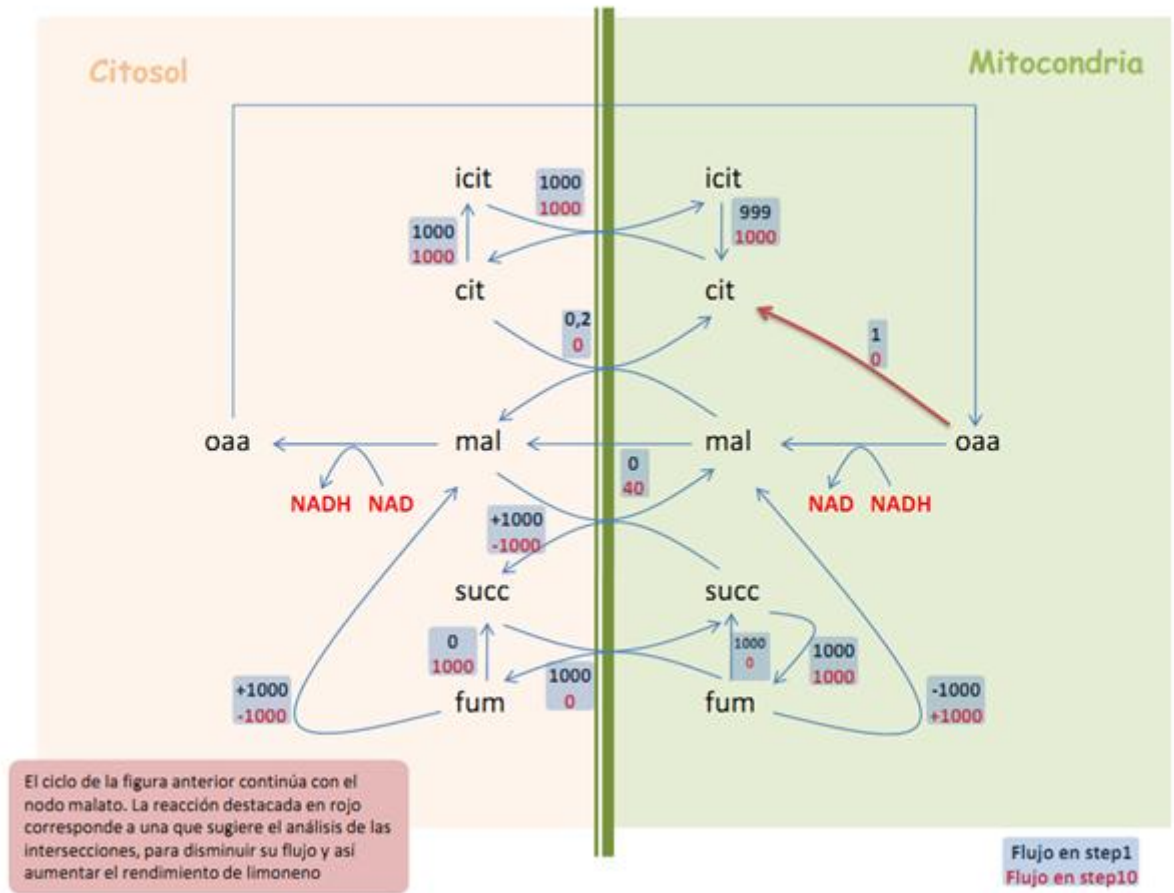
TERPENOID BACKBONE BIOSYNTHESIS



ANEXO F: CICLO DONDE PARTICIPA EL GLUTAMATO



akg: α-cetoglutarato; asp: aspartato; glu: glutamato; mal: malato; oaa: oxaloacetato; tyr: tirosina.



cit: citrato; fum: fumarato; icit: isocitrato; mal: malato; oaa: oxaloacetato; succ: succinato.

