



UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO



FACULTAD DE QUÍMICA

PROGRAMA DE MAESTRÍA Y DOCTORADO
EN CIENCIAS BIOQUÍMICAS

ANÁLISIS DE LA
COMPARTIMENTALIZACIÓN DIFERENCIAL
ENTRE PRODUCTOS DE GENES
PARÁLOGOS

T E S I S

QUE PARA OBTENER EL GRADO DE:

MAESTRA EN CIENCIAS

P R E S E N T A:

**LCG. MIRELLE CITLALI
FLORES VILLEGAS**

TUTORA: DRA. MARÍA ALICIA GONZÁLEZ MANJARREZ

MÉXICO, D. F.

AGOSTO, 2012



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

AGRADECIMIENTOS

A Alicia, por su enseñanza, paciencia, por permitirme estar dentro de su grupo de trabajo y colaborar en mi desarrollo y por confiar en mi.

A la UNAM por permitirme formar parte de su comunidad, que es un honor y un privilegio.

A mis sinodales, el Dr. Luis Servín y a la Dra. Guadalupe Espín, por ser una guía contante, por sus consejos y apoyo brindado desde el inicio.

A los miembros del Jurado, por sus valiosas aportaciones a este trabajo y por tomarse el tiempo de revisarlo.

A Lety, por su incomparable ayuda, por que sin ella pocos podríamos.

A mis compañeros del 301-Oriente, a Geovani por impulsar el inicio de este proyecto; a Juan Carlos y Alan por la compañía tempranera que tanto se disfruta; a José Carlos y Lalo por el cariño, la comprensión y por siempre brindarme una sonrisa y una buena plática; a Cristi por todos los consejos de trabajo y de vida, por estar siempre dispuesta a ayudar; a Maritrini por toda la ayuda brindada en el trabajo; a Karla, Gina y Luis por los buenos momentos compartidos y por ser muy buenos compañeros; y a James, Ximena, Joao por las buenas platicas de viajes y demás cosas.

A mis amigos de siempre, Mario, Miguel, Sandy, Javi, Germa, Christian por creer en mí y estar siempre en mi corazón.

A mis padres, por que siempre están presentes, por que sin su ayuda, nada de esto sería posible. Gracias por ser mis modelos a seguir, gracias por impulsarme a seguir adelante.

A mis hermanos, por todos sus cuidados y su amor.

A Augusto, por su infinita ayuda, comprensión, por todas las cosas que aprendimos juntos, por que gran parte de esto se la debo.

A Diana, por seguir colaborándome y queriéndome.

A mi Aldo, por convertirse en mi angelito de la guarda y cuidarme a cada momento.

A mi Edi, por que me ha permitido ver un mundo diferente, por ayudarme a ver siempre hacia arriba, por su comprensión, cuidados y amor, pero sobre todo por convertirse en mi inspiración.

A Papá, Mamá, Brisa, Alex, Aldo y Edí

ÍNDICE

ABREVIATURAS	p. 7
NOMENCLATURA	p. 8
LISTA DE FIGURAS	p.10
LISTA DE TABLAS	p. 12
RESUMEN	p. 13
I) INTRODUCCIÓN	
1.1 Duplicación génica	
1.2 Posible destino de los genes duplicados	
1.3 Divergencia funcional de los genes duplicados.	
1.4 Importancia de la localización subcelular.	
1.5 Predicción de la localización subcelular y características de las señales de localización	
1.6 Importación mitocondrial.	
1.7 Importación nuclear	
1.8 Evolución del genoma en levaduras.	
<u>1.8.1 <i>Saccharomyces cerevisiae</i></u>	
<u>1.8.2 <i>Lanhamella kluyveri</i></u>	
II) PLANTEAMIENTO DE LA INVESTIGACIÓN	p.28
2.1 Justificación	
2.2 Hipótesis	
2.3 Objetivo General	
2.4 Objetivos específicos	
2.5 Metas	
III) METODOLOGÍA	
	p. 31
3.1 Condiciones de crecimiento y cepas.	
3.2 Selección de proteínas candidatas.	
3.3 Análisis <i>in silico</i> de las características de las proteínas candidatas.	
3.4 Construcción de cepas.	
<u>3.4.1 Fusión de proteínas de <i>S. cerevisiae</i> con yECitrine.</u>	
<u>3.4.2 Fusión de proteínas de <i>L. kluyveri</i> con yECitrine.</u>	
<u>3.4.3 Delitto Perfetto.</u>	
<u>3.4.4 Mutantes de delección sencillas y dobles.</u>	
<u>3.4.5 Intercambio de promotores endógenos por el promotor tet.</u>	
3.5 Análisis de microscopía de fluorescencia	
3.6 Análisis filogenético.	
3.7 Curvas de crecimiento.	
3.8 Concentraciones mínimas inhibitorias de metilgloxal para <i>Glo2/Glo4</i> .	
3.9 Preparación del extracto celular y ensayo enzimático de alanina aminotransferasa.	
IV) RESULTADOS	p.40
4.1 Elección de parejas de proteínas empleadas en este trabajo y técnicas de selección.	
<u>4.1.1 Parejas candidatas y su función.</u>	
<u>4.1.2 Análisis <i>in silico</i> utilizado para determinar las características de las proteínas candidatas.</u>	
<u>4.1.3 Microscopía de fluorescencia de las parejas descartadas.</u>	
4.2 Estudio del mejor programa de predicción de localización subcelular	
4.3 Parejas finales.	
<u>4.3.1 Transaminasas de alanina <i>ScAlt1/ScAlt2</i></u>	

- 4.3.1.1 *Antecedentes de las transaminasas de alanina ScAlt1/ScAlt2.*
- 4.3.1.2 *Cepas utilizadas ScAlt1/ScAlt2.*
- 4.3.1.3 *Microscopía de fluorescencia de ScAlt1-yEC y ScAlt2-yEC.*
- 4.3.1.4 *Predicción de localización de LkAlt1.*
- 4.3.1.5 *Localización de la proteína LkAlt1-yEc.*
- 4.3.1.6 *Relocalización de ScAlt1 y ScAlt2.*
- 4.3.1.7 *Actividad específica de ScAlt1R-70 y ScAlt1-84.*
- 4.3.1.8 *Determinación del fenotipo de las mutantes de los genes ScALT1 y ScALT2*
 - 4.3.1.7.1 *Curvas de crecimiento en glucosa*
 - 4.3.1.7.2 *Curvas de crecimiento en etanol*
- 4.3.1.9 *Conclusiones ScAlt1/ScAlt2.*
- 4.3.1.10 *Perspectivas ScAlt1/ScAlt2.*
- 4.3.2 Aminotransferasas de aminoácidos de cadena ramificada ScBat1/ScBat2
 - 4.3.2.1 *Antecedentes de las aminotransferasas de aminoácidos de cadena ramificada ScBat1/ScBat2.*
 - 4.3.2.2 *Cepas utilizadas ScBat1/ScBat2.*
 - 4.3.2.3 *Microscopía de fluorescencia ScBat1/ScBat2.*
 - 4.3.2.4 *Predicción de localización de LkBat1 y LkBat2.*
 - 4.3.2.5 *Localización de la proteína LkBat1.*
 - 4.3.2.6 *Determinación del fenotipo de las mutantes de los genes ScBAT1 y ScBAT2*
 - 4.3.2.6.1 *Curvas de crecimiento en glucosa*
 - 4.3.2.6.2 *Curvas de crecimiento en etanol*
 - 4.3.2.8 *Conclusiones ScBat1/ScBat2.*
 - 4.3.2.9 *Perspectivas ScBat1/ScBat2.*
- 4.3.3 Glioxilasas de tipo II ScGlo2/ScGlo4
 - 4.3.3.1 *Antecedentes de las glioxilasas de tipo II ScGlo2/ScGlo4.*
 - 4.3.3.2 *Cepas utilizadas ScGlo2/ScGlo4.*
 - 4.3.3.3 *Microscopía de fluorescencia ScGlo2/ScGlo4.*
 - 4.3.3.4 *Predicción de localización de LkGlo2.*
 - 4.3.3.5 *Localización de la proteína LkGlo2.*
 - 4.3.3.6 *Relocalización de ScGlo2 y ScGlo4.*
 - 4.3.3.7 *Propiedades fisiológicas.*
 - 4.3.3.8 *Perspectivas ScGlo2/ScGlo4.*
 - 4.3.3.9 *Conclusiones ScGlo2/ScGlo4.*
- 4.3.4 Trr1/Trr2
 - 4.3.4.1 *Antecedentes ScTrr1/ScTrr2.*
 - 4.3.4.2 *Cepas utilizadas ScTrr1/ScTrr2.*
 - 4.3.4.3 *Microscopía de fluorescencia ScTrr1/ScTrr2.*
 - 4.3.4.4 *Predicción de localización de LkTrr.*
 - 4.3.3.3 *Construcciones de la pareja ScTrr1 y ScTrr2*

V) DISCUSIÓN GENERAL	p.75
VI) CONCLUSIONES GENERALES	p.78
VII) PERSPECTIVAS GENERALES	p.80
VIII) REFERENCIAS	p.81
IX) ANEXOS	p.87

ABREVIATURAS

aa	Aminoácidos
AACR	Aminoácidos de Cadena Ramificada
BLAST	Herramienta de Búsqueda por Alineamiento Local Básico
c	Citoplasma
cDNA	DNA complementario
clo	Cloroplasto
CORE	Reportero Contra seleccionable
DNA	Ácido desoxirribonucleico
DO	Densidad Óptica
Exp	Expresión
I	Isoleucina
IROS	Oligonucleótidos Recombinantes Integrativos
L	Leucina
m	Mitocondria
MA	millones de años
MM	Medio mínimo de levadura
MP	Membrana plasmática
mRNA	RNA mensajero
n	Núcleo
NLS	Señal de localización nuclear
NPC	Complejos de Poros Nucleares
NAHR	Recombinación Homóloga no Alélica
NHEJ	Recombinación por unión de extremos no homólogos
Nom_sis	Nombre sistemático
PCR	Reacción en cadena de la polimerasa
perox	Peroxisoma
presec	presecuencia
RE	Retículo Endoplásmico
RNA	Ácido ribonucleico
rpm	revoluciones por minuto
RSP	Relocalización subcelular de proteínas
<i>Sc</i>	<i>Saccharomyces cerevisiae</i>
sec	Proteína secretada
<i>Sk</i>	<i>Lanchancea kluyveri</i>
TA	Temperatura Ambiente
tet	Promotor constitutivo de Tetraciclina
V	Valina
Va	Vacuola
LIV	Valina, Isoleucina, Leucina
WGD	Duplicación total del Genoma
yEC/yEcitrine	Proteína verde fluorescente yECitrine
YPD	Medio rico de levadura

NOMENCLATURA

<i>his3</i>	Gen inactivo de <i>HIS3</i>
His3	Fosfato de imidazolglicerol dehidratasa
<i>HIS3</i>	Gen que codifica para la fosfato de imidazolglicerol dehidratasa His3
KanMX	Módulo que confiere resistencia a Geneticina
<i>leu</i>	Gen inactivo de <i>LEU2</i>
<i>LkAlt1</i>	Transaminasa de alanina de <i>Lanchancea kluyveri</i>
<i>LkALT1</i>	Gen que codifica para la transaminasa de alanina de <i>L. kluyveri</i>
<i>LkBat1</i>	Transaminasa de aminoácidos de cadena ramificada de <i>L. kluyveri</i>
<i>LkBAT1</i>	Gen que codifica para la transaminasa de aminoácidos de cadena ramificada de <i>L. kluyveri</i>
<i>LkGlo2</i>	Glioxilasa tipo II de <i>Lanchancea kluyveri</i>
<i>LkGLO2</i>	Gen que codifica para la glioxilasa tipo II de <i>Lanchancea kluyveri</i>
<i>MATa</i>	Determinación del tipo sexual a
<i>MATα</i>	Determinación del tipo sexual α
<i>met15</i>	Gen inactivo de <i>MET15</i>
NAT	Nourseotricina
<i>ScAlt1</i>	Transaminasa de alanina mitocondrial de <i>S. cerevisiae</i>
<i>ScALT1</i>	Gen que codifica para la transaminasa de alanina mitocondrial ScAlt1
<i>ScAlt1R-70</i>	La proteína ScAlt1 que no tiene los primeros 70 aminoácidos
<i>ScAlt1R-84</i>	La proteína ScAlt1 que no tiene los primeros 84 aminoácidos
<i>ScAlt1-yEC</i>	Fusión de la proteína yECitrine al extremo carboxilo del gen <i>ScALT1</i>
<i>Scalt1Δ</i>	Delección del gen <i>ScALT1</i>
<i>ScAlt2</i>	Transaminasa de alanina citoplásmica de <i>S. cerevisiae</i>
<i>ScALT2</i>	Gen que codifica para la transaminasa de alanina citoplásmica ScAlt2
<i>ScAlt2R+70</i>	La proteína Alt2 que tiene los primeros 70 aminoácidos de ScAlt1
<i>ScAlt2R+84</i>	La proteína Alt2 que tiene los primeros 84 aminoácidos de ScAlt1
<i>ScAlt2-yEC</i>	Fusión de la proteína yECitrine al extremo carboxilo del gen <i>ScALT2</i>
<i>Scalt2Δ</i>	Delección del gen <i>ScALT2</i>
<i>ScBat1</i>	Transaminasa de aminoácidos de cadena ramificada mitocondrial
<i>ScBAT1</i>	Gen que codifica para la transaminasa de aminoácidos de cadena ramificada mitocondrial Bat1
<i>ScBat1R-17</i>	La proteína Bat1 que no tiene los primeros 17 aminoácidos
<i>ScBat1-yEC</i>	Fusión de la proteína yECitrine al extremo carboxilo del gen <i>ScBAT1</i>
<i>Scbat1Δ</i>	Delección del gen <i>ScBAT1</i>
<i>ScBat2</i>	Transaminasa de aminoácidos de cadena ramificada citoplásmica
<i>ScBAT2</i>	Gen que codifica para la transaminasa de aminoácidos de cadena ramificada citoplásmica Bat2
<i>ScBat2R+17</i>	La proteína Bat2 que tiene los primeros 17 aminoácidos de Bat1
<i>ScBat2-yEC</i>	Fusión de la proteína yECitrine al extremo carboxilo del gen <i>ScBAT2</i>
<i>Scbat2Δ</i>	Delección del gen <i>ScBAT2</i>
<i>ScGlo2</i>	Glioxilasa tipo II citoplásmica de <i>S. cerevisiae</i>
<i>ScGLO2</i>	Gen que codifica para la Glioxilasa tipo II citoplásmica ScGlo2

<i>ScGlo2R+11</i>	La proteína Glo2 que tiene los primeros 11 aminoácidos de ScGlo4
<i>ScGlo2-yEC</i>	Fusión de la proteína yECitrine al extremo carboxilo del gen <i>ScGLO2</i>
<i>Scglo2Δ</i>	Delección del gen <i>ScGLO2</i>
<i>ScGlo4</i>	Glioxilasa tipo II mitocondrial de <i>S. cerevisiae</i>
<i>ScGLO4</i>	Gen que codifica para la Glioxilasa tipo II mitocondrial Glo4
<i>ScGlo4R-11</i>	La proteína Glo4 que no tiene los primeros 11 aminoácidos
<i>ScGlo4-yEC</i>	Fusión de la proteína yECitrine al extremo carboxilo del gen <i>ScGLO4</i>
<i>Scglo4Δ</i>	Delección del gen <i>ScGLO4</i>
<i>ScTrr1</i>	Reductasa de tiorredoxina mitocondrial
<i>ScTRR1</i>	Gen que codifica para reductasa de tiorredoxina mitocondrial <i>ScTrr1</i>
<i>ScTrr1-yEC</i>	Fusión de la proteína yECitrine al extremo carboxilo del gen <i>ScTRR1</i>
<i>Scrrr1Δ</i>	Delección del gen <i>ScTRR1</i>
<i>ScTrr2</i>	Reductasa de tiorredoxina citoplásmica
<i>ScTRR2</i>	Gen que codifica para reductasa de tiorredoxina citoplásmica <i>ScTrr2</i>
<i>ScTrr2-yEC</i>	Fusión de la proteína yECitrine al extremo carboxilo del gen <i>ScTRR2</i>
<i>Scrrr2Δ</i>	Delección del gen <i>ScTRR2</i>
<i>ura3</i>	Gen inactivo de <i>URA3</i>
<i>Ura3</i>	Fosfato 5' orotidina decarboxilasa
<i>URA3</i>	Gen que codifica para la fosfato 5' orotidina decarboxilasa <i>Ura3</i>
Δ	Delección del gen

LISTA DE FIGURAS

- Figura 1. Fases que determinan la preservación de un gen duplicado.
- Figura 2. Posible destino de los genes duplicados.
- Figura 3. Posibles escenarios de relocalización subcelular posteriores a la duplicación.
- Figura 4. Logos de secuencia que ayudan a determinar las posibles presecuencias.
- Figura 5. Vías principales del importe a la matriz mitocondrial y a la membrana interna.
- Figura 6. Árbol filogenético que contiene la familia de los Ascomycotas y Basidiomycotas.
- Figura 7. Metodología denominada Delitto Perfetto.
- Figura 8. Ensayo enzimático utilizado para medir la actividad de alanina aminotransferasa.
- Figura 9. Localización subcelular de las parejas eliminadas.
- Figura 10. Reacción de las transaminasas de alanina.
- Figura 11. Localizaciones de ScAlt1 y ScAlt2.
- Figura 12. Análisis filogenético de ScAlt1, ScAlt2 y SkAlt.
- Figura 13. Localización mitocondrial de LkAlt1.
- Figura 14. Localización subcelular de ScAlt1R-70 y ScAlt1R-84.
- Figura 15. Velocidades de crecimiento en glucosa como fuente de carbono y amonio o alanina como fuente de nitrógeno.
- Figura 16. Velocidades de crecimiento en etanol como fuente de carbono y amonio como fuente de nitrógeno ScAlt1-ScAlt2.
- Figura 17. Velocidades de crecimiento en etanol como fuente de carbono y alanina como fuente de nitrógeno.
- Figura 18. Compartimentalización de la vía de síntesis y degradación de los aminoácidos de cadena ramificada en *S. cerevisiae*.
- Figura 19. Localizaciones de ScBat1 y ScBat2.
- Figura 20. Análisis filogenético de ScBat1, ScBat2, LkBat1 y LkBat1.
- Figura 21. Localización mitocondrial de LkBat1.
- Figura 22. Localización de ScBat1R-17.
- Figura 23. Velocidades de crecimiento en medio con glucosa y amonio, LIV o LIV más amonio como fuente de nitrógeno.
- Figura 24. Velocidades de crecimiento en medio con glucosa y valina isoleucina o leucina como fuente de nitrógeno.

Figura 25. Velocidades de crecimiento en medio con etanol y amonio para *ScBat1* y *ScBat2*.

Figura 26. Vía del metilglioxal.

Figura 27. Resultados de la microscopía para las glioxilasas tipo II.

Figura 28. Análisis filogenético de *ScGLO2*, *ScGLO4* y *LkGLO2*.

Figura 29. Localización subcelular de *LkGlo2*.

Figura 30. Localización subcelular de *ScGlo2R+11*.

Figura 31. Los sistemas de GSH-glutaredoxina y tiorredoxina.

Figura 32. Localizaciones de *ScTrr1* y *ScTrr2*.

Figura 33. Análisis filogenético de *ScTRR1*, *ScTRR2* y *LkTRR1*.

LISTA DE TABLAS

Tabla 1. Parejas de proteínas candidatas y su función.

Tabla 2. Características de las parejas candidatas.

Tabla 3. Cepas construidas para la determinación de localización de las parejas eliminadas.

Tabla 4. Predicción de localización subcelular.

Tabla 5. Actividades específicas de alanina transaminasas (*ScAlt1/ScAlt2*)

Tabla 6. Cepas utilizadas para estudiar la parejas *ScAlt1/ScAlt2*

Tabla 7. Actividades específicas de alanina aminotransferasa (*ScAlt1* de tipo silvestre y relocalizada).

Tabla 8. Cepas utilizadas para estudiar la pareja *ScBat1/ScBat2*

Tabla 9. Cepas utilizadas para estudiar la pareja *ScGlo2/ScGlo4*.

Tabla 10. Concentraciones mínimas inhibitorias de metilglioxal para las mutantes de *ScGLO2* y *ScGLO4*

Tabla 11. Cepas utilizadas para estudiar la pareja *ScTrr1/ScTrr2*

RESUMEN

La duplicación génica ha sido considerada el proceso evolutivo más importante para el origen de nuevos genes. Al producir copias de genes preexistentes, la duplicación provee el material para que las mutaciones, la deriva génica y la selección natural puedan actuar sobre estas copias (Byun-MacKay, 2007). Los posibles procesos evolutivos que guían el destino de las nuevas copias pueden ser los siguientes: pseudogenización (pérdida de la función), subfuncionalización (repartición de las funciones del gen ancestral) o neofuncionalización (la aparición de una nueva función).

La relocalización subcelular de proteínas (*PSR*, por sus siglas en inglés) contribuye a la diversificación funcional de los productos de una duplicación génica, los cuales pueden adquirir nuevas funciones relocalizando sus productos en diferentes compartimentos dentro de una misma célula. La adquisición de una nueva función puede deberse a que hay un nuevo grupo de interactores, concentraciones diferentes de sustratos y/o condiciones fisicoquímicas distintas en el nuevo compartimento. Esta relocalización se puede dar por la alteración del péptido señal, por un error en la duplicación que altere el 5' del gen o a través de mutaciones puntuales (Byun-MacKay 2007).

En *Saccharomyces cerevisiae* el 37% de las proteínas parálogas tiene un duplicado que se localiza al menos en un compartimento diferente. Lo que ha llevado a investigar si las funciones están influenciadas por su localización. Como ejemplo de lo anterior, en el laboratorio de la Dra. González Manjarrez se determinó que los productos parálogos *ScBat1* y *ScBat2* presentan una localización diferencial, de tal manera que *ScBat1* se localiza en la mitocondria, en tanto que *Bat2* se localiza en el citosol (Colón, 2010). Dichos genes participan en el último paso de la biosíntesis de aminoácidos de cadena ramificada (leucina, isoleucina y valina, LIV) y en el primer paso de su catabolismo. Posteriormente, se encontró que la proteína ortóloga de *Kluyveromyces lactis* está localizada principalmente en núcleo y citoplasma (datos no publicados). Estos resultados indican que, durante el proceso de retención y divergencia funcional de estos genes parálogos, la relocalización pudo haber desempeñado un papel importante en su evolución y función biológica, ya que esta propicia el contacto con nuevos interactores y sustratos, lo que podría afectar selectivamente la especificidad y otras propiedades necesarias para llevar a cabo la función. Además de lo anterior, se han encontrado otras parejas de genes parálogos con localización diferencial en el metabolismo de *S. cerevisiae*, tales como *ScAlt1/ScAlt2*, *ScCit1/ScCit2*, *ScGlo2/ScGlo4*, etc.

El trabajo de maestría que aquí se presenta se centró en la importancia de la localización subcelular y en los procesos evolutivos posteriores a la duplicación de cuatro parejas de genes parálogos, que muestran una localización diferencial, los cuales son: *ScBat1/ScBat2* (transaminasas de aminoácidos de cadena ramificada), *ScAlt1* (transaminasa de alanina)/ *ScAlt2* (no se conoce su función aunque tiene un porcentaje de identidad de 67% con *ScAlt1*), y *ScGlo2/ScGlo4* (Glioxilasas tipo II) y *ScTrr1/ScTrr2* (Triorredoxin Reductasas) se comparó su localización con el de la única enzima ortóloga de la levadura tipo ancestral *Lanchancea kluyveri*.

A partir de este trabajo se definió que al igual que *ScBat1*, *ScAlt1* es exclusivamente mitocondrial, en tanto que *ScGlo4* comparte la localización citoplásmica y mitocondrial. Por su parte, *ScAlt2*, *ScBat2* y *ScGlo2* son citoplásmicas y nucleares. También se determinó que *SkAlt* y *SkBat* se localizan en la mitocondria, y que el proceso evolutivo que llevaron a cabo las proteínas parálogas fue la neolocalización ya que *ScAlt2* y *ScBat2* se encuentran en el citoplasma y en el núcleo. *SkGlo* se localiza en el citoplasma, hasta el momento no se ha podido descartar que se encuentre en otro compartimento celular, por lo que no se ha podido definir el proceso evolutivo que sufrieron las proteínas parálogas de *S. cerevisiae*.

Para determinar la importancia de la localización en la diversificación funcional de las proteínas duplicadas, se eliminaron o añadieron presecuencias de importación mitocondrial para intercambiar la región subcelular de las proteínas parálogas y se observó la repercusión de la localización en la función de las proteínas. En cuanto a las propiedades fisiológicas, se encontró que la relocalización de *ScAlt1*, de mitocondria a citoplasma y núcleo, no tiene repercusión alguna en la velocidad de crecimiento, independientemente de la condición de estudio. Para determinar la repercusión de la relocalización de *ScAlt2* a la mitocondria es fundamental determinar primero su función.

Para la pareja *ScBat1* y *ScBat2*, pudimos encontrar que la función sí está determinada por la localización, ya que se observó que al relocalizar a *ScBat1* de la mitocondria al citoplasma y núcleo ésta tiene capacidad de catabolizar de forma similar a la cepa silvestre cuando se encuentra sola en el citoplasma (*ScBat1R-17 Scbat2Δ*), mientras que cuando se encuentra en la mitocondria (*ScBAT1 Scbat2Δ*) no es 100% eficiente para catabolizar LIV y usarlo como fuente de nitrógeno.

Este trabajo muestra por primera vez que existe una relación entre la relocalización de las proteínas parálogas (*ScBat1* y *ScBat2*) y su diversificación funcional, ya que el compartimento les proporciona un carácter catabólico o biosintético.

I) INTRODUCCIÓN

1.1 Duplicación génica

La duplicación génica ha sido considerada el proceso evolutivo más importante para el origen de nuevos genes, al producir copias de genes preexistentes, la duplicación provee el material para que las mutaciones, la deriva génica y la selección puedan actuar sobre estas copias (Byun-MacKay & Geeta, 2007). El destino de los genes duplicados puede ser: la pseudogenización, la subfuncionalización o la neofuncionalización. Sin los eventos de duplicación génica la plasticidad del genoma en las especies que se adaptan a ambientes cambiantes sería muy limitada (Zhang, 2003).

Los cambios genéticos que ocurren en un genoma atraviesan cuatro etapas: 1) origen a través de una mutación, 2) fase de fijación, 3) segregación en la población, y 4) preservación. Las duplicaciones génicas tienen esta misma trayectoria, pero poseen una característica adicional y crucial para su preservación, desde el momento de la aparición de la nueva copia, la adquisición de diferencias genéticas entre las copias resultantes altera la probabilidad de que alguna de las copias se preserve. Esta fase puede ser llamada la fase de “determinación del destino” (Figura 1).

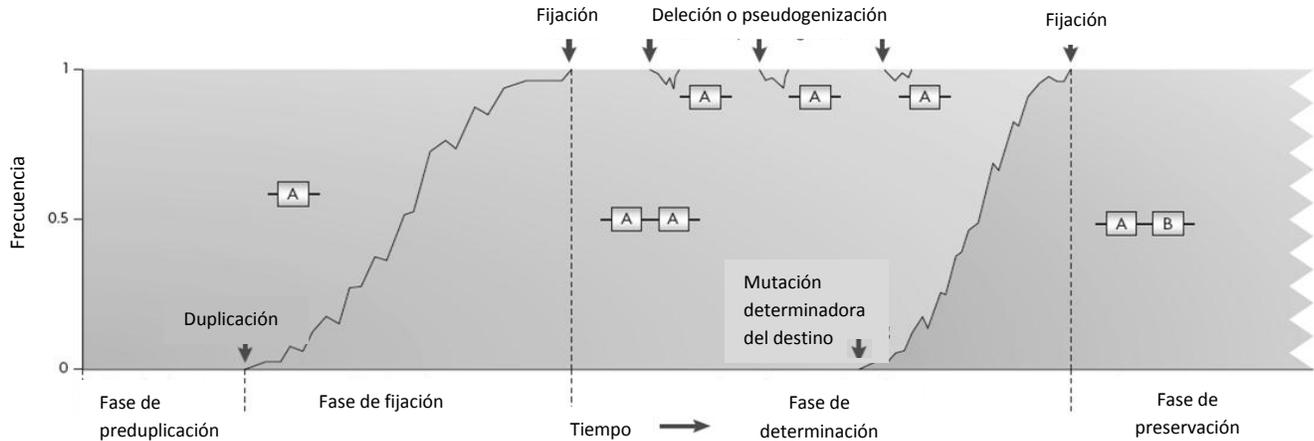


Figura 1. Fases que determinan la preservación de un gen duplicado. Primera, el origen a través de una mutación, segunda, la fase de fijación, tercera segregación en la población, y como último paso se encuentra la preservación. Obtenido de Innan & Kondrashov, 2010.

Se asume que el nuevo duplicado es indistinguible funcionalmente del original. La probabilidad de fijación de los duplicados y la longitud de la fase de fijación depende de la fuerza de selección para la pareja de parálogos (Innan & Kondrashov, 2010).

La duplicación génica puede ser el resultado de la operación de cuatro mecanismos: 1) Entrecruzamiento desigual que depende del lugar en donde ocurra el entrecruzamiento y por tanto puede duplicarse una parte de un gen, el gen entero o varios genes. El resultado de este evento son las duplicaciones en tándem que no necesariamente tienen la misma orientación, lo que explica que algunos genes duplicados tengan orientaciones opuestas. 2) Por poliploidización, en donde hay un cambio numérico en la serie completa de cromosomas, resultado de una división celular anormal o del apareamiento irregular entre dos organismos (cruza entre dos células diploides). La poliploidización tiene un efecto muy fuerte sobre la evolución de los genomas, lo cual se verá más adelante. 3) Por transposición duplicativa del DNA, esto es, mediante dos principales vías que son la recombinación homóloga no alélica (NAHR) y por unión de extremos no homólogos (NHEJ). 4) Retrotransposición que ocurre cuando un RNA mensajero (mRNA) es retrotranscrito a DNA complementario (cDNA) e insertado al genoma, cabe mencionar que dentro de esta duplicación hay pérdida de intrones y secuencias regulatorias (Hahn, 2009). Cuando se pierden secuencias regulatorias, al gen duplicado frecuentemente le faltan elementos necesarios para la transcripción y se convierte en un pseudogén, a menos que se haya insertado río abajo de una secuencia promotora no relacionada (Long, 2003; Hahn, 2009).

En el proceso de poliploidización, las redundancias genómicas existen en varios niveles: en el genoma completo, cromosomas, genes y regiones regulatorias. Cada nivel de redundancia está sujeto a procesos de mutación y deriva génica aleatoria, la cual, puede llevar a la pérdida de función del gen o de la secuencia regulatoria. Si los cromosomas duplicados pierden diferentes genes, entonces la viabilidad de los organismos dependería de que los cromosomas se complementen y guarden copias funcionales presentes antes de la duplicación (Zhang, 2003).

1.2 Posible destino de los genes duplicados

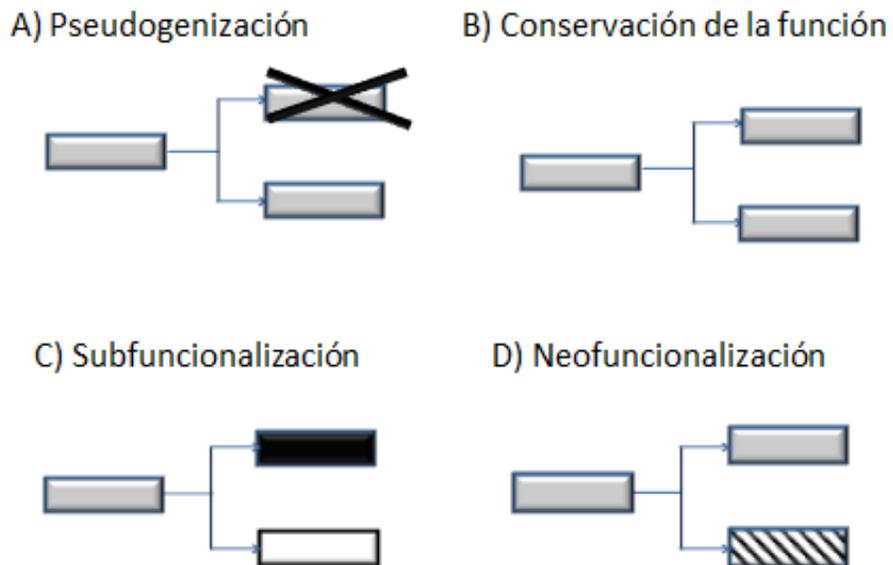
Posterior a la duplicación hay cuatro posibles destinos evolutivos que pueden sufrir cualquiera de las dos copias son (Figura 2):

A) Pseudogenización. Se relaciona con mutaciones que modifican la secuencia y, por lo tanto, la función de uno de los genes duplicados. Estas mutaciones tienen generalmente, un efecto deletéreo, es decir, que pueden llevar al cese de la expresión o de la función; posteriormente el pseudogén puede ser eliminado del genoma. Se ha determinado que a través de la evolución este es el proceso más frecuente, ya que las mutaciones generalmente son deletéreas, lo que explica que los organismos que tuvieron una duplicación total del genoma sólo mantengan una pequeña parte de sus genes duplicados. (Kimmura & King, 1979; Ohno, 1972).

B) Conservación de la función. La presencia de genes duplicados llega a ser benéfica simplemente porque hay una cantidad extra de proteína o de productos de RNA que presentan una gran demanda, como histonas o RNAs ribosomales. La otra explicación de la conservación de las dos copias, es que el segundo gen puede proveer redundancia funcional, y en caso de que el locus original se deshabilite por una mutación, se tendría esa segunda copia para “contrarrestar” el daño (Ohno, 1970).

C) Neofuncionalización. Fue descrita por Ohno en los 70's. Este modelo señala que, anterior a la duplicación, una de las dos copias es suficiente para cumplir la función requerida, por lo que las copias extra son redundantes. Si la copia nueva se fija en la población, la copia original mantendrá su función, y la nueva será liberada de la selección negativa, por lo que podrá acumular mutaciones y convertirse en un pseudogén, pero ocasionalmente, acumulará sustituciones que la lleven a tener una nueva función que será retenida por selección. Este modelo predice que la tasa de evolución (mutaciones por generación) se acelerará en la copia y se mantendrá, igual que antes de la duplicación, en la original (Ohno, 1970; Zhang, 2003).

Figura 2. Posible destino de los genes duplicados. Pseudogenización (pérdida de la función o expresión de alguna de las copias), Conservación (permanecen iguales las dos proteínas), Neofuncionalización (Ganancia de una función no conocida por el ancestro) y subfuncionalización (partición de las funciones de la proteína ancestral). Modificada de Conrad, 2007.



D) Subfuncionalización. Piatigorsky (2007) propuso que la duplicación se encuentra precedida por un periodo de ‘gene sharing’, en el cual una sola proteína lleva a cabo dos funciones distintas; posterior a la duplicación se tienen dos genes que codifican para proteínas funcionalmente complementarias a la proteína ancestral. La especialización también puede ser alcanzada por un

cambio en los patrones de expresión temporal de uno o ambos genes (Hughes, 1994; Force, 1999; Taylor & Raes, 2004).

1.3 Divergencia funcional de los genes duplicados

Anteriormente se planteó que posterior a la duplicación las copias pueden tener varios destinos, de los cuales sólo la neofuncionalización y la subfuncionalización presentan una divergencia funcional, pero, ¿Cómo se logra la divergencia funcional?

Se ha demostrado, a través de experimentos llevados a cabo en múltiples especies, que puede haber mutaciones en diferentes regiones, las primeras se localizan en la región codificante, donde la función puede variar haciendo cambios en el sitio activo, de reconocimiento del sustrato, o en la región de interacción con otras proteínas. Los segundos se relacionan con cambios en la expresión que pueden llevarse a cabo en la región cis-regulatoria o en genes reguladores.

Los cambios en la región cis-regulatoria y los cambios de la región codificante son independientes y son mucho más frecuentes en la región regulatoria, por lo que se puede encontrar, en una pareja de genes duplicados, mayor divergencia en los niveles de expresión o en los patrones de regulación.

También se ha planteado la idea de que los productos de los genes duplicados pueden tener cambios en la localización subcelular y que esto contribuye a la diversificación funcional. De manera general, se ha propuesto que la alteración de un péptido localizado en la región N-terminal en las proteínas, produce cambios en localización subcelular de una proteína duplicada, lo que puede ser un proceso que lleve a que las proteínas tengan nuevas funciones. Esta relocalización se puede dar no sólo por la alteración del péptido señal, sino también por un error en la duplicación que altere la región 5' del gen o a través de mutaciones puntuales.

1.4 Importancia de la localización subcelular

La localización subcelular de una proteína es la clave para su función dentro de la célula, dado que las distintas funciones celulares se localizan en compartimentos específicos. De tal forma, que el cambio en la localización de una proteína para cierto compartimento puede ser un proceso evolutivo que lleva al origen de nuevas funciones (Byun-McKay & Geeta, 2007). Lo anterior puede deberse a que hay un nuevo grupo de interactores, cantidades diferentes de sustratos y/o condiciones fisicoquímicas distintas en el nuevo compartimento.

Se ha demostrado que los péptidos señal evolucionan el doble de rápido que el resto de la proteína, por lo que se puede alterar la localización subcelular de una proteína duplicada fácilmente (Williams, 2000), lo que permite la adquisición de un nuevo patrón de expresión subcelular y posiblemente una nueva función bioquímica.

Un ejemplo donde la función que lleva a cabo una proteína está altamente influenciada por su localización subcelular, lo presenta la enzima que degrada la insulina (*IDE*) de humano, la cual, es una Aminopeptidasa de zinc que se expresa altamente en el citosol. En un estudio se relocalizó esta proteína artificialmente hacia la mitocondria y mostró tener nuevos blancos, degradando cierto tipo de péptidos (Lessring et al., 2004). En algunos casos, una proteína que se dirige a un nuevo compartimento subcelular cambia su eficiencia y no su función enzimática. Por ejemplo, la Liasa de metilmercurio (*MerB*) es una proteína bacteriana que ayuda a reducir los efectos tóxicos del metilmercurio, al introducir esta proteína a *Arabidopsis* mostró tener 70 veces más actividad cuando se encontraba en el retículo endoplásmico (RE) comparado con su localización original, el citoplasma. Los autores de este trabajo concluyeron que el ambiente hidrofóbico proveía mejores condiciones para la reacción catalizada por esta proteína y por eso su actividad aumentaba en el nuevo compartimento (Bizily et al., 2003). Otro ejemplo de que la localización es un factor importante para la función de una proteína es el factor transcripcional *LEAFY*, de *Arabidopsis thaliana* que es capaz de relocalizarse del núcleo al citoplasma donde puede estar involucrado en la señalización intercelular (Sessions et al., 2000).

En *Saccharomyces cerevisiae* el 37% de las proteínas parálogas tienen un duplicado que se localiza al menos en un compartimento diferente (Marques et al., 2008). Uno esperaría que las funciones ancestrales impusieran límites con respecto al potencial de diversificación subcelular de los duplicados. Lo que llevó a investigar si las funciones estaban influenciadas por su localización. Ahora se sabe que las proteínas del catabolismo están localizadas en promedio en 1.47 compartimentos, mientras que las proteínas de la biosíntesis ocupan en promedio 1.35 compartimentos (Marques et al., 2008).

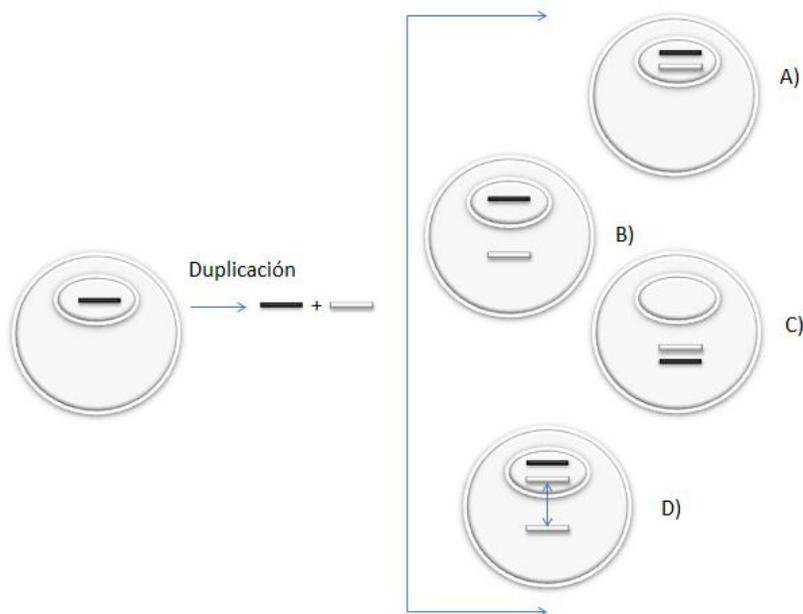
La translocación de las proteínas generalmente demanda un transporte activo dependiente de ATP o GTP o por potencial de membrana. La señal de reconocimiento, en la mayoría de los casos, se presenta y es detectable a nivel de secuencia primaria, y predominantemente se ubica en el extremo N-terminal. Los péptidos señal generalmente son de 13 a 36 aminoácidos y muestran poca identidad de secuencia entre sí, aunque también existen proteínas que tienen inmersas en su

región codificante la señal de localización. Las proteínas que no tienen señal de localización permanecen en el citoplasma. (Bannai et al., 2002).

Posterior al traslado de las proteínas a su lugar nativo, los péptidos señal localizados en el extremo amino, generalmente son cortados y degradados, por lo que se piensa que no tienen participación directa en la función.

Posterior a una duplicación génica puede ocurrir uno de los siguientes modelos: 1) que las dos proteínas mantengan la misma localización que tenían antes de la duplicación, 2) que una de las dos se relocalice a un compartimento diferente, 3) que las dos cambien su localización anterior y 4) que una de ellas permanezca en el mismo compartimento y que la otra proteína pueda tener dos localizaciones simultáneas (Figura 3). Se han acuñado términos que ayudan a diferenciar entre los procesos evolutivos que corresponden a la localización, como sublocalización cuando las regiones subcelulares del ancestro se reparten entre la pareja de duplicados, o la neolocalización, cuando un componente de la pareja paróloga tiene una localización distinta de la del ancestro (Marques, et al., 2008).

Figura 3. Posibles escenarios de relocalización subcelular posteriores a la duplicación: A) Las dos proteínas permanecen en el mismo compartimento, B) Una de las proteínas desarrolla la capacidad de encontrarse en un nuevo compartimento y su pareja continúa en el compartimento ancestral, C) Las dos proteínas se relocalizan, y D) Una permanece en el compartimento ancestral y la otra comparte dos localizaciones.



Una vez que una proteína se relocaliza el efecto sobre la adecuación₁ del organismo puede ser el siguiente:

1) Que no tenga un efecto sobre la adecuación, 2) que se tenga un efecto negativo en la adecuación y 3) que haya un efecto positivo en la adecuación del organismo.

1.- Adecuación. Es una característica típica del genotipo que describe la habilidad de sobrevivir y reproducirse y es igual a la proporción de genes del individuo presentes en la siguiente generación.

En el primer caso, el resultado puede ser que el gen duplicado se fije en la población, y mutaciones subsecuentes en el péptido señal puedan resultar en una segunda relocalización, dando una nueva oportunidad para que esta proteína pueda adquirir una nueva función. En el segundo escenario, el efecto negativo producido por la relocalización del duplicado puede resultar en su eliminación por selección purificadora. Y en el último caso, la proteína podría adquirir una nueva función bioquímica, como resultado de esta localización, la selección purificadora actuaría previniendo mutaciones que puedan destruir o minimizar su función (Byun-MacKay & Geeta, 2007).

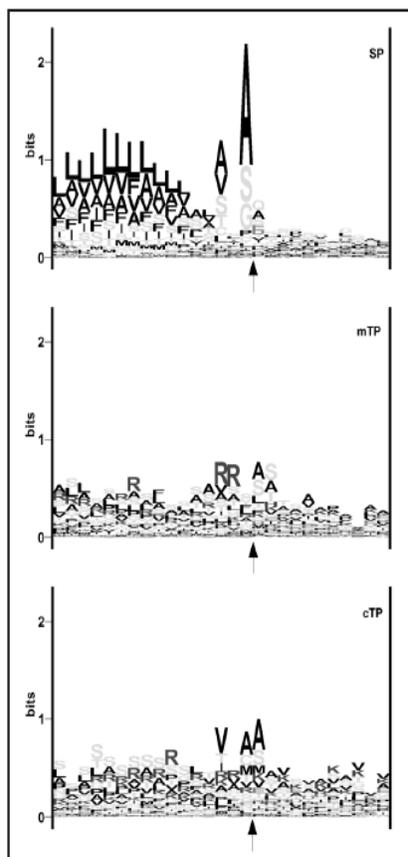


Figura 4. Logos de secuencia que ayudan a determinar las posibles presecuencias, forma de visualizar un alineamiento múltiple de secuencias y específicamente el grado de la conservación de los aminoácidos. El eje de las 'y' se mide en bits, y para secuencias de proteínas el valor máximo es de 4.3 que corresponde a una posición totalmente conservada. Obtenida de Emanuelsson, 2008.

1.5 Predicción de localización subcelular y características de las señales de localización

Debido a que a los péptidos señal no tienen secuencia consenso, el destino de la proteína se puede predecir en base a características encontradas en la secuencia primaria, como alta o baja cantidad de aminoácidos básicos o la presencia de α hélices.

Los métodos predictivos de localización subcelular utilizan redes neuronales artificiales o modelos ocultos de Markov que basándose en la construcción manual reglas específicas sobre señales encontradas hasta el momento. (Figura 4) (Emanuelsson, 2002). A continuación se mencionan las características de las señales de localización de algunos compartimentos:

Mitocondria. Los péptidos señal de mitocondria están enriquecidos en residuos cargados positivamente (arginina en su mayoría), carecen de residuos negativos y tienen la capacidad de formar α -hélices anfipáticas. Las estructuras anfipáticas son importantes para la unión a los receptores de la membrana externa mitocondrial y la carga positiva neta puede ser requerida durante el importe dirigido por cambio de voltaje a través de la membrana interna (Emanuelsson & Von Heijne, 2001).

Membrana. No hay una señal consenso para la localización en las diferentes membranas en la célula, pero la mayoría de las regiones transmembranales están presentes en forma de α -hélice, con la excepción de las proteínas con forma de barril β . Para proteínas transmembranales el hecho de que haya regiones más hidrofóbicas en la secuencia ha sido el indicador utilizado para la predicción de localización.

Núcleo. Las proteínas transportadas al núcleo están en su forma plegada. Se han detectado algunas vías de importación, las cuales, involucran a proteínas acarreadoras (importinas) que forman un complejo que es subsecuentemente translocado al poro nuclear, donde se disocia y la importina retorna al citoplasma para su reutilización. Las señales de importación al núcleo se pueden encontrar en cualquier región de la proteína, y se pueden identificar dos tipos de secuencias, las monopartitas, que tienen aminoácidos básicos (PKKKRK) o bipartitas (KRPAATKAGQAKKKLKD) (Gorlich & Kutay, 1999).

Peroxisoma. Proteínas que se localizan en el peroxisoma tienen uno de dos tipos de señales: las que se encuentran en la región N-terminal y las que se componen de tres residuos C-terminal. La secuencia consenso de la esta última es Serina-Lisina-Leucina (Emanuelsson, 2002).

1. 6 Importación mitocondrial

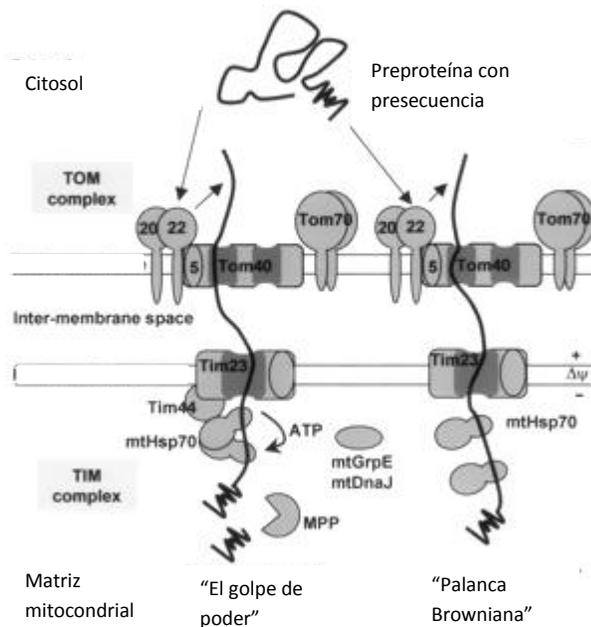


Figura 5. Vías principales del importe a la matriz mitocondrial y a la membrana interna. Obtenido de Matsuoka, 2005

Las células eucariontes están divididas en numerosos organelos, compartimentos rodeados por membranas. A pesar de esto, la mayoría de las proteínas son sintetizadas en un sólo compartimento, el citoplasma. Alrededor del 50% de estas proteínas tienen que ser transportadas a algún compartimento celular para llevar a cabo su función de la manera adecuada.

La mitocondria es un organelo que interviene en distintos procesos celulares incluyendo metabolismo energético, apoptosis, y vías de síntesis de lípidos, aminoácidos y hierro. Estos organelos tienen una membrana externa, una interna, que forma invaginaciones o crestas, que dan lugar a dos compartimentos acuosos, la matriz y el espacio intermembranal.

La mitocondria contiene alrededor de mil proteínas diferentes (en *S. cerevisiae*), de las cuales sólo el 1% se sintetiza en los ribosomas de la matriz mitocondrial, las restantes se sintetizan en los ribosomas del citosol (Bolender et al., 2008).

El 99 % de las proteínas necesarias para llevar a cabo las funciones de la mitocondria son sintetizadas en el núcleo, esto se debe a que hubo una transferencia de genes provenientes del ancestro procarionte (que dio lugar a la mitocondria) hacia el núcleo (Schmidt, 2010).

Ciertos organismos, tales como la levadura, que presentan un metabolismo facultativo, pueden vivir sin la cadena respiratoria, aunque la maquinaria de importación mitocondrial es esencial para la viabilidad (Neupert & Herrmann, 2007).

Para determinar la localización dentro de la mitocondria también se tienen ciertos códigos, por ejemplo, las proteínas que se dirigen a la membrana interna adicionalmente contienen residuos hidrofóbicos. Se ha sugerido que sólo el 60% de las proteínas que se localizan en este compartimento necesitan una extensión en la región N- terminal.

Las proteínas “precursoras” sintetizadas en el citoplasma permanecen en un estado sin estructura o con estructura parcial unida a chaperonas, como Hps70. Posteriormente el precursor se une a un receptor localizado al lado de una proteína translocadora, que lo ayudará a pasar a través de la membrana externa de la mitocondria.

Hay dos modelos principales para la importación de proteínas a la mitocondria (el paso del citoplasma hacia la matriz mitocondrial) que son los siguientes (Figura 5): “La palanca Browniana y “el golpe de poder”. En el primero, las preproteínas, después de la inserción mediada por el cambio de voltaje al canal TIM23, se mueven hacia adelante y hacia atrás al azar debido al movimiento browniano y la chaperona Hsp70 mitocondrial (Hsp70mt) se une a la preproteína, previniendo el movimiento Browniano y favoreciendo la translocación a la matriz. En el segundo modelo, Tim44 funciona como un sitio de anclaje de la Hps70mt al complejo TIM23 y la hidrólisis del ATP llevada a cabo por el dominio de ATPasa de la Hsp70mt induce un cambio conformacional de esta chaperona, el cual genera una fuerza de atracción en la preproteína y lleva a la liberación de Hsp70mt del complejo Tim44 (Matsuoka et al., 2005; Liu et al., 2003).

1.7 Importación nuclear

La separación espacial entre la transcripción y la traducción proporciona a los eucariontes una poderosa forma de controlar la expresión de los genes, pero también se necesita que haya un

transporte selectivo entre el citoplasma y el núcleo. Cabe mencionar que se han encontrado diversas proteínas que dentro del núcleo cumplen una función de regulador transcripcional, mientras que en el citoplasma cumplen otras funciones (Peláez, 2011; Sessions et al., 2000; Jauhainen et al., 2012).

Las macromoléculas que son mayores de 40kD son transportadas a través de la envoltura nuclear por medio de NPCs (Complejos de Poros Nucleares, por sus siglas en inglés) usando factores solubles de transporte o moléculas acarreadoras que tienen un ciclo entre el citoplasma y el núcleo. Los NPCs son enormes ensamblados que perforan la envoltura nuclear y que conducen un intercambio bidireccional de las moléculas entre el citoplasma y el núcleo a través de un canal central que tiene un diámetro de entre 25 y 30nm (Feldherr, 2001). Los Complejos de Poros Nucleares se construyen de copias múltiples de alrededor de treinta proteínas que colectivamente se denominan nucleoporinas (Rout & Went, 1994; Rout et al., 2000).

Hay muchas vías de transporte nuclear, pero tienen en común ciertos aspectos: los cargamentos (proteínas) son reconocidos en el citoplasma, translocados a través de los NPCs y liberados al núcleo. En cada una de las vías, los cargamentos son llevados a su importe nuclear debido a que tienen motivos de señales de localización nuclear (NLS). Hay diferentes clases de motivos NLS, cada uno es reconocido por los diferentes componentes de las diversas vías de transporte nuclear. La mayoría de estas vías utiliza una familia homóloga de moléculas acarreadoras que colectivamente se llaman β -carioferinas junto con importinas y exportinas (Görllich & Kutay, 1999; Macara, 2001; Stewart, 2007).

1.8 Evolución del genoma de las levaduras.

El género *Saccharomyces* incluye dos grupos de especies; *Saccharomyces sensu stricto*, que tienen un metabolismo similar a *S. cerevisiae* y *Saccharomyces sensu lato*, que son especies más alejadas a *S. cerevisiae*. Las levaduras que componen el grupo de las *sensu stricto* son *Saccharomyces cerevisiae*, *S. bayanus*, *S. cariocanus*, *S. kudriavzevii*, *S. mikatae* y *S. paradoxus* (Raineri, 2008). El ancestro de las levaduras *sensu stricto* tuvo una duplicación total del genoma seguida de la pérdida del 90% de sus duplicados. La especie *sensu lato* incluye a *S. barnettii*, *S. castellii*, *S. dairenensis*, *S. exiguus*, *S. humaticus*, *S. kunashirensis*, *S. martiniae*, *S. naganishii*, etc. (Mikata, 2001; Spírek, 2003).

Se ha estimado que *S. cerevisiae* divergió del ancestro común de *S. paradoxus* y *S. cariocanus* hace 5-10 millones de años (MA) y que *S. cariocanus* y *S. paradoxus* divergieron posteriormente. En cuanto al tiempo de divergencia de *S. cerevisiae* con *S. mikatae*, con *S.*

kudriavzevii y con *S. bayanus* fue de 10-15, 15-20 y 20 MA, respectivamente (Figura 6). Lo que sugiere que este grupo es relativamente joven en la evolución (Raineri et al., 2003; Replansky et al., 2008).

Los organismos que se pueden considerar de tipo ancestral, debido a que no tuvieron una duplicación total del genoma, son los que se encuentran en el género *Lanchacea* y *Kluyveromyces* (Figura 6). Se ha considerado que la especie más cercana al ancestro de *S. cerevisiae* es *L. kluyveri* por su metabolismo, su conservación génica, etc.

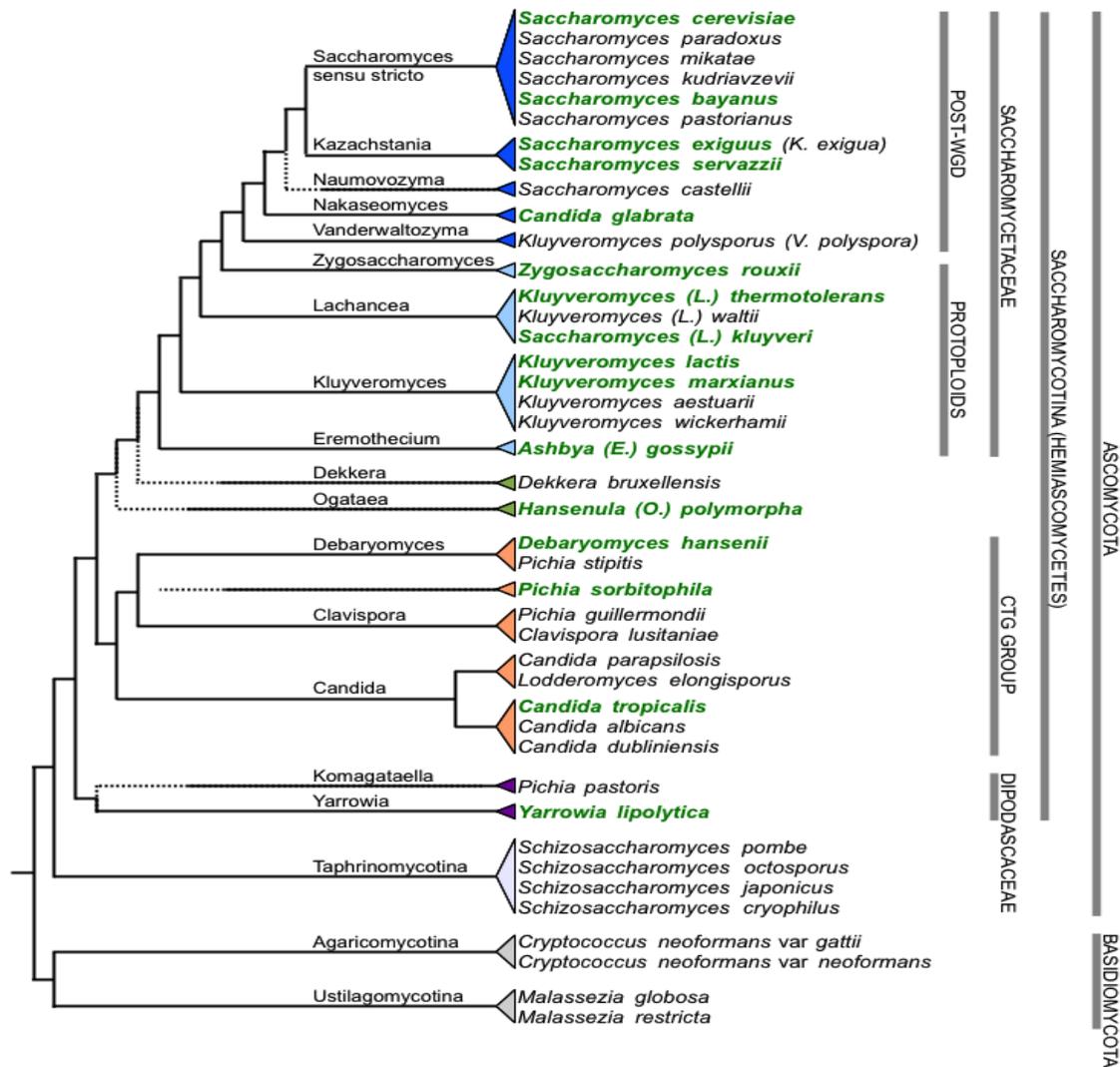


Figura 6. Árbol filogenético que contiene la familia de los Ascomycotas y Basidiomycotas.

En este trabajo se estudiarán dos organismos, *Saccharomyces cerevisiae* y *Lanchanea kluyveri*, este último se tomará como ancestro, las características de ambos son las siguientes:

1.8.1 *Saccharomyces cerevisiae*

Esta levadura es un hongo levaduriforme que es utilizado industrialmente en la fabricación del pan, cerveza y vino. Tiene un ciclo de vida alterno, haploide y diploide. Ambas formas se reproducen de forma asexual por gemación. En condiciones apropiadas dos formas haploides se fusionan, generándose un diploide. Durante la meiosis se forma un saco, llamado asca, que contiene 4 ascosporas haploides.

Este organismo es uno de los organismos modelo más estudiado, ya que es un eucarionte unicelular, tiene un crecimiento rápido, tiene una gran facilidad para dispersar las células y aislar mutantes. Por último, se ha creado un sistema muy versátil de transformación, ya que su tasa de recombinación homóloga es mucho mayor que en cualquier otro *ascomiceto*.

Mantiene dos tipos de metabolismos muy bien diferenciados. Por una parte, en condiciones en las que existen altas concentraciones de azúcares, tales como la glucosa, fructosa o maltosa, la levadura tiende a realizar una fermentación alcohólica, teniendo como producto final etanol. Una vez que los azúcares se encuentran en menor cantidad se produce la respiración, vía ciclo de Krebs. Evolutivamente este es un proceso que se ha perfeccionado a través del tiempo y con la ventaja de que su genoma se duplicó hace 100 MA.

Contiene un genoma con 16 cromosomas completamente caracterizados, cuyos tamaños varían desde 200 a 2,200 kb y que componen un genoma de un total de 12 Mb (Sherman et al., 2002).

1.8.2 *Lanchancea kluyveri*.

El hábitat de *L. kluyveri* no ha sido caracterizado completamente, hasta el momento se han encontrado alrededor de 30 aislados naturales. A pesar de esto, esta especie parece estar altamente dispersa en el ambiente.

Tiene varias diferencias cuando se compara con *S. cerevisiae*: 1) Utiliza la fermentación y degrada los azúcares sólo en la ausencia de oxígeno; 2) su genoma tiene una organización “más simple” ya que es una de las especies anteriores a las que tuvieron la duplicación total del genoma. Esta levadura tiene un genoma de aproximadamente 11.3 Mb, organizado en 8 cromosomas; 3) Su uso se ha expandido en el área de producción de proteínas, ya que su producción de biomasa es mucho mayor. 4) Tiene una tasa de recombinación homóloga mucho menor que la *S. cerevisiae*. 5)

Puede usar pirimidinas y sus productos de degradación como única fuente de nitrógeno (Gojkovic et al., 1998; Neuvéglise et al., 2000).

II) PLANTEAMIENTO DE LA INVESTIGACIÓN

2.1 Justificación

Alrededor del 37% de las proteínas parálogas de *Saccharomyces cerevisiae*, poseen una localización subcelular diferencial. Es decir, cada uno de los productos parálogos se encuentran localizados en diferentes compartimentos. Esto ha llevado a proponer que los cambios en la localización subcelular de proteínas parálogas son determinantes para la diversificación funcional de las mismas (sea esta neofuncionalización o subfuncionalización); lo que se ha denominado como PSR (Relocalización subcelular de proteínas); sin embargo, hasta el momento, no existe un trabajo que compruebe la existencia de una relación entre la relocalización de las proteínas parálogas y la diversificación de la función de las mismas.

Por lo tanto, el interés fundamental de este trabajo es resolver la siguiente pregunta:

1) ¿Cuál es la implicación de la relocalización subcelular en la diversificación funcional de las proteínas parálogas?

Para contestar a esta pregunta se intercambiará la localización subcelular entre los componentes de tres parejas de parálogos, de tal manera que si en una cepa silvestre el primer componente se encuentra localizado en la mitocondria y el segundo componente se localiza en el citoplasma, en las cepas mutantes el primer componente se localizará en el citoplasma y el segundo en la mitocondria.

La forma de evidenciar la implicación de la relocalización subcelular de proteínas se ha de llevar a cabo, analizando los cambios fisiológicos que se observan en *S. cerevisiae* como respuesta al intercambio de localización entre las proteínas parálogas. Para identificar el proceso evolutivo llevado a cabo por las proteínas parálogas, es necesaria la comparación de la localización subcelular entre dichas proteínas y sus ortólogos, provenientes de un organismo tipo ancestral. Por lo que una pregunta subsecuente es:

2) ¿Cuál es la localización de las proteínas parálogas de *S. cerevisiae*? y ¿cuál es la localización subcelular de la proteína ortóloga de un organismo tipo ancestral como *L. kluyveri*?

Esto último es relevante ya que nos ayudará a determinar los procesos evolutivos que se refieren a la localización que han sufrido las proteínas parálogas, es decir si hubo una neolocalización (ganancia de una localización no existente en la proteína ancestral) o una

sublocalización (repartición de las localizaciones subcelulares presentadas por la proteína ancestral entre las dos copias parálogas).

2.2 Hipótesis

Si existe una relación entre la relocalización de las proteínas parálogas y la diversificación de la funcional (neofuncionalización o subfuncionalización) de las mismas, entonces al momento de intercambiar su localización habrá diferencias fenotípicas que pueden ser analizadas.

2.3 Objetivo General

Determinar las implicaciones fisiológicas y evolutivas de la compartimentalización diferencial de proteínas parálogas en *Saccharomyces cerevisiae*.

2.4 Objetivos Específicos

- 1) Determinar la localización subcelular de un conjunto de proteínas parálogas en *S. cerevisiae* y de su ortólogo en *L. kluyveri*.
- 2) Intercambiar la localización subcelular de las proteínas parálogas.
- 3) Caracterizar fisiológicamente las consecuencias de la relocalización de las proteínas parálogas de *S. cerevisiae*.

2.5 Metas:

-Seleccionar parejas de proteínas parálogas que presenten al menos una localización diferencial (i.e una isoforma citoplásmica y otra mitocondrial).

-Determinación *in silico* de presecuencias involucradas en los procesos de compartimentalización.

-Realizar un análisis filogenético para determinar la dinámica evolutiva de las presecuencias en proteínas a estudiar.

-Fusionar los parálogos seleccionados con la proteína yECitrine –variante de la GFP- en *S. cerevisiae* y determinar mediante microscopía de fluorescencia la localización subcelular de las proteínas parálogas a estudiar.

-Sobrexpresar mediante promotor tet07 (bacteriano) aquellas proteínas que no se han podido visualizar bajo promotores endógenos.

-Remover/Adicionar las presecuencias para expresar simultáneamente en el mismo compartimento las proteínas productos de genes parálogos –relocalización artificial-.

-Fusionar las proteínas de *Lanchancea kluyveri* con yECitrine y determinación de localización subcelular.

-Caracterizar fisiológicamente las parejas relevantes.

III) METODOLOGÍA

3.1 Condiciones de crecimiento y cepas

Las cepas se crecieron en medio rico YPD (1% extracto de levadura, 2% peptona de caseína y 2% glucosa) y en medio mínimo con sales, elementos traza y vitaminas siguiendo la fórmula la base nitrogenada de levadura (Difco). Se usó como fuente de carbono glucosa (2%), etanol (2%) ó glicerol (3%). Dependiendo de cada experimento, se usó como fuente de nitrógeno: sulfato de amonio (40mM), alanina (7mM), valina (150mg/ml), isoleucina (30mg/ml), leucina (100mg/ml) o una combinación de éstos. Se complementaron auxotrofías, según correspondía, con uracilo (20mg/ml), metionina (20 mg/ml), histidina (20 mg/ml) y/o leucina (100mg/ml). Todas las cepas se crecieron a 30°C, los cultivos se agitaron a 250rpm.

Las cepas silvestres utilizadas en este trabajo fueron las siguientes: 1) *S. cerevisiae* BY4741 (*MATa his3 leu2 met15 ura3*), 2) *S. cerevisiae* Cla1 (*MATa ura3 leu2::LEU2*), 3) *S. cerevisiae* R1158 (*MATa URA::CMV-tTA his3 leu2 met15*), 4) *L. kluyveri* Y156 (*MATa ura3*). La lista de construcciones hechas se detallará en cada una de las secciones correspondientes.

3.2 Selección de proteínas candidatas

Se tomaron 88 de las 237 parejas del análisis de Marques, et al. (2008), de esas 88 parejas de proteínas parálogas fueron seleccionadas 13 en base a los siguientes criterios:

1) Que tuvieran una localización diferencial.

1) Se buscó preferentemente que las proteínas tuvieran actividad enzimática para posteriormente llevar a cabo la caracterización fisiológica (curvas de crecimiento) y bioquímica (cinética enzimática).

2) Porcentaje de identidad mayor al 40%, valores por arriba de este porcentaje son un buen parámetro para determinar si dos genes provienen de una duplicación.

3) Que una de las parejas presentara una presecuencia o un punto de relocalización claro.

3.3 Análisis *in silico* de las características de las proteínas candidatas

La secuencia de las proteínas parálogas de *S. cerevisiae*, junto con sus funciones y el nombre sistemático se obtuvieron de *Saccharomyces Genome Database* (Cherry et al., 1997).

Los datos de expresión dados por el análisis cuantitativo SDS/PAGE Western-blot (Ghaemmaghami, 2003) y la localización de las proteínas de *S. cerevisiae* fueron obtenidos de la base de datos llamada *Yeast GFP Fusion Localization Database* (Huh et al., 2003).

Para obtener la predicción de localización subcelular de los duplicados en *S. cerevisiae* se usaron las siguientes servidores: BaCeLlo (Pierleoni et al., 2006), Cello (Yu et al., 2006), pTarget (Guda, 2006), Targetp (Emanuelsson et al., 2007) y Mitoprot (Claros & Vincens, 1996).

La determinación de las presecuencias se hizo a través del alineamiento de las proteínas parálogas por medio de Geneious, los parámetros fueron los siguientes: penalidad de apertura de gap de 12, penalidad de extensión de gap de 3. Se utilizó una matriz de peso Blosum62, y se hicieron 5 iteraciones de refinamiento. Además se sometieron al programa Mitoprot, que predice también la posible presecuencia de localización mitocondrial.

3.4 Construcción de cepas

El procedimiento utilizado para la transformación de las cepas se obtuvo de Gietz & Schiestl publicado en el 2007, al cual se le llamó “High Efficiency Transformation in yeast”. El protocolo es el siguiente:

1. Pre-inocular en 10 ml de YPD en un Falcon de 50 ml en agitación a 30°C por toda la noche.
2. Del cultivo anterior hacer un nuevo inóculo en un matraz con 50ml de YPD a una DO=.380.
3. Incubar a 30°C por un periodo de 3 a 5 horas, hasta que llegue a una DO de 1.04. (La eficiencia permanece alta todavía con uno que esté en 2).
4. Centrifugar a 2500rpm por 5 minutos
5. Tirar el sobrenadante. Resuspender en 25 ml de agua estéril.
6. Centrifugar nuevamente.
7. Tirar el agua. Resuspender las células en 1 ml de LiAc 100mM y transferir la suspensión en un tubo eppendorf de 2ml.
8. Centrifugar por 15s a 13000 rpm. Remover el LiAc con una pipeta.
9. Resuspender las células en 400ul de LiAc 100mM.

*** si la densidad fue mayor de 1, el volumen de acetato de litio debe de ser ajustado. (DO 1.1=440µl)

10. Tomar 50ul de las muestras para cada transformación en un tupo de 1.5ml.
11. Centrifugar y posteriormente remover el LiAc.
12. Añadir 10µl de ssDNA, 2ug de muestra (1000ng=1µg) y vortexear.
13. Añadir 276ul a cada muestra de la mezcla PEG-LiAc (PEG 50%-240µl y 36µl de LiAc 1M). Con cuidado, en una de las paredes.
14. Aplicar el vórtex hasta que se deshaga la pastilla.
15. Incubar por 30 min a 30°C a 650 rpm.
16. Dar choque térmico a 42°C (lleva una leve agitación cada 4-5min).
17. Si hay resistencia: Centrifugar, quitarle el sobrenadante, añadir 1ml de YPD e incubar por 2 o 3 horas a 30°.
18. Centrifugar. Remover sobrenadante.
19. Resuspender la pastilla en 600µl de agua estéril.
20. Platear 150ul de la suspensión anterior en cajas de selección.

Los siguientes subtemas describen cómo se construyeron las cepas utilizadas en este trabajo, se puede observar que el protocolo utilizado para la transformación de las cepas fue el anteriormente mencionado.

3.4.1 Fusión de proteínas de *Saccharomyces cerevisiae* con yECitrine

Para la determinación de la localización subcelular se fusionaron las proteínas a estudiar con la proteína yECitrine, una variante de la GFP, que tiene una vida media mayor y una fluorescencia superior a ésta última, de la siguiente manera:

Se amplificó por PCR un módulo que contenía la proteína yECitrine y Ura3 (o His3) como marcador de selección, el cual tiene extremos de reconocimiento hacia el extremo C-terminal de la proteína a fusionar sin incluir su codón de término, ya que para que la fusión sea exitosa, éste codón tiene que ser remplazado. Los oligonucleótidos que se utilizaron se muestran en la tabla 1A (Anexos). El programa consistió de 5 minutos a 98°C, seguido de treinta ciclos de 1' a 98°C, 1' a 60°C y 1'30'' a 72°C, con una extensión final a 72°C por 10 minutos. Se utilizó Phusion polimerase marca NEB y un termociclador Eppendorf Mastercycler Gradient.

El vector utilizado para la amplificación de los módulos mencionados fue el pKT175 (Sheff & Thorn, 2004) o el PKT175-HIS que contienen el gen de la proteína fluorescente yECitrine más el gen de *URA3(pKAT175)* o *HIS5(pKT175HIS)* junto con su promotor, entre otros (Sheff & Thorn, 2004). El módulo completo se transformó con el protocolo "High Efficiency Transformation in

Yeast” (Gietz & Schiestl, 2007) y para su confirmación se extrajo DNA genómico de tres colonias aisladas de cada transformación se realizó la confirmación por PCR con los oligonucleótidos mostrados en la Tabla 2 de Anexos (Tabla 2A).

3.4.2 Fusión de proteínas de *L. kluyveri* con *yECitrine*

Para determinar la localización de las proteínas de tipo ancestrales se construyeron fusiones con la proteína *yECitrine* de la siguiente forma: para intercambiar el codón de término con la proteína fluorescente se construyeron tres módulos, los dos primeros eran regiones de un tamaño aproximado de 600 pb que fueron homólogos a la región codificante (antes del codón de término) y en la región 3'UTR (después del codón de término) y el último modulo estaba compuesto de la proteína *yECitrine* más *URA3* junto con sus promotores, y unos segmentos que se empalmaban 60pb con los primeros dos módulos.

Los tres PCRs independientes consistieron de 5 minutos a 98°C, seguido de treinta ciclos de 1' a 98°C, 1' a 60°C y 1'30'' a 72°C, con una extensión final a 72°C por 10 minutos. Para llevar a cabo el PCR la mezcla tenía, para los dos primeros módulos, el genómico de *L. kluyveri* y para el tercero el plásmido pKT175. Se utilizó la enzima anteriormente mencionada y el mismo termociclador. Posteriormente se unieron los tres módulos, usándolos como templados y añadiendo la misma cantidad de cada uno de los tres en una nueva reacción de PCR. La diferencia con la reacción anterior es que el tiempo de extensión fue de 3'. El módulo completo fue transformado con el protocolo anteriormente mencionado y se confirmó su correcta inserción por PCR. Los oligos de amplificación de los módulos se muestran en la tabla 3A y los de confirmación en la tabla 4A.

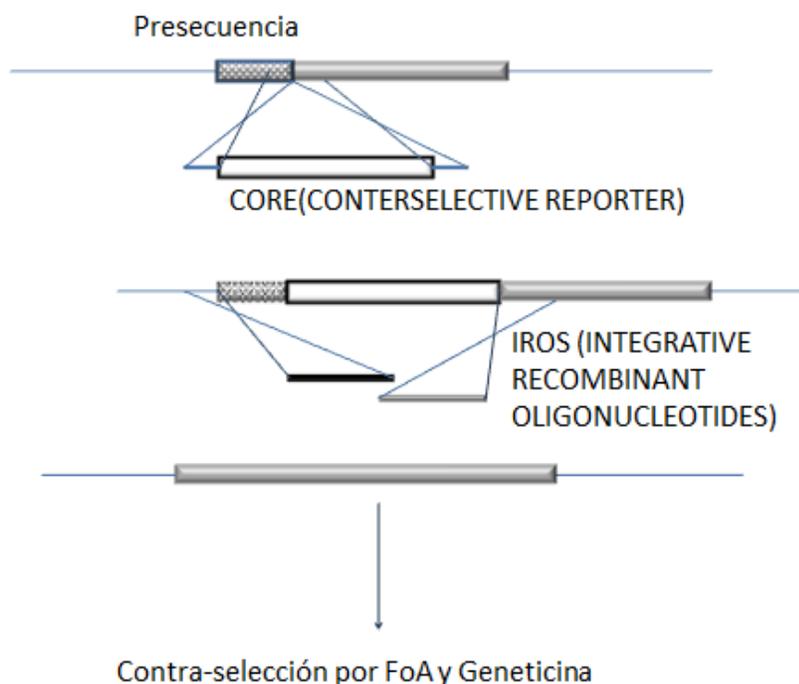
3.4.3 *Delitto Perfetto*

Esta metodología tiene como propósito la inserción, delección o intercambio de una secuencia sobre el genoma, presenta una muy baja tasa de mutaciones, lo que permite hacer cambios muy específicos. El primer paso del *Delitto Perfetto* (Figura 7) consiste en la inserción de un módulo que contiene el gen de *URA3* y el gen de resistencia a Geneticina (al cual se le llama *CORE*). Posteriormente el *CORE* es removido por la acción de dos oligonucleótidos llamados *IROS* que contienen una región de empalme entre ellos y la región que se planea insertar o que recombinan con las regiones adyacentes a las que se desea eliminar, según sea el caso.

El primer paso consistió en obtener el módulo que contenía un reportero compuesto de *URA3* y KanMX proveniente del plásmido pCORE (Storici et al., 2001) los oligonucleótidos

utilizados para la inserción del CORE se muestran en la tabla 5ª (Anexos). El módulo adicionalmente contenía regiones homólogas al promotor de los genes y a su región codificante. Se llevó a cabo de la siguiente manera: de 5 minutos a 98°C, seguido de treinta ciclos de 1' a 98°C, 1' a 63°C y 2'30'' a 72°C, con una extensión final a 72°C por 10 minutos con la enzima Phusion polymerase. Este módulo se transformó con la metodología anteriormente mencionada, y se confirmó por PCR su correcta inserción con los oligos de la tabla 7A.

Figura 7. Metodología denominada *Delitto Perfetto*. Se muestra un ejemplo donde el gen tiene una Presecuencia que posteriormente será removida. Modificada de Storicini, 2003



Posteriormente se amplificó el módulo conformado por los IROS (oligos en tabla 6A) con el siguiente protocolo: de 5 minutos a 95°C, seguido de treinta ciclos de 1' a 95°C, 30'' a 63°C y 30'' a 72°C, con una extensión final a 72°C por 10 minutos con la enzima Expand High Fidelity PCR System marca Roche. Se transformó como ya se había mencionado, se seleccionó con FoA (ácido fluoroacético) de la marca Toronto Research Chemicals y se contraseleccionó por la prototrofia a uracilo y la resistencia a Geneticina.

Para determinar que el reportero contraseleccionable ya no se encontrara insertado, es decir, que la transformación de los *IROS* fuera exitosa, se llevó a cabo una confirmación por PCR, con los oligos mostrados en la tabla 7A (Anexos).

Cabe mencionar que todas las cepas que se construyeron de ésta manera se mandaron a secuenciar por el método de Sanger en la unidad de Biología Molecular del Instituto de Fisiología celular para determinar la correcta inserción o deleción de la presecuencia. Utilizando como módulo el PCR de confirmación que fue llevado a cabo con los oligonucleótidos de confirmación (Figura 7A) Fw (en sentido) y la reacción se llevó a cabo sólo en una dirección.

3.4.4 Mutantes de deleción sencillas y dobles.

Las mutantes deletéreas se construyeron reemplazando al gen de interés con el gen de KanMX, que confiere resistencia a Geneticina. Para generar las mutantes sencillas *Scglo2Δ* y *Scglo4Δ* se amplificó el casete de KanMX, flanqueado por unas regiones de homología del 5' y el 3' de éstos genes de 60pb y se utilizó como templado el plásmido pFA6a (Wach, 1994). Con el programa 10 minutos a 98°C , seguido de treinta ciclos de 1' a 98°C, 1' a 58°C y 2' a 72°C, con una extensión final a 72°C por 10 minutos con la enzima Phusion se construyó el módulo y transformó con el protocolo de alta eficiencia (Gietz & Schiestl, 2007).

Para generar la doble mutante (*Scglo2Δ Scglo4Δ*) se tomó la cepa *Scglo2Δ* y se sustituyó el gen KanMX por el módulo clonNat, que se obtuvo mediante la digestión del plásmido p4339 con la enzima EcoRI. Posteriormente sobre el fondo de las mutantes *Scglo2Δ::NAT* se transformó el módulo de KanMX con regiones homólogas al gen *ScGLO4*, sustituyéndolo de la misma manera que la mutante sencilla.

Los oligos utilizados para generar estas mutantes, junto con los oligos de confirmación se encuentran en la tabla 8A.

3.4.5 Intercambio de promotores endógenos por el promotor tet

Para poder observar bajo el microscopio de fluorescencia las proteínas *ScAlt2*, *ScGlo4* y *ScTrr2*, sobre la cepa R1158, se intercambió el promotor endógeno de los genes *ScALT2*, *ScGLO4* y *ScTRR2* con el de tetO7 para sobreexpresarlos (Yen, 2003). El módulo se amplificó a partir del plásmido pCM325 que contiene el gen de resistencia a Geneticina (KanMX) con lo siguientes pasos: 10 minutos a 98°C , seguido de treinta ciclos de 1' a 98°C, 1' a 58°C y 2' a 72°C, con una extensión final a 72°C por 10 minutos con la enzima Phusion. Se transformó con 2ug de éste módulo con el protocolo de alta eficiencia. Los oligos utilizados para amplificar el módulo y los de confirmación se encuentran en la tabla 9A.

Posteriormente se fusionó a estas tres proteínas la proteína fluorescente yECitrine, con la misma metodología mencionada en el capítulo 3.4.1 para determinar su localización.

3.5 Análisis de microscopía de fluorescencia

En tubos Falcon de 50 ml se hizo un preinóculo de las cepas marcadas con yEC en un volumen de 10mL de YPD. Al cabo de alrededor de 16 horas las células se lavaron dos veces en agua y se tomó una alícuota para hacer un inóculo en 0.05 de densidad óptica (DO_{600}), en medio mínimo glucosa-amonio, y se dejó crecer hasta 0.5 aproximadamente. Posteriormente se centrifugaron por 5 minutos a 14000 rpm y se colocó 3ul de células en un porta objetos junto con 15ul de agarosa de bajo punto de fusión a una concentración de 1.25%. Posterior a ello, las muestras se analizaron mediante microscopía de fluorescencia con los siguientes parámetros:

Objetivo 100X, ganancia de 2, tiempo de exposición no mayor a 1 segundo.

3.6 Análisis filogenético

Se obtuvieron de Génolevures (Sherman et al., 2009) las secuencias de las proteínas ortólogas provenientes de la levadura tipo ancestral *Lanchancea kluyveri*, una especie de levadura relacionada a *Saccharomyces cerevisiae* que divergió antes de la duplicación total del genoma, la cual ocurrió 100 millones de años atrás.

Se alinearon las secuencias de los duplicados de *S. cerevisiae* con el gen de *L. kluyveri* con los parámetros anteriormente mencionados.

Para los alineamientos de tres secuencias los árboles se llevaron a cabo con el constructor de alineamientos de Geneious, con un modelo de distancia genética Jukes-Cantor (Jukes & Cantor, 1969) por UPGMA (Unweighted Pair Group Method with Arithmetic Mean).

3.7 Curvas de crecimiento

Para cada cepa se inocularon, al menos, tres colonias aisladas en 10ml de YPD y se dejaron crecer toda la noche. Posteriormente las células se lavaron dos veces con agua y se inocularon a una DO_{600} de .05 en el medio correspondiente. Las curvas de glucosa se llevaron a cabo en el crecinómetro Bioscreen C Tipo FMP00C por 48 horas, midiendo la densidad óptica a 600nm cada 30 minutos a una temperatura de 30°C, con agitación máxima. Las curvas de etanol se realizaron en matraces de 125ml con 50ml de medio por 4 días seguidos y se midió cada 3 horas, con una agitación de 250rpm y a una temperatura de 30°C.

Para determinar la velocidad específica de crecimiento primero se calcula el valor de la pendiente en la fase exponencial de crecimiento. Y se aplican las siguientes fórmulas:

Tiempo de duplicación (h) = $\log 2/m$, siendo m la pendiente

y Velocidad específica de crecimiento (μ) = $\ln 2/h$, donde $\mu = h^{-1}$

3.8 Concentraciones mínimas inhibitorias de metilglioxal para *Glo2/Glo4*.

Para definir las implicaciones fisiológicas de las mutantes sencillas y doble (*Scglo2Δ Scglo4Δ*), además de las cepas con las proteínas relocalizadas, llevamos a cabo la determinación de la concentración mínima inhibitoria (MIC), que se define como la concentración más baja a la que un compuesto inhibe el crecimiento de los organismos. Se llevó a cabo como se detalla en Andrews et. al., 2001. Se hicieron diluciones seriales para metilglioxal, donde la concentración más alta fue de 10mM y la mínima de 1mM, se tomó como control positivo donde se tenía una concentración de 0mM de metilglioxal. Se efectuó a cabo el protocolo anterior para curvas que crecen en glucosa como fuente de carbono, con la excepción del metilglioxal añadido.

3.9 Preparación de extracto celular y ensayo enzimático de alanina aminotransferasa.

Para la preparación de extracto las células fueron cosechadas centrifugando 100 mL de medio mínimo glucosa-amonio a una DO_{600} de 1.0. Después se lavaron una vez con agua destilada previamente refrigerada a 4° C. Para los casos en donde se utilizaron cultivos de un volumen distinto se realizó un ajuste proporcional debido a que la estandarización del protocolo fue hecha para preparar extracto a partir de 2.3×10^9 ($DO_{600} = 1$) células aproximadamente. Tras lavar las células, éstas se suspendieron en buffer inhibitor de proteasas (50mM Hepes, 1mM PMSF, 1mM DTT y 1mM EDTA) adicionando la 700 μ l de este buffer; luego se dieron 12 ciclos de un minuto en vórtex, un minuto en hielo, para después centrifugar a 5000 rpm a 4° C por 10 min y recuperar el sobrenadante, al cual se le adicionó lo necesario para conseguir 100 μ M de piridoxal-5-fosfato. La muestra, denominada extracto, se mantuvo en hielo mientras se prepararon los ensayos o se almacenó a -20° C en caso de que no se fuera a ensayar en el momento posterior a la preparación del extracto.

El ensayo enzimático consiste en inferir la producción de piruvato a partir del decaimiento de la concentración de NADH observado, debido al acoplamiento de la reacción:

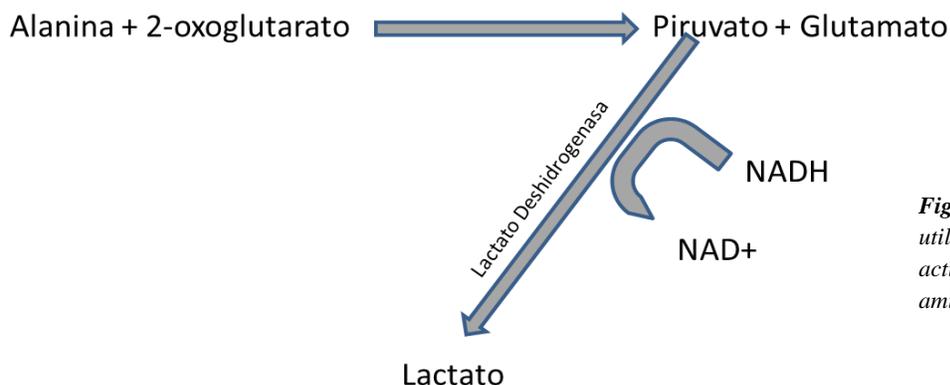


Figura 8. Ensayo enzimático utilizado para medir la actividad de alanina aminotransferasa.

La absorbancia del NADH fue medida a 340nm en un espectrofotómetro (Cary 50 Bio) Los ensayos se hicieron en celdas de 1mL conteniendo 400 mM alanina, 24mM 2-oxoglutarato para Alt1, 250 μ M NADH, 40 μ M piridoxal-5-fosfato y 10 U de Lactato deshidrogenasa de conejo, aforando al volumen final con buffer de ensayo (50mM Tris, 4mM MgCl₂ y 150 mM KCl a pH 8).

Tras preparar las celdas de ensayo y previo a adicionar el extracto se consiguió un equilibrio térmico con la celda después de mantenerlas por 5 min a 30° C. Se midieron 4 concentraciones de proteína distintas, siendo la segunda el doble de la primera y la mitad de la tercera, asimismo se realizó un control sin alanina, sin Lactato Deshidrogenasa y sin alanina ni 2-oxoglutarato por cada condición de concentraciones establecida. Ello con el objetivo de obtener un evento cuyos valores fueran sustraídos a los de nuestras muestras problema despreciando la incidencia de factores externos a la reacción de interés.

Para concluir se realizó determinación de proteína mediante el método de Lowry (Lowry, 1951) utilizando como control una curva estándar de BSA.

IV) RESULTADOS

4.1 Elección de parejas de proteínas empleadas en este trabajo y técnicas de selección

4.1.1 Parejas candidatas y su función

Para determinar las parejas de estudio, se analizó un grupo de 89 proteínas con base en los datos reportados por Marques (Marques et al., 2008), de las cuáles se tomaron 13, que presentaron una localización diferencial y que cumplieron adicionalmente con los siguientes criterios: que mostraran un porcentaje de identidad mayor al 40%, que preferencialmente presentaran una actividad enzimática, que alguna de las dos proteínas tuviera una presecuencia. En la tabla 1 se muestran las proteínas seleccionadas para este estudio y con su función.

TABLA 1. PAREJAS DE PROTEÍNAS CANDIDATAS Y SU FUNCIÓN.

P_1	Función	P_2	Función
ScAap1	Aminopeptidasa se alanina/arginina	ScApe2	Aminopeptidasa yscII
ScAcs1	Acetil coA sintasa	ScAcs2	Acetil coA sintasa
ScAld2	Aldehído deshidrogenasa	ScAld5	Aldehído deshidrogenasa
ScAlt1	Transaminasa de alanina	ScAlt2	Transaminasa de alanina
ScBat1	aminotransferasa de aa de cadena ramificada	ScBat2	aminotransferasa de aa de cadena ramificada
ScCit2	Citrato sintasa	ScCit3	Citrato sintasa
ScGlo2	Glioxilasa	ScGlo4	Glioxilasa
ScGpd1	Glicerol 3 fosfato deshidrogenasa	ScGpd2	Glicerol 3 fosfato deshidrogenasa
ScGrx1	Oxidoreductasa	ScGrx2	Oxidoreductasa
ScHta1	Remodelación de cromatina	ScHta2	Remodelación de cromatina
ScSds23	Separación celular durante la gemación	ScSds24	Separación celular durante la gemación
ScTrr1	Reductasa de tiorredoxina	ScTrr2	Reductasa de tiorredoxina
ScUra5	Fosforibosiltransferasa de orotato	ScUra10	Fosforibosiltransferasa de orotato

Tabla 1. En esta tabla se muestran los nombres de cada una de las proteínas parálogas (P_1 y P_2) junto con su función.

La pareja *ScAap1/ScApe2* tiene una localización diferencial, pero en la literatura no está definida en concreto la función de éstas proteínas, por lo que, no serviría para los estudios subsecuentes, motivo por el cual fue descartada.

4.1.2 Análisis in silico utilizado para determinar las características de las proteínas candidatas

La expresión de cada uno de los genes, junto con la localización subcelular se obtuvieron de la *Yeast GFP database*, mientras que la identidad y la región de la presecuencia fueron el resultado de los alineamientos entre los dos parálogos. La identificación de la presecuencia es fundamental debido a que esta región es suficiente para dirigir las proteínas hacia cierto compartimento, por lo que al ser eliminada (o insertada) la proteína podría relocalizarse.

La relación directa entre concentración de las proteínas y luminiscencia es básica para la determinación de la localización subcelular. La primera está relacionada con la expresión de los genes mismos, y la segunda con la sensibilidad del microscopio de fluorescencia. Una relación adecuada de estos factores permite la identificación de la correcta localización, es por esto que se inició verificando la expresión reportada en la tabla 2.

TABLA 2. CARACTERÍSTICAS DE LAS PROTEÍNAS CANDIDATAS.

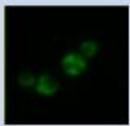
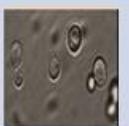
Scce	Nom_sis	Exp	Loc	Iden	Prot_presec	aa_presec
ScAcs1	YAL054C	2890	c, n	53.4	acs1	(37-67)
ScAcs2	YLR153C	225000	N			
ScAld2	YMR170C	2070	c, n	44.7	ald5	(2-24)
ScAld5	YER073W	23300	m			
ScAlt1	YLR089C	9960	-	62	alt1	(4/70)
ScAlt2	YDR111C	3480	c,n			
ScBat1	YHR208W	87300	-	76.9	bat1	(2-18)
ScBat2	YJR148W	25900	c,n			
ScCit1	YNR001C	NV	m	78.2	cit1	(2-17)
ScCit2/3	YCR005C	2310	C			
ScGlo2	YDR272W	13700	c, n	58.9	glo4	(2-11)
ScGlo4	YOR040W	tec_prob	m			
ScGpd1	YDL022W	807	pc, perox	72.4	gpd2	(2-49)
ScGpd2	YOL059W	8970	c, n			
ScGrx1	YCL035C	3030	c, n	63.3	grx2	(2-34)
ScGrx2	YDR513W	31400	N			
ScHta1	YDR225W	NV	V	98.5	puntual	
ScHta2	YBL003C	32100	N			
ScSds23	YBR214W	2840	c, n	59.4	sds24	(5-21)
ScSds24	YGL056C	21200	C			
ScTrr1	YDR353W	292000	c, n	85.6	trr1	(2-24)
ScTrr2	YHR106W	414	m			
ScUra5	YML106W	39300	c, n	73.7	ura10	
ScUra10	YMR271C	815	C			

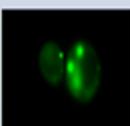
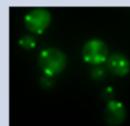
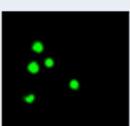
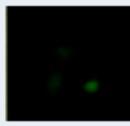
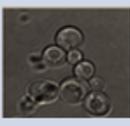
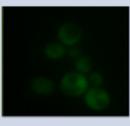
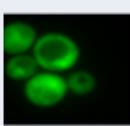
Tabla 2. *Scce*.- nombre común de los genes de *S. cerevisiae*, *Nom_sis*.- nombre sistemático, *Exp*.- expresión (Ghaemmaghami, et. al., 2003), *Loc*.- Localización subcelular (Huh, 2003), *Iden*.- Identidad entre la pareja de duplicados. Las dos últimas columnas indican cuál de las dos proteínas contiene la presecuencia y los aminoácidos que componen esta presecuencia. Citoplasma(c), mitocondria(m), núcleo(n), cloroplasto(clo), proteína secretada (sec), membrana plasmática (PM), vacuola(v), peroxisoma (perox), retículo endoplásmico (RE).

4.1.3 Microscopía de fluorescencia de las parejas descartadas

La estrategia utilizada para seleccionar las parejas finales fue la comprobación de la localización reportada y la posibilidad de llevar a cabo observaciones precisas con el microscopio de fluorescencia. En la siguiente figura se muestran las imágenes obtenidas para las parejas descartadas en este estudio.

Figura 9. Localización subcelular de las parejas eliminadas. Del lado izquierdo se muestran las localizaciones determinadas en este trabajo, del lado derecho las imágenes del microscopio de fluorescencia. Cada imagen está compuesta de dos partes, la primera es la imagen de campo claro y la segunda es la fluorescencia.

Prot1	C_claro1	Fluoresc1	Loc1	Prot2	C_claro2	Fluoresc2	Loc2
ScAld2			cito	ScAld5			mit
ScCit1			mit	ScCit3			mit
ScGpd1			cito, nuc	ScGpd2			cito, nuc

Prot1	C_claro1	Fluoresc1	Loc1	Prot2	C_claro2	Fluoresc2	Loc2
ScGrx1			perox	ScGrx2			perox
ScHta1			nuc	ScHta2			nuc
ScSds23			cito	ScSds24	NBS	NBS	NBS
ScUra5			cito, nuc	ScUra10	NBS	NBS	NBS

Como se puede observar, las parejas *ScGpd1/ScGpd2* y *ScGrx1/ScGrx2* se encuentran en el mismo compartimento subcelular, razón por la cual las descartamos. Estos resultados no están de acuerdo con lo reportado por la base de datos *Yeast GPF Database*.

En los casos siguientes, una de las proteínas de las parejas (marcada con asterisco) *ScAld2/ScAld5**, *ScSds23*/ScSds24*, *ScUra5/ScUra10** y *ScHta1/ScHta2** no tiene expresión fuerte, en el medio que las estudiamos, por lo que no pudimos definir con certeza su localización, y por lo tanto, se eliminaron.

Las parejas que finalmente quedaron incluidas en nuestro estudio fueron las siguientes: *ScAlt1/ScAlt2*, *ScBat1/ScBat2*, *ScGlo2/ScGlo4* y *ScTrr1/ScTrr2*. Las parejas *ScAlt1/ScAlt2*, *ScBat1/ScBat2* han sido estudiadas en el grupo de investigación y conocemos su papel fisiológico. Para el caso de *ScGlo2/ScGlo4* y *ScTrr1/ScTrr2* fue necesario desarrollar ensayos que nos permitieran valorar su comportamiento fisiológico, tomando en cuenta lo reportado con anterioridad (Bitto et al., 1997; Trotter & Grant, 2005).

TABLA 3. CEPAS CONSTRUIDAS PARA LA DETERMINACIÓN DE LOCALIZACIÓN DE LAS PAREJAS ELIMINADAS.

Cepa	Genotipo	Fuente
ScBY4741	<i>MATα his3 ura3 met15 leu2</i>	Open Biosystems
ScAap1-yEC	<i>MATα AAP1-yECitrine- SpHIS5 ura3 met15 leu2</i>	Este trabajo
ScApe2-yEC	<i>MATα APE2-yECitrine- SpHIS5 ura3 met15 leu2</i>	Este trabajo
ScAcs1-yEC	<i>MATα ACS1-yECitrine- SpHIS5 ura3 met15 leu2</i>	Este trabajo
ScAcs2-yEC	<i>MATα ACS2-yECitrine- SpHIS5 ura3 met15 leu2</i>	Este trabajo
ScAld2-yEC	<i>MATα ALD2-yECitrine- SpHIS5 ura3 met15 leu2</i>	Este trabajo
ScAld5-yEC	<i>MATα ALD5-yECitrine- SpHIS5 ura3 met15 leu2</i>	Este trabajo
ScCit1-yEC	<i>MATα CIT1-yECitrine- SpHIS5 ura3 met15 leu2</i>	Este trabajo
ScCit2-yEC	<i>MATα CIT2-yECitrine- SpHIS5 ura3 met15 leu2</i>	Este trabajo
ScCit3-yEC	<i>MATα CIT3-yECitrine- SpHIS5 ura3 met15 leu2</i>	Este trabajo
ScGpd1-yEC	<i>MATα GPD1-yECitrine- SpHIS5 ura3 met15 leu2</i>	Este trabajo
ScGpd2-yEC	<i>MATα GPD2-yECitrine- SpHIS5 ura3 met15 leu2</i>	Este trabajo
ScSds23-yEC	<i>MATα SDS23-yECitrine- SpHIS5 ura3 met15 leu2</i>	Este trabajo
ScSds24-yEC	<i>MATα SDS24-yECitrine- SpHIS5 ura3 met15 leu2</i>	Este trabajo
ScUra5-yEC	<i>MATα URA5-yECitrine- SpHIS5 ura3 met15 leu2</i>	Este trabajo
ScUra10-yEC	<i>MATα URA10-yECitrine- SpHIS5 ura3 met15 leu2</i>	Este trabajo

Tabla 3. Lista de cepas utilizadas para la determinación de la localización subcelular de las proteínas de estudio. El primer panel contiene el nombre conferido a las cepas construidas, posteriormente su genotipo y por último la fuente.

4.2 Estudio del mejor programa de predicción de localización subcelular.

Para predecir la localización subcelular de las proteínas del ancestro (*L. kluyveri*) se llevó a cabo un análisis utilizando 5 programas bioinformáticos y se comparó el resultado con la localización reportada anteriormente por el grupo de O'Shea Huh et al., (2003) y con la localización encontrada en el estudio que aquí se presenta. En la tabla 4 se muestran los resultados, al final de la tabla se encuentra un puntaje que representa la cantidad de aciertos del programa. Se puede observar que Cello, fue el servidor global que mostró tener más similitud a la localización establecida por el grupo de O'Shea y por este trabajo, por lo que se tomó los datos reportados en este servidor para predecir la localización de las proteínas de *L. kluyveri*. Por otra parte, se logró determinar que mitoprot, el cual reporta la probabilidad de que una proteína se encuentre en la mitocondria, tuvo un puntaje muy alto de confirmación, debido a que con base a su predicción se encontraron 7 de las 10 proteínas que su localización subcelular estaba reportada como mitocondrial.

TABLA 4. PREDICCIÓN DE LOCALIZACIÓN SUBCELULAR POR 5 DIFERENTES PROGRAMAS. BACELLO, CELLO, PTARGET, TARGETP, Y MITOPROT.

Sc	Loc_Huh	Loc_lab	BaCello	Cello	ptarget	targetp	mitoprot
ScAap1	c,n	c,n	C	C	c	-	0.19
ScApe2	c,m	c,m	C	c, m	c	m	0.9958
ScAcs1	c, n	N	C	P	p	-	0.0226
ScAcs2	N	N	C	C	p	-	0.03
ScAld2	c, n	C	M	c, m	RE	-	0.056
ScAld5	M	M	M	m	m	m	0.9987
ScAlt1	-	M	M	m	c	m	0.977
ScAlt2	c,n	c, n	M	c	c	-	0.0035
ScBat1	-	M	M	m	m	m	0.83
ScBat2	c,n	c, n	M	c	c	-	0.04
ScCit1	M	M	M	m	m	m	0.96
ScCit2	C	C		p	p	-	0.1919
ScGlo2	c, n	c, n	C	c	c	-	0.02
ScGlo4	M	m	M	m	c	m	0.76
ScGpd1	pc, perox	perox	M	c	m	m	0.2623
ScGpd2	c, n	perox	C	m	m	m	0.9925
ScGrx1	c, n	c, n	C	c	c	-	0.1993
ScGrx2	N	c, n	C	extra	c	S	0.009
ScHta1	V	N	M	n	m	m	0.9956
ScHta2	N	N	M	n	m	m	0.9956
ScSds23	c, n	c, n	C	n	n	-	0.02
ScSds24	c	-	C	n	n	m	0.04

ScTrr1	c, n	c, n	sec	c	m	-	0.0798
ScTrr2	m	M	M	c, m	m	m	0.9324
ScUra5	c, n	c, n	C	PM, Clo	c	-	0.086
ScUra10	c	-	C	Clo	c	-	0.17
	Total		16	18	16	6	

Tabla 4. Predicción de localización subcelular por cinco diferentes servidores. Los_huh, contiene la localización reportada por el grupo de Huh, mientras que Loc_lab es la localización determinada en este estudio. Los programas utilizados son: BaCeLlo (Pierleoni et al., 2006), Cello (Yu, et al., 2006), pTarget (Guda, 2006), Targetp (Emanuelsson et al., 2007) Mitoprot (Claros, 1996). BaCeLlo (Pierleoni et al., 2006), Cello (Yu, et al., 2006), pTarget (Guda, 2006), Targetp (Emanuelsson et al., 2007) Mitoprot (Claros, 1996). Al final de la tabla se muestra la identidad que tiene con la localización experimental. El sombreado significa que en al menos una localización el resultado del programa coincide con la localización probada experimentalmente

El conjunto de métodos de selección anteriores resultaron en la selección de 4 parejas de proteínas que fueron el objeto de estudio de este trabajo de maestría.

Cada capítulo siguiente contiene los antecedentes de cada par de parálogos, junto con la localización subcelular obtenida en el proyecto, la relocalización de las proteínas y la implicación fisiológica de esta relocalización determinada por los cambios fisiológicos resultantes.

4.3 Parejas finales

4.3.1 Transaminasas de alanina ScAlt1 y ScAlt2.

4.3.1.1 Antecedentes de las transaminasas de alanina ScAlt1 y ScAlt2.

La alanina, el segundo aminoácido más abundante en la célula, se obtiene mediante la conversión del glutamato más piruvato a alanina y α -cetoglutarato (Figura 10).



Figura 10. Reacción de las transaminasas de alanina.

En *S. cerevisiae* la enzima encargada de llevar a cabo la reacción anterior es la transaminasa de alanina, codificada por dos genes, *ScALT1* y *ScALT2*. Las isozimas *ScAlt1* y *ScAlt2* presentan una identidad de secuencia del 67%, y están localizadas diferencialmente, la primera se encuentra

en la mitocondria y la segunda en el citoplasma y núcleo. *ScAlt1* tiene 84 aminoácidos extras en el amino terminal, lo que se piensa que puede ser una señal de localización mitocondrial.

Antecedentes del laboratorio indican que la función de las dos transaminasas no es redundante en las condiciones hasta ahora estudiadas. Las mutantes sencillas de *Scalt1Δ* y *Scalt2Δ* y la doble mutante *Scalt1Δ Scalt2Δ* presentan un fenotipo silvestre en glucosa-amonio, lo que indica la presencia de una ruta alternativa para la biosíntesis de alanina. Hasta el momento, la(s) ruta(s) alterna(s) de la biosíntesis de alanina se desconoce, aunque se ha propuesto que la transaminasa de glutamina pudiera ser la responsable de la biosíntesis de alanina en glucosa-amonio (García-Campusano et al., 2009). En presencia de etanol como fuente de carbono y amonio como fuente de nitrógeno, la mutante *Scalt2Δ* presenta un fenotipo silvestre, en tanto que la cepa *Scalt1Δ* y la doble mutante *Scalt1Δ Scalt2Δ* no crecen, indicando que en condiciones respiratorias *ScAlt1* constituye la única vía para la biosíntesis de alanina. En presencia de alanina como única fuente de nitrógeno y de glucosa o etanol, la mutante *Scalt1Δ* y la doble mutante *Scalt1Δ Scalt2Δ* no crecen, en tanto que la mutante *Scalt2Δ* presenta un fenotipo silvestre. Estos resultados muestran que *ScAlt1* juega un papel importante en el catabolismo de alanina. Cuando se midieron las actividades enzimáticas de alanina aminotransferasa en glucosa amonio (Tabla 5), los resultados mostraron que la cepa silvestre y la mutante *Scalt2Δ* presentan la misma actividad, indicando que *ScAlt2* no tiene actividad de alanina aminotransferasa (Peñalosa, et al., en revisión).

En presencia de alanina como única fuente de nitrógeno, la expresión de *ScALT2* se reprime. Esta represión depende de la presencia del regulador de la transcripción *ScNrg1*, por lo que se construyó una cepa que tiene deletado el gen *ScNRG1*. Sin embargo, en estas condiciones *ScAlt2* tampoco posee actividad de alanina aminotransferasa. Este resultado se puede explicar de las siguientes maneras: 1) *ScAlt2* no tiene actividad de alanina aminotransferasa, 2) el ensayo para determinar actividad no es el adecuado y 3) la cantidad de proteína no es suficiente para detectar actividad. Con el propósito de sobre-expresar *ScALT2*, se construyó una mutante en la que se intercambió el promotor de *ScAlt2* por un promotor bacteriano llamado tet07 que contiene 7 operadores repetidos a los cuales se une el activador del promotor llamado tTa y que en presencia de doxiciclina éste se disocia. Estas construcciones se realizaron en la cepa R1158, debido a que ésta contiene el gen tTa. Esta cepa no mostró actividad, por lo que se concluyó que *ScAlt2* no tiene actividad de alanina aminotransferasa o que el ensayo utilizado no es el adecuado.

TABLA 7. ACTIVIDADES ESPECÍFICAS DE ALANINA AMINOTRANSFERASA (Alt1 y Alt2). LA ACTIVIDAD SE EXPRESA EN $\mu\text{mol}/\text{min}^{-1}/\text{mg}$ DE PROTEÍNA⁻¹

CEPA	ACTIVIDAD ESPECÍFICA
<i>ScALT1 ScALT2</i>	0.104±0.030
<i>ScALT2 Scalt1Δ</i>	0.011 ±0.003
<i>ScALT1 Scalt2Δ</i>	0.117 ±0.030
<i>Scalt1Δ Scalt2Δ</i>	0.023 ±0.000
<i>Scalt1Δ ScALT2 Scnrg1Δ</i>	0.015 ±0.004
<i>ScR1158 Scalt1Δ</i>	0.019 ±0.008
<i>ScR1158 Scalt1Δ tet07-ScALT2</i>	0.022 ±0.003

Tabla 5. Actividad específica de Alt1 y Alt2. Se muestra la actividad específica reportada en $\mu\text{mol}/\text{min}-1/\text{mg}$ de proteína de las cepas que tienen a las dos proteínas (*ScALT1* y *ScALT2*), las que tienen sólo una, el control negativo. Y posteriormente las cepas que tienen a *ALT2* sobreexpresado.

4.3.1.2 Cepas utilizadas *ScAlt1/ScAlt2*

TABLA 6. CEPAS UTILIZADAS PARA ESTUDIAR LA PAREJA ALT1/ALT2

Cepa	Genotipo	Fuente
<i>ScCLA1-2</i>	<i>MATa ura3 leu2::LEU2</i>	Quezada, 2008
<i>ScBY4741</i>	<i>MATa his3 ura3 met15 leu2</i>	Open Biosystems
<i>ScR1158</i>	<i>MATa URA3:: CMV-tTA his3 met15 leu3</i>	Hughes, 2000
<i>L. kluyveri</i> Y156	<i>MATa ura3</i>	Gojkovic, 2001
<i>Scalt1Δ</i>	<i>MATa alt1Δ::kanMX ura3 leu2::LEU2</i>	Park, 1999
<i>Scalt2Δ</i>	<i>MATa alt2Δ::kanMX ura3 leu2::LEU2</i>	Park, 1999
<i>Scalt1Δ Scalt2Δ</i>	<i>MATa alt2Δ::NAT alt1Δ::kanMX ura3 leu2::LEU2</i>	Hinnebusch, 2005
<i>ScAlt1-yEC</i>	<i>MATa his3 met15 leu2 ALT1-yECitrine-URA3</i>	Este trabajo
<i>ScAlt2-yEC</i>	<i>MATa his3 met15 leu2 ALT2-yECitrine-URA3</i>	Este trabajo
<i>ScAlt1R-70</i>	<i>MATa ura3 leu2::LEU2 ALT1-70Δ</i>	Este trabajo
<i>ScAlt1R-70 alt2Δ</i>	<i>MATa alt2Δ::NAT ura3 leu2::LEU2 ALT1-70Δ</i>	Este trabajo
<i>ScAlt1R-70-yEC</i>	<i>MATa leu2 ALT1-70Δ-yECitrine-URA3</i>	Este trabajo
<i>ScAlt1R-84</i>	<i>MATa ura3 leu2::LEU2 ALT1-84Δ</i>	Este trabajo
<i>ScAlt1R-84 alt2Δ</i>	<i>MATa alt2Δ::NAT ura3 leu2::LEU2 ALT1-84Δ</i>	Este trabajo
<i>ScAlt1R-84-yEC</i>	<i>MATa leu2 ALT1-84Δ-yECitrine-URA3</i>	Este trabajo
<i>ScAlt2R+70</i>	<i>MATa ura3 leu2::LEU2 ALT2-70ALT1</i>	Este trabajo
<i>ScAlt2R+70 -yEC</i>	<i>MATa leu2 ALT2-70ALT1-yECitrine-URA3</i>	Este trabajo
<i>ScAltR+84</i>	<i>MATa ura3 leu2::LEU2 ALT2-84ALT1</i>	Este trabajo
<i>ScAlt2R+70 -yEC</i>	<i>MATa leu2::LEU2 ALT2-84 ALT1-yECitrine-URA3</i>	Este trabajo

ScAlt1R-70 Alt2R-70	<i>MATa leu2::LEU2 ura3 ALT1-70Δ ALT2-70ALT1</i>	Este trabajo
LkAlt-yEC	<i>MATα ALT-yECitrine::Ura3</i>	Este trabajo
ScR1158 tetO7-Alt2-yEC	<i>MATα tetO7::kanMX-ALT2-SpHIS5 URA3::CMV-tTA met15 leu3</i>	Este trabajo

Tabla 6. Lista de cepas utilizadas para determinar el efecto de la localización de ScAlt1 y ScAlt2 en la fisiología de *S. cerevisiae*. El primer panel contiene el nombre conferido a las cepas construidas, posteriormente su genotipo y por último la fuente.

4.3.1.3 Microscopía de fluorescencia de ScAlt1-yEC y ScAlt2-yEC.

Se construyeron cepas con fusiones independientes de *ScALT1* y *ScALT2* a yECitrine (Metodología). Al observar estas cepas con microscopía de fluorescencia se encontró que ScAlt1-yEC se localiza en la mitocondria, mientras que ScAlt2-yEC se localiza en el citoplasma y núcleo. La localización de ScAlt2 se confirmó intercambiando el promotor endógeno por el de tetO7 para poder sobreexpresarlo, debido a que basalmente no se expresa en grandes cantidades y por lo tanto no visualiza correctamente.

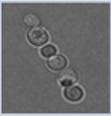
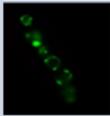
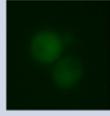
Prot1	C_claro1	Fluoresc1	Loc1	Prot2	C_claro2	Fluoresc2	Loc2
ScAlt1			mit	ScAlt2			Cito, nuc
				ScAlt2 tet			Cito, nuc

Figura 11. Localizaciones de ScAlt1 y ScAlt2. A) resultados de la microscopía para las transaminasas de alanina. B) ScpTet-Alt2. Cada imagen está compuesta de dos partes, la primera es la imagen de campo claro (C_claro) y la segunda es la fluorescencia (Fluoresc).

4.3.1.4 Predicción de localización de LkAlt1.

Se realizó un análisis filogenético para predecir la localización de la alanina aminotransferasa de *L. kluyveri* (Figura 12). Para esto, se realizó un alineamiento, que mostró una posible presecuencia. Para verificar esto, se construyó un árbol filogenético que evidenció que *LkALT1* era más parecida a su ortólogo en *S. cerevisiae* (*ScALT1*), esto indica que hay una alta probabilidad de que esté localizada en la mitocondria (Figura 12).

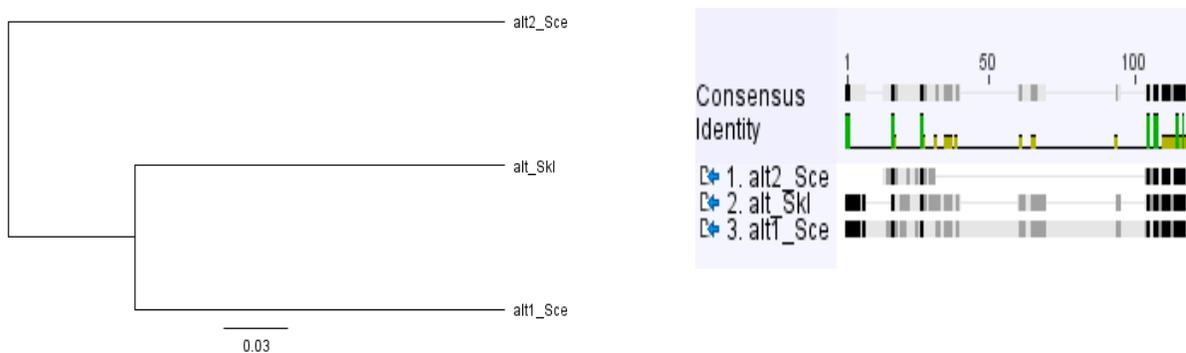


Figura 12. Análisis filogenético de ScAlt1, ScAlt2 y SkAlt. A) Se muestra el árbol filogenético que muestra la relación que tienen las tres proteínas. Se observa que Alt1 de *kluveri* es más cercana a ScAlt1 de *cerevisiae* que a ScAlt2, probablemente esto se deba a que hay una presecuencia. B) Se muestra el inicio del alineamiento realizado para las tres proteínas, se puede observar que ScAlt1 y LkAlt1 tienen una secuencia que no comparten con ScAlt2.

Dos pruebas *in silico* fueron realizadas, Cello especificó que las localizaciones subcelulares más probables eran mitocondria y citoplasma, la segunda prueba fue a través de mitoprot, la cual indicó que la probabilidad para ser exportada a la mitocondria era de 99.23%. Con los datos anteriores pudimos predecir que *LkAlt1* tendría una localización mitocondrial.

4.3.1.5 Localización de la proteína *LkAlt1*-yEC.

La ubicación de la proteína *LkAlt1* es la mitocondria, de acuerdo con lo predicho por el alineamiento, Cello y mitoprot.



Figura 13. Localización mitocondrial de *LkAlt1*. Se observa esta imagen en campo claro (C_claro) y en fluorescencia (Fluoresc).

Con los datos obtenidos por la predicción de localización y su demostración experimental se concluye que la proteína de tipo ancestral *LkAlt1* se localiza en la mitocondria y las proteínas parálogas de *S. cerevisiae* se localizan en la mitocondria (*ScAlt1*) y en el citoplasma y núcleo (*ScAlt2*). Se puede concluir que *ScAlt1* conservó la localización “tipo ancestral”, en tanto que *ScAlt2* se neo-localizó, tal y como lo ha propuesto Marques para otras proteínas.

4.3.1.6 Relocalización de *ScAlt1* y *ScAlt2*

La presecuencia de Alt1 se definió de dos maneras, en la primera se hizo un alineamiento entre ScAlt1 y ScAlt2 y viendo la diferencia en la región N-terminal, y en la segunda se determinó el sitio de corte por mitoprot.

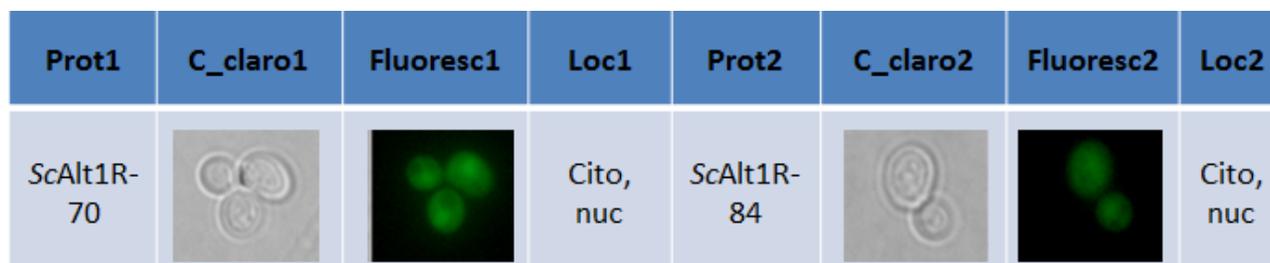


Figura 14. Localización subcelular de ScAlt1R-70 y ScAlt1R-84. Posterior a la eliminación de los primero 70 u 84 aminoácidos ScAlt1 ahora se encuentra en el citoplasma. Las imágenes están compuestas del campo claro (C_claro) y la fluorescencia (Fluoresc).

Con el alineamiento determinamos que la presecuencia consistió de 70 aminoácidos y era la siguiente:

MLSLSAKNHFTVNSITHVIKSYHIRTLTSSAEKMPHITTPFSTASSTKLKAFRKVRPVLQR
HSSSWIV

Mitoprot predijo que el corte se encontraba en el aminoácido 84 y que la presecuencia probablemente era:

MLSLSAKNHFTVNSITHVIKSYHIRTLTSSAEKMPHITTPFSTASSTKLKAFRKVRPVLQR
HSSSWIVAQNHRRLSGQSSL

Se decidió tomar como correctas las dos presecuencias, por lo que trabajó con ambas. Al removerlas, por medio de *Delitto Perfetto* se generaron 2 cepas que llamaremos ScAlt1-70 y ScAlt1-84. Estas cepas se utilizaron para construir derivadas portadoras de la secuencia de yECitrine en el carboxilo terminal. Las observaciones al microscopio de fluorescencia indicaron que la localización de ScAlt1-70-yEC y de ScAlt1-84-yEC era citoplasma y núcleo (Figura 14). Esto demuestra que remoción de las presecuencias predichas es suficiente para cambiar la localización de la proteína ScAlt1. Estos datos sugieren que podría haber una posible secuencia de internalización al núcleo, enmascarada por la señal de localización a la mitocondria, ya que se ha observado que la señal mitocondrial puede excluir la posibilidad de localizar a las proteínas en el núcleo, aunque éstas contengan una señal de localización nuclear (Otterlei, 1998).

Se logró construir las cepas que contienen las presecuencias de *ScAlt1* (70 y 84) sobre el gen de *ScALT2*, modificando la metodología de Delitto Perfetto, lo cual consistió en utilizar colas de recombinación de 600 pares de bases. Lo que evidenció que para insertar un segmento de gran tamaño es necesario aumentar la región de recombinación sobre el gen de interés. Para verificar la correcta inserción de las presecuencias se mandaron secuenciar estas cepas, el resultado fue que todas contenían una delección que producía un codón de paro alrededor de la base 100, por lo que estas cepas no se utilizaron para los experimentos subsecuentes. Este cambio en la pauta de lectura pudo deberse a un error en la amplificación (PCR) de la región codificante. Posteriormente se reintentarán estas construcciones, con nuevos productos de PCR.

4.3.1.7 Actividad específica de *ScAlt1R-70* y *ScAlt1-84*.

Como se había mencionado en la introducción (capítulo 1.4), las presecuencias de localización mitocondrial tienen una longitud de entre 13 y 36 aminoácidos, por lo que 70 u 84 aminoácidos podrían no contener sólo la presecuencia, sino que también pudieran ser sitios importantes para la función de esta proteína.

Para analizar el papel de la presecuencia en la actividad enzimática, se midió la actividad de transaminasa de alanina en la cepa silvestre y en las relocalizadas (*ScAlt1-70*, *ScAlt1-84*)

TABLA 7. ACTIVIDADES ESPECÍFICAS DE ALANINA AMINOTRANSFERASA (ALT1 DE TIPO SILVESTRE Y RELOCALIZADA). LA ACTIVIDAD SE EXPRESA EN $\mu\text{mol}/\text{min}^{-1}/\text{mg}$ DE PROTEÍNA⁻¹

CEPA	ACTIVIDAD ESPECÍFICA
<i>ScALT1 Scalt2Δ</i>	.070±0.009
<i>ScALT1R-70 Scalt2Δ</i>	.111±0.01
<i>ScALT1R-84 Scalt2Δ</i>	.111±0.02

Tabla 7. Comparación de la actividad específica entre la cepa que tiene a Alt1 en el citoplasma y las que tienen a Alt1 en la mitocondria (ALT1R-70), las tres cepas se encuentran en un fondo alt2Δ.

Se puede observar que, bajo nuestras condiciones, las cepas que contienen la proteína relocalizada tienen una actividad específica más alta que la cepa silvestre. Este resultado puede explicarse de la siguiente manera: 1) la proteína *ScAlt1* sólo se encuentra activa cuando pierde la presecuencia, por lo tanto las mutantes *ScAlt1-70* o *ScAlt1-84*, que ya no tiene presecuencia, constituyen una población homogénea, mientras que para la *Alt1* silvestre, la población procesada y

no procesada puede constituir una mezcla que da como resultado un menosprecio de la actividad enzimática. 2) Hay una diferencia en el plegamiento de la proteína dentro de la mitocondria con respecto a la del citoplasma, lo que podría resultar en una diferencia en actividad.

4.3.1.8 Determinación del fenotipo de las mutantes de los genes *ScAlt* y *ScAlt2*

Para determinar la repercusión fisiológica de la relocalización se determinaron las velocidades de crecimiento en diferentes condiciones:

4.3.1.8.1 Curvas de crecimiento en glucosa.

En medio con glucosa, como fuente de carbono, y sulfato de amonio, como fuente de nitrógeno (Figura 15), se encontró que la mutante sencilla *Scalt1Δ* crece ligeramente menos que la cepa silvestre, mientras que la mutante sencilla *Scalt2Δ* crece igual que la cepa silvestre. Con esto, se confirma que Alt1 pudiera tener una cierta repercusión en la biosíntesis de alanina, mientras que Alt2 no parece tener ningún papel en este proceso. Con el fin de comprobar esta propuesta, se determinaron las pozas endógenas de alanina en cultivos de las cepas WT (*ScCla1*), *Scalt1Δ*, *Scalt2Δ* y *Scalt1Δ Scalt2Δ*. Éstos resultados muestran que *ScAlt1* provee el 80% de la poza de alanina, en tanto que *ScAlt2* no contribuye en absoluto a la biosíntesis de este aminoácido (Datos no mostrados). Si la localización mitocondrial de *ScAlt1* determina su papel en la biosíntesis de alanina al remover la presecuencia de 70 u 84 aminoácidos, se comportará como la mutante sencilla *Scalt1Δ*. Se observó que las cepas *ScAlt1R-70* y *ScAlt1R-84* (barras 5 y 6) y las *ScAlt1-70 Scalt2Δ* y *ScAlt1-84 Scalt2Δ* (barras 7 y 8) tienen una pequeña disminución en la velocidad de crecimiento (con un 95% de confiabilidad). Como el fenotipo de crecimiento no se afecta de manera importante en las mutantes *Scalt1Δ*, será necesario determinar las pozas de alanina para valorar con exactitud si la relocalización de *ScAlt1* afecta la biosíntesis de alanina.

En presencia de alanina como única fuente de nitrógeno, podemos determinar la contribución de *ScAlt1-70* y *ScAlt1-80* en el catabolismo de alanina. Como se ha reportado (Peñalosa et al., en revisión), Alt1 es la única enzima capaz permitir el uso de alanina como fuente de nitrógeno la levadura *S. cerevisiae*, ya que la mutante sencilla (*Scalt1Δ*) o la doble (*Scalt1Δ Scalt2Δ*) no son capaces de crecer en presencia de alanina como única fuente de nitrógeno. Si el catabolismo de alanina estuviera influenciado por la localización, las cepas que tienen a *ScAlt1* relocalizada no serían capaces de crecer, o presentarían una menor velocidad de crecimiento. Como se muestra en la Figura 15, las mutantes relocalizadas presentan un fenotipo equivalente al de una

cepa silvestre o una mutante *Scalt2Δ*. Por lo tanto, se puede concluir que la localización mitocondrial de *Alt1* no juega un papel determinante en el catabolismo de alanina.

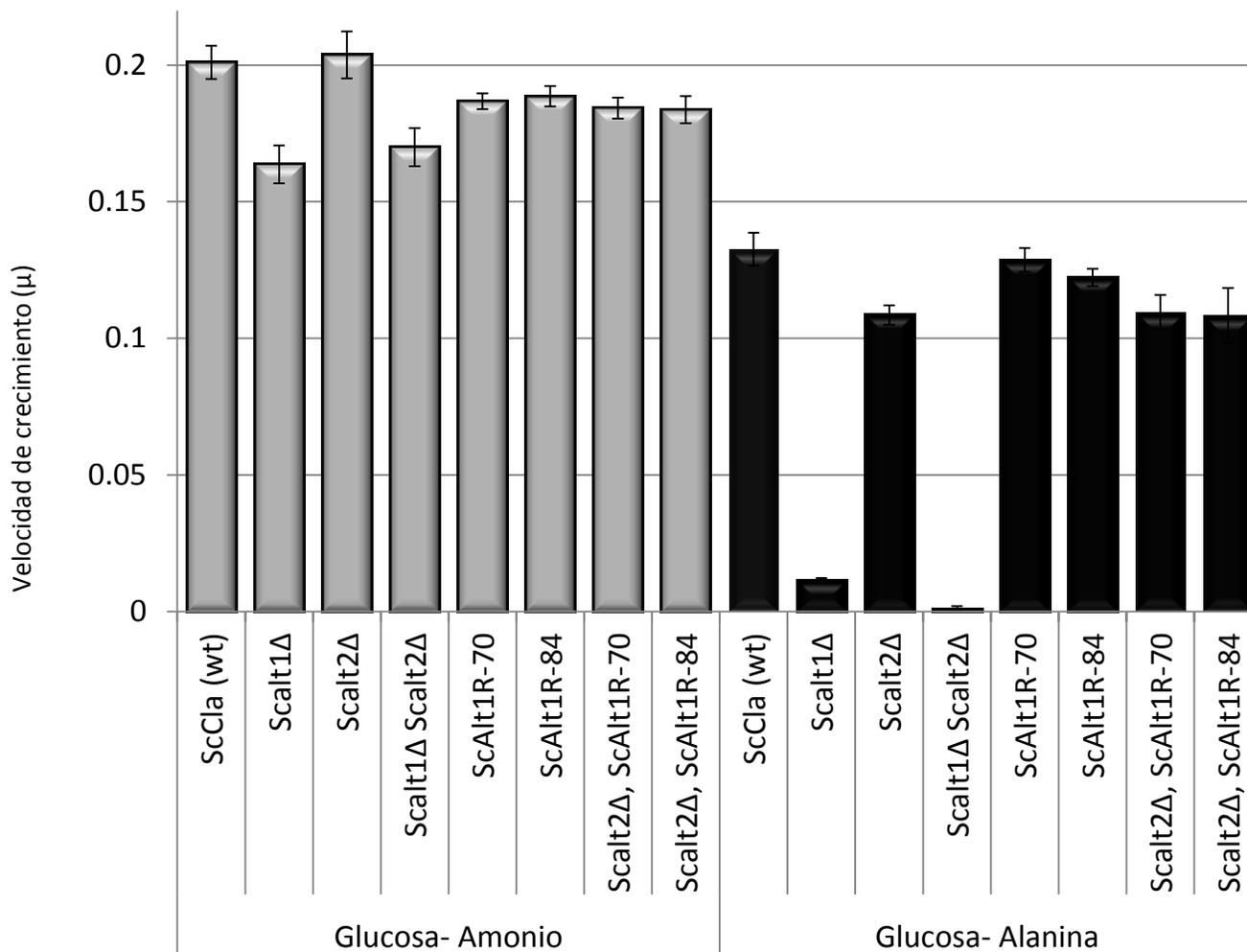


Figura 15. Velocidades de crecimiento en glucosa como fuente de carbono y amonio o alanina como fuente de nitrógeno. Se muestra la cepa silvestre (*ScCla*), las mutantes sencillas y doble (*Scalt1Δ*, *Scalt2Δ* y *Scalt1Δ Scalt2Δ*), las cepas que tienen a *ScAlt1* relocalizada en un fondo silvestre (*ScAlt1R-70* y *ScAlt1R-84*) y en un fondo con *ScALT2* deletado (*alt2 ΔScAlt1R-70* y *Scalt2Δ ScAlt1R-84*)

4.3.1.8.1 Curvas de crecimiento en etanol.

Para determinar si la localización de *ScAlt1* o *ScAlt2* es fundamental en la biosíntesis de alanina, en el metabolismo respiratorio, se confirmaron los fenotipos de las mutantes sencillas

(*Scalt1Δ* y *Scalt2Δ*) y la doble (*Scalt1Δ Scalt2Δ*) en un medio con etanol como fuente de carbono y amonio como fuente de nitrógeno (Figura 16). La mutante sencilla *Scalt1Δ* y la doble mutante (*Scalt1Δ Scalt2Δ*) no crecen, por lo que se confirma que en el metabolismo respiratorio Alt1 es fundamental para la producción de alanina. También podemos observar que la mutante *Scalt2Δ* exhibe una disminución significativa en la velocidad específica de crecimiento, por lo que, al parecer, en este medio tiene un papel en la biosíntesis de alanina.

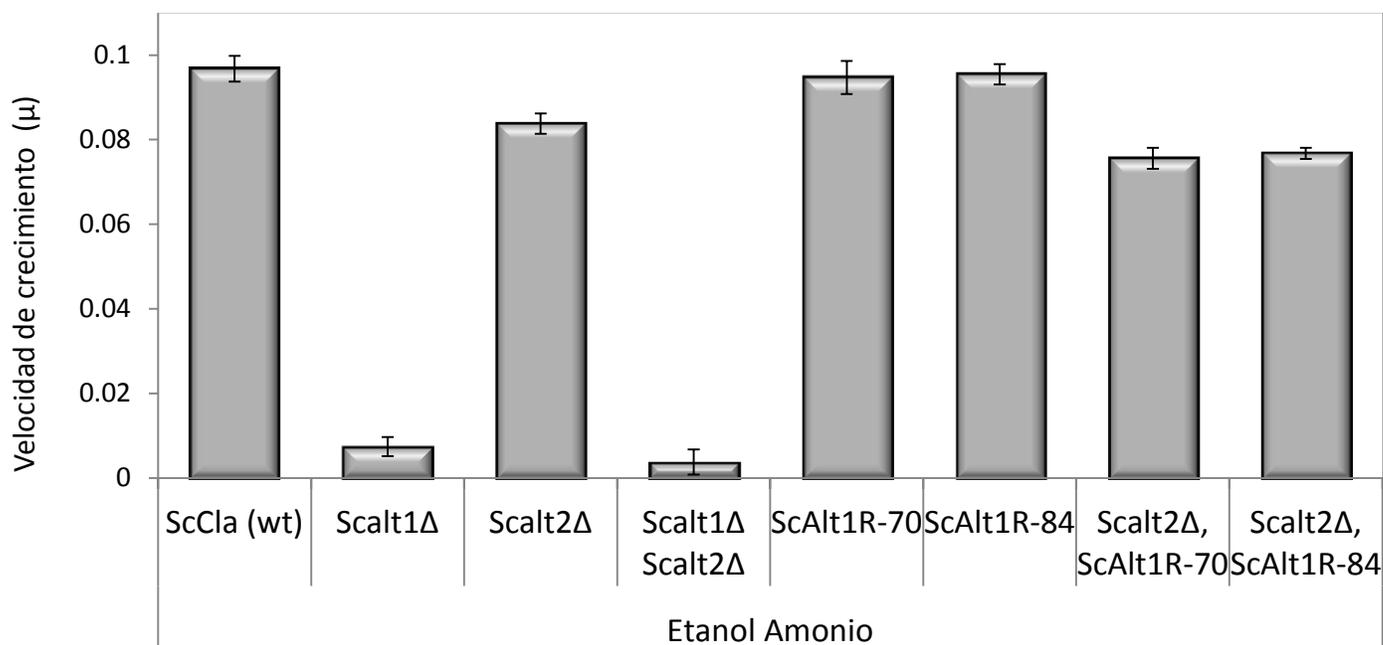


Figura 16. Velocidades de crecimiento en etanol como fuente de carbono y amonio como fuente de nitrógeno *ScAlt1-ScAlt2*. Se muestra la cepa silvestre (*ScCla*), las mutantes sencillas y doble (*Scalt1Δ*, *Scalt2Δ* y *Scalt1Δ Scalt2Δ*), las cepas que tienen a *ScAlt1* relocalizada en un fondo silvestre (*ScAlt1R-70* y *ScAlt1R-84*) y en un fondo con *ScALT2* deletado (*alt2 ΔScAlt1R-70* y *Scalt2Δ ScAlt1R-84*)

Posteriormente se analizó la repercusión de la relocalización de la proteína Alt1 en el metabolismo respiratorio y observamos que en un fondo silvestre, las cepas *ScAlt1-70* y *ScAlt1-84* tienen un crecimiento similar al de la cepa silvestre (Figura 16). Se determinó la velocidad de crecimiento de las cepas que tienen a Alt1 relocalizada en un fondo *Scalt2Δ*, se pudo observar que hay una pequeña disminución en la velocidad de crecimiento, que no es estadísticamente significativa, por lo que bajo estas condiciones la localización no es importante para la función de *ScAlt1*.

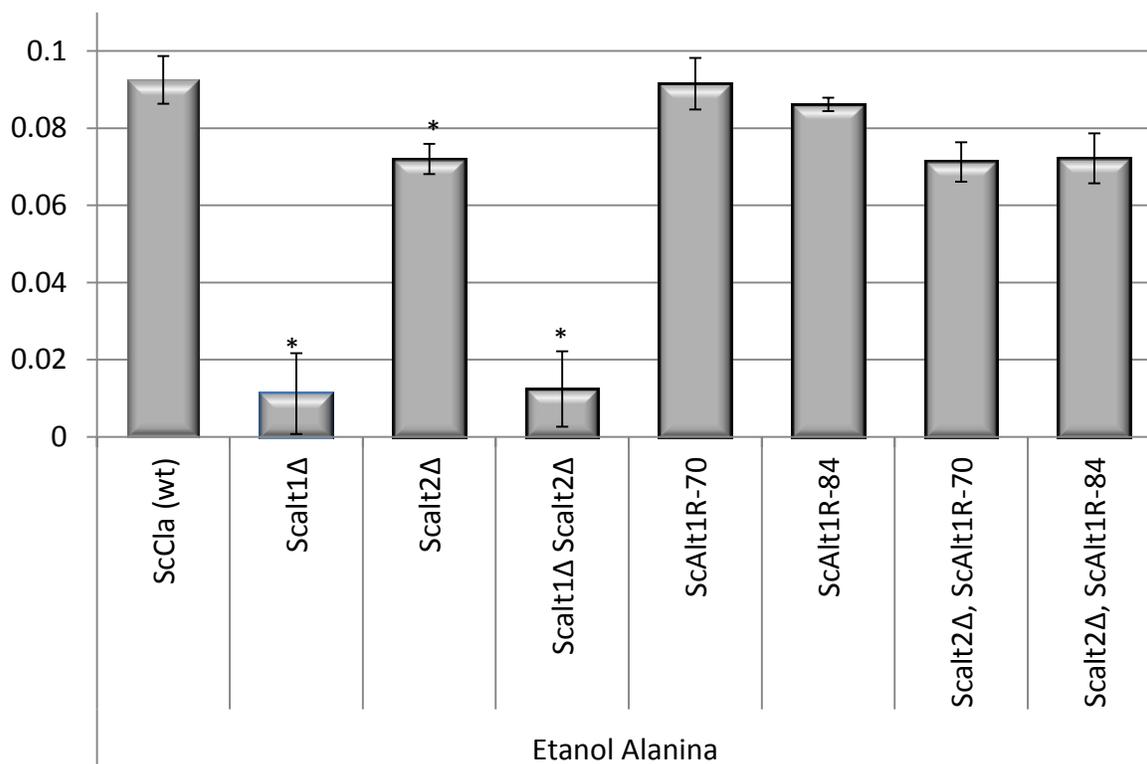


Figura 17. Velocidades de crecimiento en etanol como fuente de carbono y alanina como fuente de nitrógeno. Se muestra la cepa silvestre (*ScCla*), las mutantes sencillas y doble (*ScAlt1Δ*, *ScAlt2Δ* y *ScAlt1Δ ScAlt2Δ*), las cepas que tienen a *ScAlt1* relocalizada en un fondo silvestre (*ScAlt1R-70* y *ScAlt1R-84*) y en un fondo con *ScALT2* deletado (*alt2 ΔScAlt1R-70* y *ScAlt2Δ ScAlt1R-84*)

En la figura 17 se puede ver la contribución de *ScAlt1* y *ScAlt2* y de las mutantes relocalizadas en el catabolismo de la alanina en condiciones respiratorias. Como se había reportado, las cepas que no tienen *ScAlt1* no crecen, debido a que esta enzima constituye la única vía de catabolismo de alanina tanto en etanol como en glucosa. Se observó que *ScAlt2* podría tener una pequeña contribución en el catabolismo de alanina en condiciones respiratorias, similar a lo que se encontró en glucosa (Figura 15). Encontramos también que las cepas que contienen a *Alt1* relocalizada (*ScAlt1R-70*, *ScAlt1R-84*), en cualquiera de los dos fondos (con o sin *ScAlt2*), tienen un comportamiento similar a su cepa silvestre o a la mutante *ScAlt2Δ*. Por lo que para el catabolismo de alanina tampoco la localización de *ScAlt1* es determinante.

4.3.1.9 Conclusiones *ScAlt1/ScAlt2*

- *ScAlt1* se localiza en la mitocondria, mientras que *ScAlt2* se localiza tanto en el citoplasma como en el núcleo.

- La proteína “tipo ancestral” está localizada en la mitocondria, por lo que el proceso evolutivo, con respecto a la localización, que sufrió *ScAlt2* es la neolocalización, en tanto que *ScAlt1* conservó la localización “tipo ancestral”.
- La actividad específica de *ScAlt1* aumenta ligeramente cuando la proteína no contiene los primeros 70 u 84 aminoácidos.
- La falta de *ScAlt2* disminuye ligeramente la velocidad de crecimiento en glucosa-alanina, etanol-amonio y etanol-alanina.
- Hay una disminución en la velocidad de crecimiento en las cepas que tienen a *ScAlt1* relocalizada en condiciones fermento-respiratorias (glucosa), en comparación con la velocidad de la cepa silvestre. Este fenotipo es similar al de una mutante *Scalt1Δ*. Es importante hacer notar que dado que existe una ruta adicional para la síntesis de alanina, el efecto de la falta de *ScAlt1* (*Scalt1Δ*) como de la mutación que elimina la secuencia de localización mitocondrial (*ScAlt-70* y *ScAlt-84*) es difícil de valorar. Por lo tanto, resultará interesante identificar la ruta alternativa y obtener mutantes. En un fondo genético *Scalt2Δ* XΔ (mutante X, responsable de llevar a cabo la ruta alterna de la biosíntesis de alanina) será posible valorar correctamente el valor de *ScAlt1* y de su localización mitocondrial en la biosíntesis de este aminoácido.
- La velocidad de crecimiento de la cepas que contienen a *Alt1* relocalizada (*ScAlt1R-70*, *ScAlt1-84*), en un fondo silvestre o con *ScALT2* deletado en glucosa amonio, es equivalente a la de la cepa *ScALT1* en condiciones respiratorias.
- *Alt1* participa tanto en la biosíntesis como en el catabolismo de alanina en condiciones fermento-respiratorias y respiratorias. En tanto que *ScAlt2* no participa en el metabolismo de alanina. Por lo tanto *ScAlt1* conservó las dos capacidades. En estas circunstancias, la localización mitocondrial no determina la función y es posible que la concentración de pirúvico y alanina en el citoplasma y la mitocondria sea similar.

4.3.1.10 Perspectivas *ScAlt1/ScAlt2*

- Identificar la ruta *ScAlt1-ScAlt2*-independiente que participa en la biosíntesis de alanina, obtener mutantes y determinar su papel en la biosíntesis y catabolismo de alanina.
- Determinar la función de *ScAlt2*, lo que podría ayudar a identificar un fenotipo para estudiarlo en la cepa que tiene a *ScAlt2* relocalizada.

- Determinar pozas de alanina en la cepa en la que se relocizó *ScAlt1*.
- Analizar la contribución de la proteína ancestral para el metabolismo de alanina en *L. kluyveri*, para ello se deberá realizar la mutante *Skalt1Δ* y se definirá su efecto en la velocidad de crecimiento y en las pozas de alanina.

4.3.2 Aminotransferasas de aminoácidos de cadena ramificada ScBat1/ScBat2

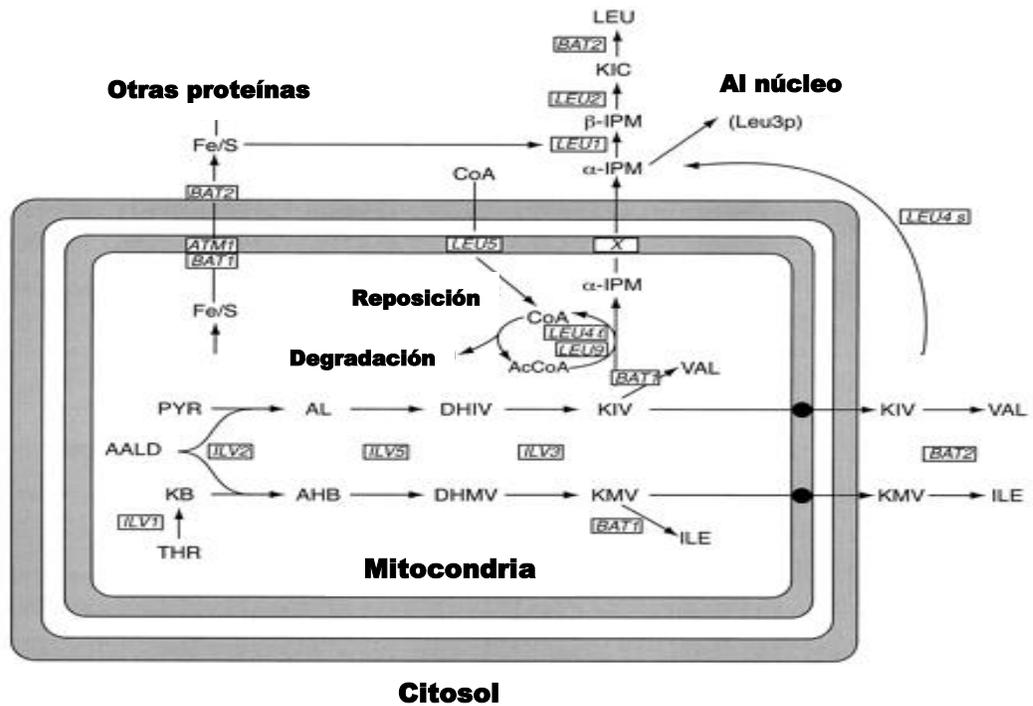
4.3.2.1 Antecedentes de las aminotransferasas de aminoácidos de cadena ramificada ScBat1/ScBat2

Los genes *ScBAT1* y *ScBAT2* codifican para dos enzimas que tienen una identidad de secuencia del 77% entre ellas, se sabe que la localización de Bat1 es mitocondrial, mientras que la de *ScBat2* es citoplásmica. Catalizan la transaminación reversible de aminoácidos de cadena ramificada, como son valina (V), isoleucina (I) y leucina (L) a sus α -cetoácidos, α -cetoisocaproato (KIC), α -ceto- β -metilvalerato (KMV) y α -cetoisovalerato (KIV), respectivamente (Kispal et al., 1996). Además, participan en la vía de Ehrlich, de catabolismo de aminoácidos para la formación de los alcoholes fusel, compuestos que tienen un gran impacto sobre el sabor y olor de los productos de la fermentación de *S. cerevisiae* (Eden et al., 2001). Los componentes principales de la biosíntesis y el catabolismo de LIV se muestran en la figura 18. Estos tres aminoácidos son los más hidrofóbicos y por tal motivo juegan un papel crucial en la estructura terciaria de las proteínas, particularmente en la proteínas que se localizan en la membrana (Brosnan & Brosnan, 2006).

La expresión de *ScBAT1* es 4 veces mayor que la de *ScBAT2* en medio que contiene glucosa y amonio (condiciones biosintética), mientras en medio con glucosa y LIV (condiciones catabólicas) *ScBAT2* se expresa 12 veces más que *ScBAT1*. Las diferencias en la expresión de *ScBAT1* y *ScBAT2* se reflejan en la actividad enzimática de ScBat1 y ScBat2, de suerte que Bat1 presenta una capacidad principalmente biosintética, en tanto que ScBat2 posee una actividad primordialmente catabólica (Colón et al., 2011).

Adicionalmente se han observado diferencias en las velocidades de crecimiento en las mutantes sencillas (*Scbat1Δ* y *Scbat2Δ*) y en la doble (*Scbat1Δ Scbat2Δ*) bajo diferentes condiciones fisiológicas:

Figura 18.
Compartimentalización de la vía de síntesis y degradación de los aminoácidos de cadena ramificada en *S. cerevisiae*. Modificada de Kohlhaw, 2003.



1) En Glucosa- Amonio (un medio que propicia la biosíntesis de LIV) la mutante *Scbat1Δ* tiene una disminución en la velocidad de crecimiento con respecto a la cepa silvestre, mientras que la *Scbat2Δ* presenta un fenotipo silvestre. La mutante doble es auxótrofa de LIV, por lo que se llegó a la conclusión que las dos únicas enzimas capaces de biosintetizar LIV son *ScBat1* y *ScBat2*.

2) En Glucosa-LIV la mutante *Scbat1Δ* crece de manera similar que la cepa silvestre, mientras que la mutante *Scbat2Δ* tiene una disminución en el crecimiento, lo que puede dar idea de la importancia de *ScBat2* para llevar a cabo el catabolismo de LIV. La doble mutante, *Scbat1Δ Scbat2Δ* no crece bajo estas condiciones, debido a que no hay otra enzima que pueda catabolizar los aminoácidos de cadena ramificada para su utilización como fuente de nitrógeno y el crecimiento sólo se recupera cuando se añade otra fuente de nitrógeno al medio, como amonio.

4.3.2.2 Cepas utilizadas *ScBat1/ScBat2*

TABLA 8. CEPAS UTILIZADAS PARA ESTUDIAR LA PAREJA BAT1/BAT2

Cepa	Genotipo	Fuente
<i>ScCLA1-2</i>	<i>MATa ura3 leu2::LEU2</i>	Quezada et al., 2008
<i>ScBY4741</i>	<i>MATa his3 ura3 met15 leu2</i>	Open Biosystems

<i>L. kluyveri</i> Y156	<i>MATa ura3</i>	Gojkovic et al., 2001
<i>Scbat1Δ</i>	<i>MATa bat1Δ::kanMX ura3 leu2::LEU2</i>	Este trabajo
<i>Scbat2Δ</i>	<i>MATa bat2Δ::kanMX ura3 leu2::LEU2</i>	Este trabajo
<i>Scbat1Δ/Scbat2Δ</i>	<i>MATa bat1Δ::NAT bat2Δ::kanMX ura3 leu2::LEU2</i>	Este trabajo
<i>ScBat1-yEC</i>	<i>MATa BAT1-yECitrine-URA3 his3 met15 leu2</i>	Este trabajo
<i>ScBat2-yEC</i>	<i>MATa BAT2-yECitrine-URA3 his3 met15 leu2</i>	Este trabajo
<i>ScBat1R-17</i>	<i>MATa ura3 leu2::LEU2 BAT1-17Δ</i>	Este trabajo
<i>ScBat1R-17-yEC</i>	<i>MATa BAT1-17Δ-yECitrine-SpHIS5 ura3 leu2::LEU2</i>	Este trabajo
<i>ScBat1R-17 Scbat2Δ</i>	<i>MATa bat2Δ::NAT ura3 leu2::LEU2 BAT1-17Δ</i>	Este trabajo
<i>ScBat2R+17</i>	<i>MATa ura3 leu2::LEU2 BAT2-17BAT1</i>	Este trabajo
<i>ScBat2R+17 Scbat1Δ</i>	<i>MATa bat1Δ::NAT ura3 leu2::LEU2 BAT2-17BAT1</i>	Este trabajo
<i>ScBat2R-17 yEC</i>	<i>MATa ura3 leu2::LEU2 BAT2-17BAT1-yECitrine-SpHIS5</i>	Este trabajo

Tabla 8. Lista de cepas utilizadas para determinar el efecto de la localización de *Bat1* y *Bat2* en la fisiología de *S. cerevisiae*. El primer panel contiene el nombre conferido a las cepas construidas, posteriormente su genotipo y por último la fuente.

4.3.2.3 Microscopía de fluorescencia *ScBat1-yEC* y *ScBat2-yEc*

El resultado de la microscopía de fluorescencia fue que estas transaminasas de aminoácidos de cadena ramificada tienen la siguiente localización: *Bat1* se encuentra en la mitocondria y *Bat2* en el citoplasma y en el núcleo (Figura 19). En virtud de que la expresión de *BAT1* y *BAT2* es alta, todas las observaciones se hicieron en la cepa silvestre.

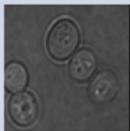
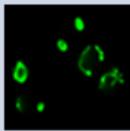
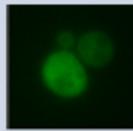
Prot1	C_claro1	Fluoresc1	Loc1	Prot2	C_claro2	Fluoresc2	Loc2
<i>ScBat1</i>			mit	<i>ScBat2</i>			Cito, nuc

Figura 19. Localizaciones de *ScBat1* y *ScBat2*. Resultados de la microscopía para las transaminasas de aminoácidos de cadena ramificada. Mientras que *Bat1* se localiza en la mitocondria, *Bat2* se encuentra en el citoplasma. Cada imagen está compuesta del campo claro (*C_claro*) y de la fluorescencia (*Fluoresc*).

4.3.2.4 Predicción de localización de *LkBat1* y *LkBat2*.

El resultado del BLAST realizado para encontrar los genes de *L. kluyveri* que corresponden a las BCATs nos mostró dos marcos de lectura, que podrían corresponder a dos transaminasas de cadena ramificada. Una de éstas dos posibles proteínas, a la que llamaremos *LkBat2*, al parecer no

es funcional, ya que se encuentra truncada en el aminoácido 126. El segundo marco corresponde a una secuencia (*LkBAT1*) que probablemente contenga una presecuencia mitocondrial.

En el árbol filogenético presentado en la figura 20 se observa que *LkBat1* se agrupa con *Bat1* y que la proteína trunca codificada por *LkBat2*, se aleja tanto de *ScBat1* y *ScBat2* como de *LkBat1*. Este árbol filogenético sirvió para predecir la localización subcelular; dado que la probabilidad dada por mitoprot, para *LkBat1* es de 0.988 y para *LkBat2* es de 0.279 (un resultado de 1 significa que es completamente probable que se encuentre en la mitocondria), se concluyó que la primera se puede localizar en la mitocondria, mientras que la segunda en el citoplasma. Debido a que la proteína *SkBat2* es trunca y no parece ser funcional, no fue considerada para los siguientes experimentos.

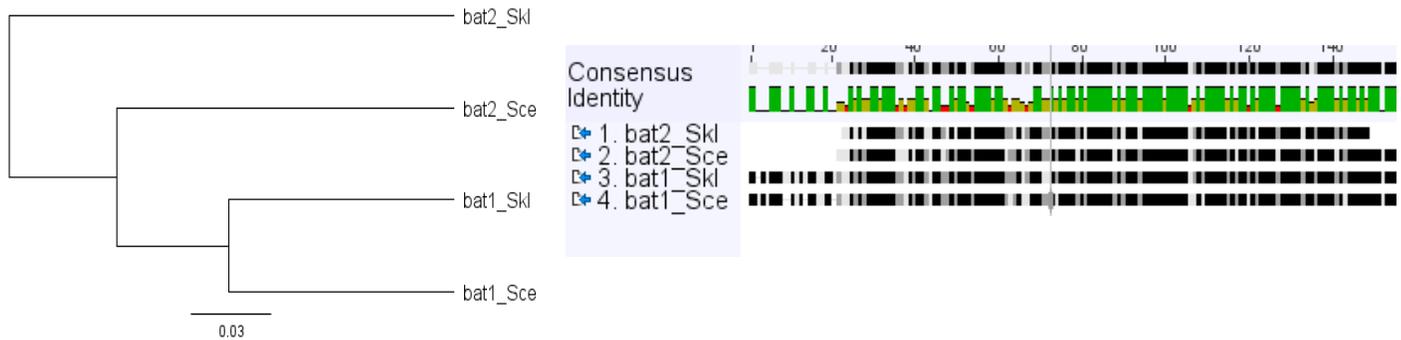


Figura 20. Análisis filogenético de *ScBat1*, *ScBat2*, *LkBat1* y *LkBat2*. A) Se muestra el árbol filogenético que muestra la relación que tienen las cuatro proteínas. Se observa que *LkBat1* es más cercana a *ScBat1*, que *ScBat2* se encuentran dentro del mismo grupo por más alejados y que *LkBat2* se aleja de las otras tres proteínas, probablemente esto se deba a que es una proteína trunca. B) Se muestra el inicio del alineamiento realizado para las cuatro proteínas, se puede observar que *ScBat1* y *LkBat1* tienen una presecuencia que no comparten con *ScBat2* y *LkBat2*.

4.3.2.5 Localización de la proteína *LkBat1*.

Al fusionar *LkBat1* con la proteína *yECitrine* se confirmó la predicción de localización, por lo que pudimos observar que esta proteína se encuentra en la mitocondria (Figura 21). Se puede concluir que *ScBat2* sufrió una neolocalización, encontrándose en citoplasma y núcleo, en tanto que *ScBat1* conservó la localización “tipo ancestral”

Prot	C_claro	Fluoresc	Loc
<i>LkBat1</i>			mit

Figura 21. Localización mitocondrial de *LkBat1*. La imagen está compuesta del campo claro y de la fluorescencia transmitida.

4.3.2.6 Relocalización de *ScBat1* y *ScBat2*

Para definir la presecuencia de *ScBat1* se tomaron los datos determinados por el grupo de Riego (Riego et. al. datos no publicados) en donde se fusionaron los primeros 17 aminoácidos a la proteína rojo fluorescente (RFP) y se determinó que éstos eran suficientes para localizar a *ScBat1* en la mitocondria, en concordancia con los resultados de mitoprot. Los 17 aminoácidos son los siguientes:

MLQRHSLKLGKFSIRTL.

Se removió la presecuencia de *ScBat1* y simultáneamente se fusionó con la proteína yECitrine, como resultado podemos observar que la proteína se localiza ahora en citoplasma y núcleo al igual que la proteína *ScBat2*.

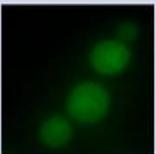
Prot	C_claro	Fluoresc	Loc
<i>ScBat1R</i> -17			Cito, nuc

Figura 22. Localización de *ScBat1R-17*. Se observa que al eliminar los primeros 17 aminoácidos *ScBat1* se localiza ahora en el citoplasma y en el núcleo.

No se logró relocalizar a *ScBat2* a la mitocondria, por lo que se intentará aumentar las secuencias de recombinación para lograr la inserción de la presecuencia de *ScBat1* (17 aa) sobre la secuencia de *ScBAT2*.

4.3.2.7 Determinación del fenotipo de las mutantes de los genes *ScBAT1* y *ScBAT2*

4.3.2.7.1 Curvas de crecimiento en glucosa.

Como se había mencionado anteriormente, las proteínas *ScBat1* y *ScBat2* tienen un papel en la biosíntesis y en el catabolismo de los aminoácidos de cadena ramificada (AACR), leucina, isoleucina y valina.

Se confirmó que en un medio que contiene glucosa como fuente de carbono (Figura 23), *Bat1* tiene un papel primordial en la biosíntesis de los aminoácidos LIV, por lo que la mutante sencilla (*Scbat1Δ*) crecida en medio con amonio como fuente de nitrógeno tiene una velocidad de crecimiento menor que la cepa silvestre. La mutante *Scbat2Δ* presenta un fenotipo silvestre, indicando que su participación en la biosíntesis de LIV es menor que la de *ScBat1*. La doble

mutante no crece, debido a que no existe otra enzima capaz de sintetizar los aminoácidos de cadena ramificada. Si la localización mitocondrial de *ScBat1* fuera determinante para su función, la mutante *ScBat1-17* se comportaría como una mutante sencilla *Scbat1Δ*; podemos observar que las cepas que tienen *ScBat1-17* (Figura 23 barra 5) en fondo silvestre o *Scbat2Δ*, presentan una disminución significativa en la velocidad de crecimiento, que se revierte al añadir LIV al medio de cultivo. Esto indica que la relocalización de *ScBat1* al citoplasma restringe su capacidad de biosintetizar LIV. Esta disminución podría deberse a que en el citoplasma la concentración de los cetoácidos pertinentes es baja o a que *ScBat1* presenta una mayor capacidad catalítica cuando se localiza en la mitocondria.

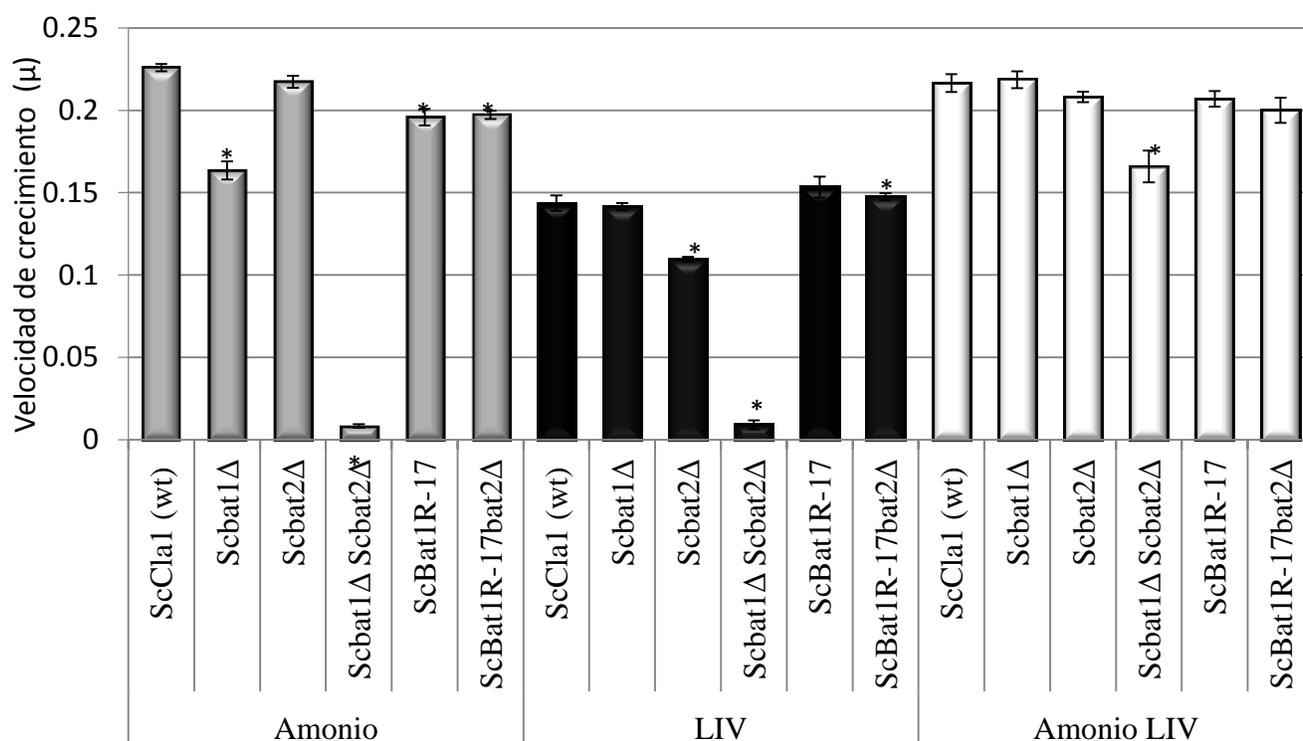


Figura 23. Velocidades de crecimiento en medio con glucosa y amonio, LIV o LIV más amonio como fuente de nitrógeno. Se muestra la cepa silvestre (*ScCla*), las mutantes sencillas y doble (*Scbat1Δ*, *Scbat2Δ* y *Scbat1Δ Scbat2Δ*), las cepas que tienen a *ScBat1* relocalizada en un fondo silvestre (*ScBat1R-17*) y en un fondo con *ScBAT2* deletado (*batt2Δ ScBat1R-17*)

En un medio que tiene LIV como fuente de nitrógeno, observamos que, en general, todas las cepas tienen un crecimiento menor que cuando el medio se suplementa con amonio, ya que estos tres aminoácidos no son una buena fuente de nitrógeno. La mutante sencilla *Scbat1Δ* no muestra una diferencia en el crecimiento con respecto a la cepa silvestre, mientras que la mutante *bat2Δ* tiene una disminución en la velocidad de crecimiento, lo que sugiere que *ScBat2* tiene un papel en el catabolismo de LIV, y que en ausencia de ésta, *ScBat1* cumple con esta función sólo

parcialmente. El fenotipo se ve aumentado en una mutante doble, ya que no mostró crecimiento. Podemos observar que la mutante portadora de la proteína Bat1 relocalizada al citoplasma crece ligeramente mejor que la silvestre, y esto se puede deber a que se encuentra junto con ScBat2 y juntas son más capaces de catabolizar LIV. Para comprobar esto se observó el crecimiento de la cepa que tiene a Bat1 en el citoplasma en un fondo *Scbat2Δ*, encontrándose un incremento en la velocidad de crecimiento con respecto a la BAT1 *Scbat2Δ*. Por lo tanto se puede concluir que cuando ScBat1 se encuentra en el citoplasma es capaz de llevar a cabo el catabolismo de LIV, sin importar que ScBat2 se encuentre o no en estado silvestre. Esto es muy relevante, ya que con estos datos se puede asegurar que el catabolismo de los aminoácidos de cadena ramificada se mejora cuando Bat1 está en el citoplasma (Figura 23), al punto que puede suplir por completo la falta de ScBat2 (*ScBat1-17 Scbat2Δ*). Es posible suponer, que esto se debe a que la disponibilidad de sustratos para el catabolismo de LIV es mayor en el citoplasma que dentro de la mitocondria.

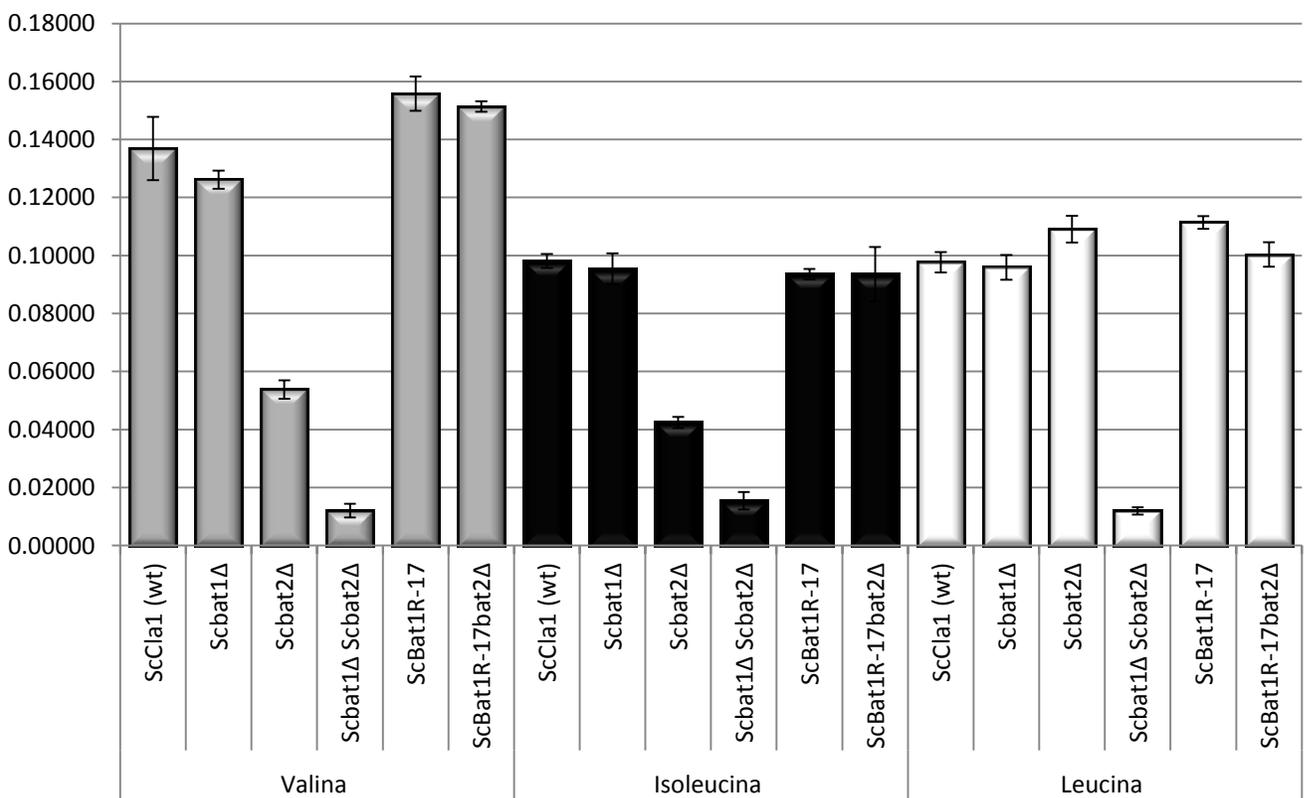


Figura 24. Velocidades de crecimiento en medio con glucosa y valina isoleucina o leucina como fuente de nitrógeno. Se muestra la cepa silvestre (ScCla), las mutantes sencillas y doble (Scbat1Δ, Scbat2Δ y Scbat1Δ Scbat2Δ), las cepas que tienen a ScBat1 relocalizada en un fondo silvestre (ScBat1R-17) y en un fondo con ScBAT2 deletado (batt2Δ ScBat1R-17).

En la figura 24 se muestra la velocidad de crecimiento de las diferentes cepas con valina, isoleucina o leucina como única fuente de nitrógeno. Se confirmó que *ScBat2* tiene un papel esencial en la utilización de valina o isoleucina, mientras que en la de leucina no es tan importante, ya que crece de manera similar a la cepa silvestre. La figura 24 muestra que una mutante *Scbat2Δ* presenta una deficiencia en la capacidad para utilizar valina o isoleucina como únicas fuentes de nitrógeno, defecto que se corrige en una doble mutante *ScBat1-17 Scbat2Δ*. Estos datos indican que la localización mitocondrial de *ScBat1* limita su capacidad para catabolizar LIV, y que el citoplasma aumenta su capacidad de catabolizar estos tres aminoácidos.

4.3.2.7.2 Curvas de crecimiento en etanol

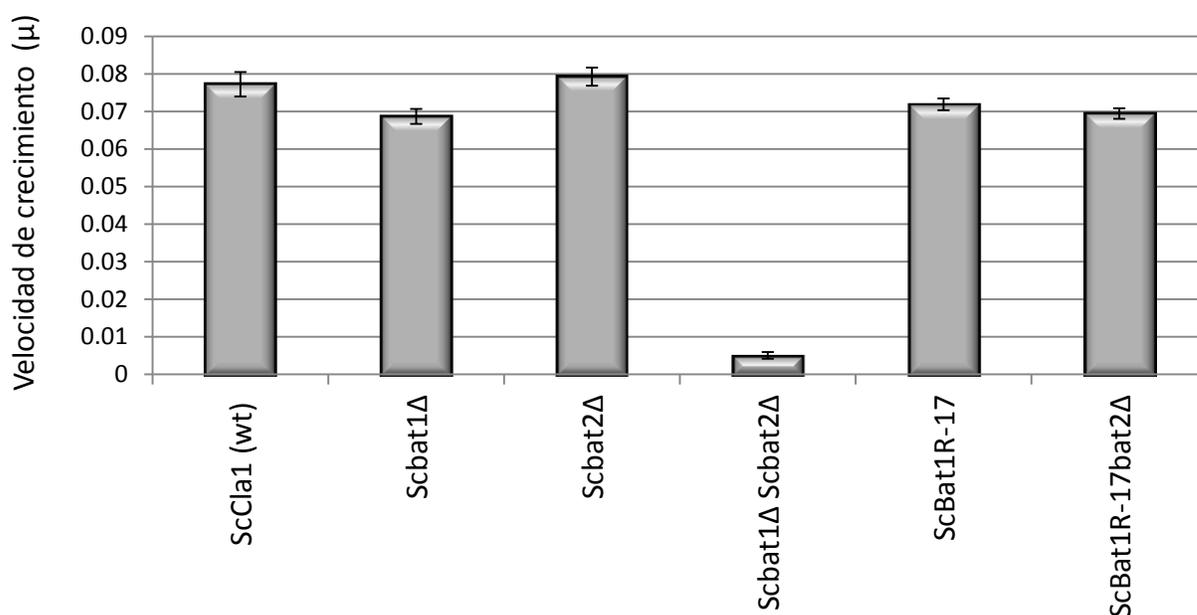


Figura 25. Velocidades de crecimiento en medio con etanol y amonio para *ScBat1* y *ScBat2*. Se muestra la cepa silvestre (*ScCla*), las mutantes sencillas y doble (*Scbat1Δ*, *Scbat2Δ* y *Scbat1Δ Scbat2Δ*), las cepas que tienen a *ScBat1* relocalizada en un fondo silvestre (*ScBat1R-17*) y en un fondo con *ScBAT2* deletado (*batt2Δ ScBat1R-17*).

En la figura 25 se puede observar que *ScBat1* conserva su papel como enzima biosintética en condiciones respiratorias, ya que la mutante sencilla *Scbat1Δ* muestra una disminución en la velocidad de crecimiento. La mutante *Scbat2Δ* muestra un fenotipo silvestre. La doble mutante, como ya habíamos visto en el metabolismo fermentativo, no crece debido a que necesita de los aminoácidos de cadena ramificada para su crecimiento. La cepa que tiene a *Bat1* relocalizada muestra un fenotipo igual al de la mutante *Scbat1Δ*. Por lo que podemos concluir que la localización citoplásmica de *Bat1* en condiciones respiratorias restringe su capacidad para la biosintetizar LIV.

En un medio que tiene etanol, como fuente de carbono, y LIV como única fuente de nitrógeno, las cepas no son capaces de crecer, indicando que estos aminoácidos son pobremente catabolizados (Colón et al., 2011).

4.3.2.8 Conclusiones *ScBat1/ScBat2*

- La localización de *ScBat1* es mitocondrial y la de *ScBat2* es citoplásmica y nuclear.

-La localización de la proteína ancestral *LkBat1* es mitocondrial, por lo de podemos concluir que *ScBat2* sufrió una neolocalización, en tanto que *ScBat1* conservó la localización “tipo ancestral”.

-En condiciones fermento-respiratorias (glucosa) la relocalización de *ScBat1-17* al citoplasma: 1) reduce su capacidad biosintética e 2) incrementa su capacidad catabólica, de suerte que es capaz de substituir el papel de *ScBat2* en el catabolismo de LIV. Por lo tanto, la compartimentalización de *ScBat1* en la mitocondria favorece su papel biosintético y reduce su capacidad para catabolizar LIV.

-En condiciones respiratorias (etanol) la relocalización de *ScBat1-17* al citoplasma: 1) reduce su capacidad biosintética. Por lo tanto, la compartimentalización de *Bat1* en la mitocondria favorece su papel biosintético.

4.3.2.9 Perspectivas *ScBat1/ScBat2*

- Relocalizar a *ScBat2*.
- Determinar la implicación fisiológica de la relocalización de *ScBat2*.
- Determinar pozas de aminoácidos de las proteínas relocalizadas.
- Definir si *ScBat2* tiene otra función.
- Analizar la contribución de la proteína ancestral para el metabolismo de LIV en *L. kluyveri*.
- Relocalizar *LkBat* al citoplasma.

4.3.3 Glioxilasas de tipo II *ScGlo2/ScGlo4*

4.3.3.1 Antecedentes de las glioxilasas de tipo II *ScGlo2/ScGlo4*

El sistema de glioxilasa está compuesto de dos enzimas, glioxilasa I y glioxilasa II. Este sistema usa glutatión como cofactor, convierte los 2-oxoaldehídos en sus ácidos 2-hidroxi correspondientes. El sustrato más importante parece ser el metilglioxal. En la mayoría de los organismos este compuesto tóxico es formado por el producto de la glicólisis a través de la acción de la triosa fosfato isomerasa (Richard, 1993). Pero la mayor cantidad de metilglioxal producido por *S. cerevisiae* y otros microorganismos es por la acción de la dihidroxiacetona fosfato que es convertida a metilglioxal por la metilglioxal sintetasa.

Una vez formado el metilglioxal éste modifica irreversiblemente moléculas, formando unos productos llamados AGE (Advanced Glycation End-products). El metilglioxal reacciona con los nucleótidos guanyl en los ácidos nucleicos para formar aductos de imidazopurinas y reacciona con la lisina y la arginina de las proteínas formando derivados del imidazol.

La glioxilasa I actúa sobre el hemitioacetal (HTA) producido de forma no enzimática a partir del 2-oxoaldehído y el glutatión y lleva a la formación de S-2 hidroxiacilglutatión. La glioxilasa II cataliza la hidrólisis del tioéster correspondiente al 2-(R)-hidroxiácido, generando glutatión libre (Figura 26) (Martins, 2001).

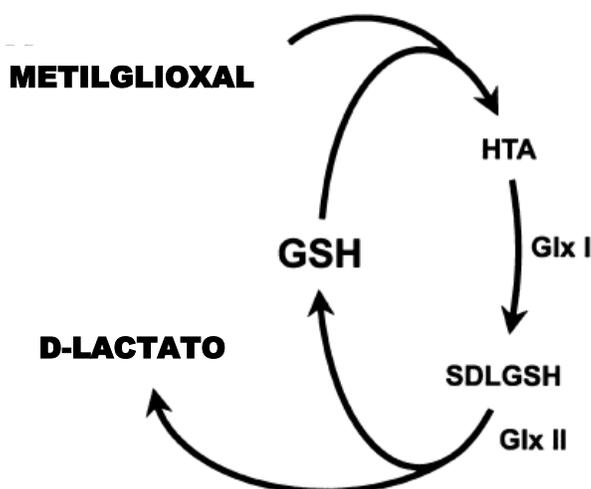


Figura 26. Vía del metilglioxal. El metilglioxal reversiblemente reacciona con el glutatión (GSH) formando hemitioacetal, éste es isomerizado a SDLGSH en una reacción catalizada por la glioxilasa I. Este tioéster es hidrolizado a D-lactato y GSH por la acción de la glioxilasa II. Obtenida de Margarida, 2001.

En *S. cerevisiae* existen dos genes cuyos productos son las glioxilasas II (*ScGLO2* y *ScGLO4*), la región codificante de *ScGLO2* corresponde a una proteína de 274 aminoácidos y a una masa molecular de 31,360 daltones. El gen de *ScGLO4* codifica para una proteína de 285

aminoácidos y una masa molecular de 32,325 daltones. Mientras que el gen de *ScGLO2* se expresa en glucosa y en glicerol, *ScGLO4* sólo se encuentra activo en glicerol (Bito et al., 1997).

La proteína *ScGlo2* se localiza en el citoplasma mientras que *ScGlo4* en la mitocondria y en el citoplasma, cuyos primeros 11 aminoácidos son responsables de su transporte (Bito et al., 1999).

La tasa de crecimiento de las mutantes sencillas y dobles disminuye conforme aumenta la concentración de metilglioxal en el medio. El crecimiento de la cepa silvestre se inhibe completamente a 8mM de metilglioxal en glucosa y 6mM en glicerol. La concentración mínima inhibitoria (MIC, por sus siglas en inglés) fue 6mM de metilglioxal para la mutante *Scglo4Δ* en ambas fuentes de carbono. El crecimiento de la mutante en *Scglo2Δ* se inhibe totalmente a una concentración de 4mM en glucosa y 6mM en glicerol. La MIC de la doble mutante fue de 2mM en glucosa y 4mM en glicerol.

4.3.3.2 Cepas utilizadas *ScGlo2/ScGlo4*

TABLA 9. CEPAS UTILIZADAS PARA ESTUDIAR LA PAREJA GLO2/GLO4

Cepa	Genotipo	Fuente
ScBY4741	<i>MATα his3 ura3 met15 leu2</i>	Open Biosystems
ScR1158	<i>MATα URA3::CMV-tTA his3 met15 leu2</i>	Hughes, 2000
LkY156	<i>MATα ura3</i>	Gojkovic, 2001
Scglo2Δ	<i>MATα glo2Δ::kanMX his3 ura3 met15 leu2</i>	Este trabajo
Scglo4Δ	<i>MATα glo4Δ::kanMX his3 ura3 met15 leu2</i>	Este trabajo
Scglo2Δ/glo4Δ	<i>MATα glo2Δ::NAT glo4Δ::kanMX his3 ura3 met15 leu2</i>	Este trabajo
ScGlo2-yEC	<i>MATα GLO2-yECitrine-URA3 his3 met15 leu2</i>	Este trabajo
ScGlo4-yEC	<i>MATα GLO4-yECitrine-URA3 his3 met15 leu2</i>	Este trabajo
ScGlo2R+11	<i>MATα GLO2-11GLO4 his3 ura3 met15 leu2</i>	Este trabajo
ScGlo2R+11 glo4Δ	<i>MATα glo42Δ::NAT his3 ura3 met15 leu2 GLO2-11GLO4</i>	Este trabajo
ScGlo4R-11	<i>MATα GLO4-11 his3 ura3 met15 leu2</i>	Este trabajo
ScGlo4R-11 glo2Δ	<i>MATα glo2Δ::NAT his3 ura3 met15 leu2 GLO4-11</i>	Este trabajo
R1158 tetGLO4-yEC	<i>MATα tetO7::kanMX-GLO4-SpHIS5 URA3::CMV-tTA met15 leu3</i>	Este trabajo

Tabla 9. Lista de cepas utilizadas para determinar el efecto de la localización de *Glo2* y *Glo4* en la fisiología de *S. cerevisiae*. El primer panel contiene el nombre conferido a las cepas construidas, posteriormente su genotipo y por último la fuente.

4.3.3.3 Microscopía de fluorescencia *ScGlo2/ScGlo4*

Se pudo confirmar, por medio del análisis de fluorescencia, que la glioxilasa II codificada por *ScGLO2* se encuentra en el citoplasma y en el núcleo. Dado que *ScGlo4* en medio mínimo glucosa-amonio no se expresa altamente, se introdujo el promotor *tet* para sobreexpresarlo. Pudimos concluir que *ScGLO4* comparte una de las localizaciones, el citoplasma, y se diferencia en que también se encuentra en la mitocondria (Figura 27).

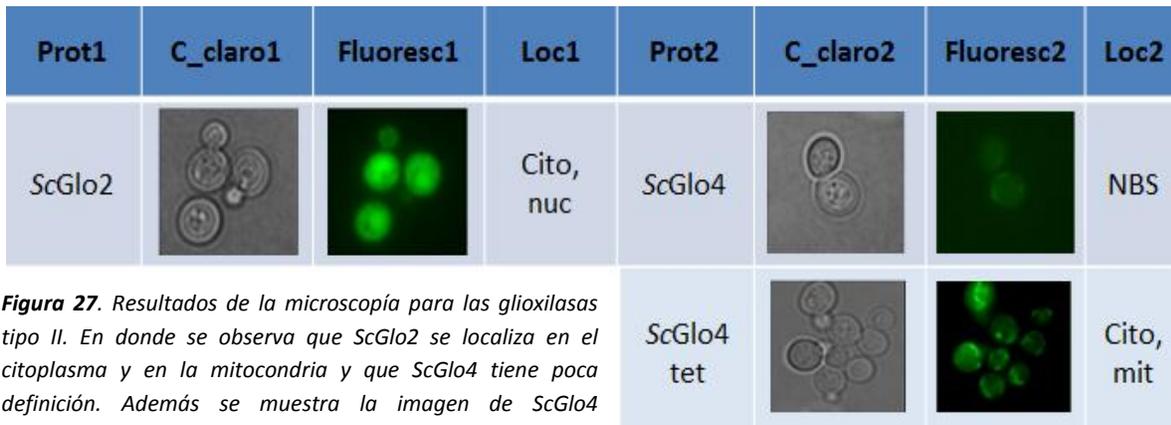


Figura 27. Resultados de la microscopía para las glioxilasas tipo II. En donde se observa que *ScGlo2* se localiza en el citoplasma y en la mitocondria y que *ScGlo4* tiene poca definición. Además se muestra la imagen de *ScGlo4* sobreexpresada (*pTet-ScGlo4*) que muestra que se encuentra en el citoplasma y en la mitocondria. Las imágenes están compuestas del campo claro y de la fluorescencia.

4.3.3.4 Predicción de localización de *LkGlo2*.

Para determinar la posible localización de la única proteína ancestral (*SkGlo*) se llevó a cabo el siguiente análisis de secuencias:

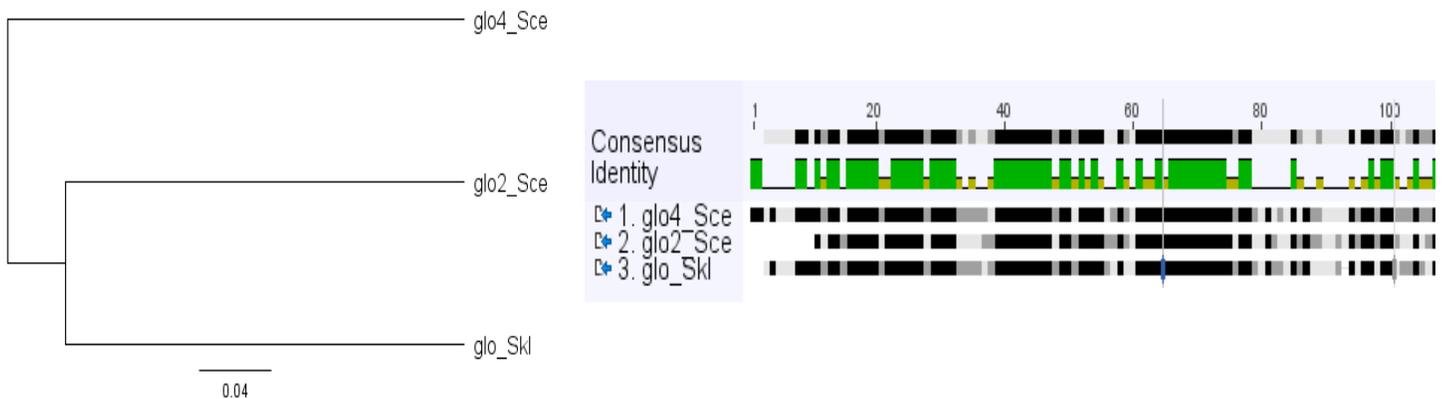


Figura 28. Análisis filogenético de *ScGLO2*, *ScGLO4* y *LkGLO*. A) Se muestra el árbol filogenético que muestra la relación que tienen las tres proteínas. Se observa que *ScGlo2* de *kluveri* es más cercana a *ScGlo2*, mientras que *ScGlo4* se alja de estas dos. Probablemente esto se deba a que *ScGlo4* tiene una presecuencia mitocondrial. B) Se muestra el inicio del alineamiento realizado para las tres proteínas, se puede observar que *ScGlo4* tiene una presecuencia que no comparten con *ScGlo2* y *LkGlo2*.

Tomando en cuenta que la probabilidad de localizarse en la mitocondria, determinada por mitoprot, de *LkGlo2* es de 0.6679 y que la proteína *LkGlo2* se agrupa con *ScGlo2*, llegamos a la conclusión de que no se puede predecir una localización con los datos anteriores.

4.3.3.5 Localización de la proteína *LkGlo2*.

Como se puede observar, la localización de la proteína de tipo ancestral *LkGlo* se localiza en el citoplasma. Hasta el momento no se puede descartar la posibilidad de que también se encuentre en la mitocondria o en el núcleo, para comprobar esto se marcará esta cepa con myc, se aislarán las mitocondrias además del núcleo y se comprobará por Western-Blot su presencia o ausencia dentro de éstos compartimentos.

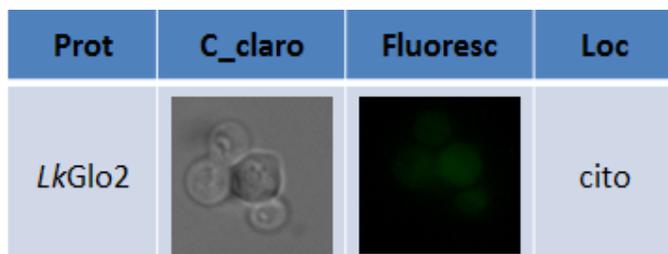


Figura 29. Localización subcelular de *LkGlo2*, se encuentra en el citoplasma. Figura compuesta del campo claro y de la fluorescencia

Las condiciones de expresión del ancestro son parecidas a las reportadas para *ScGLO4* de *S. cerevisiae*, ya que ésta sólo se expresa en un medio que tiene glicerol como fuente de carbono. Por lo anterior se puede afirmar, que en cuanto a la regulación, el gen de *ScGLO2* se comporta de manera similar a *ScGlo4* y no a *ScGlo2*, que se expresa en glucosa.

4.3.3.6 Relocalización de *ScGlo2* y *ScGlo4*

La presecuencia de *Glo4* se encontraba reportada por el grupo de Breitenbach (Bito et al., 1997) en donde se definió que era de 10 aminoácidos y estaba compuesta por los aminoácidos: MKFLLQQIRN. Al eliminar esta presecuencia, obtuvimos la cepa *ScGlo4R-11*. Para determinar la localización subcelular de esta proteína se fusionó con *yEctrine*, pero no se pudo determinar su localización, debido a que, bajo las condiciones estudiadas, este gen no se expresa, por lo que se intercambiará su promotor para confirmar a detalle la localización.

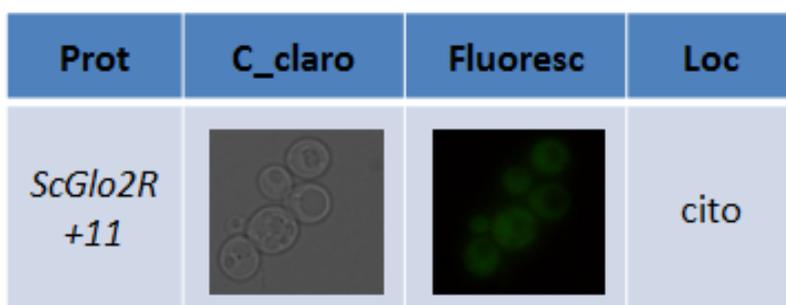


Figura 30. Localización subcelular de *ScGlo2R+11*. Se observa q esta proteína se sigue localizando en el citoplasma. No se puede determinar si se encuentra en el núcleo o en la mitocondria. Compuesta del campo claro y de la fluorescencia.

Para determinar si cambió la localización subcelular de ScGlo2 al insertarle la presecuencia de ScGlo4 se fusionó con la proteína yEctrine y se observó bajo el microscopio de fluorescencia. La señal de fluorescencia por lo que no podemos determinar si realmente hay una parte de la proteína que se localiza en la mitocondria, por lo que para esto, se necesitarán extraer organelos y definir por Western-Blot la presencia de esta proteína. Para confirmar que la presecuencia de ScGlo4 se hubiera insertado en el lugar correcto y que las construcciones no hubieran modificado de alguna manera la secuencia, se mandaron secuenciar por el método de Sanger. El resultado de este experimento demostró que efectivamente la presecuencia estaba en el lugar correcto y que no había alguna modificación extra.

4.3.3.7 Determinación del fenotipo de las mutantes de los genes ScGLO2 y ScGLO4

Como anteriormente se ha reportado, las mutantes sencillas y dobles de las glioxilasa de tipo II son más sensibles a metilglioxal, por lo que para determinar la repercusión de la localización se llevaron a cabo concentraciones mínimas inhibitorias (MIC's) de este compuesto.

TABLA 10. DETERMINACIÓN DE LAS CONCENTRACIONES MÍNIMAS INHIBITORIAS DE METILGLIOXAL.

Cepa	Concentración Mínima Inhibitoria
BY4741 (GLO2 GLO4)	5mM
<i>Scglo2Δ</i> (GLO4)	3mM
<i>Scglo4Δ</i> (GLO2)	4mM
<i>Scglo2Δ Scglo4Δ</i>	2mM
ScGlo2R+11	4mM
ScGlo4R-11	5mM
ScGlo4R-11 <i>Scglo2Δ</i>	3mM

Tabla 10. Concentraciones mínimas inhibitorias de metilglioxal para las cepas mutantes de ScGLO2 y ScGLO4.

En la tabla anterior se puede observar que la concentración mínima inhibitoria de la cepa que tiene la mutante sencilla *Scglo2Δ* es menor que la mutante que tiene deletado el gen de *ScGLO4*, este resultado puede deberse a que ScGlo2 tiene una mayor capacidad de llevar a cabo la reacción de la hidrólisis del S-2-hidroxiacilglutathione que ScGlo4.

Se puede observar que la cepa que tiene a ScGlo2 y ScGlo4 en el citoplasma (*ScGlo4R-11*) acepta la misma cantidad de metilglioxal que la cepa silvestre (5mM), por lo que se puede descartar que las condiciones de los diferentes organelos afecte la eficacia de ScGlo4 para llevar a cabo su correcta función.

La cepa que tiene *ScGlo4* en el citoplasma, en un fondo *Scglo2Δ*, se comporta de igual manera que cuando sólo está deletado el gen de *ScGLO2*, por lo que nos podemos inclinar hacia la hipótesis de que hay una diferencia en la capacidad de hidrolizar el S-2-hydroxiacilglutation entre las dos proteínas y que la disponibilidad de sustratos no es diferente entre el citoplasma y la mitocondria.

Para comprobar lo anterior, se analizaron las concentraciones mínimas inhibitorias en la cepa *ScGlo2R+11*, en donde *ScGlo2* tiene la presecuencia de *ScGlo4* y en donde se asumió que se parte de esta proteína se encuentra en la mitocondria. El resultado fue que a una concentración de 4mM esta cepa deja de crecer, esto puede deberse a que una parte de *ScGlo2*, que tiene una mayor capacidad de llevar a cabo la reacción de hidrólisis que *ScGlo4*, se encuentra en la mitocondria y en este organelo esta proteína no funciona a toda su capacidad, por lo que se llegó a la conclusión de que si hay una repercusión de la relocalización de *ScGlo2*.

Cabe mencionar que éstos resultados son diferentes a los anteriormente reportados, donde la MIC de la cepa silvestre era de 8mM, 6 mM para *Scglo4Δ* y 4mM para *Scglo2Δ* (Bito et al., 1997). Esta diferencia puede deberse a que el fondo genético no es el mismo.

4.3.1.8 Conclusiones *Glo2/Glo4*.

- La localización subcelular de *Glo2* es el citoplasma y el núcleo, mientras que la de *ScGlo4* es la mitocondria y el citoplasma.
- Es menor la concentración mínima inhibitoria de las mutante sencilla *Scglo2Δ* que en la mutante *Scglo4Δ*, lo que indica que *ScGlo2* tiene mayor actividad que *ScGlo4*.
- La localización de la proteína tipo ancestral *LkGlo2* es el citoplasma, para determinar qué proceso evolutivo siguió la localización de las proteínas es necesario, por medio de Western-Blot, definir si esta proteína se encuentra en algún otro compartimento.

4.3.3.9 Perspectivas *Glo2/glo4*

- Determinar las condiciones de expresión de *ScGlo4*.
- Construir las dobles relocalizadas.
- Definir con exactitud la localización de *LkGlo2*, *ScGlo2R+11* y *ScGlo4R-11*.

4.3.4 Reductasas de Tioredoxinas *ScTrr1/ScTrr2*

4.3.4.1 Antecedentes de las reductasas de tioredoxinas *ScTrr1/ScTrr2*

La reductasa de tioredoxina es un miembro de la familia de las nucleótido-disulfuro pirimidina-oxidoreductasas capaces de reducir un enlace disulfuro.

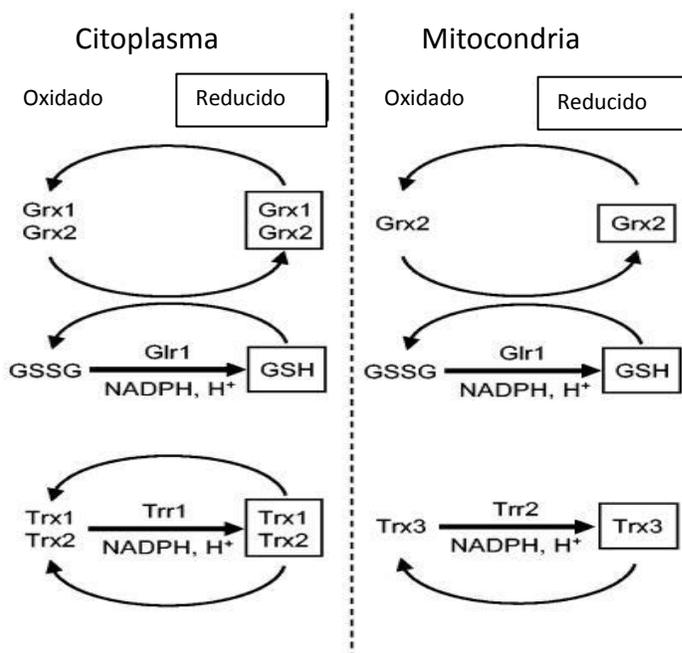


Figura 31. Los sistemas de GSH-glutaredoxina y tioredoxina. Obtenido de Trotter, 2005

El sistema de tioredoxinas y el de glutatión/tiorredoxina ayudan a mantener el ambiente reducido en la célula y juega un papel importante en la defensa del estrés oxidativo (Figura 31). Alternativamente, se ha propuesto que este sistema es importante para la biosíntesis de purinas y pirimidinas o para el plegamiento de proteínas, etc.

Las tioredoxinas son proteínas que actúan como antioxidantes, facilitando la reducción de otras proteínas a través de un intercambio tiol sulfuro en una de las cisteínas inmersas en ella.

Se caracterizan a nivel de secuencia de aminoácidos por la presencia de dos cisteínas vecinas, que son capaces de formar un motivo CXXC. Estas cisteínas son importantes para la capacidad de estas proteínas de reducir a otras. Existen en su forma reducida como tioredoxina (SH)₂ o en su forma oxidada como tioredoxinas S₂ cuando forman un puente disulfuro intramolecular (Holmgren, 1989).

Cepas que tienen deletado el gen de *ScTRR1* son hipersensibles a peróxido de hidrógeno, sensibles a la temperatura y son auxótrofos de metionina (Machado, 1997). Mientras que la cepa *trr2Δ* son mucho más sensibles a estreptomycin y tiene una resistencia a estrés hiperosmótica menor.

4.3.3.2 Cepas utilizadas *ScTrr1/ScTrr2*

TABLA 10. CEPAS UTILIZADAS PARA ESTUDIAR LA PAREJA *ScTRR1/ScTRR2*

Cepa	Genotipo	Fuente
BY4741	MATα <i>his3 ura3 met15 leu2</i>	Open Biosystems

ScTrr2Δ	MATα trr2Δ::kanMX his3 ura3 met15 leu2	Este trabajo
ScTrr1-yEC	MATα TRR1-yECitrine-URA3 his3 met15 leu2	Este trabajo
ScTrr2-yEC	MATα TRR2-yECitrine-URA3 his3 met15 leu2	Este trabajo

Tabla 11. Lista de cepas utilizadas para determinar el efecto de la localización de ScTrr1 y ScTrr2 en la fisiología de *S. cerevisiae*. El primer panel contiene el nombre conferido a las cepas construidas, posteriormente su genotipo y por último la fuente.

4.3.3.3 Microscopía de fluorescencia ScTrr1/ScTrr2

La localización de las proteínas ScTrr1 y ScTrr2 se muestran en las siguientes imágenes:

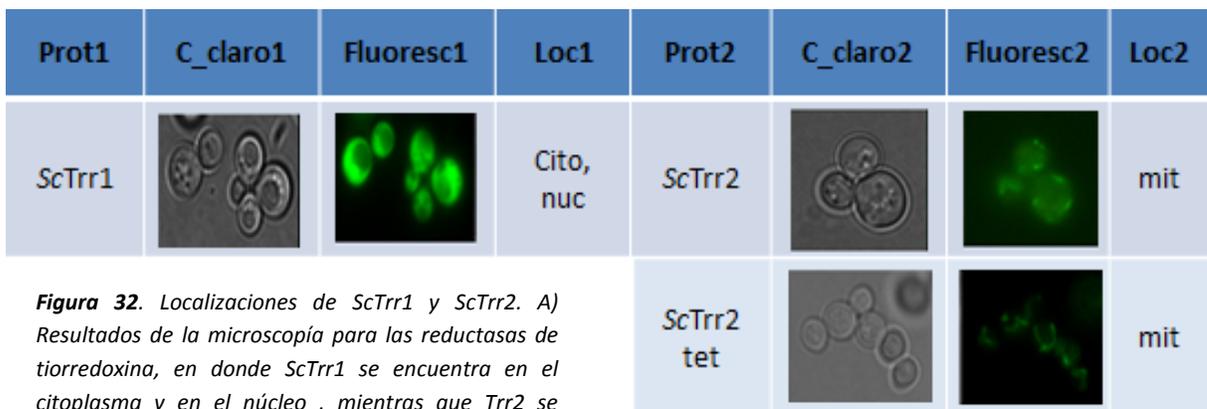


Figura 32. Localizaciones de ScTrr1 y ScTrr2. A) Resultados de la microscopía para las reductasas de tiorredoxina, en donde ScTrr1 se encuentra en el citoplasma y en el núcleo, mientras que Trr2 se localiza en la mitocondria B) ScpTet-Trr2. Se confirmó que se localiza en la mitocondria mediante la sobreexpresión de ScTrr2.

Podemos observar que ScTrr1 se encuentra en el citoplasma, mientras que ScTrr2 se encuentra en la mitocondria.

Al analizar la secuencia de *L. kluyveri* que corresponde al gen ancestral de la reductasa de tiorredoxina (*LkTRR1*), nos dimos cuenta de que aunque se agrupa con la ScTrr1 de *S. cerevisiae* tiene una probabilidad de encontrarse en la mitocondria de 0.80, por lo de predecimos que se encuentra en la mitocondria (Figura 33).

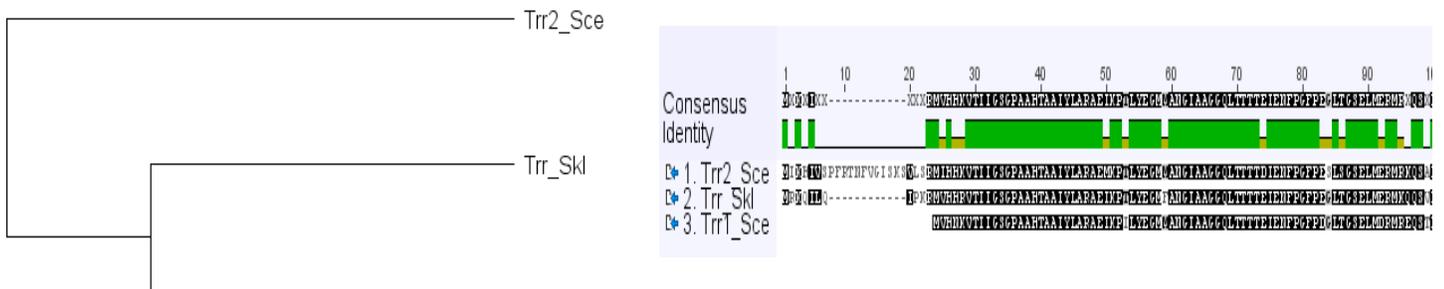


Figura 33. Análisis filogenético de ScTrr1, ScTrr2 y LkTrr1. A) Se muestra el árbol filogenético que muestra la relación que tienen las tres proteínas. Se observa que LkTrr1 es más cercana a ScTrr1, mientras que ScTrr2 se aleja de estas dos. Probablemente esto se deba a que ScTrr1 y LkTrr1 tienen una presecuencia mitocondrial. B) Se muestra el inicio del alineamiento realizado para las tres proteínas, se puede observar que ScTrr1 y LkTrr1 tienen una presecuencia que no comparten con ScTrr2.

4.3.3.3 Construcciones de la pareja *ScTrr1* y *ScTrr2*

Algunos grupos han propuesto que la mutante sencilla *Sctrr1Δ* no es viable mientras que en otros grupos si se han logrado tener mutantes de esta enzima, esto puede deberse a que las mutantes se han hecho en diferentes cepas. En este trabajo se intentó tener las dos mutantes sencillas *Sctrr1Δ* y *Sctrr2Δ* y sólo pudimos obtener la segunda. En el caso de la introducción del *CORE*, que podría ser vista como una interrupción del gen, tuvimos los mismos resultados, pudimos introducir el *CORE* en *ScTRR2* pero no en *ScTRR1*, por lo que se decidió dejar de trabajar con estas proteínas, ya que no habríamos podido tener el set completo de las propiedades fisiológicas de las mutantes y de las relocalizadas.

V) *DISCUSIÓN GENERAL*

La duplicación génica es un rasgo que se comparte en casi todas las especies, ésta constituye una fuente de nuevos genes que están sujetos a cuatro posibles destinos evolutivos, los cuáles son pseudogenización, conservación de la función, subfuncionalización o neofuncionalización. Estos dos últimos representan una forma de diversificación funcional que puede ser alcanzada por cambios en la región codificante, o en la región regulatoria. Este trabajo se centra en los cambios de la región codificante que dan lugar a diferencias en la localización subcelular. Por lo que el objetivo de este trabajo de Maestría fue determinar si la compartimentalización diferencial de los productos de genes parálogos tiene un efecto que determine o influya la diversificación funcional de los mismos.

Para llevar a cabo este análisis, se seleccionaron parejas de genes parálogos cuyos productos se presumía estaban localizados diferencialmente. En la primera parte de este trabajo, se determino la localización de las proteínas parálogas. En los casos en las que cada una de las proteínas tenía una localización peculiar (mitocondrial vs citoplásmica), se intercambió la localización de las mismas eliminando o adicionando presecuencias de internalización mitocondrial, por el método de Delitto Perfetto. Este método tiene como ventajas los cambios puntuales dentro del genoma de un organismo sin dejar huella o rastro alguno; esto implica que los genes modificados se encuentran en contexto, es decir, cumplen con los mismos patrones de expresión, debido a que están regulados por mismo el promotor y se localizan en la misma región con respecto al centrómero que la secuencia silvestre. Posterior a la relocalización, se comprobó el efecto de esta sobre la fisiología de *Saccharomyces cerevisiae*, por medio de curvas de crecimiento en medios que evidenciaban la función de las proteínas estudiadas.

En este trabajo, se utilizaron como modelos de estudio cuatro parejas que presentaban una localización diferencial, las cuáles fueron *ScAlt1/ScAlt2*, *ScBat1/ScBat2*, *ScGlo2/ScGlo4* y *ScTrr1/ScTrr2*. De estas proteínas sólo algunas se lograron relocalizar, debido a que el protocolo de Delitto Perfetto necesita algunas modificaciones para insertar secuencias grandes; por lo que sólo se logró determinar la importancia de la localización en las proteínas a las que se les removió la presecuencia mitocondrial (*ScAlt1*, *ScBat1* y *ScGlo4*), o a la que se le insertó una presecuencia de pequeña (*ScGlo2* de 11 aminoácidos). Cabe mencionar que en el caso de *ScTrr1/ScTrr2*, no se construyeron las cepas, ya que la delección de *Trr1* y la inserción del CORE (paso necesario para llevar a cabo el Delitto Perfetto) mostraban un fenotipo inviable.

En el caso de *ScGlo2* y *ScGlo4*, se determinó que en las condiciones estudiadas, hay una implicación de la relocalización de *ScGlo2*, ya que la cepa que tiene relocalizada a esta proteína muestra un fenotipo de tipo intermedio entre la cepa *Scglo2Δ* y la cepa silvestre. Por otro lado, al relocalizar *ScGlo4* no se observó una diferencia en la fisiología de *S. cerevisiae*, en este caso se deberá determinar el efecto del intercambio de localización en otras condiciones de crecimiento.

El caso de *ScAlt1* los experimentos realizados demuestran que no hay una diferencia fisiológica entre las cepas que tienen a *ScAlt1* relocalizada en el citoplasma al compararlas con aquellas que presentan la localización nativa en la mitocondria. Dado que *Alt1* juega un papel en la síntesis de alanina, será de suma importancia determinar la concentración intracelular de alanina en éstas dos cepas (*ScAlt1* mitocondrial vs *ScAlt1* citoplásmica) para determinar si hay una repercusión fisiológica de la relocalización. Vale la pena hacer notar que en este caso, la pareja de *ScAlt1*; es decir *ScAlt2* no presenta actividad de alanino aminotransferasa y por lo tanto pudiera ser el caso que aun cuando *ScALT1* proviene de la duplicación genómica de *S. cerevisiae*, en realidad no tenga un parálogo funcional (Peñalosa et al., aceptado). Los estudios funcionales de *ScAlt1* han mostrado que la aminotransferasa codificada por este gen constituye la única vía de catabolismo de alanina (Peñalosa et al., aceptado). Tal vez esto resulte en que su actividad no dependa de la localización de la proteína.

Se sabe que *ScBat1* y *ScBat2* sufrieron una subfuncionalización con respecto a la proteína de *Kluyveromyces lactis*, ya que *KIBat1* tiene un carácter principalmente biosintético, mientras que *ScBat2* es catabólica y *KIBat1* es biosintética, por lo que se considera que sufrieron una subfuncionalización (Colón et al., 2011). Este trabajo muestra que en el caso de *Bat1*, la relocalización tiene como consecuencia disminuir el carácter biosintético de la enzima y aumentar el carácter catabólico de la misma. Por lo tanto la *Bat1* citoplásmica es capaz de utilizar como única fuente de nitrógeno los aminoácidos de cadena ramificada (LIV) con una mayor eficacia que la cepa que tiene a *ScBat1* localizada en la mitocondria.

Las enzimas *ScLys20/ScLys21* (homocitato sintasas) y *ScGdh1/ScGdh3* (glutamato deshidrogenasas) son productos de genes duplicados que han sido ampliamente estudiados en el grupo de trabajo de la Dra. González y provienen de la duplicación total del genoma. La primera pareja (*ScLys20/ScLys21*) interviene en el primer paso de la biosíntesis de lisina, catalizando la condensación de α -cetoglutarato y acetil CoA para formar homocitrato. Las propiedades cinéticas de éstas isoenzimas y presentan una inhibición alostérica diferencial, *ScLys21* se inhibe con menores concentraciones de lisina que *ScLys20*. Otro nivel de regulación presente en estas dos

enzimas es que en condiciones respiratorias el homocitrato es producido principalmente por *ScLys21*, en tanto que en condiciones fermentativas *ScLys20* y *ScLys21* juegan un papel redundante (Quezada et al., 2008).

ScGdh3 y *ScGdh1* son isoenzimas dependientes de NADP^+ que catalizan la síntesis del glutamato a partir de amonio y de α -cetoglutarato. La expresión de los genes *ScGDH1* y *ScGDH3* es diferencial y está influenciada por la fuente de carbono. La expresión de *ScGDH3* se induce en etanol y se reprime en glucosa, caso contrario a *ScGDH1* que se encuentra fuertemente inducida en ambas fuentes de carbono. Además de esto, ha habido cambios en la región codificante que permite que la velocidad de consumo de α -cetoglutarato sea mayor para *ScGdh1* que para *ScGdh3* (DeLuna et. al, 2001).

Los cambios que llevan a la diversificación de las dos parejas mencionadas anteriormente son cambios en la región regulatoria o en la región codificante (sitio activo, sitio de unión al sustrato, etc.). Es importante mencionar que no tienen una regulación localización-dependiente, ya que *ScGdh1* y *ScGdh3* se localizan las dos en el citoplasma y *ScLys20* y *ScLys21* en el núcleo.

Las parejas que se utilizaron en este proyecto tienen una regulación adicional que es la dependiente de la localización, por lo que este trabajo aporta la primera evidencia experimental que indica que la localización diferencial de proteínas parálogas constituye uno de los factores que determinan la diversificación funcional de las proteínas parálogas.

VI) CONCLUSIONES GENERALES

- *ScAlt1* y *ScBat1* se localizan en la mitocondria, mientras que *ScAlt2* y *ScBat2* se localizan tanto en el citoplasma como en el núcleo, en tanto que la localización subcelular de *ScGlo2* es el citoplasma y el núcleo, mientras que la de *Glo4* es la mitocondria y el citoplasma.

-La localización de la proteína ancestral *LkBat1* y *LkAlt1* es mitocondrial, por lo que podemos concluir que *ScAlt2* y *ScBat2* experimentaron una neolocalización, en tanto que *ScBat1* y *ScAlt1* conservaron la localización “tipo ancestral”. La localización de la proteína tipo ancestral *LkGlo* es el citoplasma, es necesario definir en qué otro organelo se localiza.

-En condiciones fermento-respiratorias (glucosa) la relocalización de *ScBat1-17*, es decir, *ScBat1* sin los primeros 17 aminoácidos y que adicionalmente se localiza en el citoplasma: 1) reduce ligeramente su capacidad biosintética y 2) incrementa su capacidad catabólica, de suerte que es capaz de substituir el papel de *ScBat2* en el catabolismo de LIV. Por lo tanto, la compartimentalización de *ScBat1* en la mitocondria favorece su papel biosintético y reduce su capacidad para catabolizar LIV.

-En condiciones respiratorias (etanol) la relocalización de *ScBat1-17* al citoplasma: 1) reduce su capacidad biosintética. Por lo tanto, la compartimentalización de *Bat1* en la mitocondria favorece su papel biosintético.

- La actividad específica de *ScAlt1* aumenta ligeramente cuando la proteína no contiene los primeros 70 u 84 aminoácidos.

- La falta de *ScAlt2* disminuye ligeramente la velocidad de crecimiento en glucosa-alanina, etanol-amonio y etanol-alanina.

- Hay una disminución en la velocidad de crecimiento en las cepas que tienen a *ScAlt1* relocalizada en condiciones fermento-respiratorias (glucosa), en comparación con la velocidad de la cepa silvestre. Este fenotipo es similar al de una mutante *Scalt1Δ*. Es importante hacer notar que dado que existe una ruta adicional para la síntesis de alanina, el efecto de la falta de *ScAlt1* (*Scalt1Δ*) como de la mutación que elimina la secuencia de localización mitocondrial (*ScAlt-70* y *ScAlt-84*) es difícil de valorar. Por lo tanto, resultará interesante identificar la ruta alternativa y obtener mutantes. En un fondo genético *Scalt2Δ XΔ* (mutante X, responsable de llevar a cabo la ruta alterna de la biosíntesis de alanina) será posible valorar correctamente el valor de *ScAlt1* y de su localización mitocondrial en la biosíntesis de este aminoácido.

- La velocidad de crecimiento de las cepas que contienen a *ScAlt1* relocalizada (*ScAlt1R-70*, *ScAlt1-84*), en un fondo silvestre o con *ScALT2* deletado en glucosa amonio, es equivalente a la de la cepa *ScALT1* en condiciones respiratorias.
- La concentración mínima inhibitoria de las mutante sencilla *Scglo2Δ* es menor que en la mutante *Scglo4Δ*.

VII) PERSPECTIVAS GENERALES

- Definir si alguna de las proteínas estudiadas tiene otra función

Probablemente estas proteínas tienen una función diferente a la reportada y la importancia de la localización es mucho más significativa para la otra función probable.

-Determinar si podemos encontrar sintéticos letales de ScAlt1 o ScBat2 relocalizados.

De los 6 genes con los que se trabajaron sólo ScALTI y ScBAT2 tienen sintéticos letales con otro gen, por lo que podríamos encontrar que al relocalizar éstas dos proteínas y mutar ese otro gen tendríamos un sintético letal, en caso de que la localización sea relevante para esa función.

-Medir pozas de aminoácidos para las cepas que tienen las proteínas relocalizadas.

Como anteriormente se ha determinado, en ciertas ocasiones, al deletar cierto gen no hay un cambio en la velocidad de crecimiento, pero las pozas de aminoácidos si cambian, esto es debido a que los aminoácidos se encuentran en exceso, por lo que se considera necesario medir éstas pozas en las cepas que tienen las proteínas relocalizadas.

- Determinar si existe una relocalización dependiente de medio de las parejas seleccionadas.

Se ha demostrado que hay ciertas proteínas cuya localización es dependiente del medio o del estadio de crecimiento (Barwise, 1996), queremos establecer si cualquiera de nuestras proteínas tiene este tipo de patrón.

- En caso de existir lo anterior, determinar si el ancestro cumple con los mismos patrones relocalización dependiente de medio.

-Determinar la importancia en el metabolismo de alanina, LIV y metilglioxal de las proteínas ortólogas en L. kluyveri.

- Relocalizar al ancestro y determinar sus implicaciones fisiológicas.

Se pueden comparar las implicaciones fisiológicas de la relocalización del ancestro contra las implicaciones de la relocalización en *S. cerevisiae*, ya que no tiene las dos localizaciones encontradas en las parejas de parálogos.

VIII) REFERENCIAS

- Bannai, H., Tamada, Y., Maruyama, O., Nakai, K. & Miyano, S. (2002) Extensive feature detection of N-terminal protein sorting signals. *Bioinformatics*. 18(2): 298-305.
- Bito, A., Haider, M., Hadler, I., Breitenbach, M. (1997) Identification and phenotypic analysis of two glioxilase II encoding genes from *Saccharomyces cerevisiae*, GLO2 and GLO4, and intracellular localization of the corresponding proteins. *The Journal of Biological Chemistry*. 272(34) : 21509-21519.
- Bito, A., Haider, M., Briza, P., Strasser, P., Breitenbach, M. (1999) Heterologous expression, purification, and kinetic comparison of the cytoplasmic and mitochondrial glioxilase II enzymes, Glo2p and Glo4p, from *Saccharomyces cerevisiae*. *Protein Expression and Purification* 17: 456-464.
- Bizily, S.P., Kim, T., Kandasamy, M.K. & Meagher, R.B. (2003) Subcellular Targeting of Methylmercury Lyase Enhances Its Specific Activity for Organic Mercury Detoxification in Plants. *Plant Physiology*. 131(2): 463-471.
- Bolender, N., Sickmann, A., Wagner, R., Meisinger, C. & Pfanner, N. (2008) Multiple pathways for sorting mitochondrial precursor proteins. *EMBO Reports*. 9(1): 42-49.
- Brosnan J.T. & Brosnan, M.E. (2006) Branched-Chain amino acids: metabolism, physiological function and application. *Journal of Nutrition*. 136: 207S-211S.
- Casadio, R., Martelli, P.& Pierleoni, A. (2008) The prediction of protein subcellular localization from sequence: a shortcut to functional genome annotation. *Briefings in Functional Genomics and Proteomics*. 7(1): 63-73.
- Cherry JM, Ball C, Weng S, Juvik G, Schmidt R, Adler C, Dunn B, Dwight S, Riles L, Mortimer RK, Botstein D (1997) Genetic and physical maps of *Saccharomyces cerevisiae*. *Nature*. 387(6632 Suppl):67-73.
- Claros, M. G., & Vincens, P. (1996). Computational method to predict mitochondrially imported proteins and their targeting sequences. *European journal of biochemistry / FEBS*, 241(3), 779-86.
- Colón, M., Hernández, F., López, K., González, J., Quezada, H., Lopez, G., Aranda, C., González, A. (2011) *Saccharomyces cerevisiae* Bat1 and Bat2 Aminotransferases Have Functionally Diverged from the Ancestral-Like *Kluyveromyces lactis* Orthologous Enzyme. *PLOS ONE*. 6(1): e16099.
- Dayhoff, M., R. Schwartz, & B. Orcutt. (1978). A model of evolutionary change in proteins. *Atlas of protein sequence and structure*. 5: 345–352.

- Eden, A., Van Nederveelde, L., Drukker, M., Benvenisty, N., Debouq A.m(2001) Involvement of branched-chain amino acid aminotransferases in the production of fusel alcohols during fermentation in yeast. *Applied Microbiology and Biotechnology* 55(3): 296-300.
- Emanuelsson, O., & von Heijne, G. (2001). Prediction of organellar targeting signals. *Biochimica et biophysica acta*, 1541(1-2), 114-9.
- Emanuelsson, O. (2002). Predicting protein subcellular localisation from amino acid sequence information. *Briefings in bioinformatics*, 3(4), 361-76.
- Emanuelsson, O., Brunak, S., von Heijne, G., & Nielsen, H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools. *Nature protocols*, 2(4), 953-71. doi: 10.1038/nprot.2007.131.
- Feldherr, C. M., Akin, D. & Cohen, R. J. (2001) Regulation off functional nuclear pores size in fibroblasts. *Journal of Cell Science*.114:4621–4627.
- Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783–791.
- García-Campusano, F., Anaya, V.H., Robledo-Arratia, L., Quezada, H., Hernández, H., Riego, L. & González, A., (2009) ALT1-encoded alanine aminotransferase plays a central role in the metabolism of alanine in *Saccharomyces cerevisiae*. *Canadian Journal of Microbiology*. 55: 1-7.
- Ghaemmaghami, S., Huh, W., Bower, K., Howson, R. W., Belle, A., Dephoure, N., et al. (2003). Global analysis of protein expression in yeast. *Nature*, 425(6959), 737-41.
- Gietz, R.D. & Schiestl, R.H. (2007) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols*. 2: 31-34.
- Gojkovic, Z., Paracchini, S. & Piskur, J. (1998) A new model organism for studying the catabolism of pyrimidines and purines. *Advances in Experimental Medicine and Biology*. 431:475-479.
- Görlich, D. & Kutay, U., Transport between the cell nucleus and cytoplasm. *Annual Review of Cell and Developmental Biology*. 15:607-660.
- Guda, C. (2006). pTARGET: a web server for predicting protein subcellular localization. *Nucleic acids research*, 34(Web Server issue), W210-3.
- Guindon, S., & Gascuel, O. (2003). A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. *Systematic Biology*, 52(5), 696-704.
- Hahn, M. W. (2009). Distinguishing among evolutionary models for the maintenance of gene duplicates. *The Journal of heredity*, 100(5), 605-17.
- Holmgren, A. (1989) Thioredoxin and glutaredoxin systems. *Journal of Biological Chemistry* 264(24): 13963-13966.

- Hughes, A.L. (1994). The evolution of functionally novel proteins after gene duplication. *Proceedings of the Biology Sciences of London*. 256, 119-124.
- Huh, W. K., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S., et al. (2003). Global analysis of protein localization in budding yeast. *Nature*, 425, 686-691.
- Innan, H., & Kondrashov, F. (2010). The evolution of gene duplications: classifying and distinguishing between models. *Nature reviews. Genetics*, 11(2), 97-108.
- Jauhiainen, A., Thomsen, C., Strömbom, L., Grundevik, P., Andersson, C., Danielsson, A., Andersson, M.K., Nerman, O., Römkvist, L., Stahlberg, A. & Aman, P. (2012) Distinct Cytoplasmic and Nuclear Functions of the Stress Induced Protein DDIT3/CHOP/GADD153. *Plos One*. 7(4): e33208- e33208.
- Jukes, T., & Cantor, C. (1969). Evolution of protein molecules. 21–132, in *Mammalian protein metabolism*, Volume III (H. Munro, ed.). *Academic Press, New York*.
- Kerppola, T. K. (2006). Design and implementation of bimolecular fluorescence complementation (BiFC) assays for the visualization of protein interactions in living cells. *Nature protocols*, 1(3), 1278-86.
- Kispal, G., Steiner, H., Courts, D.A., Rolisnki, B., Lill R., Mitochondrial and cytosolic branched-chain amino acid transaminases from yeast, homologs of the myc oncogene-regulated Eca39 protein. *The Journal of Biological Chemistry* 271(40):24458-24464.
- Kimura, M. & King J. L. (1979) Fixation of a deleterious allele at one of two duplicate loci by mutation pressure and random drift. *Proceedings of the National Academy of Science. USA*. 76, 2858-2861.
- Leissring, M. A., Farris, W., Wu, X., Christodoulou, D. C., Haigis, M. C., Guarente, L., et al. (2004). Alternative translation initiation generates a novel isoform of insulin-degrading enzyme targeted to mitochondria. *The Biochemical journal*, 383(Pt. 3), 439-46.
- Liu, Q., D´Silva, P., Walter, W., Marszalek, J. & Craig, E. A. (2003) Regulated Cycling of Mitochondrial Hsp70 at the Protein Import Channel. *Science*. 300: 139-141.
- Long, M. (2003) The origin of new genes: glimpses from the young and old. *Nature Review Genetics*. 4, 865–875.
- Lowry, O. H., Rosebrough, N.J., Farr, A. R. & Randall R.J. (1951) Protein measurement with the Folin-Phenol reagents. *Journal of Biological Chemistry*. 193: 265-275.
- Macara, I. G. (2001) Transport into and out of the nucleus. *Microbiol. Molecular Biology Reviews*. 65, 570–594.

- Machado, A.K., Morgan, B. & Merril, G. F. (2007) Thioredoxin Reductase-dependent Inhibition of MCB Cell Cycle Box Activity in *Saccharomyces cerevisiae*. *The Journal of Biological Chemistry*. 27(4): 17045-17054.
- Marques, A. C., Vinckenbosch, N., Brawand, D., & Kaessmann, H. (2008). Functional diversification of duplicate genes through subcellular adaptation of encoded proteins. *Genome biology*, 9(3), R54. doi: 10.1186/gb-2008-9-3-r54.
- Martins, A.M., Mendes, P., Cordeiro, C., Freire, A.P., In situ kinetic analysis of glyoxalase I and glyoxalase II in *Saccharomyces cerevisiae*. *European Journal of Biochemistry*, 268(14):3930-3936.
- Matsuoka, T., Wada, J., Hashimoto, I., Zhang, Y., Eguchi, J., Ogawa, N., Shikata, K., Kanwar, Y., & Makino, H. (2005) Gene Delivery of Tim44 Reduces Mitochondrial Superoxide Production and Ameliorates Neointimal Proliferation of Injured Carotid Artery in Diabetic Rats. *Diabetes*. 54: 2882-2890.
- Mikata, K., Ueda-Nishimura K. & Hisatomi. T. (2001) Three new species of *Saccharomyces* sensu lato van der Walt from Yaku Island in Japan: *Saccharomyces naganishii* sp. nov., *Saccharomyces humaticus* sp. nov. and *Saccharomyces yakushimaensis* sp. nov. *International Journal of Systematic and Evolutionary Microbiology*. 51, 2189–2198.
- Nakai, K. (2000). Protein sorting signals and prediction of subcellular localization. *Advances in Protein Chemistry*. Vol. 54, pp. 277–344.
- Neupert, W. & Herrmann, J.M. (2007) Translocation of proteins into mitochondria. *Annual Reviews of Biochemistry*. 76:723-49.
- Neuvéglise, C., Bon, E., Lépingle, A., Wincker, P., Artiguenave, F., Gaillardin, C. & Casaregola, S. (2000) Genomic exploration of the hemiascomycetous yeasts: 9. *Saccharomyces kluyveri*. *FEBS Letters* 487(1):56-60.
- Ohno, S. (1970) Evolution by gene duplication. *Springer Verlag*, Berlin.
- Petersen, R. F., Nilsson-Tillgren, T. & Piskur J. (1999) Karyotypes of *Saccharomyces* sensu lato species. *International Journal of Systematic Bacteriology*. 49: 1925–1931.
- Piatigorsky, J., Gene sharing and evolution. (2007) Harvard University Press. Cambridge, MA.
- Pierleoni, A., Martelli, P. L., Fariselli, P., & Casadio, R. (2006). BaCelLo: a balanced subcellular localization predictor. *Bioinformatics (Oxford, England)*, 22(14), e408-16. doi: 10.1093/bioinformatics/btl222.
- Rainieri, S., Zambonelli, C. & Kaneko, Y., (2003) *Saccharomyces* sensu stricto: Systematics, Genetic Diversity and Evolution. *Journal of Bioscience and Bioengineering*, 96, 1-9.
- Replansky, T., Koufopanou, V., Greig, D. & Bell G. (2008) *Saccharomyces* sensu stricto as a model system for evolution and ecology. *Trends in Ecology*. 23(9), 494-501.

- Richard, J.P. (1993) Mechanism for the formation of methylglyoxal from triosephosphates. *Biochemical Society Transactions* 21: 549–553.
- Rout, M. P., Aitchison, J.D., Suprpto, A., Hjertaas, K., Zhao, Y. & Chait, B.T. (2000) The yeast nuclear pore complex: composition, architecture, and transport mechanism. *Journal of Cell Biology*. 148:635–651.
- Rout, M. P. & Wentz, S. R. (1994) Pores for thought: nuclear pore complex proteins. *Trends in Cell Biology*. 4, 357–365.
- Schmidt, O., Pfanner, N., Meisinger, C. (2010) Mitochondrial protein import: from proteomics to functional mechanisms. *Nature Reviews Molecular Cell Biology*. 11: 655-667.
- Sessions, A., Yanofsky, M. & Weigel, D. (2000) Cell-Cell Signaling and Movement by the Floral Transcription Factors LEAFY and APETALA1. *Science*. 289(5480): 779-781.
- Sheff, M. A., & Thorn, K. S. (2004). Optimized cassettes for fluorescent protein tagging in *Saccharomyces cerevisiae*. *Yeast*, 21, 661-670.
- Sherman, D. J., Martin, T., Nikolski, M., Cayla, C., Souciet, J., Durrens, P., et al. (2009). Genolevures: protein families and synteny among complete hemiascomycetous yeast proteomes and genomes. *Nucleic Acids Research*, 37: D550-D554.
- Spírek, M., Yang, J., Groth, C., Petersen, R.F., Langkjaer R.B., Naumova E.S., Sulo P., Naumov G.I. & Piskur J. (2003) High-rate evolution of *Saccharomyces sensu lato* chromosomes. *FEMS Yeast Research*. Jun;3(4):363-73
- Stewart, M. (2007) Molecular mechanism of the nuclear protein import cycle. *Nature Reviews Molecular Cell Biology*. 8:195-208.
- Storici, F., Lewis, L. K., & Resnick, M. a. (2001). In vivo site-directed mutagenesis using oligonucleotides. *Nature biotechnology*, 19(8), 773-6.
- Taylor, J. S., & Raes, J. (2004). Duplication and divergence: the evolution of new genes and old ideas. *Annual review of genetics*, 38, 615-43.
- Trotter, E.W. & Grant, C. (2005) Overlapping roles of the cytoplasmic and mitochondrial redox regulatory systems in the yeast *Saccharomyces cerevisiae*. *Eukaryotic Cell*. 4(2): 392-400.
- Wach, A., Brachat, A., Poehlmann, R. & Philippsen, P. (1994). New heterologous modules for classical or PCR-based gene disruptions in *Saccharomyces cerevisiae*, *Yeast* 10, 1793-1808.
- Williams, E. J., Pal, C., & Hurst, L. D. (2000). The molecular evolution of signal peptides. *Gene*, 253(2), 313-22.
- Yen, K., Gitsham, P., Wishart J, Oliver, S.G., Zhang, N. (2003) An improved tetO promoter replacement system for regulating the expression of yeast genes. *Yeast* 20: 1255 - 1262

- Yu CS, Chen YC, Lu CH, Hwang JK. (2006). Prediction of protein subcellular localization. *Proteins: Structure, Function and Bioinformatics*, 64:643-651.
- Zhang, J. (2003). Evolution by gene duplication: an update. *Trends in Ecology & Evolution*, 18(6), 292-298.

IX) ANEXOS

Contenido

Tabla 1A. Oligonucleótidos para la amplificación de la proteína yECitrine para la fusión con las proteínas de *S. cerevisiae*

Tabla 2A. Oligonucleótidos para la confirmación de la fusión de la proteína yECitrine con las proteínas de *S. cerevisiae*

Tabla 3A. Oligonucleótidos para la amplificación de la proteína yECitrine para la fusión con las proteínas de *L. kluyveri*.

Tabla 4A. Oligonucleótidos para la confirmación de la fusión de la proteína yECitrine con las proteínas de *L. kluyveri*.

Tabla 5A. Oligos para insertar el reportero contraseleccionable.

Tabla 6A. IROS utilizados para remover/adicionar las presecuencias.

Tabla 7A. Oligonucleótidos de confirmación de 1) introducción del CORE. 2) Remoción/Adición de la presecuencia dada por los IROS.

Tabla 8A. Tabla 8A. Oligonucleótidos para deletar los genes GLO2 y GLO4 y para confirmación de la transformación.

Tabla 9A. Oligonucleótidos empleados para el intercambio de los promotores endógenos por el de tet y para su confirmación.

Figura 1A- 4A. Resultados de secuenciación de las cepas que se sometieron a *Delitto Perfetto* alineadas con las secuencias obtenidas en la base de datos SGD.

Tabla1A

Nombre	Secuencia
aap1YFPfo	GAGATCGTGAAATCGTTGCTACCTACTTGCCTGAGCATGAGTATGATCAAGGTGACGGTGCTGGTTTA
aap1YFPpre	CATTGGCATATGTATATATTTGTGTGTAGGTGCTCGTGAAAAAAAAAAGTCGATGAATTCGAGCTCG
ape2YFPfo	AAAGTTTGGCTCAATCACTGGACACTATTACTTCGAAAGCGCAGTGGGGTGGTGACGGTGCTGGTTTA
ape2YFPpre	AATAGTAACCATTTTCCCTCAAATACTTGTGACGACATCACGGTCTCTGTCGATGAATTCGAGCTCG
ald2YFPfo	ATACCTATCTTCAAACAAAAGCAGTTACATAAAATCTCTCTTTGGACAACGGTGACGGTGCTGGTTTA
ald2YFPpre	AGGCTAAGCTATTTAGGTACTTACGCACCTTACATAATGATAACTATCACATCGATGAATTCGAGCTCG
ald5YFPfo	ACTACACTCAAACAAAATCTGTGCAAAATTGCCATTGACAAGCCAATTCGTGGTGACGGTGCTGGTTTA
ald5YFPpre	TGTTTATCATAACATCTTCAATGAGCAGTCAACTCGGGCCTGAGTACTTCGATGAATTCGAGCTCG
alt1FwTag	TTAAGAAATGGGAAAAGTTTCCATAAAGAATTTTTTGACCAATACCGTGACGGTGACGGTGCTGGTTTA
alt1RvTag	ATAAACTAGCTATTTAAATGTTTATTGAAGACTGTTCTGCCCTTTTATTCGATGAATTCGAGCTCG
alt2FwTag	TTCAAGACTGGAAAGAATTTTCATCAAGATTTCTTCAGCAAGTATCGTAATGGTGACGGTGCTGGTTTA
alt2RvTag	ATGGATGCAAAAACAAATAGAAAGCCATTATCATTAGTTTTTTTTTTTCATTCGATGAATTCGAGCTCG
bat1FwTag	ACGGTAGAGTCAATTATGGTAACTGGTCAAAAACCTGTTGCCGACTTGAACGGTGACGGTGCTGGTTTA
bat1RvTag	TTCTCGTTTTTTTTTTTTGGGGGGGAGGGGATGTTACCTTCATTATCATCGATGAATTCGAGCTCG
bat2FwTag	ATGGCGAGACTGAGCATGGCAATTGGTCAAGGGTGTACTGATTTGAACGGTGACGGTGCTGGTTTA
bat2RvTag	AGTTTTATTCTTTTTAACTTTTAATTACTTTACGTAGCAATAGCGATACTTCGATGAATTCGAGCTCG
cit1Fw Tag	TCTCCACCGAAAAATACAAGGAGTTGGTAAAGAAAATCGAAAAGTAAGAACGGTGACGGTGCTGGTTTA
cit1Rv Tag	AAATACGTGTTTGAATAGTCGCATACCCTGAATCAAAAATCAAATTTTCCTCGATGAATTCGAGCTCG
cit2Fw Tag	ATTCTACTGAGAAATACAAGGAATTGGTCAAAAACATTGAAAGCAAACCTAGGTGACGGTGCTGGTTTA
cit2Rv Tag	ATGAGGAAAGAAAAATATGCAGAGGGGTGTAAAAGTAGGATGTAATCCAATCGATGAATTCGAGCTCG
cit3Fw Tag	ACTTGGAGGGTCTGGAAGCACTTACCAAAGCAAGCAATGTTAACAAGTTGGGTGACGGTGCTGGTTTA
cit3Rv Tag	ATTTCAATATTGTAATAGCAAACGAAGCATTCTTGTAAATTTGAACTGCGTCGATGAATTCGAGCTCG
glo2YFPfo	GGGACAGAGCTCAAATTATGGATAAATTGAGAGCTATGAAAAACCGTATGGGTGACGGTGCTGGTTTA
glo2YFPpre	AGGGTATGCTAAAAGAATATTTTAAAGTATATTACTCTGTTGGCTTTCTCGATGAATTCGAGCTCG
glo4YFPfo	ATCCCAGATCCGTAGTGATGCAAGAGTTAAGAAAGCTTAAAAACCGTATGGGTGACGGTGCTGGTTTA
glo4YFPpre	TATCATTAGTCCTCAGACATACTTTGCAAATATTTAGACTGCATGCCAGTCGATGAATTCGAGCTCG
gpd1YFPfo	CAATGAAGAACCTGCCGGACATGATTGAAGAATTAGATCTACATGAAGATGGTGACGGTGCTGGTTTA
gpd1YFPpre	CCTCGAAAAAGTGGGGGAAAGTATGATATGTTATCTTTCTCCAATAAATTCGATGAATTCGAGCTCG
gpd2YFPfo	GCATGGAAGACCTACCGGAGATGATTGAAGAGCTAGACATCGATGACGAAGGTGACGGTGCTGGTTTA
gpd2YFPpre	GGAAAAAGAGGCAACAGGAAAGATCAGAGGGGGAGGGGGGGGAGAGTGTTCGATGAATTCGAGCTCG
grx1YFPfo	TGAGGGAGACTGGTGAATTGGAGGAATTGTTAGAACCCTATTCTTGCAAATGGTGACGGTGCTGGTTTA
grx1YFPpre	TATATAAGCTCATTATAAACCTGTGTGCATGGAAAAAATTTGTCTGCCCTCGATGAATTCGAGCTCG
grx2YFPfo	CTTTGAAGAAAAATGGCAAGTTAGCTGAAATATTGAAGCCGGTATTTCAAGGTGACGGTGCTGGTTTA
grx2YFPpre	AAAGATATATATATGTAATATTATGAAGGGGATATTAGCGTAATTTAAATTCGATGAATTCGAGCTCG
hta1YFPfo	ACTTGTGGCAAAGAAGTCTGCCAAGGCTACCAAGGCTTCTCAAGAATTAGGTGACGGTGCTGGTTTA
hta1YFPpre	GGAGAAGCAGTTTAGTTCCTTCCGCCTTCTTTAAAATACCAGAACCGATTCGATGAATTCGAGCTCG
hta2YFPfo	ACTTGTGGCAAAGAAGTCTGCCAAGACTGCCAAGCTTCTCAAGAAGTGGGTGACGGTGCTGGTTTA
hta2YFPpre	CCCAAATGACAAGAATGTTGATTTGCTTTGTTTCTTTCAACTCAGTTCGATGAATTCGAGCTCG

sds23YFPfo	ATAAGGAAATTGATCCGCAGATGGCAAGAAAGCAAAGAGGGCATATAGGCGGTGACGGTGCTGGTTTA
sds23YFPpre	ATATATGTATATGTAGTTAAATGAAGCCAGTTTTTTTTTCGTCTGGTTTGTTCGATGAATTCGAGCTCG
sds24YFPfo	AAAAATCTTTTACAATGACTAAAAGAAAATATTGTTTCATTTATATGGTTTGGTGACGGTGCTGGTTTA
sds24YFPpre	ATATAAATGTTTTTCAGAACAAGTATATACGCTTAAATGTCTACAAAATACTCGATGAATTCGAGCTCG
trr1YFPfo	CTGGTTGTATGGCCGCTTTGGATGCTGAGAAATACTTAACTTCCCTAGAAGGTGACGGTGCTGGTTTA
trr1YFPpre	GATAAGTATACAAAAATTTGAGTGTATCTATTTTATAATGGAAAATTCATTCGATGAATTCGAGCTCG
trr2YFPfo	CCGGATGCATTGCTGCTTTGGATGCAGAACGGTACCTAAGTGCCCAAGAGGGTGACGGTGCTGGTTTA
trr2YFPpre	TACGTAACACAATCGAACAGTAAAACAAAAATAAACCGGTTTTTTTACGTCGATGAATTCGAGCTCG
ura5YFPfo	AAGAGAAAAGCAAGATCGAACAAATACCTTCAAACCTACGGTGCCTCCGCTGGTGACGGTGCTGGTTTA
ura5YFPpre	TAATAGTTCTTAAAAGAGATAAAATAAATCATTTAATTAATAAAAACTGATTTTCGATGAATTCGAGCTCG
ura10YFPfo	CACCAGAGCAAAAATCAGCGATTGAAAACCTACCGTAAGGCCTATGGTATAGGTGACGGTGCTGGTTTA
ura10YFPpre	CTTTCCTATACAAACAATGATAATCAAATGGTCCATGTCAGTTCTATAGTTCGATGAATTCGAGCTCG

Tabla 1A. Oligonucleótidos utilizados para fusionar las proteínas candidatas de *S. cerevisiae* con yECitrine. En negritas y cursivas se muestran los pares de bases que ayudan a la amplificación del módulo y el resto es la parte que recombina en *S. cerevisiae*

Tabla 2A.

Nombre Oligo	Secuencia	Nombre Oligo	Secuencia
aap1CONFfo	ACGCCAATGGGAACAAACAGG	grx2CONFfo	GCCCTCTTGATTAACGGACAC
aap1CONFpre	CGCACGCCAGTCAATATTCAC	grx2CONFpre	AGACGGAAATTTAGCGGGTCTC
ape2CONFfo	TGCAGCGGTAATAAGGAAGCC	hta1CONFfo	CGGTGGTAAAGGTGGTAAAGC
ape2CONFpre	ATTTCCAGTGTGCACGGGTTG	hta1CONFpre	TGGGCGAATAAGGACAGAAGC
ald2CONFfo	ACTGATGTCCCACAAACATCG	hta2CONFfo	AAGAGGTAACACTACGCCAGAG
ald2CONFpre	ATGCACTAAAGGGCGTGGAAG	hta2CONFpre	GGGCAATTCTGGCTGTAAACG
ald5CONFfo	ATGGGTTAGCCGCAGGTATTC	sds23CONFfo	ACTCCACCCGTAACGACATTG
ald5CONFpre	GGGTGTACAGTCCCGATAATG	sds23CONFpre	TCATGGGACCGATTCTGAAGG
glo2CONFfo	TATTAGGACGCCCTGCCATAC	sds24CONFfo	CGCTCAATCGTCTGCTAATGG
glo2CONFpre	TTCTTTCAGCCACGCGAAGAC	sds24CONFpre	ATGGTCGTGGAGGTTAATGCTG
glo4CONFfo	GTAAAGGGCTGTTGGTGAGAC	trr1CONFfo	ATTGAAATCCTTTACAACACT
glo4CONFpre	CACAGCTCAACGCTATCTTCC	trr1CONFpre	AACGCATCGTGTAGGTCTGTG
gpd1CONFfo	CAAAGAGTCGGTTTGGGTGAG	trr2CONFfo	CGTTATTGGTGGTGGTGACTC
gpd1CONFpre	TACATCCTTGTGAGCCTTGG	trr2CONFpre	TCAGACTGAACGGTGACTTGG
gpd2CONFfo	TCAGGCGGTAGAAACGTCAAG	ura5CONFfo	GGTCCTGCCTACAAAGGTATC
gpd2CONFpre	GCCAAATGCGACATGAGTCAC	ura5CONFpre	TGAAGCAGCTACAGGCAAGAG
grx1CONFfo	GCACGTCAAGGACCTTATTGC	ura10CONFfo	GAGCATCGCTTGAAGACAAGAG
grx1CONFpre	AGAGGGTATGAAGGCCATCTG	ura10CONFpre	TTTCGTCGTCCCAGAAATCAGC

Tabla 2A. Oligonucleótidos para la confirmación de la fusión de la proteína yECitrine con las proteínas de *S. cerevisiae*

Nombre	Secuencia
altSKfw1	ACTGACTGGATCAAAGCTTGGGAAGAATTTACAAAAAGTTTTACAATGAGTACCGTGAC GGTGACGGTGCTG TTTTTTTTCTGTTTTCTGCACAAGCGTCTTAAAAAAAAGTTCTGAGACTGGATATGT TCGATGAATTC
altSKrv1	GAGCTCG
altSKfw2	ACATATCCAGTCTCAGAACCTTTTTTTTAAAGACGCTTGTGCAAGAAAACAGAAAAATAAA
altSKrv2	GTCACGGTACTCATTGTAACCTTTTTGTGAAATTCTCCCAAGCTTGGATCCAGTCAGT
altSKfw3	CCATCTTATCTCCAGAGTCCATC
altSKrv3	CCAAGCATGGGTTTAACACCG TTGCTTCTTGGATCGCTGATATCCAATACGGTAGGGAAAAACACGGAACTGGTCCAAGGTTGTAGC
bat1_T_SkFw	TGACCTAAACAACGGTGACGGTGCTGGTTA TTAAATAACTCTATACAAAACGTTGTGGTCCCATTTATCTTTGGTTCACCTTGTTTCTGTCTTTAAT
bat1_T_SkRe	TTACAATGCT TCGATGAATTCGAGCTCG
BatSKfw2	AAGAAACAAGGTGAACCAAAGATAAATGGGAACCACAACGTTTTGTATAGAGTTATTTAA
BatSKrv2	CCTGGACCAGTTCCGTGTTTTCCCTACCGTATTGGATATCAGCGATCCAAGAAGCAA
bat2_C_SkFw	TATGTCAGCTTCGCTAGGCTC
BatSKrv3	GAGGCTCTATGATTCATCTGTCC GAAAGTCGGTTGGTGATGCTCAAGGTGCTTGGCCTAAAGCCCGAGTGATGGACAAGTTGAGGCAGA
glo_T_SkFw	TGAAAAACCATAT GGTGACGGTGCTGGTTA TCATCTGTGTAATAAATAAACTTGAATCCGGTGGTTTTGCCTCCGCATCAACTAGGATCGCCCCCT
glo_T_SkRe	ACTTAAAGACCT TCGATGAATTCGAGCTCG
gloSKfw2	TCCTAGTTGATGCGGAGGCAAAAACCACCGGATTTCAAGTTTATTATTACACAAGATGA
gloSKrv2	TCAACTTGCCATCACTCGGGCTTTAGGCCAAGCACCTTGAGCATCACCAACCGACTTTC
glo_C_SkFw	AGATTGGCAGCATCGAAGCGG
gloSKrv3	ATGTCCTCGTCATTAGCAGCC

Tabla 3A. Oligonucleótidos para la fusión de la proteína *yECitrine* con las proteínas de *S. kluveri*

Nombre	Secuencia
alt1SKConfFw	aaaccggtgttctgattccaatc
alt1SKConfRv	AGAGAGTGGTACACCTATGGC
bat1SkConfFw	ttctccagtggtccatactac
bat1SkConfRv	GCACTCTTGCATCCGTTTCT
glo2SkConfFw	GATCCTGCCGAACCTCTAGAA
glo2SkConfRv	TGGTAACGGTGGACGCTTTTCA

Tabla 4A. Oligonucleótidos para la confirmación de la fusión de la proteína *yECitrine* con las proteínas de *S. cerevisiae*

Tabla 5A.

Nombre_oligo_CORE Secuencia

alt1_CORE_FW	AAGTCATATCATATAAGGACTCTCACTTCAAGCGCAGAAAAATGCCACAGAGCTCGTTTTCGACTGG
alt1_CORE_RV	AAGCTTTTAACTTTGTACTACTGGCTGAGGTAGAAAAAGGAGTAGTGATATCCTTACCATTAAGTTGATC
bat1_CORE_FW	GAACCTTAGCATACTAAAACATGTTGCAGAGACATTCCTTGAAGTTGGGGGAGCTCGTTTTCGACTGG
bat1_CORE_RV	TTGGATGCATCTAATGGGGCACCAGTAGCGAGTGTCTGATGGAGAATTTTCCTTACCATTAAGTTGATC
bat2_core_FW	ATTTTAGAAAATTTAAGGGAAAGCATCTCCACGAGTTTTAAGAACGATATGGAGCTCGTTTTCGACTGG
bat2_core_RV	TGTTGTGTGGTAGTTATCTTAACTTTGGAGGCGTCTAGGGGTGCCAAGGTTTCCTTACCATTAAGTTGATC
glo2Corefw	GGAAGACTGCAGGACAACACAAAAACACAAGAAAAGGAGGACAAAAATGGAGCTCGTTTTCGACTGG
glo2Corerv	CAGTAATTAACACCACCAGACTCCCATCTCATTTTAATCGATTTAACCTGTCCTTACCATTAAGTTGATC
glo4Corefw	AGCAAATTTCTTTCGGCATGAAGTTTTTACTGCAGCAAATAAGAAATATGGAGCTCGTTTTCGACTGG
glo4Corerv	CTATAATTGACACCACCTGTCAACCATCGCATTTTTATAGGCTTAACATGTCCTTACCATTAAGTTGATC
alt2COREFw	ATAAAAGATAAGAATAAAGGGAAGGAGCGCCACAGTAGTGTGTTTTGTAAAGTGGTAAACACAGA AGAAGTAAGAACACGAGCTCGTTTTCGACTGG
alt2CORERv	TTAGTAATCTTGCCAGCGGGCTTAAATCCAAATCCTTTGCGGTGAACACACCTTTCAAATCCTGTTG GTGTGTCATTGTCCTTACCATTAAGTTGATC

Tabla 5A. Oligonucleótidos para insertar el reportero contraseleccionable (CORE) en los genes candidatos finales. Las letras en cursivas y negritas son las necesarias para amplificar el módulo de CORE y el resto es la parte que recombina con el genoma.

Nombre_IROS	Secuencia
alt1_D70_FW	TTTAATTGTTTCTGCTTCTCAATTGAACGCATATAAAATATATTTCCCAGTCTTTATTTTG CTCTCTCCACGAT GGCTCAAATCATAGACGTTTATTAT
alt1_D70_RV	CGTTATTCGAAGTTTTCAACGTGTGGTGTGGAAAGCGATTCAAATGACGCAGGTCGTTTA GCGAAGATTGA CCAGATAATGAACGTCTATGATTTTGAGC
alt1_D84_FW	TTTAATTGTTTCTGCTTCTCAATTGAACGCATATAAAATATATTTCCCAGTCTTTATTTTG CTCTCTCCACGAT GAACGACCTGCGTCATTTGAATCGCT
alt1_D84_RV	CATTACGTCTTCCAAAGTCAATTGTTCGGCGGGATAAAACTCGTTATTTCGAAGTTTTCA ACGTGTGGTGTG GAAAGCGATTCAAATGACGCAGGTCGTT
alt2DPFw2	TGGTAAACACAGAAGAAGTAAGAACACATGTTATCACTGTCTGCCAAAAATCAC

alt2DPRv2	GTGATTTTTGGCAGACAGTGATAACATGTGTTCTTACTTCTTCTGTGTTTACCA
alt270Fw	AGAGACATAGCTCTTCCTGGATTGTTACAATGACACACCAACAGGATTTGAA
alt270Rv	TTCAAATCCTGTTGGTGTGTCATTGTAACAATCCAGGAAGAGCTATGTCTCT
alt284Fw	GTTCATTATCTGGTCAATCTTCGTAACAATGACACACCAACAGGATTTGAA
alt284Rv	TTCAAATCCTGTTGGTGTGTCATTGTTAGCGAAGATTGACCAGATAATGAAC
alt2IROFw1	AGTATAAAGCGGATACTGCGGAATT
alt2IORv1	CCCATACCTCAACTACCGCCA ACTACATGTTTTCGTTAGAATAAATCACCCTATAAACGCAAAATCAGCTAGAACCCTTAGC ATACTAAAACAT
bat1_D17_FW	GACTGGTGCCCCATTAGATGCATCCAAA GGTGAATGTCTGGCCGAACACTAATTCTTCATTTGGTCTTGGCTTGGATGGATTTGGGTT TCTAGTAATTTT
bat1_D17_RV	TAGTTTGGATGCATCTAATGGGGCACCA
bat2Rcsecrbat1_Fw	CTAACTAAAATTTTAGAAATTTAAGGGAAAGCATCTCCACGAGTTTTAAGAACGATATG TTGCAGAGACATTccttgagttgggaaattctccat
bat2Rcsecbat1_RV	AGCGAGTGTCTGATGGAGAATTTCCCAACTTCAAGGGGCTTAGATGCATGTTGTGTGG TAGTTATCTTAACTTTGGAGGCGCTAGGGGTGCCAAGGT AAAAGCTGTTTGAAGACTAATCTGGAAGTACTGAAAAGAGGAGATTTACTCATTCTTTG GTGCAGCAAATT
iroglo4fw	TCTTTCGGCATGCATGTTAAGCCTATAAA GGTCAATCAGCCATGAGTTTCTCCTGTCTCAGTGCTTAGTAAATAACTATAATTGACAC CACCTGTCAACC
iroglo4rv	ATCGCATTTTTATAGGCTTAACATGCAT TAAGTTGCTAGATTTACCAATATCAAAGGGAAAGACTGCAGGACAACACAAAAACACAA GAAAAGGAGGA
iroglo2fw	CAAAAAATGAAGTTTTTACTGCAGCAAATAA TTTACTGTCCGACAATAGATAGCAGTAATTAACACCACCAGACTCCCATCTCATTTTAAT CGATTTAACCTGC
iroglo2rv	ATATTTCTTATTTGCTGCAGTAAAAAC ACAAGTGCTATACAGCAAATAGCGAACAGTACGAAAGTAAACATCATATTATCAATAAT GATAAAACATATAGTTTCGCCATTCAGGACGAATTTTGTTG
irotrr1fw	GCGGCGGTGTGTGCAGCTGGACCTGAACCAATGATAGTAACTTTGTTGTGAACCATCCTT GACAGCACGGACTTGTCTGATGCCAACAAAATTCGTCCTGA
irotrr1rv	CGCAATGAAAACTTTTAGGGGATCGTCCAACATTAAGGAAGAACGAAACGGACTCCAC AGTTTCTAATATAAATAAACAatgattcatcacaaggttac
irotrr2fw	CCTCATATAATGTGGGCTTCATCTCTGCTCTTGCCAAGTATATAGCAGCGGTGTGGGCAG CGGGGCCAGAACCTATGATTgtaacctgtgatgaatcat
irotrr2rv	

Tabla 6A. IROS utilizados para remover/adicionar las presecuencias.

Tabla 7A.

Nombre_ConfDP	Secuencia
alt1_CONF_FW	TTCTGACCTGCTGTGGTTGAC
alt1_CONF_RV	ATTGGGATGGCACCTCTAACG
bat1_CONF_FW	CAGATTCATACCGGCTGTGC

bat1_CONF_RV	CACAAGCAGATGGGTCAAGAG
bat2_confbat1_FW	GGTACCGGCCTTTAAGCTTTG
bat2_confbat1_RV	TCTTATCCGGACGGAACATGG
alt2confcorefw	GCTGATGTCCTCCACACGTCT
alt2confcorerv	GCTTCTGATCCAGTTGCTGCG
glo2DPconf_fw	ATGCTCGCCGGTTTAGCGATG
glo2DPconf_rv	ACTCTGGCAGAACCTCAGGAG
glo2DPconf_fw	ATGCTCGCCGGTTTAGCGATG
glo2DPconf_rv	ACTCTGGCAGAACCTCAGGAG
glo4DPconf_fw	CATCGCGTGGAGACTGCCTAT
glo4DPconf_rv	CCTCCCACAACCGGCAATGAA
trr1DPconf_fw	GACAGTGCAGCGCAACTAGTG
trr1DPconf_rv	GGCAGCAATACCGTTCGCCAT
trr2DPconf_fw	CCC CTT CAG TGC AGT GGT ACA
trr2DPconf_rv	GTA TAT AGC AGC GGT GTG GGC

Tabla 7A. Oligonucleótidos de confirmación de la introducción del CORE y su posterior remoción.

Tabla 8A.

Nombre	Secuencia
glo2_delfw	ACCAATATCAAAGGGAAGACTGCAGGACAACACAAAAACACAAGAAAAGGAGGACAAAAACGTACGCTGCAGGTCGAC
glo2_delrv	TTGAAATAAAAAGGGTATGCTAAAAAGAATATTTTTAAGTATATTATACTCTGTTGGCTTTTCATCGATGAATTCGAGCTCG
glo4_delfw	TCTGGAAGTACTGAAAAGAGGAGATTTACTCATTCTTTGGTGCAGCAAATTTCTTTTCGGCCGTACGCTGCAGGTCGAC
glo4_delrv	CTGACCCACCTATCATTAGTCCTTCAGACATACTTTGCAAATATTTAGACTGCATGCCAGATCGATGAATTCGAGCTCG
glo2del_Cofw	CGGGCTATTGCAGGTGAGCAA
glo2del_Corv	TCTACTCGAGCCGCCGTATT
glo4del_Cofw	CGCGTGGAGACTGCCTATGAA
glo4del_Corv	GCTGGGGGTGGCTACTTAACT

Tabla 8A. Oligonucleótidos para deletar los genes GLO2 y GLO4 y para confirmación de la transformación.

Tabla 9A.

Nombre	Secuencia
Alt2Fw/pTet	CTATTGTTTCTACTAATGTGCTGCGGGCTGATGTCCTCCACACGTCTTCA CGTACGC TGCAGGTCGACGG
Alt2Rv/pTet	AAATCCTTTGCGGTGAACACACCTTTCAAATCCTGTTGGTGTGTCATTGT CATAGGC CACTAGTGGATCTG
Glo4Fw/pTet	CTCAACTGACAAGGTGGCAAGAACTACAACGACTCAAGTTAGTTGAGAAA CGTAC GCTGCAGGTCGACGG

Glo4Rv/pTet	ATTTTATAGGCTTAACATGCATATTTCTTATTTGCTGCAGTAAAACTT CATAGGC CACTAGTGGATCTG
Trr2Fw/pTet	ACGATACTCCGATTTTCGTCAGATTACGAGAGAATTTGCTCACAAGAATCA CGTACG CTGCAGGTCGACGG
Trr2Rv/pTet	TTGCTGATGCCAACAAAATTCGTCTGAATGGCGAACTATATGTTTTAT CATAGG CCACTAGTGGATCTG
Alt2Fw/pTetConf	GGTAAAAGCGGTTGCTCCTG
Glo4Fw/pTetConf	CAGTTGTTACCGGTTGGTTTGC
Trr2Fw/pTetConf	GTCAGCGATTGCACCATAGTC
pMC325Rv/Conf	ATGCTTGATGGTCGGAAGAGG

ALT1 -70

Región 3' UTR

NoName	811	ccatactaggaaaagctcttccggtactactactgctttctcgctatcttactctagttgttcatactagtccacacggtgctcgagaaca
NoName	49	ccatactaggaaaagctcttccggtactactactgctttctcgctatcttactctagttgttcatactagtccacacggtgctcgagaaca
NoName	901	ggaccagcttctttgtttctccccgcatttaattgttctgcttctcaattgaacgcataataatatttcccagctctttatcttgc
NoName	139	ggaccagcttctttgtttctccccgcatttaattgttctgcttctcaattgaacgcataataatatttcccagctctttatcttgc
NoName	991	tctctccacatgctatacactgtctgccaaaaatcacttcacagtgagtaattctataactcagcttattaagtcatacatataaggac
NoName	229	tctctccacatgctatacactgtctgccaaaaatcacttcacagtgagtaattctataactcagcttattaagtcatacatataaggac
NoName	1081	tctcacttcaagcgcagaaaaatgccacatatcactactcctttttctacctcagccagtagtacaaagttaaaagcttccaggaaagt
NoName		-----
NoName	1171	tagaccogtctacagagacatagctcttctctggattgttgctcaaaatcatagacggtcattatctgggtcaatcttcgctaaacgacct
NoName	242	-----gctcaaaatcatagacggtcattatctgggtcaatcttcgctaaacgacct
NoName	1261	gogtcatttgaatcgctttccacaccacacggttgaaaacttogaataacgagttttatcccgccgaacaattgactttggaagacgtaaa
NoName	292	gogtcatttgaatcgctttccacaccacacggttgaaaacttogaataacgagttttatcccgccgaacaattgactttggaagacgtaaa
NoName	1351	tgaaaatgtcttgaaggctaagtacgccggttagaggtg--ccatcccaatgagagctgaagaattgaaagctcaactggagaaggatcctc
NoName	382	tgaaaatgtcttgaaggctaagtacgccggttagagngnccatccnaat-----

Codón de inicio

Región codificante

ALT1 -84

NoName 31 -----aacgatcgt
 NoName 701 ctaacgattttcgaccttcggtgagacaacttcacgtactcttctgacctgctgtggtgacgtatttagagttgctgcgccctttgtaaacgatcgt

NoName 40 gacccgcgaccatactagaaaagctcttccgttactactactgctttctcgtctattttactctagtttcatactagtcacacggtgctcgagaaca
 NoName 801 gacccgcgaccatactagaaaagctcttccgttactactactgctttctcgtctattttactctagtttcatactagtcacacggtgctcgagaaca

Región 3' UTR

NoName 140 ggaccagcttctttgtttctccccgcatttaattgtttctgcttctcaattgaaacgatataaaatatttccccagctctttattttgctctctccag
 NoName 901 ggaccagcttctttgtttctccccgcatttaattgtttctgcttctcaattgaaacgatataaaatatttccccagctctttattttgctctctccag

Codón de inicio

NoName 240 atg
 NoName 1001 atgcttactactgctgccaataatcacttcacagtgagtaattctataactcaggttattaagtcatatcatataaggactctcacttcaagcgagaaa

NoName -----
 NoName 1101 aaatgccacatatcactactcctttttctacctcagccagtagtacaagttaaaagctttcaggaaagttagaccgctcctacagagacatagctcttc

Región codificante

NoName 243 -----aacgacctgctcatttgaatcgctttccacaccacacggtgaaaact
 NoName 1201 ctggattgttgctcaaaatcatagacggtcattatctggtcaatcttcgctaaacgacctgctcatttgaatcgctttccacaccacacggtgaaaact

NoName 291 tcgaataacgagttttatccccgcgaacaattgactttggaagacgtaaatgaaaatgtcttgaaggctaagtacgcgcttagaggtgcnntnnaat--
 NoName 1301 tcgaataacgagttttatccccgcgaacaattgactttggaagacgtaaatgaaaatgtcttgaaggctaagtacgcgcttagaggtgcatcccaatga

Bat1-17

YHR208W 701 gaatgatcttactgatagagccaggcggttgatactttgtgcagatttcataccggctgctgctatttactgatgaattggctctcttttgttaatggccatccgatccatattgccttcttactgactcaaaat
 NoName 1 -----ctatttactgatga-ttgctctcttttgttaatggccatccgatccatattgccttcttactgactcaaaat

Región 3' UTR

YHR208W 841 tctattgtgtttgccggtaccggcctttaagctttgaaaaaaaaaagcatctgaaaaaaaaatggcactataaagagagctagtggttaacaactacatgttttcggttagaataaatcacctataaacgcaaaatcagcta
 NoName 77 tctattgtgtttgccggtaccggcctttaagctttgaaaaaaaaaagcatctgaaaaaaaaatggcactataaagagagctagtggttaacaactacatgttttcggttagaataaatcacctataaacgcaaaatcagcta

Codón de inicio

YHR208W 981 gaaccttagcactactaaaatcatgtgcagagacattccttgaagttgggaaattctccatcagaacactcgctactggtgccccattagatgcatccaaactaaaaatctactagaaccccaaatccatccaagccaaga
 NoName 217 gaaccttagcactactaaaatcatgt-----actggtgccccattagatgcatccaaactaaaaatctactagaaccccaaatccatccaagccaaga

Región codificante

YHR208W 1121 ccaaatgaagaattagtggtcgccagacattcaccgatcatatgttgaccattccttggtcagcacaagaagggtggggcactccacacatcaagccttacggtaatctttctcttgaccatctgcttggattcca
 NoName 306 ccaaatgaagaattagtggtcgccagacattcaccgatcatatgttgaccattccttggtcagcacaagaagggtggggcactccacacatcaagccttacggtaatctttctcttgac-----

Alt2+70

```

NoName      1  -----nnnttgcttctc
NoName      541 tttatgggttctcaaagtacttcgtcataaattacttcattatacctaagaaaataatttcatttttttctgaatttttagcagaaaaa
NoName      13  tttatgggttctc-aagtacttcgtcataaattacttcattatacctaagaaaataatttcatttttttctgaatttttagcagaaaaa

NoName      631 aaaaaagttcaagtttggcctcattcgcaatattttccatggaaacggcttgatcttgaggcttctcctactgtttggcgatttta
NoName      102 aaaaaagttcaagtttggcctcattcgcaatattttccatggaaacggcttgatcttgaggcttctcctactgtttggcgatttta
                                     Región 3' UTR

NoName      721 tttatttttcctactcttatggcccataatgtacatctacatatctcttcaaattaccaatacacttttaggatagctaaatttactac
NoName      192 tttatttttcctactcttatggcccataatgtacatctacatatctcttcaaattaccaatacacttttaggatagctaaatttactac

NoName      811 ttogtggggaactatattctgttcacttttcttttgatcctttcttgactgacatcatccttttctcattcagtaaatacccaatat
NoName      282 ttogtggggaactatattctgttcacttttcttttgatcctttcttgactgacatcatccttttctcattcagtaaatacccaatat

NoName      901 ttctcatatctctagaggcataaaagataagaataaaggaaggagcgccacagtagtggtgttttgtaagtggttaaacacagaagaa
NoName      372 ttctcatatctctagaggcataaaagataagaataaaggaaggagcgccacagtagtggtgttttgtaagtggttaaacacagaagaa

NoName      991 gtaagaacacatggttatcactgtctgccccaaaatcacttcacagtgagtaattctataactcaggttattaagtcatatcatataaggac
NoName      462 gtaagaacacatggttatcactgtctgccccaaaatcacttcacagtgagtaattctataactcaggttattaagtcatatcatataaggac
                                     Codón de inicio
                                     Región codificante

NoName      1081 tctcacttcaagcgcagaaaaaatgccacatatcactactcctttttctacctcagccagtagtacaagttaaaagctttcaggaaagt
NoName      552 tctcacttcaagcgcagaaaaaatgccacatatcactactcctttttctacctcagccagtagtacaagttaaaagctttcaggaaagt

NoName      1171 tagaccgctcctacagagacatagctcttctctggattggttacaatgacacaccaacaggatttgaaagggtgtgttcaccgcaaaggattt
NoName      642 tagaccgctcctacagagacatagctcttctctggattggttacaatgacacaccaacaggatttgaaagggtgtgttcaccgcaaaggattt
    
```

Alt2+84

```

NoName      541 tttatgggttctcaaagtacttcgtcataaattacttcattatacctaagaaaataatttcatttttttctgaatttttagcagaaaaa
NoName      13  tttatgggttctca-agtacttcgtcataaattacttcattatacctaagaaaataatttcatttttttctgaatttttagcagaaaaa

NoName      631 aaaaaagttcaagtttggcctcattcgcaatattttccatggaaacggcttgatcttgaggcttctcctactgtttggcgatttta
NoName      102 aaaaaagttcaagtttggcctcattcgcaatattttccatggaaacggcttgatcttgaggcttctcctactgtttggcgatttta
                                     Región regulatoria

NoName      721 tttatttttcctactcttatggcccataatgtacatctacatatctcttcaaattaccaatacacttttaggatagctaaatttactac
NoName      192 tttatttttcctactcttatggcccataatgtacatctacatatctcttcaaattaccaatacacttttaggatagctaaatttactac

NoName      811 ttogtggggaactatattctgttcacttttcttttgatcctttcttgactgacatcatccttttctcattcagtaaatacccaatat
NoName      282 ttogtggggaactatattctgttcacttttcttttgatcctttcttgactgacatcatccttttctcattcagtaaatacccaatat

NoName      901 ttctcatatctctagaggcataaaagataagaataaaggaaggagcgccacagtagtggtgttttgtaagtggttaaacacagaagaa
NoName      372 ttctcatatctctagaggcataaaagataagaataaaggaaggagcgccacagtagtggtgttttgtaagtggttaaacacagaagaa

NoName      991 gtaagaacacatggttatcactgtctgccccaaaatcacttcacagtgagtaattctataactcaggttattaagtcatatcatataaggac
NoName      462 gta-gaacaacatggttatcactgtctgccccaaaatcacttcacagtgagtaattctataactcaggttattaagtcatatcatataaggac
                                     Codón de inicio
                                     Errores
                                     Región codificante

NoName      1081 tctcacttcaagcgcagaaaaaatgccacatatcactactcctttttctacctcagccagtagtacaagttaaaagctttcaggaaagt
NoName      551 tctcacttcaagcgcagaaaaaatgccacatatcactactcctttttctacctcagccagtagtacaagttaaaagctttcaggaaagt

NoName      1171 tagaccgctcctacagagacatagctcttctctggattggttgcataaatcatagacggttcattatctgttcaatcttcgctaacaatgac
NoName      641 tagaccgctcctacagagacatagctcttctctggattggttgcataaatcatagacggttcattatctgttcaatcttcgctaacaatgac
    
```