

INSTITUTO POTOSINO DE INVESTIGACIÓN CIENTÍFICA Y TECNOLÓGICA, A.C.

POSGRADO EN CIENCIAS EN BIOLOGIA MOLECULAR Análisis de las secuencias que regulan el procesamiento del transcrito del gen a1 de *Candida glabrata*

Tesis que presenta

Alba Saucedo Fuentes

Para obtener el grado de

Maestro(a) en Ciencias en Biología Molecular

Director de la Tesis: Dra. Irene Beatriz Castaño Navarro

San Luis Potosí, S.L.P., Noviembre de 2011



Constancia de aprobación de la tesis

La tesis "Análisis de las secuencias que regulan el procesamiento del transcrito del gen a1 de Candida glabrata" presentada para obtener el Grado de de Maestro(a) en Ciencias en Biología Molecular fue elaborada por Alba Saucedo Fuentes y aprobada el 01 de Noviembre de 2011 por los suscritos, designados por el Colegio de Profesores de la División de Biología Molecular del Instituto Potosino de Investigación Científica y Tecnológica, A.C.

Dra. Irene Beatriz Castaño Navarro

Dra. Margarita Rodríguez y Domínguez Kessler

Dr. Samuel Lara González



Créditos Institucionales

Esta tesis fue elaborada en el Laboratorio de Microbiología Molecular de la División de Biología Molecular del Instituto Potosino de Investigación Científica y Tecnológica, A.C., bajo la dirección de la Dra. Irene Castaño Navarro, apoyada por el proyecto No. CB-2005-48304.

Durante la realización del trabajo el autor recibió una beca académica del Consejo Nacional de Ciencia y Tecnología (No. de registro 230936) y del Instituto Potosino de Investigación Científica y Tecnológica, A. C.



Instituto Potosino de Investigación Científica y Tecnológica, A.C.

Acta de Examen de Grado

El Secretario Académico del Instituto Potosino de Investigación Científica y Tecnológica, A.C., certifica que en el Acta 089 del Libro Primero de Actas de Exámenes de Grado del Programa de Maestría en Ciencias en Biología Molecular en la opción de Biologia Molecular está asentado lo siguiente:

En la ciudad de San Luis Potosí a los 22 días del mes de noviembre del año 2011, se reunió a las 18:00 horas en las instalaciones del Instituto Potosino de Investigación Científica y Tecnológica, A.C., el Jurado integrado por:

IPICYT
i

a fin de efectuar el examen, que para obtener el Grado de:

MAESTRA EN CIENCIAS EN BIOLOGÍA MOLECULAR

sustentó la C.

Alba Saucedo Fuentes

sobre la Tesis intitulada:

Análisis de las secuencias que regulan el procesamiento del transcrito del gen al de Candida glabrata

que se desarrolló bajo la dirección de

Dra. Irene Beatriz Castaño Navarro

El Jurado, después de deliberar, determinó

APROBARLA

Dándose por terminado el acto a las 20:30 horas, procediendo a la firma del Acta los integrantes del Jurado. Dando fe el Secretario Académico del Instituto.

A petición de la interesada y para los fines que a la misma convengan, se extiende el presente documento en la ciudad de San Luis Potosí, S.L.P., México, a los 22 días del mes de noviembre de 2011.

Dr. Marcial Bonilla Marí Secretario Académico Mtra. Ivonne Lizette cuevas Velez Jefa del Departamento del Posgrado



Dedicatoria

A Natalia Maruri

Agradecimientos

Al Consejo Nacional de Ciencia y Tecnología

Al Instituto Potosino de Investigación Científica y Tecnológica, A. C.

A la Dra. Irene Castaño por la dirección en este trabajo.

A la Dra. Margarita Rodriguez y Dominguez Kessler y al Dr. Samuel Lara por las aportaciones en el escrito.

Contenido

Constancia de aprobación de la tesis	ii
Créditos institucionales	iii
Acta de examen	iv
Dedicatorias	v
Agradecimientos	vi
Contenido	vii
Lista de tablas	ix
Lista de figuras	х
Abreviaturas	xi
Resumen	xii
Abstract	xiii
1. Resumen	1
2. Introducción	2
3. Resultados	5

La proteína a 1 se ubica en el núcleo.	8
---	---

La distancia entre el ORF del gen **a**1 y la secuencia 12 flanqueante 3' es importante para el procesamiento del transcrito primario del gen **a**1 La expresión de la fusión traduccional **a**1::GFP MTL1 de 12 manera extracromosomal requiere de secuencias importantes presentes en el cromosoma necesarias para el procesamiento del mensajero **a**1

*El transcrito a*1 *del locus MTL*2 *se procesa en muy poca* 13 *cantidad*

4.DISCUSION	18
5.Materiales y métodos	25
6. Referencias	30
7. Datos suplementarios	32

Lista de Tablas

Tabla 1 . Fluorescencia emitida por la proteína a 1::GFP a partir de las fusiones a 1::GFP con secuencias flanqueantes5' y 3' de los <i>loci MTL1</i> y <i>MTL2</i> integradas al cromosoma y episomales.		
Tabla 2. Oligonucleótidos utilizados en este estudio.	33	
Tabla 3. Plásmidos utilizados y generados en este estudio.	34	
Tabla 4. Cepas utilizadas y generadas en este estudio.	36	

Lista de figuras

<u>Figura 1</u>	Regiones idénticas entre a 1 en <i>MTL1 y</i> a 1 en <i>MTL2</i> . El <i>ORF</i> del gen a 1 en <i>MTLI</i> es idéntico al de <i>MTL2.</i>	6
<u>Figura 2</u>	Esquema de integración de las fusiones traduccionales a 1::GFP en los <i>loci MTLI y MTL2.</i>	7
<u>Figura 3</u>	Esquema de las fusiones traduccionales de expresión extracromosomales a 1::GFP <i>MTLI y</i> a 1::GFP <i>MTL2.</i>	9
<u>Figura 4</u>	La proteína a 1 colocaliza con el núcleo cuando el gen se transcribe a partir del <i>locus MTLI.</i>	10
<u>Figura 5</u>	Es detectable la proteína a 1 cuando la transcripción se efectúa a partir del l <i>ocus MTLI.</i>	11
<u>Figura 6</u>	La secuencia flanqueante 3' es determinante importante para el procesamiento completo del RNAm del gen a 1.	14
<u>Figura 7</u>	La fase logarítmica de crecimiento en medio mínimo, representa un ambiente propicio para procesar eficientemente la fusión traduccional a partir de <i>MTLI</i> .	16
<u>Figura 8</u>	Las secuencias clonadas, que flanquean la fusión traduccional a 1::GFP en el plásmido replicativo,	16
<u>Figura 9</u>	Existen señales desconocidas en que determinan el procesamiento del RNA mensajero a 1.	24

Abreviaturas

(por sus siglas en inglés)

- CAA <u>C</u>asamino <u>A</u>cids
- **FRT** <u>Flp R</u>ecombination <u>T</u>argets
- *GFP* <u>Green Fluorescent Protein</u>
- *hph* <u>hygromycin ph</u>osphotransferase
- *MTL* <u>mating type-like locus</u>
- ORF <u>Open Reading Frame</u>
- PGK1 3-PhosphoGlycerateKinase
- **RFU** <u>R</u>elative <u>F</u>luorescence <u>U</u>nits
- **UTR** <u>Untranslated Region</u>
- YPD <u>Yeast extract-Peptone-Dextrose</u>
- **5'ss** 5' <u>Splice Site</u>
- **3'ss** 3' <u>Splice Site</u>

Resumen

"Análisis de las secuencias que regulan el procesamiento del transcrito del gen *a1* de *Candida glabrata*. "

Candida glabrata posee tres loci MTL que codifican genes ortólogos a los que confieren identidad celular sexual en Saccharomyces cerevisiae. Los loci MTL y los ortólogos de S. cerevisiae (MAT, HML y HMR) codifican los genes a1, α1 y α2 que dan lugar a proteínas con actividad de factores transcripcionales. En S. cerevisiae y C. glabrata el gen a1 posee dos intrones que se escinden del transcrito primario para dar lugar al producto protéico. En nuestro laboratorio determinamos que en C. glabrata, el procesamiento parcial y completo del transcrito solo ocurre cuando este se origina a partir del locus MTL1, pero no cuando se transcribe a partir de MTL2, aún cuando ambos ORFs del gen a1 son idénticos. En este trabajo analizamos la participación de las secuencias flanqueantes del gen a1 en el procesamiento diferencial del transcrito en C. glabrata, mediante fusiones traduccionales del gen a1 de MTL1 con la proteína GFP, y flanqueado con las secuencias al 5' y al 3' provenientes de MTL1 o de MTL2. Encontramos que se detecta una mayor producción de fluorescencia cuando la fusión a1::GFP está flanqueada por las secuencias propias de MTL1 que cuando está flanqueada por las de MTL2. Lo que sugiere que el gen a1 expresado a partir de MTL2 no está silenciado y no se procesa posiblemente debido a una estructura más compacta de la cromatina en este locus.

PALABRAS CLAVE: *Candida glabrata*, gen **a**1, procesamiento alternativo, secuencias flanqueantes 5' y 3', *MTL1,MTL2*.

Abstract

"Analysis of the role of the sequences flanking the *a1* gene in the processing control of the *a1* transcript in *C. glabrata*."

Candida glabrata contains three mating type-like loci (MTL), which contain orthologous genes to those that confer sexual cell-type identity in Saccharomyces cerevisiae The three MTL loci, like the orthologous loci in S. cerevisiae MAT, HML and *HMR*, contain the **a**₁, α ₁ and α ₂ genes that encode transcription factors. The a1 gene in both species contains two introns that have to be spliced out from the primary transcript in order to produce a functional **a**1 protein. We found that C. glabrata partial and complete processing of the primary **a**1 transcript, only occurs when it is transcribed from MTL1 locus, but not from MTL2, even though both *ORF*s are identical. In this paper we analyzed the participation of the sequences flanking the **a**1 gene in the differential processing of the **a**1 transcript in C. glabrata. To this end we constructed translational fusions of the **a**1 gene (from *MTL*1) with GFP and flanked by 5' and 3' MTL1 or MTL2 sequences. We found that fluorescence of GFP is increased significantly when the fusion is flanked by sequences from the MTL1 locus compared to when the fusion is flanked by MTL2 sequences. Our data suggests that *a*¹ gene from *MTL*² is not silenced and not processed possibly because of a more compact chromatin structure at this locus.

Key words: *Candida glabrata*, **a**1 gene, differential processing, flanking 5' and 3' sequences, *MTL1*, *MTL2*.

Análisis de las secuencias que regulan el procesamiento del transcrito del gen *a*1 de *Candida glabrata*.

ALBA SAUCEDO FUENTES, IRENE CASTAÑO *

IPICYT. Instituto Potosino de Investigación Científica y Tecnológica. Camino a la Presa San José #2055 Lomas 4a sección. San Luis Potosí, San Luis Potosí, SLP, México

* Autor de correspondencia: icastano@ipicyt.edu.mx, Fax: (52) 444-834-2039

Resumen

Candida glabrata posee tres loci MTL que codifican genes ortólogos a los que confieren identidad celular sexual en Saccharomyces cerevisiae. Los loci MTL y los ortólogos de S. cerevisiae (MAT, HML y HMR) codifican los genes a1, α1 y α2 que dan lugar a proteínas con actividad de factores transcripcionales. En S. cerevisiae y C. glabrata el gen a1 posee dos intrones que se escinden del transcrito primario para dar lugar al producto protéico. En nuestro laboratorio determinamos que en C. glabrata, el procesamiento parcial y completo del transcrito solo ocurre cuando este se origina a partir del locus MTL1, pero no cuando se transcribe a partir de MTL2, aún cuando ambos ORFs del gen a1 son idénticos. En este trabajo analizamos la participación de las secuencias flanqueantes del gen a1 en el procesamiento diferencial del transcrito en C. glabrata, mediante fusiones traduccionales del gen a1 de MTL1 con la proteína GFP, y flanqueado con las secuencias al 5' y al 3' provenientes de MTL1 o de MTL2. Encontramos que se detecta una mayor producción de fluorescencia cuando la fusión a1::GFP está flangueada por las secuencias propias de MTL1 que cuando está flanqueada por las de MTL2. Lo que sugiere que el gen a1 expresado a partir de MTL2 no está silenciado y no se procesa posiblemente debido a una estructura más compacta de la cromatina en este locus.

PALABRAS CLAVE: *Candida glabrata*, gen **a**1, procesamiento alternativo, secuencias flanqueantes 5' y 3', *MTL1*, *MTL2*.

1

Introducción

Candida glabrata es la segunda especie más frecuente del género Candida que forma parte de la flora normal humana, solo después de C. albicans (Fidel, et al., 1999).Esta levadura es filogenéticamente más cercana a S. cerevisiae, pero comparte características similares a C. albicans entre las que destaca su capacidad patogénica (Barns, et al., 1991). C. glabrata es el comensal más prevalente en adultos mayores (Lockhart, et al., 1999). Además del estado inmunocomprometido del hospedero, la evolución de una infección por C .glabrata se debe en cierta medida a su resistencia innata a los azoles (Fortun, et al., 1997). En condiciones de crecimiento limitado de nutrientes por tiempos prolongados, C. glabrata lleva a cabo una transición de la forma de levadura a pseudohifa (Csank & Haynes, 2000). C. glabrata, S. cerevisiae, y C. albicans se clasifican de acuerdo con el tipo de genes de apareamiento expresados en determinada condición y estadío celular. La levadura modelo S. cerevisiae posee genes de apareamiento en los loci HML, HMR y MAT. Los dos primeros están silenciados, mientras que el locus MAT se encuentra transcripcionalmente activo. El locus MAT puede expresar información tipo a (genotipo MATa), es decir, se expresa el gen a 1 o bien se expresa información α contenida en los genes $\alpha 1$ y $\alpha 2$ (genotipo *MAT* α). Sin embargo, no expresa información **a** e información α al mismo tiempo. Una levadura MATa puede aparearse con una levadura MATα. Al igual que S. cerevisiae, C. albicans y C. glabrata poseen genes ortólogos a los genes del locus MAT, y están codificados en los loci denominados MTL por sus siglas en inglés mating type-like locus. Debido a su naturaleza diploide, solo las levaduras homocigotas de *C. albicans MTL***a** o *MTL*α, pueden aparearse in vivo (Hull, et al., 2000) e in vitro (Magee & Magee, 2000). En C. glabrata se han identificado tres *loci MTL:* donde *MTL1* puede contener informacion tipo **a** o tipo α según la cepa analizada, MTL2 generalmente contiene informacion tipo a y el locus MTL3 de tipo α . Dentro de su *locus* correspondiente, los genes de apareamiento están separados por una región promotora intergénica compartida (Srikantha, et al., 2003). El gen a1 de S. cerevisiae posee dos intrones; de éstos, el intrón localizado

cerca de la secuencia que codifica la región carboxilo terminal puede ser escindido o bien ser retenido dentro de la secuencia del transcrito y dar como producto dos variantes de procesamiento alternativo, la primera codifica para la proteína a1 mientras que la segunda origina una versión más corta de a1 ya que dentro de la secuencia del intrón se localizan codones de paro. Sin embargo, mediante el empleo de fusiones traduccionales para evaluar el silenciamiento del locus MAT y el transcrito del gen a1, se ha demostrado que sólo el transcrito del gen a1 generado por el procesamiento de ambos intrones (procesado completamente) da lugar a una proteína con actividad funcional (Ner & Smith, 1989). En células diploides la proteína a1 de S. cerevisiae forma un heterodímero junto con α2. Este heterodímero posee actividad represora sobre la expresión de genes que determinan la identidad de tipo sexual celular. Anteriormente en nuestro grupo se evaluó el grado de silenciamiento a lo largo de las regiones cromosómicas MTL2 y MTL3 de C. glabrata mediante inserciones del gen reportero URA3. El locus MTL2 es activo transcripcionalmente lo que difiere del locus ortólogo HMR de S. cerevisiae. El locus MTL3 está mayoritariamente silenciado (Ramirez-Zavaleta, et al., 2010). Así mismo se generó una batería de mutantes nulas sencillas, dobles y una triple mutante nula de cada locus MTL, para analizar el nivel de expresión, mediante RT-PCR del gen a1 a partir de los loci MTL1 y MTL2. Se determinó que C. glabrata expresa un transcrito no procesado del gen a1 cuando la transcripción se origina del locus MTL2, mientras que se procesa de manera parcial y completa cuando se transcribe a partir del locus MTL1. El procesamiento diferencial del transcrito del gen a1 sucede a pesar de que ambos genes son idénticos en sus ORFs y 500 pb en la region 5' flanqueante (excepto por dos cambios en la posición -375 y -360). Hacia la region 3' flanqueante ambas secuencias son idénticas hasta 370 pb río abajo a partir del codón de paro (Ramirez-Zavaleta, et al., 2010).

Alrededor de 1% de los genes de *C. glabrata* contienen intrones. En *S. cerevisiae* el 5% del total de los genes también poseen intrones. En las levaduras, la mayoría de genes que contienen intrones, poseen solo uno. Estos genes interrumpidos por intrones tienden a ser altamente expresados y regulados.

3

Los intrones de levaduras son cortos, el de mayor longitud contiene 1,002 nucleótidos. De los intrones presentes en levaduras, los presentes en genes para proteínas ribosomales son más largos que aquellos que se encuentran en genes que codifican para proteínas no ribosomales. Los genes que contienen intrones ocupan un tercio de la transcripción celular total; tan solo 102 de los 139 genes que codifican para proteínas ribosomales (RPGs) contienen intrones (Pleiss, *et al.*, 2007).

Respecto a las secuencias que conforman a los intrones de levaduras, se puede decir que están muy conservadas, estas son: El branch point, los sitios de splicing al 5' y 3' o 5'ss y 3'ss y regiones ricas en polipirimidina (Schwartz, *et al.*, 2008).

El posicionamiento de nucleosomas y modificación de histonas mediados por la RNA polimerasa II parecen ser determinantes en la remoción y procesamiento alternativo de intrones. La ocupación de nucleosomas a lo largo de los exones es mayor respecto a la de los intrones (Andersson, *et al.*, 2009).

Las modificaciones postraduccionales como acetilación, metilación, fosforilación y ubiquitinación en los extremos amino-terminal de las histonas se presentan preferentemente en los nucleosomas presentes en los exones y los bordes exónintrón que en los de los intrones. La trimetilación H3K4 es determinante en el procesamiento alternativo (Schwartz & Ast, 2010). La asociación del snRNA U1 y de los factores reguladores de procesamiento alternativo de la familia de proteínas SR, es dependiente de cambios en la estructura de la cromatina (Schwartz, *et al.*, 2009). Dichos mecanismos podrían ser dirigidos por información contenida a lo largo de la secuencia del DNA que flanquea el transcrito **a**1 en *MTL*1 y no para el que proviene de *MTL*2. Mediante fusiones traduccionales del gen **a**1 con la proteína verde fluorescente GFP, analizamos la posible participación de las secuencias o regiones flanqueantes al gen **a**1, al 5' y al 3' o RFL 5' y RFL3', tanto de *MTL*1 como *MTL*2, en el procesamiento del pre-RNAm de **a**1. El transcrito del gen **a**1 se procesa completamente a partir del *locus MTL*1 y es clara la detección de la proteína **a**1::GFP en el núcleo celular tanto en fase estacionaria de crecimiento en medio rico YPD como durante la fase logarítmica en medio mínimo. A partir del locus *MTL2* no detectamos claramente procesamiento del transcrito **a**1 indirectamente mediante la localización celular la proteína **a**1::GFP.

Resultados

Anteriormente en nuestro grupo de trabajo, mediante RT-PCR, observamos procesamiento parcial y completo del transcrito del gen **a**1 a partir del locus *MTL*1 y una falta de procesamiento a partir de *MTL*2, durante la fase estacionaria de crecimiento en medio rico YPD (Ramirez-Zavaleta, *et al.*, 2010).

Las secuencias flanqueantes al 5' y 3' del *ORF* del gen **a**1 en los *loci MTL*1 y *MTL2,* pueden estar determinando el procesamiento diferencial del transcrito primario **a**1, es decir; que estas secuencias flanquantes, contengan elementos que dirijan la remoción de ambos intrones del RNA mensajero transcrito a partir del *locus MTL*1 y participen o influyan en la falta de procesamiento del RNA mensajero del mismo gen transcrito a partir del *locus MTL*2.

Mediante fusiones traduccionales en el extremo carboxilo-terminal del *ORF* del gen **a**1 de *MTL*1 con el *ORF* que codifica para la proteína GFP o verde fluorescente, flanqueadas por las secuencias al 5' y 3' provenientes de *MTL*1 o *MTL*2, estudiamos la influencia de estas secuencias flanqueantes en el procesamiento diferencial del RNA mensajero del gen **a**1, como se describe en los siguientes apartados

La construcción de la fusión **a**1::GFP a partir del locus *MTL1*, consideró a partir de - 641 pb de la secuencia flanqueante 5' hasta 699 pb después del codón de paro (Fig. 2A y Fig. 1 Datos suplementarios). La fusión **a**1::GFP a partir del locus *MTL2* consideró a partir de -637pb de la secuencia flanqueante 5' hasta 481pb después del codón de paro de la secuencia flanqueante 3' (Fig. 2B)

Las fusiones traduccionales se integraron en los *loci MTL1* y *MTL2* en la cepa $mtl(1,2,3)\Delta$ (CGM531) de *C. glabrata* por doble recombinación homóloga.

Figura 1.



<u>Figura 1.</u> Regiones idénticas entre **a**1 en *MTL1* y **a**1 en *MTL2*. El *ORF* del gen **a**1 en *MTL1* es idéntico al de *MTL2*. Hay una identidad 500 pb hacia la región flanqueante 5' (RFL 5'), excepto por dos cambios en la posición -375 y -360. Hacia la región flanqueante 3' (RFL 3') ambas secuencias son idénticas hasta 370 pb río abajo a partir del codón de paro y abarcan en *MTL2* la secuencia completa de un fragmento del gen *EMG1* adyacente, mientras que en *MTL1* se abarca parte de la secuencia del mismo gen que se encuentra completo en este locus. El esquema de la región *MTL1* se invirtió con respecto a la orientación real en el cromosoma B para poder mostrar directamente las regiones homólogas en la misma orientación relativa.

Figura 2.



Figura2. Esquema de integración de las fusiones traduccionales a1::GFP en los loci MTL1 y MTL2.

- A) Por doble recombinación homóloga se integró la fusión traduccional a1::GFP MTL1 hph en el locus MTL1 del cromosoma B de C. glabrata (cepa CGM1144). Al escindir el cassette de resistencia (hph) se generan cepas a1::GFP MTL1 integrantes de la fusión traduccional a1::GFP (cepa CGM 1366).
- B) De la misma manera se integró la fusión a1::GFP MTL2 hph en el locus MTL2 cromosoma E (cepa CGM 1364). La escisión de hph, coloca, a la fusión traduccional a1::GFP MTL2, cerca de su RFL3' (cepa CGM 1366).

Además ambas fusiones traduccionales **a**1::GFP flanqueadas por las secuencias 5' y 3' provenientes de *MTL1* o *MTL2* se clonaron en el vector replicativo *CEN/ARS* y se transformaron por separado en el fondo de la cepa CGM531 (Fig 3 y Tabla 3 Datos suplementarios).

La proteína **a**1 se ubica en el núcleo.

Las cepas de *C. glabrata* que contienen las fusiones **a**1::GFP flanqueadas por las secuencias al 5' y 3' provenientes de *MTL1* o *MTL2* fueron analizadas mediante microscopía de fluorescencia. Cuando el transcrito se genera en el locus *MTL1*, la proteína **a**1 se localiza en el núcleo, indicando procesamiento completo del mensajero a partir del *locus MTL1* (Fig 4A). La fluorescencia observada en la cepa con la fusión flanqueada por las secuencia al 5' y 3' a partir de *MTL2* es difusa y distribuida en toda la célula (Fig 4B).

Incluso la cepa parental $mtl(1,2,3)\Delta$ CGM531 y la cepa CGM514 (CGM1/pMC14 GFP::3' UTR) o control negativo de emisión de GFP (vector sin promotor), muestran una señal fluorescente detectable aún sin contener la fusión **a**1::GFP en el genoma (dato no mostrado); en cambio, el control positivo que consiste en la cepa parental transformada con un plásmido que contiene GFP que se transcribe a partir de un promotor fuerte (P_{CTA1}), emite una fluorescencia muy elevada, estos valores son:12,232 RFU en fase logarítmica y 7,556 RFU durante la fase estacionaria en medio rico YPD y de 10,416 RFU durante la fase logarítmica y 6426 RFU durante la fase estacionaria cuando el cultivo se lleva a cabo en medio mínimo.

Figura 3.



<u>Figura 3</u>.Esquema de las fusiones traduccionales de expresión extracromosomales a 1::GFP *MTL1* y a 1::GFP *MTL2*. Ambas fusiones traduccionales a 1::GFP *MTL1* y a 1::GFP *MTL2*, fueron subclonadas en el plásmido replicativo *CEN/ARS* pRB2.0 para estudiar el procesamiento alternativo fuera del contexto cromosómico.

Figura 4.

Α

a1::GFP secuencias 5' y 3' flanqueantes del gen a1

В

a1::GFP secuencias 5' y 3' flanqueantes del gen a1

 MTL1
 MTL2

 DAPI
 GFP
 DIC

 Image: Image:

Figura 4. La proteína a1 colocaliza con el núcleo cuando el gen se transcribe a partir del locus MTL1.

- A) Se muestran tres campos diferentes de la misma cepa, la proteína a1 se localiza en el núcleo en la cepa CGM1298 (a1::GFP-MTL1), esto indica que hay un procesamiento completo del mensajero a partir del locus MTL1.
- B) Se muestran tres campos diferentes de la misma cepa, existe poca emisión de florescencia cuando la fusión *a1*::GFP se transcribe a partir del *locus MTL2*. Esta pequeña señal de fluorescencia parece no localizar con el núcleo y podría ser una auto-florescencia intrínseca de nuestra cepa de *C. glabrata* la cual muestra esta pequeña señal detectable aún sin contener la fusión *a1*::GFP en el genoma y sin teñir (Dato no mostrado).

Figura 5.



Figura 5. La proteína **a**1 es detectable cuando la transcripción se efectúa a partir del l*ocus MTL1*. En una densidad celular correspondiente a 0.8, se detecta mayor emisión de GFP en fase logarítmica que en fase estacionaria en la cepa que contiene la fusión **a**1::GFP en *MTL1* (CGM1298). Esta emisión confirma la presencia del transcrito completamente procesado visto mediante RT-PCR (datos no mostrados y Ramirez-Zavaleta et al., 2010). La células correspondientes a la cepas del control positivo (plásmido que expresa GFP a partir del promotor de *CTA1*) emite 12232 RFU en medio rico YPD en fase logarítmica y 7556 RFU en fase estacionaria, en medio mínimo en fase logarítmica se detectan 10416 RFU y durante la fase estacionaria en medio mínimo 6426 RFU. Como control negativo se utilizó la cepa parental (*mtl*(*1*,*2*,*3*) Δ o CGM531) que emite cierta auto-fluorescencia (~2000 RFUs). La autofluorescencia emitida por la cepa parental se le restó a todas las mediciones de las demás cepas que contienen las fusiones. Todas las gráficas representan las medias obtenidas de más de tres cuantificaciones experimentales.

Cuantificamos la fluorescencia emitida por el número de células equivalente a una densidad óptica de 0.8 a 600 nm. Se cultivaron en medio rico YPD y medio mínimo las cepas que contienen las fusiones traduccionales integradas al cromosoma en *MTL1* y *MTL2*. Aquellas que contienen las fusiones en el plásmido *CEN/ARS* se cultivaron sólo en medio mínimo. Un mayor número de unidades relativas de fluorescencia (RFU) fueron detectadas en la fase logarítmica de crecimiento en medio mínimo en la cepa con la fusión **a**1::GFP-1 o a1::GFP *MTL1* (Fig. 5). Tomamos como control positivo de emisión de fluorescencia a la cepa CGM473 (GCM1/pMC18) que contiene el gen GFP que se transcribe a partir del promotor fuerte de *CTA1* y como control negativo la emisión de autofluorescencia de la cepa parental CGM531 *mtl*(*1,2,3*) Δ (Tabla 3 Datos suplementarios). Todas las gráficas representan medias obtenidas de tres o más cuantificaciones realizadas.

La distancia entre el ORF del gen **a**1 y la secuencia flanqueante 3' es importante para el procesamiento del transcrito primario del gen **a**1

Comparamos la fluorescencia de la cepa con la fusión **a**1::GFP *MTL1 hph* y la fluorescencia de la cepa con la fusión **a**1::GFP *MTL1*, ambas cultivadas en medio rico YPD hasta fase estacionaria (Fig 6). El cassette de resistencia a higromicina o *hph* de la fusión **a**1::GFP *MTL1 hph* (Fig 2A) consiste de : 3' UTR del gen *CTA1*, el promotor del gen *PGK1*, el gen *hph* y el 3' UTR del gen *HIS3*. Determinamos que en fase estacionaria la cercanía de la secuencia flanqueante 3' permite un mayor procesamiento del transcrito **a**1. En fase logarítmica alejar la secuencia flanqueante 3' permite un mayor procesamiento del transcrito del transcrito **a**1 (Tabla 1).

La expresión de la fusión traduccional **a**1::GFP MTL1 de manera extracromosomal requiere de secuencias importantes presentes en el cromosoma necesarias para el procesamiento del mensajero **a**1

Transformamos las fusiones **a**1::GFP *MTL1* <u>CEN</u> y **a**1::GFP *MTL2* <u>CEN</u> en la cepa parental CGM531 (*mtl*(1,2,3) Δ). La fluorescencias emitida por la proteína **a**1::GFP a partir de las fusiones traduccionales **a**1::GFP *MTL1* <u>CEN</u> y **a**1::GFP *MTL2* <u>CEN</u> corresponde a la misma emisión de nuestro control negativo de fluorescencia.

Este resultado puede deberse a la falta de un elemento o secuencia no incluida en las secuencias flanqueantes ya sea al 5`o 3', que puede partcipar en conjunto con la maquinaria de splicing en el procesamiento del RNA mensajero *a1* y cuya participación puede estar presente cuando se realiza el experimento a partir de las fusiones traduccionales integradas en el cromosoma, La expresión episomal disminuye significativamente el procesamiento del transcrito del gen *a1* de manera que no se detecta fluorescencia a partir del plasmido.(Fig 8 A y Fig 8B).

El transcrito **a**1 del locus MTL2 se procesa en muy poca cantidad

La detección casi nula de fluorescencia de la proteína **a**1 por parte de las cepas que contienen la fusión **a**1::GFP en *MTL2* integrada al cromosoma o fuera de este contexto, se debe al procesamiento menos eficiente del mensajero de la fusión **a**1::GFP en *MTL2* (Fig 7). La auto-fluorescencia intrínseca en la cepa parental $mtl(1,2,3)\Delta$ (CGM531) impide discernir, mediante microscopía de fluorescencia, una colocalización nuclear de la proteína **a**1::GFP en las cepas que contienen la fusión traduccional en *MTL2*.

Figura 6.



Figura 6. La secuencia flanqueante 3' es determinante importante para el procesamiento completo del RNAm del gen a1. La cepa a1::GFP-hph (CGM 1144) que contiene el gen de resistencia a higromicina (Fig 2A) fue incluida en la cuantificación de fluorescencia como control para evaluar la influencia de la cercanía de la secuencia 3' flanqueante en el procesamiento del transcrito a1. En fase estacionaria en medio rico YPD la cercanía de la secuencia 3' permite mayor procesamiento transcrito un del a1.

Figura 7.



Β

Logarítmica





La fusión tradicional **a**1::GFP en *MTL1* emite mayor fluorescencia en fase logarítmica en medio mínimo comparada con una menor fluorescencia detectada a partir de la fusión insertada en *MTL2*.Esta podría tratarse en parte, de auto-fluorescencia propia de la célula.



<u>Figura 8</u>. Las secuencias clonadas, que flanquean la fusión traduccional $a_1::GFP$ en el plásmido replicativo, no son suficientes para promover el procesamiento eficiente del transcrito a_1 . La fusión traduccional $a_1::GFP$ en el contexto cromosómico provee de las señales necesarias para llevar a cabo el procesamiento total del mensajero a_1 de *MTL1*, sin embargo, al colocar la fusión de manera episomal el procesamiento decae significativamente. La escasa emisión en la señal de GFP a partir de *MTL2* no indica que se lleve a cabo procesamiento del mensajero a partir de este locus. La auto-fluorescencia intrínseca en la cepa parental $mtl(1,2,3)\Delta$ (CGM531) impide discernir, mediante microscopía de fluorescencia, una posible localización nuclear de la proteína $a_1::GFP$ en las cepas que contienen la fusión traduccional en *MTL2*.

En la siguiente tabla se resumen las unidades relativas de fluorescencia emitidas por la proteína **a**1::GFP a partir de las fusiones traduccionales **a**1::GFP integradas en los *loci MTL1* y *MTL2* (*a*1::GFP *MTL1* y *a*1::GFP *MTL2*) además de la fluorescencia emitida por la cepa que contiene la fusión *a*1::GFP-*hph* donde la secuencia 3' se encuentra alejada por el cassette *hph*, así como de las cepas que portan las fusiones traduccionales **a**1::GFP expresadas a partir del plásmido episomal CEN/ARS (**a**1::GFP *MTL1 CEN* y **a**1::GFP *MTL2 CEN*).

Tabla 1. Fluorescencia emitida por la proteína de fusión **a**1::GFP a partir de las cepas que contienen integrada la fusion *a*1::GFP ya sea en el *locus MTL1* o en *MTL2* integradas al cromosoma y episomales.

Сера	Medie	o rico YPD	MM ó Médio	Mínimo
Genotipo	Fase logarímica (<i>log</i>)	Fase estacionaria (<i>sp</i>)	Fase logarímica (<i>log</i>)	Fase estacionaria (<i>sp</i>)
a1::GFP MTL1	*ND	723	1145	37
a1::GFP-hph	5894	450	ND	294
a1::GFP MTL2			491	ND
a1::GFP MTL1 CEN			150	
a1::GFP MTL2 CEN			474	

RFU (Unidades relativas de fluorescencia)

*ND:No detectable

DISCUSIÓN

El *locus MTL2* de *C. glabrata* generalmente contiene información tipo **a**, que consiste en el gen **a**1 que posee dos intrones. Además, entre el 20 y el 50% de los aislados clínicos de *C. glabrata* tienen una copia idéntica del mismo gen **a**1 en el *locus MTL*1.

El RNA mensajero del gen **a**1 en *C. glabrata*, se procesa parcial y totalmente cuando se transcribe a partir del locus *MTL*1, pero no cuando se transcribe a partir de *MTL*2. Aún cuando los *ORFs* del gen **a**1 en *MTL*1 y *MTL*2 son idénticos, no es posible detectar procesamiento del transcrito primario del gen **a**1 a partir del *locus MTL*2. La presencia de codones de paro dentro de los dos intrones que posee el gen **a**1, hace necesaria su escisión total para obtener el producto proteico completo. Las secuencias al 5' y 3' del *ORF* **a**1 provenientes de *MTL*1 o *MTL*2, podrían determinar el procesamiento diferencial del transcrito primario del gen **a**1.

Para empezar a entender las bases moleculares por las cuales el transcrito del gen **a**1 se procesa diferencialmente cuando la transcripción del gen se origina en el locus *MTL1* o en *MTL2*, construimos fusiones traduccionales del gen **a**1 de *MTL1* con la proteína GFP. Estas fusiones se flanquearon por las secuencias al 5' y 3' provenientes de *MTL1* o *MTL2* y se integraron en los *loci MTL1* y *MTL2* en la cepa que no contiene información de apareamiento en ninguno de los tres *loci MTL* (cepa *mtl*(1,2,3) Δ). Las fusiones traduccionales también se subclonaron y expresaron en plásmidos replicativos *CEN/ARS* para estudiar la participación de las regiones flanqueantes y/o la estructura de la cromatina en el procesamiento del transcrito **a**1 cuando la transcripción se realiza fuera del contexto cromosómico (en plásmidos replicativos).

Utilizando las fusiones traduccionales: 1) Detectamos fluorescencia a partir de la fusión traduccional flanqueada por las secuencias al 5' y 3' provenientes de *MTL1*, durante la fase estacionaria de crecimiento en medio rico YPD y durante la fase logarítmica de crecimiento en medio mínimo, lo que indica procesamiento del

transcrito **a***1* en estas dos condiciones (Fig.5 y Tabla 1). Además la proteína **a**1 que se sintetiza a partir de esta fusión, se localiza principalmente en el núcleo (Fig. 4), 2) Detectamos fluorescencia difusa y no colocalizable con el núcleo a partir de la fusión traduccional integrada y flanqueada por las secuencias al 5' y 3' provenientes de *MTL2* (Fig. 7). 3) No detectamos fluorescencia significativa, ni logramos definir la localización de esta pequeña señal, cuando el procesamiento se da a partir de las fusiones traduccionales expresadas episomalmente.

En *S. cerevisiae* la proteína **a**1 participa en procesos distintos a solamente mantener y establecer una identidad de tipo celular sexual. La proteína **a**1 y α 2 en *S. cerevisiae* forman un heterodímero con actividad represora en la expresión de genes de identidad de tipo celular. Debido a que *C. glabrata* no mantiene una identidad específica de tipo celular (Ramírez-Zavaleta et al., 2010; Mueller et al., 2008) la proteína **a**1 detectada y localizada en el núcleo durante la fase estacionaria de crecimiento en medio rico YPD, nos hace suponer que **a**1 está participando en procesos celulares distintos al de otorgar una identidad específica

Los datos obtenidos en este trabajo sugieren que existen señales en el medio de cultivo, así como la fase de crecimiento que influyen en la eficiencia de procesamiento del transcrito del gen **a**1 (Fig. 5). Posiblemente, estas señales se transmiten a través de las secuencias localizadas al 3' del *ORF* del gen **a**1 (Fig. 6). Además, encontramos que también la estructura de la cromatina (o la eficiencia de la transcripción) se requieren para el procesamiento de este transcrito (Fig. 8).

Mediante la herramienta bioinformática YEASTRACT o Yeast Search for Transcriptional Regulators And Consensus Tracking (*www.yeastract.com*) basada en información obtenida a partir de la levadura *S. cerevisiae*, buscamos secuencias de unión conservadas para proteínas ortólogas modeladoras de la estructura del DNA y factores moduladores de la transcripción presentes en *C. glabrata*, que estuvieran localizadas en las secuencias flanqueantes 5' y 3' del *ORF* del gen **a**1 tanto en *MTL*1 como en *MTL*2. En la secuencia flanqueante 5' de

a1 en ambos *loci*, se presentan probables sitios de unión para la proteína ortóloga a Xbp1 de S. cerevisiae: 4 sitios de unión en MTL1 y uno solo en MTL2. Xbp1 funciona como represor transcripcional con afinidad a secuencias promotoras de genes de ciclinas y su expresión es inducida bajo condiciones de estrés y deprivación de nutrientes durante la mitosis tardía (Mai & Breeden, 1997). La proteína ortóloga putativa a Xbp1 en C. glabrata presenta 35% de identidad respecto a Xbp1 de S. cerevisiae, esta identidad está conservada en el sitio de unión a DNA. En la secuencia flangueante 3' en ambos loci encontramos sitios de unión conservados para la proteína ortóloga Ash1 de S. cerevisiae que posee actividad de regulador negativo de la transcripción del gen HO (Sil & Herskowitz, 1996). A lo largo del ORF del gen **a**1 tanto en MTL1 como en MTL2 existen sitios de unión para las proteínas Ash1, Fkh1 y Fkh2. Estos últimos factores no caracterizados pero presentes en C. glabrata, podrían regular la etapa de elongación de ciertos transcritos ya que en S. cerevisiae se les ha econtrado que participan en la fosforilación de las serinas 5 y 2 del carboxilo terminal de la RNA polimerasa II (Schwartz, et al., 2009). Existen sitios de unión para Fkh1 y Fkh2 en los tres exones y el segundo intrón del gen *a*1 en *MTL*1 y *MTL*2. La eficiencia con la que estos componentes actúan podría explicar el procesamiento diferencial del transcrito de **a**1 en ambos *loci*.

En *S. cerevisiae* la ruta de control de RNA mensajero o procesamiento alternativo mediante el reconocimiento de codones de paro prematuros o vía Nonsensemediated decay (NMD) considera la distancia entre un codón de paro y un sitio especial aún no bien ubicado que parece encontrarse en el 3' UTR o en algunos casos estar cerca o hacia el 3' UTR, para marcar y controlar la degradación de transcritos primarios (Metzstein & Krasnow, 2006). De manera interesante, nuestra fusión traduccional **a**1::GFP construida con secuencias flanqueantes a partir del locus *MTL1* de *C. glabrata* en la que se aleja la secuencia flanqueante 3' por 2 kb (Fig. 6A), puede observarse un aumento significativo en la cantidad de transcrito procesado durante la fase logarítmica en medio rico YPD. El procesamiento del transcrito de la fusión **a**1::GFP es totalmente ausente o de nula detección cuando la misma fusión traduccional tiene la secuencia flanqueante 3' ubicada en su lugar inmediatamente después del *ORF* del gen **a**1 tal y como se dispone en el cromosoma de *C. glabrata.* Los componentes principales que participan en la vía NMD en *S. cerevisiae* se encuentran presentes y conservados en *C. glabrata.* En *C. glabrata* la proteína ortóloga a la proteína Upf2 de *S. cerevisiae* es 55% idéntica y 73% similar a Upf2 de *S. cerevisiae* la cual posee actividad de helicasa de RNA. Upf3 conserva 35% de identidad y 50% de similitud respecto a la proteína Upf3 de *S. cerevisiae*, cuya función es dirigir nucleasas y a las enzimas que eliminan el capping 5' de transcritos primarios.

La velocidad de crecimiento vegetativo también determina el procesamiento. En medio mínimo, durante la fase logarítmica donde la tasa de crecimiento es alta, hay señales no determinadas aún, que podrían determinar cambios en la estructura de la cromatina mediante cambios posiblemente en los patrones de acetilación, pero principalmente de metilación en las histonas, este cambio tiene efecto en la llegada, acceso y posicionamiento de la RNA polimerasa en el promotor del gen a1 del locus MTL1. Observamos que alejar la secuencia flangueante 3' impide el procesamiento del RNA mensajero a1. Cuando la secuencia flanqueante 3' está justo cerca al ORF, se observa procesamiento alto del RNA mensajero. En la fase logarítmica pero en medio rico YPD, se presenta la dinámica contraria, aquí la cercanía de la secuencia flanqueante 3' impide el procesamiento del transcrito. Aparentemente en medio mínimo la cromatina parece más abierta cuando se aleja la secuencia 3', es decir localmente en el *locus MTL1* y en la vecindad del gen **a**1. Además en el medio mínimo la cromatina parece ser más compacta cuando la secuencia flangueante 3' se encuentra cerca al ORF a1 anulando la escisión de ambos intrones.

En fase estacionaria donde el crecimiento es menor, en medio mínimo la lejanía de la secuencia flanqueante 3' mejora el procesamiento del transcrito primario *a*1, mientras que medio rico YPD la cercanía de esta secuencia flanqueante hace posible la escisión de intrones del RNA mensajero *a*1.

21

Se podria determinar el nivel y tipo de modificaciones post-traduccionales en las histonas con el fin de establecer si existe una correlación entre los cambios en la estructura de la cromatina y el grado de procesamiento completo del transcrito del gen *a*1 en el *locus MTL*1. Esto permitiria correlacionar el acceso de la maquinaria de transcripción con el de los factores que controlan el procesamiento que permiten la elongación de transcrito primario y el ensamblaje de la maquinaria de splicing. Scherrer & Spingola, mencionan que el procesamiento puede ser ineficiente en las levaduras, durante el crecimiento vegetativo debido a que los bordes exón-intrón la mayoría de las veces, son sitios de splicing subóptimos (Scherrer & Spingola, 2006).

El hecho de que el transcrito proveniente de *MTL2* no se procese, de alguna manera compensa la falta de silenciamiento en el locus *MTL2* de *C. glabrata*, el cual, a diferencia de lo que ocurre en *S. cerevisiae*, no está silenciado. Si el transcrito de **a**1 se procesara completamente cuando proviene de *MTL2*, esto tendría la consecuencia de que las cepas que contienen información tipo α en *MTL1*, podrían expresar ambos tipos de información de apareamiento en células haploides (información α a partir de *MTL1* e información **a** a partir de *MTL2* que es activo transcripcionalmente). Posiblemente se ha seleccionado en *C. glabrata* el mecanismo por el cual no se procesa el transcrito de **a**1 procedente de *MTL2*, precisamente para evitar que se expresen ambos tipos de información (caracterísitico de células diploides) en células haploides.

Los datos presentados en este trabajo apuntan a que hay probablemente varias señales que producen el procesamiento diferencial del transcrito de **a**1. En el laboratorio continuaremos con estos estudios para encontrar cuales son las señales del medio que disparan un procesamiento eficiente, y cuáles son las secuencias específicas, probablemente localizadas al 3' del *ORF*, así como la estructura de la cromatina que se requiere para procesar completamente el transcrito del gen **a**1.

22

Para continuar con este trabajo, se propone utilizar una fusión del gen **a**1 seguida de tres copias del gen de la proteína verde fluorescente o GFP, que se ha encontrado que en *C. glabrata* aumenta significativamente la señal de fluorescencia por encima de la autofluorescencia intrínseca de la mayoría de las cepas de *C. glabrata*.



<u>Figura 9</u>. Existen señales desconocidas en que determinan el procesamiento del RNA mensajero *a1*. Mediante la cuantificación de unidades relativas de fluorescencia o RFU emitidas por la proteína *a1::GFP* generada a partir del RNA mensajero *a1::GFP* con secuencias flanqueantes al 5' y 3' a partir del *locus MTL1* podemos distinguir factores determinantes del procesamiento de intrones del transcrito *a1:* Moléculas aún desconocidas que componen el medio de cultivo que junto con variantes como la velocidad de crecimiento, cercanía y lejanía de la secuencia flanqueante 3' condicionan el procesamiento del RNA mensajero *a1*.

24

Materiales y métodos

Medios de cultivo

Medios de cultivo para C. glabrata:

YPD contiene: Extracto de levadura 10 g/L, peptona 20 g/L, uracilo 50 mg/ml, glucosa 2%. <u>Agar YPD</u>: Al YPD líquido se le adiciona 20 g/L de agar. <u>Agar YPG</u>: Extracto de levadura 10 g/L, peptona 20 g/L, uracilo 50 mg/ml, glicerol 3%, agar 20g/L. <u>El medio casaminoácidos (CAA)</u> contiene: base nitrogenada de levadura 1.7 g/L, sulfato de amonio 5 g/L, casaminoácidos 6 g/L y glucosa 2%. <u>Agar CAA</u>: Al medio CAA líquido se le adiciona agar 20 g/L y glucosa 2% (p/v).

E. coli se creció en medio Luria-Bertani (LB) que contiene: Extracto de levadura 5 g/L, triptona 10 g/L y NaCl 5g/L. Agar LB: Adicionar 15g/L de agar al LB líquido. <u>Medio SOC</u>: Extracto de levadura 5g/l, triptona 20 g/L, glucosa 0.4 % (p/v), NaCl 10 mM, KCL 2.5 mM, MgSO4 10 mM, MgCl2 10 mM. <u>Medio LB-carbenicilina</u>: Añadir 1mL de carbenicilina (A.G.Scientific®) a una concentración de 50mg/mL por litro de medio (LB-Cb 50).El medio <u>LB-cloramfenicol-sacarosa</u> contiene: Triptona 10g/L, extracto de levadura 5g/L,NaCl 10g/L y sacarosa 50 g/L. Se añadieron 20 µg/mL de concentración final de cloramfenicol (Cm₂₀) para mantener presión de selección plasmídica. Para medio sólido en cajas se agregó agar 15g/L. Todas las construcciones plasmídicas se transformaron en la cepa DH10 por electroporación.

Todos los cultivos, tanto de *E. coli* como de *C. glabrata*, se crecieron a una temperatura de incubación de 30°C.

Clonación de productos de PCR y construcción de fusiones traduccionales **a**1::GFP para integración a los loci MTL1 y MTL2

Fusión traduccional a 1::GFP MTL1 :

El fragmento de 1.2 Kb que contiene el secuencia flanqueante 5' y el ORF del gen **a**1 a parir del locus *MTL1* se amplificó con los oligonucleótidos 769 y 771 (Figura 1 y Tabla 1 Material suplementario) usando como molde DNA genómico de la cepa CGM413 (*mtl*2 Δ) y clonado en el sitio *Stu*l de pMB11 para generar el plásmido pSF1 (Tabla 2 Material suplementario). El fragmento que contiene la secuencia flanqueante 3' del gen **a**1 del locus *MTL1* de 699 pb se amplificó con los oligonucleótidos 772 y 773 (Figura 1 y Tabla 1 Material Suplementario) usando como molde DNA genómico de la cepa CGM413 (*mtl*2 Δ) y clonado en el sitio *Stu*l de pMB11 para generar el plásmido pSF3 (Tabla 2 Material suplementario). pSF1 se cortó con *Bam*HI y se recuperó la banda de 1.214 Kb que se clonó en el sitio *Bam*HI en el plásmido integrativo pOZ16. Esta clonación dio lugar al plásmido intermediario pSF5 (tabla 2 Material suplementario). pSF3 se digirió con *Hind* III y *Xho* I y se recuperó la banda de 699pb que se clonó en los sitios *Hind* III y *Xho* I de pSF5 para generar la fusión traduccional **a**1::GFP flanqueada por secuencias del locus *MTL1* en el plásmido integrativo pSF8 (Tabla 2 Material suplementario).

Fusión traduccional a1::GFP MTL2 :

El fragmento de 481 pb que contiene el secuencia flanqueante 3' del gen **a**1 del locus *MTL2* se amplificó con los oligonucleótidos 776 y 777 (Figura 1 y Tabla 1 Material Suplementario) usando como molde DNA genómico de la cepa CGM390 (*mtl1* Δ) y clonado en el sitio *Stul* en pMB11 para generar el plásmido pSF11 (Tabla 2 Material suplementario). El fragmento que contiene la secuencia flanqueante 5' del gen **a**1 del locus *MTL2* de 637 pb se amplificó con lo oligonucleótidos 774 y 775 (Tabla 1 Material Suplementario) usando como molde DNA genómico de la cepa CGM390 (*mtl1* Δ) y clonado en el sitio *Stul* en pMB11 para generar el plásmido pSF11 (Tabla 2 Material suplementario). El fragmento que contiene la secuencia flanqueante 5' del gen **a**1 del locus *MTL2* de 637 pb se amplificó con lo oligonucleótidos 774 y 775 (Tabla 1 Material Suplementario) usando como molde DNA genómico de la cepa CGM390 (*mtl1* Δ) y clonado en el sitio *Stul* en pMB11 para generar el plásmido pSF13 (Tabla 2 Datos Suplementarios). El fragmento que contiene el *ORF* del gen **a**1 de 522pb a partir del *locus MTL1* se amplificó con los oligonucleótidos 330 y 771 (Figura 1 y Tabla 1 Datos Suplementarios) usando como molde DNA genómico de la cepa CGM413 (*mtl2* Δ) y clonado en el sitio *Stul*

en pMB11 para generar el plásmido pSF16 (Tabla 2 Datos Suplementarios). pSF16 se digirió con BamHI y Xbal y el fragmento de 522 pb se clonó en los sitios BamHI y Spel del plásmido integrativo pOZ16 generando el plásmido intermediario pSF19 (Tabla 2 Datos Suplementarios). pSF13 se cortó con Sacl y Notl y se recuperó la banda de 637 pb que se clonó en los sitios Sacl y Notl de pSF19 para generar el plásmido pSF20 (Tabla 2 Datos Suplementarios). pSF11 se cortó con Sall y Xhol y se recuperó la banda de 481 pb que se clonó en el sitio Sall y Xhol en pSF20 con lo cual se generó finalmente la fusión traduccional a1::GFP flangueada por secuencias del locus MTL2 (plásmido integrativo pSF22). Las fusiones traduccionales contenidas en los plásmidos integrativos pSF8 y pSF22, se escindieron mediante la digestión con Bsgl (que genera secuencias de C. *glabrata*) y se integraron en el cromosoma de la cepa CGM531 *mtl*(1,2,3) Δ de C. glabrata por doble recombinación homóloga. Se eligieron colonias resistentes a higromicina (*hph*^R), se extrajo DNA genómico para comprobar la integración de las fusiones traduccionales tanto en MTL1 como en MTL2 mediante PCR. Las transformantes correctas se transformaron con el plásmido pMZ18 que codifica para la recombinasa Flp1 que escinde el cassette de higromicina a través del reconocimiento y recombinación de los sitios FRT, que flanguean el cassette de resistencia a higromicina. Esta escisión de *hph* coloca la fusión *a*1::GFP contigua a sus respectivas regiones flanqueantes al 3' de los loci MTL1 o MTL2. Se realizó un PCR comprobatorio de escisión del gen *hph* en el cual se espera un fragmento de 3.1 Kb para ambas fusiones a1::GFP a partir de los *loci MTL1* y *MTL2*. Tanto las clonas que contienen integradas en el cromosoma las fusiones a1::GFP en MTL1 y MTL2 con el cassette hph (Hyg^R) como aquellas en donde se hizo la resolución del mismo (Hyg^S), se utilizaron para ensayar fluorescencia.

Construcción de las fusiones traduccionales extracromosomales en plásmidos replicativos **a**1::GFP con regiones flanqueantes de los loci MTL1 y MTL2.

El plásmido pSF8 que contiene la fusión traduccional **a**1::GFP a partir de *MTL1* fue cortado con *Spel* y *Xho*l para recuperar el fragmento de 5Kb que contiene la fusión completa y se subclonó en el plásmido *CEN/ARS* pGRB2.0 en los sitios

Spel y Xhol generando el plásmido pSF23 (Tabla 2 Datos Suplementarios). El plásmido pSF22 se digirió con Sacl y Xhol y se recuperó la banda de 4.8 Kb que contiene la fusión completa a1::GFP con secuencias flangueantes de MTL2 y se subclonó en pGRB2.0 generando el plásmido pSF27 (Tabla 2 Datos suplementarios). Para escindir el cassette de higromicina de 2.3 Kb de pSF23 y pSF27, ambos plásmidos se cortaron con Xbal. Para pSF23 se recuperaron y religaron los fragmentos de 5.5 Kb, que comprende el 5' flangueante y ORF de a1 en *MTL1* y el fragmento de 1.9 Kb que contiene el 3' flanqueante de **a**1 en *MTL1*, dando lugar al plásmido pSF29 (Tabla 2 Material suplementario). De la digestión con Xbal de pSF27 se recuperaron y religaron los fragmentos de 6 Kb, que contiene el 5' y 3' flanqueantes del gen **a**1 en *MTL*2 y el fragmento de 1.2 Kb que contiene la fusión traduccional a1::GFP, dando lugar al plásmido pSF35 (Tabla 2 Datos suplementarios). pSF23 y pSF27 se transformaron en las cepas CGM1 (silvestre) y CGM531 *mtl*(1,2,3)Δ. pSF29 y pSF35 fueron transformados solo en el fondo CGM531 *mtl*(1,2,3) Δ (Tabla 3 Material suplementario). Solo las cepas que contiene a las fusiones extracromosomales **a**1::GFP en *MTL1* o *MTL2* Hyg^s fueron analizadas por fluorescencia.

Ensayos de flourescencia.

Las cepas se sembraron a partir del resguardo en glicerol a -80°C en 5 ml de medio mínimo e incubaron a 28°C con agitación continua durante toda una noche. Las cepas auxótrofas de uracilo: Cepas con la fusión traduccional *a*1::GFP ya sea en el *locus MTL1* o *MTL2*, CGM531 *mtl*(*1,2,3*) Δ y CGM514 (CGM1/pMC14 GFP::3 UTR) (Tabla 3 Material suplementario), se crecieron con uracilo adicionado al medio mínimo. A partir del cultivo crecido toda una noche, se resembraron las cepas en medio mínimo e incubaron a 25°C por dos días hasta alcanzar fase estacionaria de crecimiento que en medio mínimo corresponde a una O.D₆₀₀= 9. Una vez sincronizadas las levaduras en fase estacionaria de crecimiento se volvieron a inocular las cepas en medio mímino para inducir fase logarítmica durante cuatro duplicaciones sucesivas a partir de una O.D₆₀₀ = 0.0625 hasta alcanzar una O.D₆₀₀= 1 o fase logarítmica. Una vez alcanzada la O.D₆₀₀= 1, se

tomó 1 ml de cultivo de cada cepa tanto en fase estacionaria como logarítmica se centrifugó a 13000 rpm por 1min, se descartó el medio mínimo, se lavaron dos veces y resuspendieron en 1ml de agua MiliQ® estéril. Se tomaron 200 µl de células de cada cepa y se colocaron por separado y sucesivamente en pozos de una placa de 96 el volumen correspondiente a una O.D= 0.8, O.D= 0.4 de células en fase logarítmica de crecimiento y en el pozo contiguo, 200 µl células en fase estacionaria de cada cepa. El resto de las células se reservó para teñir con DAPI y llevar a cabo microscopía de fluorescencia. Una vez ordenadas las muestras se analizaron en el flourómetro Multimode Analysis Software 3.3.0.9 DTX 880 Beckman Coulter ® y se cuantificó la emisión de GFP de cada cultivo depositado en cada pozo con un haz excitatorio UV de 485 nm. Los valores obtenidos se expresan como unidades relativas de fluorescencia emitidas (RFU). Ya obtenidas las RFU se realizó un análisis estadístico. De las RFU que emitieron las cepas tanto integrantes de la fusión y cepas con las fusiones extracromosomales se obtuvo el promedio de emisión de GFP. La autofluorescencia emitida por la cepa parental se le restó a todas las mediciones de las demás cepas que contienen las fusiones. Todas las gráficas representan las medias obtenidas de más tres cuantificaciones experimentales. Para todos los experimentos, el control positivo fue la fluorescencia emitida por la cepa CGM 473 que posee un plásmido replicativo con la construcción P_{CgCTA1}CgCTA1::GFP que emite una señal de la proteína GFP de manera constitutiva en toda la célula. Como control negativo de emisión de fluorescencia consideramos las RFU de la cepa CGM 514 ó CGM1/pMC14 GFP::3 UTR, que por carecer de promotor para la transcripción del gen de la proteína GFP, no debería emitir señal alguna (pero emite señal de autofluorescencia). Las cepas generadas y analizadas en este trabajo tienen como fondo genético el correspondiente a la cepa CGM 531 o $mtl(1,2,3)\Delta$, nula de información en los loci de apareamiento.

Referencias

Andersson R, Enroth S, Rada-Iglesias A, Wadelius C & Komorowski J (2009) Nucleosomes are well positioned in exons and carry characteristic histone modifications. *Genome Res* **19**: 1732-1741.

Barns SM, Lane DJ, Sogin ML, Bibeau C & Weisburg WG (1991) Evolutionary relationships among pathogenic Candida species and relatives. *Journal of bacteriology* **173**: 2250-2255.

Csank C & Haynes K (2000) Candida glabrata displays pseudohyphal growth. *FEMS Microbiology Letters* **189**: 115-120.

Fidel PL, Jr., Vazquez JA & Sobel JD (1999) Candida glabrata: review of epidemiology, pathogenesis, and clinical disease with comparison to C. albicans. *CLINICAL MICROBIOLOGY REVIEWS* **12**: 80-96.

Fortun J, Lopez-San Roman A, Velasco JJ, *et al.* (1997) Selection of Candida glabrata strains with reduced susceptibility to azoles in four liver transplant patients with invasive candidiasis. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology* **16**: 314-318.

Hull CM, Raisner RM & Johnson AD (2000) Evidence for mating of the "asexual" yeast Candida albicans in a mammalian host. *Science* **289**: 307-310.

Lockhart SR, Joly S, Vargas K, Swails-Wenger J, Enger L & Soll DR (1999) Natural defenses against Candida colonization breakdown in the oral cavities of the elderly. *Journal of Dental Research* **78**: 857-868.

Magee BB & Magee PT (2000) Induction of mating in Candida albicans by construction of MTLa and MTLalpha strains. *Science* **289**: 310-313.

Mai B & Breeden L (1997) Xbp1, a stress-induced transcriptional repressor of the Saccharomyces cerevisiae Swi4/Mbp1 family. *Mol Cell Biol* **17**: 6491-6501.

Metzstein MM & Krasnow MA (2006) Functions of the nonsense-mediated mRNA decay pathway in Drosophila development. *PLoS Genet* **2**: e180.

Ner SS & Smith M (1989) Role of intron splicing in the function of the MATa1 gene of Saccharomyces cerevisiae. *MOLECULAR AND CELLULAR BIOLOGY* **9**: 4613-4620.

Pleiss JA, Whitworth GB, Bergkessel M & Guthrie C (2007) Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. *PLoS biology* **5**: e90.

Ramirez-Zavaleta CY, Salas-Delgado GE, De Las Penas A & Castano I (2010) Subtelomeric silencing of the MTL3 locus of Candida glabrata requires yKu70, yKu80, and Rif1 proteins. *Eukaryotic cell* **9**: 1602-1611.

Scherrer FW, Jr. & Spingola M (2006) A subset of Mer1p-dependent introns requires Bud13p for splicing activation and nuclear retention. *RNA* **12**: 1361-1372.

Schwartz S & Ast G (2010) Chromatin density and splicing destiny: on the cross-talk between chromatin structure and splicing. *EMBO J* **29**: 1629-1636.

Schwartz S, Meshorer E & Ast G (2009) Chromatin organization marks exon-intron structure. *Nature structural & molecular biology* **16**: 990-995.

Schwartz SH, Silva J, Burstein D, Pupko T, Eyras E & Ast G (2008) Large-scale comparative analysis of splicing signals and their corresponding splicing factors in eukaryotes. *Genome Res* **18**: 88-103. Sil A & Herskowitz I (1996) Identification of asymmetrically localized determinant, Ash1p, required for lineage-specific transcription of the yeast HO gene. *Cell* **84**: 711-722.

Srikantha T, Lachke SA & Soll DR (2003) Three mating type-like loci in Candida glabrata. *Eukaryotic cell* **2**: 328-340.

Datos suplementarios

Figura 1<u>S</u>.



<u>Figura 1S</u>. Sitios de hibridación y fragmentos amplificados para la construcción de las fusiones traduccioneles *a*1::GFP en los *loci MTL1* y *MTL2*.

 Tabla 2. Oligonucleótidos utilizados en este estudio.

NUMERO	SECUENCIA (5' a 3')	SITIO DE HIBRIDACIÓN
16	AGAAATACGCACGAACACGATATAGAGG	HIS3 FW-out
147	CACAAGCTTTTATATAATGGCCCATGCGTC	EMG1@670Hind3 Rv MTL1
150	CCGGTACCTTGTGCAGGTTAGTGTGGCACCTAAATG	Epa20@2801 Kpn Bsg Fw MTL2
181	CCTTAGTGATTCATTTGTATCC	699 twd a1
320	CGCTCTAGACTGCTGAACTGTGACGACCATC	EMG1@135XbaFwMTL1
321	CCCGAATTCGCGTCTTTCCTGTGATTATGATG	BUD5@-35RI-RvMTL1
329	GCTTCTGCAGTATCTACAATAGATAGAAGG	EMG@-42 Fw (Pst)
330	CCCTCTAGAAAAATGATGACAGTAGACCCAATACAAGATC	a1@nt1 Xba Rv
332	GGGAATTCGTCTTTCCTGTGATTATGATGTAAGTTGC	Bud5@-37R1 Rv
409	CGCAGATCTTTATTTGTACAATTCATCCATACC	GFP@@715bp Bgl II REV
410	GCGGGATCCTAATGTCTAAAGGTGAAGGTGAAGAATTATTC	GFP@1bp BamHI FW
468	GACCCAATACAAGATCTACGC	Fwd @13 A1
469	AGATCTTTCGATCTCTTGCGC	Rev @283 A1
604	AAAAATATCAAGTTCCTGAATTCG	sacB@181Fw
605	CAAAGACGATGTGGTAGCC	sacB@321pbRev
764	CCCGGATCCGGTGGTGATCTCTTGTGTGG	BUD5@-424BamRvMTL1a
769	CCCGGATCCTTGTGCAGGGTGGTGATCTCTTGTGTGG	BUD5@-424Bam Bsg Rv
770	CCCGGATCCGTAGAGTGACTTCAGATATTGTACAACTG	alfa2@-424FwBam
771	GCGGATCCAGTCACGATTGTTTAGATCTTTCGATC	a1@494noTAA-GFP-RV Bam
772	CGCAAGCTTCTATAGTTCCTCCTTACTCTTTTATAG	a1@+1HinFw
773	CGGCTCGAGTTGTGCAGCTGCTGAACTGTGACGACCATC	EMG@135Xho Bsg
774	CGCGAGCTCTTGTGCAGGCTCTTCACTCAACGTACTCC	a1MTL2@-637SacBsgFw
775	CCGCGGCCGCATTATTGATTTGTCTAAAGATTTTGAATAG	a1MTL2@-1NotRv
776	GCGGTCGACCTATAGTTCCTCCTTACTCTTTATAG	a1MTL2@+1SalFw
777	GCGCTCGAGTTGTGCAGCTGATATGATTTGGATATTGAAG	emgMTL2@-144XhoBsg Rv

 Tabla 3. Plásmidos utilizados y generados en este estudio.

PLÁSMIDO	DESCRIPCIÓN	OLIGONUCLEOTIDOS	REFERENCIA
pGR2.0	CEN/ARS Ura+	col	ección del laboratorio
pSF3	699 pb hacia 3' después del ATG de a 1 <i>MTL1 Stu I</i> pMB11	# 772 (a1MTL1@+HinFw) # 773 (EMG@135MTL1XhoBsgFw)	colección del laboratorio
pSF5	524 pb <i>ORF</i> a 1 más 666 pb bases río arriba del sitio de inicio de la transcripción de a 1 <i>Bam</i> HI pOZ16	# 769 (BUD5@-424BamBsgRvMTL1a) # 771 (a1@494noTAA-GFP-RvBam)	colección del laboratorio
pSF8	Fusión a1::GFP para integrar en <i>locus</i> <i>MTL1</i>	# 772 (a1MTL1@+HinFw) # 773 (EMG@135MTL1XhoBsgFw) # 769 (BUD5@-424BamBsgRvMTL1a) # 771	colección del laboratorio
pSF11	481 pb después de ATG de a 1 MTL2	# 776 (a1MTL2@+1SalFw) # 777 (emgMTL2@-144XhoBsgRv)	colección del laboratorio
pSF13	604 pb río arriba del sitio de inicio de la transcripción de a 1 del <i>locus MTL</i> 2	# 774 (a1MTL2@-637SacBsgFw) #775 (a1MTL2@-1NotRv)	colección del laboratorio
pSF15	gen α2 de 1.386 Kb	# 770 (alfa2@-424FwBam) # 321(BUD5@-35RI-RvMTL1)	colección del laboratorio
pSF16	522 pb gen a 1 a 494 no TAA	# 771 (a1@494noTAA-GFP-RV Bam) # 330 (a1@nt1 Xba Rv)	colección del laboratorio
pSF19	ORF a 1 Spe I-Bam HI pOZ16	# 771 (a1@494noTAA-GFP-RV Bam) # 330 (a1@nt1 Xba Rv)	colección del laboratorio
pSF20	ORF a 1 Spe I-Bam HI pOZ16 más 604 pb río arriba del sitio de inicio de la transcripción de a 1 del locus <i>MTL</i> 2	# 774 (a1MTL2@-637SacBsgFw) # 775 (a1MTL2@-1NotRv) # 771 (a1@494noTAA-GFP-RV Bam) # 330 (a1@nt1 Xba Rv)	colección del laboratorio

pSF22	Fusión a 1::GFP para integrar en <i>locus MTL2</i>	# 777 (emgMTL2@-144XhoBsg Rv # 774 (a1MTL2@-637SacBsgFw) # 775 (a1MTL2@-1NotRv) # 771 (a1@494noTAA-GFP-RV Bar # 330 (a1@nt1 Xba Rv)	ⁿ⁾ colección del laboratorio
pSF23	Fusión a 1::GFP <i>MTL1</i>	# 773 (EMG@135Xho Bsg)	colección del
	en pGR2.0	# 769 (BUD5@-424Bam Bsg Rv)	laboratorio
pSF27	Fusión a 1::GFP <i>MTL2</i>	# 774 (a1MTL2@-637SacBsgFw)	colección del
	en pGR2.0	# 777 (emgMTL2@-144XhoBsg Rv	laboratorio
pSF29	Fusión a 1::GFP <i>MTL1</i>	# 773 (EMG@135Xho Bsg)	colección del
	en pGR2.0 sin <i>hph</i>	# 769 (BUD5@-424Bam Bsg Rv)	laboratorio
pSF33	<i>ORF GFP</i> pOZ16 en	# 409 GFP@@715bp BgI II REV	colección del
	pGR2.0	# 410 GFP@1bp BamHI FW	laboratorio
pSF35	Fusión a 1::GFP <i>MTL2</i> en pGR2.0 sin <i>hph</i>	# 774 (a1MTL2@-637SacBsgFw) # 775 (a1MTL2@-1NotRv)	colección del laboratorio

Tabla 4. Cepas de utilizadas	s y generadas en este estud	io.
------------------------------	-----------------------------	-----

CEPA	GENOTIPO	FENOTIPO	BIBLIOGRAFÍA
0014			0
CGM1	BG14, Ura3Δ(-85 a +932):: I n903, derivada del alsiado clínico BG2	Kn	Согтаск еt al.1999
CGM531	$mtl(1,2,3)\Delta$	Ura	Salas, 2008
CGM473	CGM1/pMC18	Ura+	Medina,2010
CGM514	CGM1/pMC14 GFP::3 UTR	Ura+	Medina,2010
CGM1144	CGM531/pSF8 Bsgl	hph ^r	Este trabajo
CGM1145	CGM531/pSF8 Bsgl	hph ^r	Este trabajo
CGM1296	CGM1145/pMZ18 <i>mtl</i> (1,2,3) Δ a1::GFP <i>MTL1</i>	hph ^s	Este trabajo
CGM1297	CGM1145/pMZ18 <i>mtl</i> (1,2,3) ∆ a1::GFP <i>MTL1</i>	hph ^s	Este trabajo
CGM1298	CGM1144/pMZ18 <i>mtl</i> (1,2,3) Δ a1::GFP <i>MTL1</i>	hph ^s	Este trabajo
CGM1299	CGM1144/pMZ18 <i>mtl</i> (1,2,3) Δ a1::GFP <i>MTL1</i>	hph ^s	Este trabajo
CGM1342	CGM1/pSF23 a1::GFP 5' y 3' flanquentes MTL1	hph ^{r,} Ura⁺	Este trabajo
CGM1343	CGM1/pSF23 a1::GFP 5' y 3' flanquentes MTL1	hph ^{r,} Ura⁺	Este trabajo
CGM1344	CGM1/pSF27 a1::GFP 5' y 3' flanquentes MTL2	hph ^{r,} Ura⁺	Este trabajo
CGM1345	CGM1/pSF27 a1::GFP 5' y 3' flanquentes MTL2	hph ^{r,} Ura⁺	Este trabajo
CGM1346	CGM531/pSF23 a1::GFP 5' y 3' flanquentes MTL1	hph ^{r,} Ura⁺	Este trabajo
CGM1347	CGM531/pSF23 a1::GFP 5' y 3' flanquentes MTL1	hph ^{r,} Ura⁺	Este trabajo
CGM1348	CGM531/pSF27 a1::GFP 5' y 3' flanquentes MTL2	hph ^{r,} Ura⁺	Este trabajo
CGM1349	CGM531/pSF27 a1::GFP 5' y 3' flanquentes MTL2	hph ^{r,} Ura⁺	Este trabajo
CGM1358	CGM531/pSF33 ORF del gen GFP en pGRB2.0 sin 3' UTR	fluorescente ⁻ ,Ura+	Este trabajo
CGM1359	CGM531/pSF33 ORF del gen <i>GFP</i> en pGRB2.0 sin 3' UTR	fluorescente ⁻ ,Ura+	Este trabajo
CGM1360	CGM531/pSF29 a1::GFP MTL1 sin hph	Ura+	Este trabajo
CGM1361	CGM531/pSF29 a1::GFP MTL1 sin hph	Ura+	Este trabajo
CGM1362	CGM531/pSF35 a1::GFP MTL2 sin hph	Ura+	Este trabajo
CGM1363	CGM531/pSF35 a1::GFP MTL2 sin hph	Ura+	Este trabajo
CGM1364	CGM531/pSF22 Bsg I	hph ^r	Este trabajo
CGM1365	CGM531/pSF22 Bsg I	hph ^r	Este trabajo
CGM1366	CGM1364/pMZ21 mtl(1,2,3) Δ a1::GFP MTL2	hph ^s	Este trabajo
CGM1367	CGM1364/pMZ21 mtl(1,2,3) Δ a1::GFP MTL2	hph ^s	Este trabajo
DH10	F- <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) Φ80dlacZΔM15 ΔlacX74 deoR recA1 endA1 araD139 Δ(ara,leu)7697 galU galKλ- rpsL nupG	<i>Tc</i> ' Sm'	Calvin N M et al, 1988