

DEPARTAMENT DE BIOQUÍMICA I BIOLOGIA  
MOLECULAR

ESTUDIO GENÓMICO DE LA TRANSCRIPCIÓN Y DE LA  
DEGRADACIÓN DE LOS mRNAs EN *SACCHAROMYCES*  
*CEREVISIAE*

VICENTE JOSÉ PELECHANO GARCÍA

UNIVERSITAT DE VALÈNCIA  
Servei de Publicacions  
2009

Aquesta Tesi Doctoral va ser presentada a València el dia 30 de març de 2009 davant un tribunal format per:

- Dr. Fernando González Candelas
- Dr. Joaquín Dopazo Blázquez
- Dr. Sebastián Chávez de Diego
- Dr. Vicente Tordera Donderis
- Dra. Susana Rodríguez Navarro

Va ser dirigida per:  
Dr. José Enrique Pérez Ortín

©Copyright: Servei de Publicacions  
Vicente José Pelechano García

---

Dipòsit legal: V-3756-2009  
I.S.B.N.: 978-84-370-7512-9

Edita: Universitat de València  
Servei de Publicacions  
C/ Arts Gràfiques, 13 baix  
46010 València  
Spain  
Telèfon:(0034)963864115



**Estudio genómico de la  
transcripción y de la degradación  
de los mRNAs en  
*Saccharomyces cerevisiae***

Trabajo realizado en el Departamento de Bioquímica y  
Biología Molecular de la Universitat de València por Vicente  
José Pelechano García para optar al grado de Doctor en  
Bioquímica

Valencia, Febrero de 2009

JOSÉ ENRIQUE PÉREZ ORTÍN, Doctor en Ciencias Biológicas y Catedrático del Departamento de Bioquímica y Biología Molecular de la Universitat de València

CERTIFICA que Vicente José Pelechano García, Licenciado en Bioquímica por la Universitat de València, ha realizado bajo mi dirección el trabajo que bajo el título de "Estudio genómico de la transcripción y de la degradación de los mRNAs en *Saccharomyces cerevisiae*" presenta para optar al grado de Doctor en Bioquímica por la Universitat de València.

Burjassot, Febrero de 2009

Fdo. Dr. José Enrique Pérez Ortín

A la meva família,

## Agradecimientos:

En primer lugar quiero agradecer a mi director, el Dr. José Enrique Pérez, por brindarme la oportunidad de realizar la tesis en su grupo y ayudarme a empezar en este complicado mundo que es la ciencia. También quiero agradecer al Dr. José García que siempre me ha ayudado y aconsejado durante todos estos años y a la Dra. Paula Alepuz con cuya ayuda siempre he podido contar.

Quiero agradecer muy especialmente a todos los compañeros y amigos que han pasado por nuestro laboratorio. Especialmente quiero agradecerlo a Oreto, que me contestó hace ya muchos años a la pregunta “¿Pero, para que sirve esto de los chips de DNA?”, a Lorena con la que he tenido tantas discusiones científicas, y de las otras, de camino a casa, a Fany que nos ha aguantado con buena cara mientras discutíamos de ciencia en los almuerzos y a Helena que siempre ha estado ahí. Pero no quiero olvidarme de muchos otros como Toni, Peio, Priyanka, Elena, Lola, Juan, Bea, Rafa, Trini, Laura,... que tanto me han ayudado durante estos años.

También quiero agradecer las fructíferas discusiones que he tenido con toda la gente de los grupos del Dr. Sebastián Chavez y Dr. Francisco Navarro en las reuniones anuales de Baeza, tanto dentro como fuera de la sala de reuniones. Especialmente he aprendido mucho de los miembros de esos grupos que han pasado por Valencia, como Silvia, Alfonso, Fernando y Mamen.

Quiero agradecer a toda la gente que ha realizado estancias en nuestro laboratorio durante estos años compartiendo con nosotros su forma de trabajar y ver la ciencia.

La última etapa de la tesis ha sido mucho más agradable en compañía de toda la gente que hemos compartido el despacho de becarios, donde he podido seguir en directo las diferentes etapas del proceso de escritura de una tesis. Por supuesto también quiero agradecer a todos los componentes del departamento de Bioquímica que siempre han estado dispuestos a prestarme su ayuda.

También tengo que agradecer este trabajo a mis amigos de Alpeia y de la Banda de Música que me han ayudado a desconectar de la ciencia de vez en cuando.

Finalmente quiero agradecerles a mi familia. Yo no estaría escribiendo esto si no fuera por el gran apoyo de toda mi familia que, aunque no comprenden exactamente que es lo que estoy haciendo, me animan a seguir adelante.

Muchas gracias a todos.

# Índice

<b>Índice .....</b>	<b>VI</b>
<b>Abreviaturas: .....</b>	<b>X</b>
<b>Introducción.....</b>	<b>2</b>
<b>La levadura <i>S. cerevisiae</i> como organismo modelo .....</b>	<b>2</b>
<b>Estudio globales y tecnologías ómicas .....</b>	<b>2</b>
Técnicas genómicas. Los chips de DNA.....	3
<b>El proceso de la transcripción .....</b>	<b>5</b>
Transcripción por la RNA polimerasa II .....	5
<b>Técnicas genómicas para el estudio de la transcripción .....</b>	<b>8</b>
GRO: Genomic Run-on .....	8
ChIP-on-chip: estudio genómico de la localización de proteínas..	10
<b>Degradación de los mRNAs en eucariotas .....</b>	<b>10</b>
Rutas de degradación de mRNAs en eucariotas .....	11
Equilibrio entre <i>p-bodies</i> y traducción. ....	13
Técnicas clásicas para el estudio de la estabilidad de los mRNAs	13
Cinética de la degradación de los mRNAs .....	15
<b>Estrategias de expresión génica .....</b>	<b>16</b>
Diferentes estrategias respecto a las tasas de síntesis y degradación. .....	17
Estado estacionario para la expresión génica .....	17
Ruido biológico para la expresión génica.....	18
<b>Genes RP y regulón RiBi como genes altamente regulados .....</b>	<b>19</b>
<b>Silenciamiento transcripcional.....</b>	<b>21</b>
Maquinaria de silenciamiento transcripcional en levadura .....	21
El silenciamiento transcripcional como mecanismo de regulación génica.....	22



<b>Objetivos .....</b>	<b>25</b>
<b>Materiales y métodos .....</b>	<b>27</b>
<b>Técnicas microbiológicas .....</b>	<b>27</b>
Cepas de levadura .....	27
Medios y condiciones de cultivo .....	27
<b>Técnicas de Biología Molecular.....</b>	<b>28</b>
Extracción de ácidos nucleicos .....	28
Extracción de RNA .....	28
Extracción de DNA.....	28
PCR.....	29
PCR cuantitativa .....	29
<i>Northern</i> blot.....	30
Cálculo de vidas medias de mRNAs.....	31
Cuantificación de RNA poliadenilado .....	32
Oligonucleótidos utilizados en este trabajo .....	33
<b>Técnicas genómicas .....</b>	<b>33</b>
Hibridación de Chips de DNA.....	33
Cálculo genómico de la cantidad de los mRNAs .....	34
Marcaje de cDNA usando oligo dT .....	34
Marcaje de cDNA usando cebadores aleatorios .....	35
Marcaje de DNA genómico .....	35
RNA Pol II ChIP-on-Chip (RPCC) .....	36
Genomic run-on .....	39
Genomic run-on fluorescente de alta densidad.....	41
<b>Técnicas Bioinformáticas .....</b>	<b>45</b>
Análisis de imagen y normalización .....	45
Análisis genérico para los <i>macroarrays</i> de Nylon.....	45
Normalización de la señal de transcripción obtenida mediante GRO .....	46
Normalización de la señal de transcripción obtenida mediante RPCC .....	47
Análisis de los chips de embaldosado de Affymetrix .....	47
Análisis bioinformático de los resultados.....	47

**Resultados y discusión ..... 50**

**1. Estudio del estado estacionario de la expresión génica..... 51**

- 1.1 Estado estacionario de la expresión génica durante el crecimiento exponencial ..... 51
- 1.2 Reanálisis de datos genómicos de variación de mRNAs durante el crecimiento exponencial ..... 52
- 1.3 Análisis detallado de los cambios de cantidad de mRNA y tasa de transcripción durante el crecimiento exponencial. .... 56
- 1.4 Desviaciones del estado estacionario para la expresión génica58
- 1.5 Asunción del estado estacionario para la expresión génica.... 63
- 1.6 Discrepancias entre cálculos directos e indirectos para el estudio del recambio de los mRNAs ..... 65

**2 Cálculo Genómico de estabilidades de mRNA ..... 69**

- 2.1 Medida de vidas medias individuales con tiolutina..... 69
- 2.2 Comprobación del efecto inhibitorio de la tiolutina sobre la degradación de los mRNAs usando mutante *rpb1-1* ..... 73
- 2.3 Estudio del efecto de la tiolutina sobre la desadenilación de los mRNAs ..... 74
- 2.4 Efecto diferencial de la estabilización del los mRNAs a escala genómica..... 76
- 2.5 Modelo hipotético de inhibición de la degradación de mRNAs por la tiolutina..... 79
- 2.6 Uso de la tiolutina como inhibidor transcripcional ..... 80
- 2.7 Medida de vidas medias genómicas ..... 83
- 2.8 Comparación de estabilidades de mRNAs medidas con métodos genómicos diversos..... 85

**3. Desarrollo de nuevas técnicas genómicas para el estudio de la transcripción ..... 88**

**3.1. Cálculo de la transcripción basado en ChIP-on-chip ..... 88**

- 3.1.1 RPCC: RNA Pol ChIP-on-chip ..... 88
  - Diseño experimental del RPCC ..... 89
  - Cálculo de la tasa de transcripción ..... 90
  - Comparación de cálculos genómicos de la tasa de transcripción..... 93
  - Uso de diferentes anticuerpos para el RPCC ..... 99
  - RPCC: ventajas e inconvenientes ..... 103

3.1.2. Correcciones de las tasas de transcripción del GRO usando los datos de RPCC .....	103
3.1.3 Comparación entre diferentes aproximaciones genómicas para el estudio de la transcripción .....	110
Estudio funcional de las diferencias entre GRO y RPCC .....	110
Comparación RPCC/GRO en diferentes condiciones de crecimiento.....	116
Regulación de la acumulación de maquinaria transcripcional en los genes RP .....	125
Efecto de la ruta PKA en la acumulación de maquinaria transcripcional en los genes RP durante el crecimiento en glucosa. ....	129
Efecto de dominio silenciador de Rap1p en la acumulación de maquinaria transcripcional en los genes RP .....	133
Modelo de regulación génica mediante la maquinaria de silenciamiento génico.....	137
<b>3.2. GRO fluorescente de alta densidad.....</b>	<b>141</b>
Diseño experimental .....	142
Incorporación <i>in vivo</i> de diferentes nucleótidos modificados.	142
Purificación de RNA nascente.....	145
Perspectivas de aplicación del GRO fluorescente.....	152
<b>Conclusiones .....</b>	<b>155</b>
<b>Bibliografía .....</b>	<b>158</b>

# Abreviaturas:

**Ac:** anticuerpo

**ArrayCGH:** hibridación genómica comparada usando chips de DNA

**BAC:** cromosoma bacteriano artificial

**cDNA:** DNA copia producto de una retrotranscripción

**DNA pol:** DNA polimerasa

**dpm:** desintegraciones por minuto

**gDNA:** DNA genómico

**GRO:** *Genomic Run-on*

**GTF:** factores generales de transcripción

**IP:** Inmunoprecipitación

**mRNA:** RNA mensajero

**p-adj:** p-valor ajustado por test múltiple

**PcG:** proteínas del grupo Polycomb

**PKA:** protein quinasa A

**PrPP:** Pausa próxima al promotor

**p-valor:** nivel de significación  $\alpha$  más pequeño que permite rechazar la hipótesis nula

**qPCR:** PCR cuantitativa

**RA:** cantidad de mRNA

**rDNA:** región del DNA que codifica los rRNAs

**RiBi:** genes de Biogénesis de Ribosomas

**RNA pol:** RNA polimerasa

**RP:** genes de las proteínas ribosómicas

**RPCC:** RNA Pol ChIP-on-chip

**rRNA:** RNA ribosómico

**Sarkosil:** lauroilsarcosinato de sodio

**SNP:** polimorfismos de un único nucleótido

**TOR:** ruta “*target of Rapamycin*”

**TR:** tasa de transcripción

Nota: A lo largo de este trabajo se ha usado la convención anglosajona de separar los decimales con un punto.

X

# **Introducción**

# Introducción

## La levadura *S. cerevisiae* como organismo modelo

La levadura *Saccharomyces cerevisiae* es un organismo ampliamente utilizado por el hombre desde la antigüedad. Es el organismo responsable de la fermentación de las bebidas alcohólicas como la cerveza y el vino, así como el utilizado para la fabricación del pan. *S. cerevisiae* es una levadura de vida libre que pertenece a la clase de los hongos hemiascomicetos y que puede alternar entre ciclos de vida haploide y diploide. *S. cerevisiae* también es el eucariota unicelular más utilizado como organismo modelo debido a su facilidad de cultivo y manipulación genética. Además, a pesar de la gran distancia evolutiva que la separa de los eucariotas superiores, sus procesos celulares básicos están muy conservados. Todo esto hace que exista una extensa comunidad científica que la usa como modelo de estudio, no sólo por sus aplicaciones biotecnológicas sino también para comprender mejor el funcionamiento de procesos celulares básicos.

*S. cerevisiae* en la forma haploide tiene un tamaño de 4 µm de diámetro y un genoma con una secuencia completa de unas 12 Mb en las que hay unos 6000 genes. Por su importancia como organismo modelo, así como por el pequeño tamaño de su genoma, fue el primer organismo eucariota en ser secuenciado (Goffeau A. *et al* 1996). Además, desde la aparición de las técnicas genómicas que permiten un abordaje global de los procesos celulares *S. cerevisiae* se ha convertido en el principal organismo utilizado en el desarrollo de estas técnicas.

## Estudio globales y tecnologías ómicas

En los últimos años se han desarrollado un gran número de técnicas para el estudio de los problemas biológicos que se caracterizan por la obtención de una gran cantidad de datos a partir de un único experimento y el uso de herramientas bioinformáticas para el estudio posterior de los resultados. Estas estrategias, llamadas “ómicas”, se diferencian de las técnicas clásicas de biología molecular en que, en vez de estudiar con detalle un único proceso (gen, proteína,...), permiten un abordaje “global” del problema. Entre las diferentes ómicas podemos destacar la genómica (estudia los genes o secuencias de DNA en general), la proteómica (estudia las proteínas) o la metabolómica (estudia los

metabolitos), aunque cada vez hay más campos de la biología que usan este tipo de aproximaciones globales.

La principal ventaja de las tecnologías ómicas es que permiten realizar simultáneamente una gran cantidad de medidas. Sin embargo, la desventaja intrínseca asociada a este tipo de técnicas es que, aunque el error experimental asociado a cada una de las medidas individuales sea pequeño, debido al gran número de medidas, un significativo número de los resultados serán erróneos. Para disminuir este error es necesario realizar réplicas experimentales de cada experimento y usar herramientas estadísticas para analizar los resultados. Además, la gran cantidad de datos que se obtienen con cada experimento hace imprescindible el uso de herramientas bioinformáticas para la interpretación de los resultados. De esta manera, el análisis de los resultados y la integración de la ingente cantidad de datos “ómicos” se ha convertido por sí misma en una importante disciplina dentro del campo de la bioinformática.

### **Técnicas genómicas. Los chips de DNA.**

Las técnicas genómicas son aquellas que permiten analizar de forma simultánea una gran cantidad de genes u otras secuencias de DNA en un único experimento. Y dentro de este campo, una de las tecnologías más importante y ampliamente utilizada es la de los chips de DNA (también llamados micromatrices o “*microarrays*”).

Los chips de DNA consisten en un soporte sólido sobre el que, por métodos automatizados, se fijan secuencias de DNA conocidas (sondas) organizadas sobre la superficie. Posteriormente, gracias a la hibridación por complementariedad de los ácidos nucleicos, las moléculas de una muestra problema se unen preferentemente a la región que contiene sondas con secuencia complementaria. De esta forma se puede determinar que moléculas y en qué cantidad estaban presentes en una muestra.

Existe una gran variedad de chips de DNA diferentes atendiendo a su fabricación. Así, se pueden fabricar sobre diferentes soportes, siendo el más común las láminas de vidrio recubiertas con grupos reactivos, aunque también se pueden usar membranas de nylon (los llamados “*macroarrays*”). Por lo que respecta al tipo de sondas que se depositan sobre el chip pueden ser fragmentos de diferente longitud de DNA de doble o simple cadena. Además, dependiendo de la tecnología usada, las sondas se pueden sintetizar “*in vitro*” y ser depositadas posteriormente sobre el soporte o sintetizarse directamente “*in situ*” sobre el soporte.

## Introducción

Con esta gran variedad de tecnologías diferentes los chips de DNA pueden poseer desde unos pocos cientos a millones de sondas.

Según el diseño y el número de sondas en los chips de DNA se pueden encontrar desde chips con BACs como sondas que cubren grandes regiones de los cromosomas y cuya principal aplicación es la hibridación genómica comparada (*ArrayCGH*) para el estudio de anomalías cromosómicas, hasta los chips con oligos de embaldosado que han surgido en los últimos años y que cubren la totalidad del genoma solapando las sondas entre ellas y ofreciendo una resolución de hasta unos pocos pares de bases (fig 1). Sin embargo, los chips más comúnmente utilizados actualmente son los chips de cDNAs (con sondas que cubren completamente los genes del organismo) o aquellos que poseen uno o varios oligos dispuestos en el interior de cada gen. En cuanto al diseño también se puede diferenciar entre chips con sondas específicas para cada una de las hebras o con sondas de doble cadena que no permiten distinguir entre hebras.

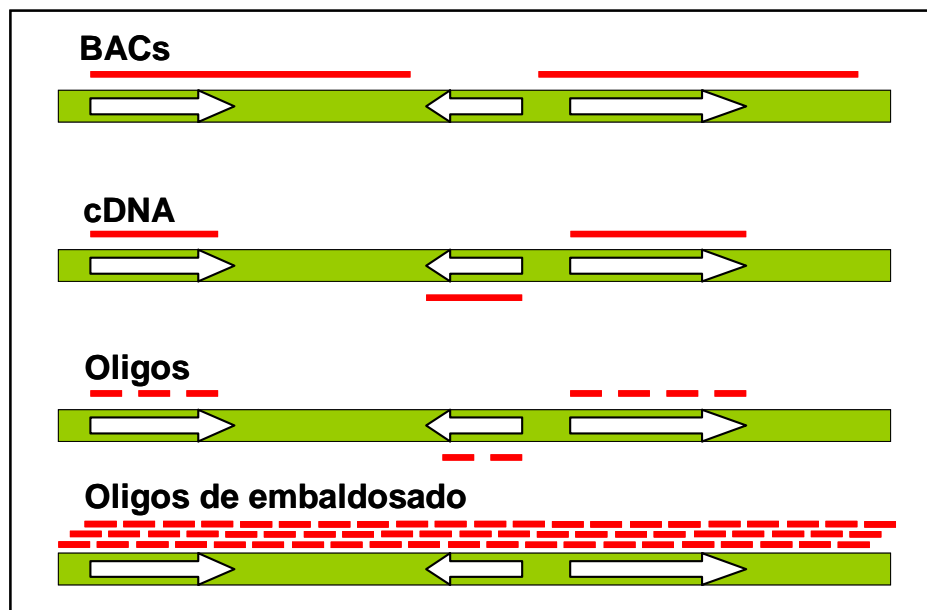


Figura 1. Clasificación de los chips de DNA según el diseño de las sondas y su disposición respecto a la anotación del genoma de estudio. En rojo se esquematizan las sondas presentes en el chip, y en blanco los genes anotados en el genoma.

Por lo que respecta a las aplicaciones de los chips de DNA éstas son múltiples y aumentan continuamente a medida que se aplican estas tecnologías en diferentes campos de investigación. Sin embargo, la aplicación más usual es la evaluación de los mRNAs presentes en una condición determinada (transcriptoma), aunque otros usos comunes van desde el estudio de anomalías cromosómicas con la hibridación



cromosómica comparada o la detección de SNPs, hasta la localización física de proteínas en regiones concretas del genoma (ChIP-on-chip) o la detección de patógenos.

## **El proceso de la transcripción**

La transcripción es el proceso por el que las células sintetizan RNA usando como molde el DNA. Este es un proceso complejo y altamente regulado y en él intervienen múltiples factores, tanto factores generales de la transcripción como factores transcripcionales específicos. En *S. cerevisiae* existen 3 polimerasas nucleares encargadas de producir los diferentes tipos de RNA. Por una parte, la RNA pol I es la encargada de la transcripción de los rRNAs y su actividad corresponde a un total del 60% de la actividad transcripcional de la célula (Warner J.R. 1999). Por otra parte la RNA pol II se encarga de transcribir los mRNAs y algunos ncRNAs, representando la actividad de esta polimerasa un 30% de la actividad total. Finalmente la RNA pol III es la que transcribe los tRNAs, el rRNA 5S y otros ncRNAs.

## **Transcripción por la RNA polimerasa II**

La RNA pol II es la encargada de producir los mRNAs que darán lugar a las proteínas. Aunque su actividad transcripcional represente sólo un pequeño porcentaje de la actividad total, la gran diversidad de moldes de DNA que debe ser capaz de reconocer (todos los genes codificantes de proteínas) y la regulación a la que se ve sujeta la convierten en la más estudiada. La RNA pol II presenta una alta conservación estructural entre todos los eucariotas, y en levadura esta formada por 12 subunidades. Rpb1 es la subunidad mayor, con unos 200 kDa. Esta subunidad presenta un dominio carboxiterminal (CTD) característico, formado por la repetición de un heptapéptido (Tyr-Ser-Pro-Thr-Ser-Pro-Ser). El número de repeticiones de este heptapéptido varía entre los diferentes eucariotas, siendo de 26 repeticiones en *S. cerevisiae*, 34 en *C. elegans*, 43 en *D. melanogaster* y 52 en humanos, lo que sugiere que el número de repeticiones puede estar relacionado con el incremento de la complejidad del genoma (Hampsey M. 1998). Además este dominio actúa como punto de anclaje para diferentes proteínas durante el proceso de la transcripción.

Durante el proceso de la transcripción por la RNA pol II podemos distinguir diferentes etapas: el reclutamiento de la polimerasa al promotor, el ensamblaje de la maquinaria general de transcripción

## Introducción

(GTFs), la iniciación de la transcripción [compuesta por el ensamblaje del complejo de preiniciación (PIC), formación del complejo abierto e iniciación], elongación transcripcional y terminación. Durante este proceso se asocia a multitud de otros complejos tanto generales como específicos de genes concretos. Además, durante las diferentes etapas el dominio CTD de la RNA pol II sufre un ciclo de fosforilaciones característico (fig 2). Estas fosforilaciones no son sólo necesarias para una correcta transcripción, sino que también son necesarias para la remodelación de la cromatina después del inicio de la transcripción y para el procesado de los RNAs nacientes. El CTD se encuentra en forma no fosforilada durante la formación del complejo de preiniciación. Sin embargo, cuando se inicia la transcripción el CTD es fosforilado en la Serina 5 por Kin28p (que forma parte de TFIIF). La fosforilación en Ser5 permite iniciar la transcripción y reclutar la maquinaria necesaria para añadir la caperuza al RNA naciente. Posteriormente es necesaria la acción de la fosfatasa Fcp1p para eliminar esta fosforilación y liberar así la maquinaria de adición de caperuza de la maquinaria transcripcional. De esta forma la fosforilación en Ser5 va disminuyendo a medida que transcurre la elongación. Posteriormente se da una fosforilación en la Ser 2 por parte de Ctk1p y Bur1p. Esta fosforilación es necesaria para producir una transición entre una etapa de transcripción temprana y la elongación. La fosforilación en la Ser2 va aumentando durante la elongación y es necesaria para una correcta terminación, permitiendo el reclutamiento de la maquinaria para el procesado 3' del mRNA naciente. Finalmente, al producirse la terminación, las fosfatasas Fcp1p y Ssu72p (entre otras) reciclan las polimerasas a su estado hipofosforilado inicial (Hirose Y. y Ohkuma, Y. 2007; Svejstrup J.Q. 2004).

Sin embargo, el punto de vista clásico de la transcripción en el que las polimerasas se encuentran exclusivamente en los promotores o dentro de los genes está cambiando en los últimos años (Koch F. *et al* 2008). Estudios de chips de DNA de embaldosado realizados en *S. cerevisiae* han permitido determinar que el 85% de su genoma se transcribe durante el crecimiento en medio rico, correspondiendo una parte importante de esas regiones a transcritos antisentido (David L. *et al* 2006; Xu Z. *et al* 2009). Además, se ha encontrado que las polimerasas no se distribuyen de forma homogénea por las regiones codificantes, sino que pueden acumularse en zonas concretas como el principio de los intrones en metazoos (Brodsky A.S. *et al* 2005). Con respecto a la localización de las polimerasas, recientemente se ha demostrado que estas pueden estar prereclutadas en multitud de genes a la espera de órdenes para entrar en

transcripción activa. Esto se ha descrito tanto en levaduras (Radonjic M. *et al* 2005) como en metazoos (Muse G.W. *et al* 2007; Radonjic M. *et al* 2005; Zeitlinger J. *et al* 2007). Sin embargo la situación no es idéntica en ambos casos. En el caso de los metazoos se ha descrito que las polimerasas pueden haber iniciado la transcripción y permanecer pausadas dentro de los 50 primeros nucleótidos en lo que se llama pausa próxima al promotor (PrPP, *promoter proximal pausing*) (Core L.J. y Lis, J.T. 2008; Margaritis T. y Holstege, F.C. 2008; Price D.H. 2008; Tamkun J.W. 2007). Siendo este mecanismo de regulación transcripcional una generalización de lo que ya se había descrito previamente para metazoos como c-Myc en mamíferos o el gen HSP70 en *Drosophila* (Bentley D.L. y Groudine, M. 1986; Rougvie A.E. y Lis, J.T. 1988). Sin embargo esta pausa próxima al promotor no se ha encontrado en levadura.

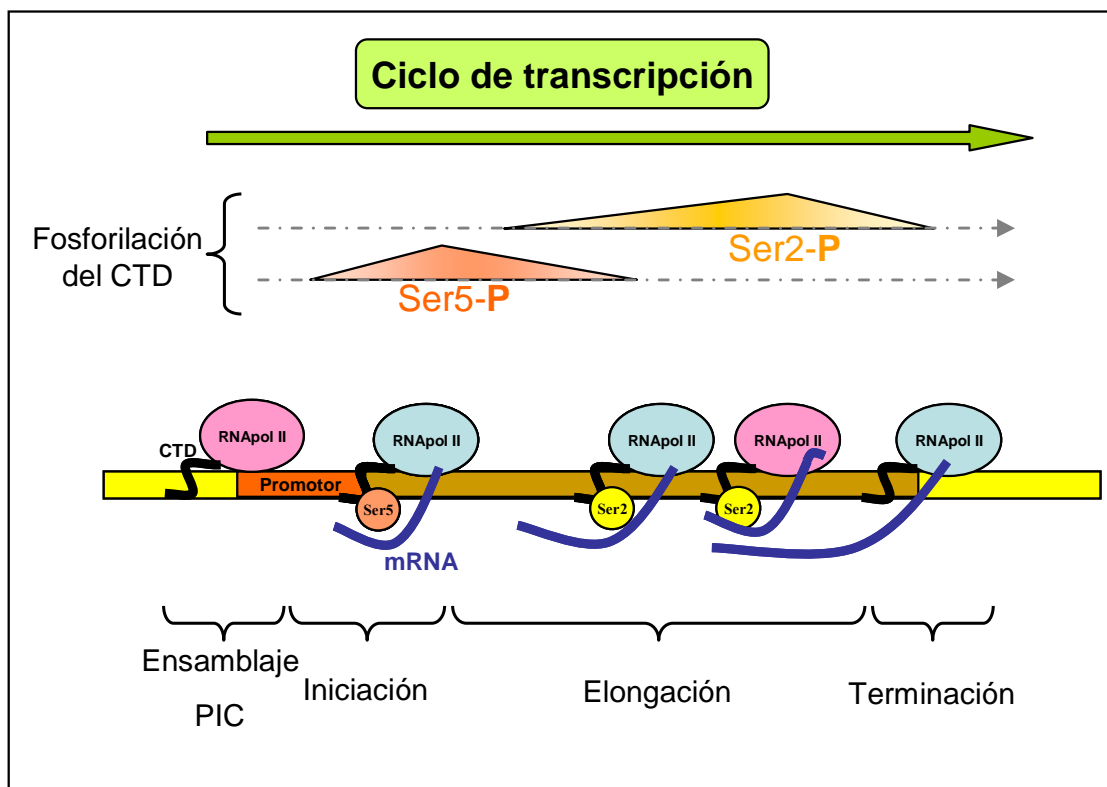


Figura 2. Ciclo de fosforilaciones del CTD a lo largo del proceso de transcripción. El dominio carboxiterminal de la RNA pol II (CTD) se fosforila en las Serinas 2 y 5 del heptapéptido (Y**S**PT**S**PS) durante el proceso de la transcripción. La polimerasa se recluta en forma hipofosforilada, durante la iniciación se fosforila en Ser5 y gradualmente va perdiendo esa fosforilación. La fosforilación en Ser2 va aumentando durante la elongación del mRNA y en la terminación el CTD vuelve al estado hipofosforilado.

## Introducción

También hay que tener en cuenta que la visión de la transcripción en la que la maquinaria transcripcional es reclutada a los genes para ejercer su función es sólo una simplificación. De hecho cada vez hay más evidencias de la importancia de la localización espacial de los diferentes genes dentro del núcleo (Dekker J. 2008; Sexton T. *et al* 2007), con respecto a las llamadas factorías transcripcionales (Marenduzzo D. *et al* 2007) (que son acumulaciones de RNA polimerasas donde se reclutan los genes para ser transcritos) o al poro nuclear (Komili S. y Silver, P.A. 2008).

## **Técnicas genómicas para el estudio de la transcripción**

### **GRO: Genomic Run-on**

La inmensa mayoría de trabajos utilizando chips de DNA estudian la cantidad de mRNA de una muestra y cómo cambia esta en diferentes condiciones o tratamientos (Sambrook J. y Russell, D.W. 2001). Sin embargo la cantidad de mRNA de la célula en una condición determinada depende de dos factores, la transcripción (producción de nuevo mRNA) y la degradación del mRNA existente (relacionado con la estabilidad del mRNA). Para comprender mejor el proceso por el que la célula responde a diferentes condiciones y poder diseccionar los dos procesos anteriores es muy útil disponer de una herramienta genómica que estudie específicamente el proceso de la transcripción.

El “*Genomic Run-on*” (*GRO*) es una técnica desarrollada en nuestro grupo de investigación (García-Martínez J. *et al* 2004) que permite el cálculo de las tasas de transcripción de los genes de levadura. Esta técnica es un escalado genómico del “*Transcription Run-on*” (*TRO*) (Hirayoshi K. y Lis, J.T. 1999). Básicamente consiste en tres pasos:

1. Permeabilización de las células mediante el uso de un detergente (Sarkosil) que desorganiza parcialmente la cromatina pero deja unida las polimerasas en transcripción activa. Sin embargo, la transcripción se detiene porque los NTPs escapan de la célula debido a la permeabilización.
2. Un pulso de elongación de los mRNAs nacientes por parte de las propias polimerasas de las células al añadir una elevada concentración de NTPs incluyendo un nucleótido radiactivo ( $^{33}\text{P}$ -UTP).

3. Purificación de los RNAs marcados e hibridación de los mismos con un *macroarray* de nylon.

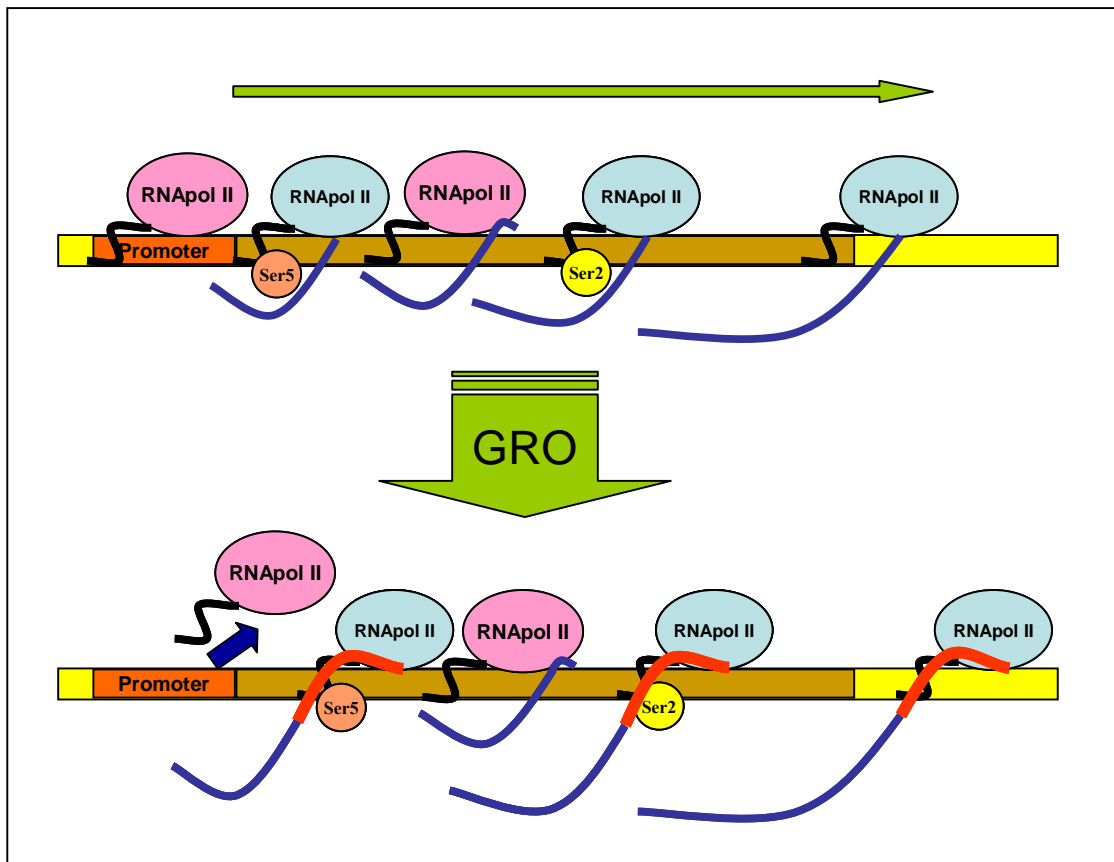


Figura 3. Modelo del marcaje de mRNAs nacientes mediante *run-on*. Aquellas polimerasas que están en transcripción activa (azules) en el momento de someter a las células a un experimento de *run-on* elongan su mRNA naciente marcándolo con el precursor radiactivo, pero aquellas que no estén en transcripción activa (aunque estén en contacto con el gen) no darán señal de GRO.

Mediante este protocolo, las RNA polimerasas que están en transcripción activa en la célula producen un fragmento de RNA naciente marcado radiactivamente (fig 3). La señal obtenida de la hibridación del RNA naciente, normalizada por una señal obtenida con la hibridación de DNA genómico, permite obtener una medida de la densidad de polimerasas transcripcionalmente activas por gen. Posteriormente, asumiendo una velocidad de transcripción constante a lo largo del gen, se puede calcular la tasa de transcripción para cada uno de los genes de levadura para un momento determinado.

Recientemente ha sido publicado un estudio en el que se aplica este mismo concepto para, mediante el uso de secuenciación masiva, realizar

## Introducción

un estudio a nivel genómico de la transcripción en humanos (Core L.J. *et al* 2008).

## **ChIP-on-chip: estudio genómico de la localización de proteínas**

Otra técnica genómica utilizada para el estudio de la transcripción es la inmunoprecipitación de cromatina. Esta técnica permite estudiar si una proteína se une específicamente a una secuencia concreta del DNA (Kuo M.H. y Allis, C.D. 1999). Su aplicación más común es la determinación de la unión de factores transcripcionales a los genes que regulan. Esta técnica consiste básicamente en entrecruzar las proteínas con el DNA usando un reactivo bifuncional como el formaldehído. Una vez unidas las proteínas entre sí y con el DNA se rompen las células y se fragmenta la cromatina mediante ultrasonidos. Posteriormente, se incuba la cromatina fragmentada con un anticuerpo específico para la proteína de interés y se obtiene una fracción enriquecida en fragmentos de DNA que estaban entrecruzados con la proteína de interés. Una vez enriquecida la muestra en cromatina se revierte el entrecruzamiento entre DNA y proteínas, y se purifican los fragmentos de DNA. En la inmunoprecipitación de cromatina “clásica”, o ChIP (*Chromatin Immuno Precipitation*), esta muestra de DNA se analiza mediante PCR para determinar si la proteína de interés se une a regiones concretas del genoma.

El ChIP-on-chip es un escalado genómico del ChIP. Pero, en vez de analizar la unión de la proteína de interés a secuencias concretas del DNA mediante PCR, sirve para estudiar simultáneamente todos los sitios de unión de una proteína sin necesidad de tener un conocimiento “*a priori*” de las zonas donde se uniría. Para ello es necesario marcar la muestra de DNA inmunoprecipitado y analizarlo usando chips de DNA (Iyer V.R. *et al* 2001; Ren B. *et al* 2000).

## **Degradación de los mRNAs en eucariotas**

La cantidad de mRNAs presentes en la célula es el producto de dos procesos opuestos. Por una parte la síntesis de nuevos RNAs (transcripción) y por otra la degradación de los mismos. Aunque tradicionalmente se ha estudiado más intensamente el proceso de la transcripción como el proceso más importante para la regulación de la cantidad de mRNAs, la degradación de los mRNAs es también un proceso importante para esta regulación en la célula. De hecho, cada vez se está dando más importancia a la etapa postranscripcional como un

factor decisivo en la regulación de la expresión génica en eucariotas (Mata J. *et al* 2005; Pérez-Ortín J.E. 2007; Pérez-Ortín J.E. *et al* 2007; Sunnerhagen P. 2007) y, por lo tanto, el estudio de la estabilidad de los mRNAs esta siendo cada vez el objeto de más esfuerzos por parte de la comunidad científica.

### **Rutas de degradación de mRNAs en eucariotas**

En eucariotas existen principalmente dos grandes rutas por las que la mayoría de los mRNAs son degradados (Meyer S. *et al* 2004; Parker R. y Song, H. 2004). Estas dos vías se diferencian por tener una degradación  $5' \rightarrow 3'$  o  $3' \rightarrow 5'$ , siendo la deadenilación o recorte de la cola poli (A) el primer paso común de ambas rutas (fig 4).

Las sucesivas etapas en la degradación de los mRNAs están altamente reguladas y las células poseen enzimas encargados de realizar las diferentes etapas. En levadura, el principal complejo enzimático con actividad deadenilasa es el complejo Ccr4-Not, aunque existen otros complejos con actividad deadenilasa como los complejos PAN y PARN. Además, cada tipo de complejo deadenilasa presenta una preferencia por un tipo de mRNPs y por lo tanto, la alteración de la función de alguno de estos complejos afecta preferentemente a una determinada población de mRNAs. El recorte de la cola poli (A) es un paso limitante en la degradación de los mRNAs, y sólo cuando ésta se ha recortado por debajo de un determinado tamaño (unos 10 nucleótidos en levadura o unos 30-60 en mamíferos) puede seguir la degradación de los mRNAs. Después del recorte de la cola poli (A), los mRNAs pueden seguir dos vías diferentes para continuar su degradación.

En la vía de degradación  $5' \rightarrow 3'$ , el primer paso consiste en la eliminación de la caperuza  $5'$  de los mRNA por parte del complejo Dcp1p-Dcp2p y posteriormente los mRNAs son degradados por la nucleasa  $5' \rightarrow 3'$  Xrn1p.

Alternativamente, los mRNAs pueden ser degradados por la vía  $3' \rightarrow 5'$ . En este caso es el exosoma citoplasmático el encargado de degradar los mRNAs. Todos estos procesos están altamente regulados y los mRNAs pueden ser degradados preferentemente por algunas de estas vías pudiendo darse una regulación diferencial de los mRNAs en función de su vía de degradación (Benard L. 2004; He F. *et al* 2003). Además también existen mecanismos específicos que pueden controlar específicamente la estabilidad de unos pocos mRNAs mediante cortes endonucleolíticos (Puig S. *et al* 2005).

## Introducción

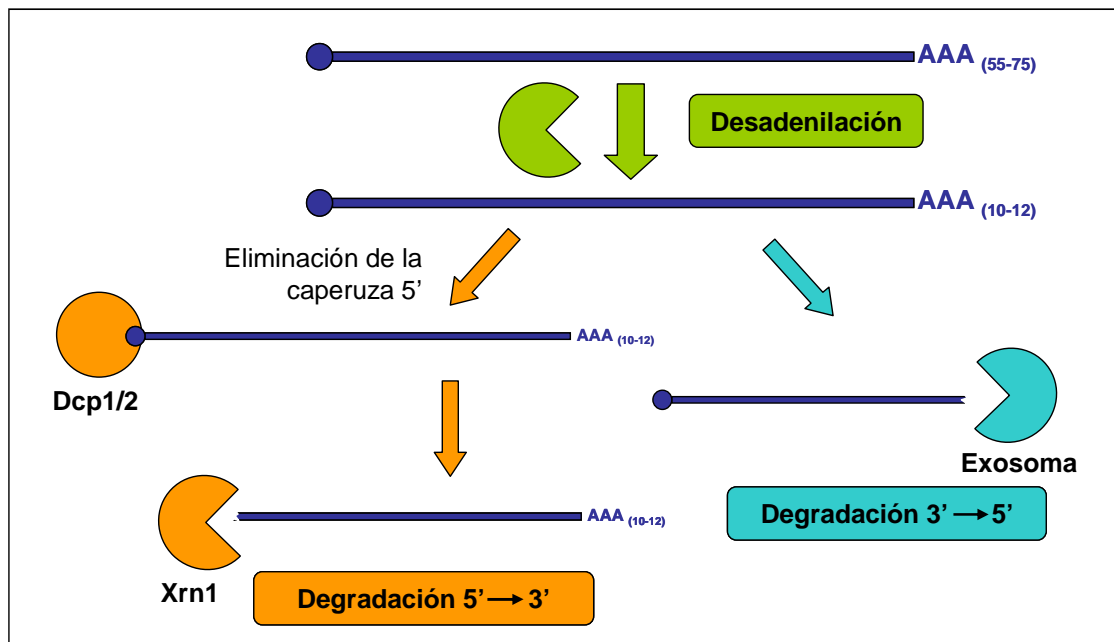


Figura 4. Principales vías de degradación de mRNAs en eucariotas. Los mRNAs después de un paso inicial de desadenilación, pueden degradarse por la vía 5'→3' después de eliminarse la caperuza 5', o por una degradación 3'→5' mediada por el exosoma citosólico. Adaptado de Parker R. y Song, H. (2004).

La degradación de los mRNAs, además de en el citoplasma, también puede darse en el núcleo. En el núcleo las vías de degradación son similares a las descritas para el citoplasma. Así, se puede encontrar una ruta de degradación 5'→3', que lleva a cabo la exonucleasa Rat1p (homóloga de la citoplasmica Xrn1p), y una ruta 3'→5', en la que el exosoma nuclear es el encargado de la degradación (Saguez C. *et al* 2005). Estas rutas eliminan principalmente mRNAs defectuosos (con intrones o truncados) o carentes del adecuado complemento proteico.

Además de estas rutas de degradación generales, las células poseen mecanismos altamente especializados para reconocer y degradar mRNAs anómalos (Parker R. y Song, H. 2004). De esta manera, existen múltiples mecanismos especializados que detectan pre-mRNAs con defectos en su procesado y los degradan. También hay mecanismos que eliminan específicamente mRNAs que contienen codones de parada prematuros (*Non-sense Mediated Decay*, NMD) o que incluso no contienen codones de parada (*Non Stop Decay*, NSD).



## **Equilibrio entre *p-bodies* y traducción.**

Un factor importante para la degradación de los mRNAs es su localización subcelular. Dependiendo de su localización los mRNAs pueden estar en transcripción activa formando parte de los polirribosomas o en lugares de almacenamiento o degradación.

Recientemente se ha descubierto que existen en eucariotas agregados discretos en el citoplasma formados por acumulaciones de mRNAs que no se encuentran en transcripción, los llamados “cuerpos de procesado” o *p-bodies* (Sheth U. y Parker, R. 2003). Los *p-bodies* están formados básicamente por proteínas implicadas en la eliminación de la caperuza 5' de los mRNAs, la degradación 5'→3' y la represión de la transcripción. Además, existe un equilibrio entre el ensamblaje de los mRNAs en la maquinaria de traducción o en *p-bodies*, de forma que según las necesidades celulares los mRNAs se pueden dirigir hacia uno u otro complejo. Además, los mRNAs que se han dirigido hacia los *p-bodies* no están exclusivamente destinados a la degradación, sino que pueden volver a salir para retornar a la maquinaria de traducción. Por lo tanto, el envío de los mRNAs hacia los *p-bodies* para su degradación no es un proceso irreversible, sino que existe un equilibrio dinámico y altamente regulado de los mRNAs entre traducción, degradación y almacenamiento (Eulalio A. *et al* 2007; Parker R. y Sheth, U. 2007).

Es importante destacar que, además de los *p-bodies*, existen otros tipos de gránulos de RNA en levadura o en otros eucariotas. Así se pueden destacar los gránulos presentes en células embrionarias que almacenan RNAs maternos, los gránulos neuronales que transportan mRNAs en las neuronas para ser traducido cerca de las sinapsis o los gránulos de estrés que almacenan mRNAs que la célula no traduce durante un estrés (Anderson P. y Kedersha, N. 2006). Además, la separación entre los diferentes tipos de gránulo no es estanca y se están descubriendo nuevas clases (Gaillard H. y Aguilera, A. 2008).

## **Técnicas clásicas para el estudio de la estabilidad de los mRNAs**

Para estudiar la estabilidad de los mRNAs en *S. cerevisiae* existen diferentes aproximaciones (Parker R. *et al* 1991). Básicamente se pueden clasificar estas técnicas en dos grandes grupos:

## Introducción

- Las basadas en técnicas de pulso y caza donde se marcan los RNAs y posteriormente se observa cómo desaparece la molécula marcada
- Las basadas en una parada de la transcripción y una cuantificación posterior de la cantidad de mRNA a diferentes tiempos para cuantificar la desaparición de las moléculas.

La aproximación más directa es realizar un pulso y caza usando precursores del RNA radiactivos. El problema de esta aproximación es que es necesario trabajar con una gran cantidad de radiactividad con la dificultad técnica que ello supone y sólo permite detectar aquellos RNAs más abundantes. Por esto, este tipo de experimentos no son muy habituales y son difícilmente utilizables a escala genómica. Consecuentemente, las técnicas más habitualmente usadas son las basadas en una parada de transcripción seguida de una determinación de la cantidad de mRNAs.

Para estudiar la estabilidad de un mRNA concreto se puede construir una cepa con el gen de interés bajo el control de un promotor regulable que pueda ser apagado, como un promotor GAL (Nam H.G. y Fried, H.M. 1986) o un promotor TetO (Belli G. *et al* 1998; Gari E. *et al* 1997). Sin embargo, esta técnica sólo permite calcular la estabilidad para el mRNAs concreto para el que se haya construido la cepa con el gen de interés bajo el control del promotor regulable. Además, en estos casos los genes se expresan a partir de un promotor diferente al suyo, y por lo tanto tienen alterada su transcripción.

Otra posibilidad es producir una parada general de la transcripción y después medir la desaparición de los mRNAs de interés. El problema general de este tipo de técnicas es que una parada general de la transcripción es un evento muy estresante para las células, y por lo tanto desencadenan una serie de respuestas celulares que puede interferir con las medidas de estabilidad de mRNAs. Por otra parte las diferentes drogas o estreses necesarios para producir una parada transcripcional pueden a su vez tener efectos secundarios sobre otros procesos metabólicos pudiendo afectar a los resultados. Además, al ser necesaria la medida de la estabilidad de los mRNAs durante un periodo prolongado tras la inhibición general de la transcripción, la propia inhibición de la transcripción puede afectar a la presencia de factores implicados en la degradación de mRNAs.

Una de las posibilidades para producir una parada general de la transcripción es usar una cepa mutante con un alelo termosensible de la RNA pol II (*rpb1-1*) para el cual la transcripción se reduce a menos del 10% a los dos minutos del cambio de 24°C a 36°C (Nonet M. *et al* 1987). Esta alternativa obliga a usar esta cepa mutante evitando así la necesidad de construir una cepa diferente para cada gen de interés. La contrapartida de este método es que, para producir la parada de la transcripción, es necesario realizar un choque térmico (con el estrés que esto produce a las células) y realizar las medidas de estabildades de los mRNAs a 37°C, que no es la temperatura habitual de crecimiento.

Otra de las posibilidades para producir una parada general de la transcripción es usar drogas inhibitoras de la transcripción como la tiolutina (Herrick D. *et al* 1990; Jiménez A. *et al* 1973) o la 1,10-fenantrolina (Santiago T.C. *et al* 1986; Scrutton M.C. *et al* 1971) que inhiben la transcripción hasta el 5% o el 10%, usando 3 µg/mL o 100 µg/mL respectivamente (Parker R. *et al* 1991). Esta aproximación tiene la ventaja de poder usarse con cualquier cepa de interés y ser técnicamente sencilla (sólo es necesario añadir la droga a la concentración adecuada). Sin embargo el uso de estas drogas también desencadenan en célula la respuesta al estrés (Adams C.C. y Gross, D.S. 1991; Grigull J. *et al* 2004).

Las técnicas basadas en una parada general de la transcripción son fácilmente adaptables a escala genómica. De hecho, el uso de un mutante termosensible de la polimerasa ha sido usado previamente con éxito para el cálculo genómico de la estabilidad de los mRNAs (Wang Y. *et al* 2002). Además, durante el desarrollo de este trabajo ha surgido un estudio genómico en el que comparaban diferentes técnicas de medida de estabilidad a escala genómica usando diferentes inhibidores transcripcionales (Grigull J. *et al* 2004). Sin embargo, ninguna de estas técnicas ha permitido determinar con éxito las estabildades para mRNAs cuya transcripción se induce como respuesta al estrés, dado que la propia técnica induce la transcripción de estos genes.

### **Cinética de la degradación de los mRNAs**

Todas las técnicas anteriormente citadas permiten medir la desaparición de los mRNAs. Sin embargo, para poder expresar sus estabildades de forma cuantitativa es necesario asumir un modelo cinético para su degradación. La degradación de los mRNAs no es un proceso sencillo, sino que consta de diferentes etapas con diferentes parámetros cinéticos (Cao D. y Parker, R. 2001). Sin embargo, se suele realizar una

## Introducción

simplificación de estos procesos asumiendo que siguen una cinética de primer orden.

De esta manera, se puede realizar un análisis simple de la variación de su concentración en función del tiempo en ausencia de síntesis de mRNA [1]. En esta ecuación  $k_d$ , que tendrá signo negativo, representa la constante de degradación para el mRNA y  $t$  el tiempo. Posteriormente, conociendo la cantidad de mRNAs a diferentes tiempos tras la inhibición se puede calcular el valor de la constante de degradación. Para ello sólo es necesario representar la relación en escala logarítmica entre la concentración de mRNA a un tiempo determinado respecto a la concentración inicial frente al tiempo. Si se ajustan esos puntos a una línea recta, la pendiente de esta será la constante de degradación ( $k_d$ ) [2]. Una vez obtenida la constante de degradación se puede calcular directamente la vida media ( $t_{1/2}$ ) de los mRNAs [3].

$$[1] \quad d[\text{mRNA}]/dt = k_d[\text{mRNA}]$$

$$[2] \quad \ln([\text{mRNA}]_t/[\text{mRNA}]_0) = k_d t$$

$$[3] \quad t_{1/2} = (\ln 2)/k_d$$

Sin embargo, hay que recordar que estos cálculos son sólo una simplificación útil de la cinética de la degradación de los mRNAs, y en el cálculo de vidas medias a partir de la disminución de la cantidad total de los mRNAs se puede infraestimar la vida media real de los mRNAs (Cao D. y Parker, R. 2001). Además, existen descripciones de curvas de degradaciones bifásicas (Brendolise C. *et al* 2002) que ponen de manifiesto que en algunos casos puede no ser correcto simplificar la degradación de los mRNAs a un proceso estocástico de primer orden.

## **Estrategias de expresión génica**

La expresión génica es el proceso mediante el cual la información contenida en un gen da lugar a un producto génico con actividad biológica (proteínas o RNAs). La expresión génica es un proceso altamente regulado, y esta regulación se puede dar a lo largo de todas las etapas necesarias para producir una actividad biológica determinada de una proteína. Así por ejemplo, la regulación puede darse tanto al nivel de síntesis o degradación de mRNAs y proteínas, como modificando su actividad o localización subcelular.

Las células pueden conseguir una determinada concentración final de proteína usando diferentes estrategias de regulación génica, entendiendo por estrategias tanto la relación entre los diferentes parámetros cinéticos

(velocidades de síntesis de mRNAs y proteínas, velocidades de degradación,...) como la regulación de cada una de estas etapas. Dependiendo de la estrategia utilizada, la célula puede elegir entre diferentes alternativas para optimizar diferentes parámetros de interés como el ahorro energético, la velocidad de respuesta o el ruido biológico. Además, en general, las células usan estrategias de expresión génica comunes para los genes involucrados en las mismas rutas fisiológicas (García-Martínez J. *et al* 2007).

### **Diferentes estrategias respecto a las tasas de síntesis y degradación.**

Por lo que respecta a las tasas de síntesis y degradación de las moléculas, si la célula usa estrategias con elevadas tasas de síntesis y degradación, puede adaptar la concentración de la molécula muy rápidamente a nuevas condiciones fisiológicas. Sin embargo, este alto recambio de moléculas lleva asociado a su vez un gran consumo energético. Por el contrario, se puede conseguir la misma concentración final de moléculas usando tasas de síntesis más bajas y disminuyendo también su velocidad de degradación (aumentando la estabilidad de las moléculas). Sin embargo, de esta manera se sacrifica la capacidad de producir cambios rápidos para adaptarse a nuevas situaciones a cambio de una reducción en el consumo energético (García-Martínez J. *et al* 2007; Pérez-Ortín J.E. *et al* 2007).

Además, aunque en general la concentración de un mRNA es directamente proporcional a la concentración de la proteína que codifica (Greenbaum D. *et al* 2003; Lu P. *et al* 2007), en algunos casos esto no es cierto lo que confirma la importancia de los mecanismos de control postranscripcionales (Pradet-Balade B. *et al* 2001).

El hecho de que las células puedan usar diferentes estrategias de expresión génica para obtener las mismas concentraciones finales de proteína y mRNAs, y la importancia que éstas tienen en cuanto a la regulación de los diferentes procesos biológicos hace que cada vez se estén realizando más estudios encaminados a estudiar conjuntamente los diferentes parámetros cinéticos que contribuyen a producir una determinada concentración de proteína (Pérez-Ortín J.E. 2007; Sunnerhagen P. 2007).

### **Estado estacionario para la expresión génica**

Una de las asunciones más habitualmente utilizadas para estudiar la cantidad de mRNAs y proteínas es la del estado estacionario para su

## Introducción

síntesis y degradación. En este caso, se asume que la concentración de un determinado producto es constante a lo largo del tiempo, de forma que sus tasas de síntesis y degradación están en equilibrio. Esto permite relacionar directamente las tasas de síntesis y degradación de una molécula para una determinada concentración. Esta asunción se puede aplicar a una población de células que se encuentren en una situación fisiológica estable (García-Martínez J. *et al* 2004). Sin embargo, no puede aplicarse, al menos para todos los genes, en aquellos casos en los que las células cambien de situación fisiológica (Pérez-Ortín J.E. *et al* 2007).

## **Ruido biológico para la expresión génica**

Otro factor importante para la elección de una determinada estrategia de expresión génica es el ruido biológico. El ruido biológico para la expresión génica hace referencia a las fluctuaciones aleatorias que producen variaciones en la cantidad final de proteína (o mRNA) para un gen determinado. Matemáticamente se define como la desviación estándar del número de moléculas (Maheshri N. y O'Shea, E.K. 2007). Esas variaciones se deben a los procesos estocásticos que ocurren a lo largo de todo el proceso de expresión génica (transcripción, traducción,...) y su frecuencia y magnitud pueden afectar a la eficacia biológica. De esta manera se puede diferenciar entre ruido *transcripcional*, que mide la variabilidad de mRNAs y el ruido *traduccional*, que parte del ruido transcripcional y lo amplifica debido al ruido que se produce durante la traducción. Por todo esto, el ruido biológico es también un parámetro sometido a presiones evolutivas para el que los seres vivos han adoptado diferentes soluciones. Así, por ejemplo, los genes cuyas proteínas forman parte de complejos macromoleculares deben de ser poco ruidosos, de manera que las concentraciones finales de las diferentes proteínas se encuentren en la proporción adecuada. Por otra parte, el ruido también puede ser beneficioso dado que la existencia de variaciones en la cantidad de productos génicos puede ser importante en la diferenciación de los seres pluricelulares y en facilitar la supervivencia en ambientes con fluctuaciones ambientales (Raser J.M. y O'Shea, E.K. 2005).

Una de las causas del ruido biológico es que el proceso de activación de los promotores (apertura de la cromatina, reclutamiento de factores transcripcionales y maquinaria de transcripción,...) es un paso lento en la producción de mRNA, de forma que la transcripción ocurre mediante pulsos o “fogonazos” de transcripción (Chubb J.R. *et al* 2006). Además,

se considera que hasta un 90% de los eventos de iniciación transcripcional son abortivos y no producen, por lo tanto, un mRNA completo (Struhl K. 2007). Por lo tanto, la propia estructura de los promotores y la forma en que se produce la transcripción de los diferentes genes puede influir en su ruido. Por ejemplo, los genes cuyos promotores sufren procesos de activación frecuentes seguidos de una transcripción menos eficiente son menos sensibles a los pulsos de transcripción y por tanto son genes menos ruidosos. En el otro extremo estarían genes muy ruidosos cuyo promotor se activa menos frecuentemente, pero una vez activos producen eficientemente mRNAs mediante pulsos de transcripción (Hofer T. y Rasch, M.J. 2005; Raser J.M. y O'Shea, E.K. 2004). Dentro de este tipo de genes, los más característicos suelen ser aquellos que poseen caja TATA en su promotor.

## **Genes RP y regulón RiBi como genes altamente regulados**

La síntesis de los ribosomas es uno de los procesos que requieren un mayor consumo energético por parte de la célula (Warner J.R. 1999) y, en consecuencia, es también un proceso altamente regulado. Es un proceso complejo que requiere la participación tanto de los rRNAs como proteínas que deben estar presentes en unas relaciones estequiométricas determinadas. Los ribosomas, además, sufren un complejo proceso de maduración. Por lo que hace referencia a la transcripción, los ribosomas necesitan productos de las tres RNA pol nucleares, la RNA pol I (transcribe el rRNA precursor 35S), la RNA pol II (transcribe los mRNAs de los genes de las proteínas ribosómicas (RP) y de las proteínas del regulón RiBi, que son necesarias para la síntesis maduración y exporte de los ribosomas, así como factores de traducción y proteínas relacionadas con las polimerasas I y III) y la RNA pol III (que transcribe el rRNA precursor 5S).

La síntesis de los ribosomas está regulada principalmente por la disponibilidad de nutrientes en el medio y la velocidad de crecimiento de las células, dado que la capacidad de crecimiento celular está directamente relacionada con la producción de nuevos ribosomas (Jorgensen P. *et al* 2004; Regenbreg B. *et al* 2006). Hay dos grandes rutas de señalización, la ruta TOR y la ruta Ras/PKA, que están implicadas en la detección de nutrientes en el medio y que regulan la síntesis de ribosomas afectando de forma coordinada a la transcripción de las tres polimerasas. Además, los genes RP y RiBi suelen presentar

## Introducción

respuestas transcripcionales a las variaciones ambientales muy similares (Gasch A.P. *et al* 2000; Miyoshi K. *et al* 2003).

La transcripción de los genes RP está altamente regulada y se conocen multitud de factores implicados en esta regulación (fig 5). Estos genes suelen tener promotores con uno o dos sitios de unión para las proteínas bifuncionales Rap1 y/o Abf1 (que también tienen funciones en el silenciamiento transcripcional). Cuando hay transcripción de los genes RP, la proteína Sfp1 se localiza en el interior del núcleo. En esa situación, interacciona con la proteína Fhl1 y su coactivador Ifh1 para activar la transcripción de los genes RP. Esta activación se da probablemente cambiando la función de Rap1p (o de Abf1p) de represor a activador. Además de estas vías de regulación deben existir otras, como por ejemplo un efecto de Sch9 sobre transcripción, probablemente fosforilando a Rap1p (Jorgensen P. *et al* 2004).

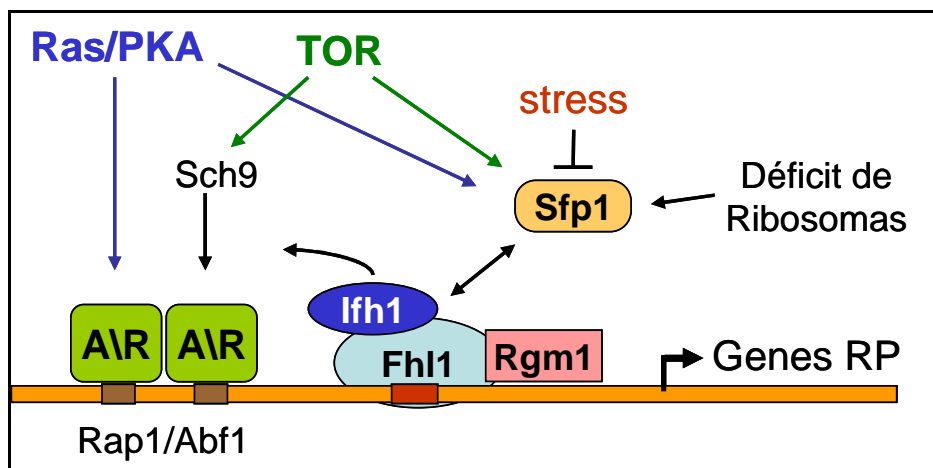


Figura 5. Modelo de regulación de la transcripción para los genes de las proteínas ribosomales. Los genes RP suelen tener uno o dos sitios de unión para Rap1 (o Abf1). Por otra parte, la localización subcelular de Sfp1 está regulada por diferentes señales, y a su vez regula la localización de Ifh1/Fhl1. Las proteínas Fhl1/Ifh1 pueden activar la transcripción cambiando el estado de las proteínas Rap1 (o Abf1) de un estado represor a uno activador. Adaptado de Jorgensen P. *et al.* (2004).

El alto grado de regulación que poseen estos genes les aporta otras características. De hecho, en un estudio genómico para determinar el ruido biológico asociado a cada gen, realizado midiendo la cantidad de miles de proteínas en células individuales, se determinó que estos genes poseen un bajo nivel de ruido biológico (Newman J.R. *et al* 2006). Este bajo nivel de ruido biológico se debe en parte a la capacidad de Rap1p y Abf1p de desplazar a los nucleosomas y dejar los promotores más



accesibles de forma que se minimizan los pulsos de transcripción así como la rápida degradación de proteínas ribosómicas libres, dado que pueden ser tóxicas para las células (Sun X.X. *et al* 2008). Por otra parte, este menor ruido es necesario para la perfecta coordinación entre la cantidad de ribosomas y el crecimiento celular.

## **Silenciamiento transcripcional**

El silenciamiento génico transcripcional es el proceso epigenético por el que se apaga la expresión de los genes de manera prolongada mediante cambios en la organización de la cromatina, como la modificación de histonas. Es importante diferenciar el silenciamiento producido por la modificación de proteína que impide la transcripción (transcripcional) del llamado silenciamiento postranscripcional producido por RNA de interferencia y otros RNAs pequeños que destruyen los mRNAs ya transcritos. Este último tipo no existe en *S. cerevisiae*.

## **Maquinaria de silenciamiento transcripcional en levadura**

En levadura, las proteínas de la familia Sir son las responsables de silenciar determinadas zonas del genoma formando una estructura cromatínica especial análoga a la heterocromatina que se forma en organismos superiores. Estas proteínas están implicadas en regular la estabilidad de la cromatina reprimiendo tanto la transcripción como la recombinación, y también han sido relacionadas con la longevidad (Gross D.S. 2001). Las proteínas Sir se asocian básicamente a tres tipos de regiones genómicas en levadura: a los telómeros, a los loci de apareamiento inactivos (HMR y HML) y a copias silenciadas de rDNA.

Las proteínas Sir no se unen directamente al DNA, sino que son reclutadas por otras proteínas. En los telómeros, es la proteína Rap1p (que también es un factor transcripcional asociado a los genes de las proteínas ribosómicas) la que recluta a Sir3p y Sir4p, que posteriormente reclutan a Sir2p. En el caso de los loci de apareamiento se reclutan a unas regiones llamadas silenciadores situadas en los flancos del locus a silenciar las proteínas Rap1p, Abf1p (que también se asocia con los promotores de los genes de las proteínas ribosómicas) y ORC (el complejo de reconocimiento de inicio de replicación). Estas proteínas junto con Sir1p reclutan al complejo Sir2/3/4 que se extiende por la cromatina entre los dos silenciadores. Por otra parte, en el caso del rDNA

## Introducción

Sir2p se recluta vía Net1p para suprimir la recombinación en esas regiones.

Sir3p y Sir4p tienen funciones estructurales que permiten que se una Sir2p, que es una histona desacetilasa encargada principalmente de desacetilar las lisinas 9 y 14 de la histona H3, y la lisina 12 de la H4. Además, la desacetilación de estos residuos se asocia con la ausencia de transcripción. Por otra parte, Sir3p y Sir4p son también necesarias para formar estructuras cromatínicas de orden superior y para anclar las zonas silenciadas a la periferia nuclear.

Clásicamente se creía que la inhibición de la transcripción durante el silenciamiento se debía básicamente a un impedimento estérico que impedía que la maquinaria de transcripción se asociara a los promotores de los genes silenciados. Sin embargo, recientemente se ha demostrado que la cromatina silenciada no está por ello libre de maquinaria transcripcional (Gao L. y Gross, D.S. 2008; Sekinger E.A. y Gross, D.S. 2001). Por lo tanto, la represión transcripcional en la cromatina silenciada se debe a una inhibición de la transcripción tras el reclutamiento de la maquinaria (Gao L. y Gross, D.S. 2006). Es decir, que la cromatina silenciada mediante Sir permite la unión de factores transcripcionales específicos y de componentes del PIC (como la TBP y la RNA pol II), pero impide su actividad transcripcional. El modo de acción de la maquinaria de silenciamiento que existe en levaduras es muy similar al que ocurre en organismos superiores, aunque las proteínas implicadas no estén relacionadas. Por ejemplo, en *Drosophila* las proteínas implicadas en el silenciamiento transcripcional son las proteínas del grupo Polycomb (PcG), y a pesar de ser una forma evolutivamente diferente de silenciamiento, parece actuar de manera similar a como lo hacen las proteínas Sir en levaduras, permitiendo también el reclutamiento de la maquinaria transcripcional pero impidiendo su actividad (Pirrotta V. y Gross, D.S. 2005).

## **El silenciamiento transcripcional como mecanismo de regulación génica.**

Recientemente, estudios genómicos han permitido descubrir la importancia de la regulación del silenciamiento de genes implicados en el desarrollo celular tanto en *Drosophila* como en células madre humanas (Bernstein B.E. *et al* 2006; Lee T.I. *et al* 2006; Tolhuis B. *et al* 2006). Además, otros estudios demuestran que puede existir un reclutamiento de PcG a ciertos genes aunque estos estén transcripcionalmente activos

(Bracken A.P. *et al* 2006). Todo esto refuerza la idea de que el silenciamiento transcripcional es un proceso mucho más dinámico de lo que originalmente se pensaba, lo que está haciendo que se empiecen a estudiar sus posibles implicaciones con la regulación transcripcional (Ringrose L. 2007).

## **Objetivos**

# Objetivos

El objetivo general de este trabajo consiste en desarrollar y aplicar nuevas técnicas genómicas para el estudio del recambio de los mRNAs en *S. cerevisiae*. Concretamente, se ha estudiado a escala genómica tanto la degradación como la síntesis de los mRNAs. Todo ello con el fin de mejorar la visión conjunta de estos dos fenómenos que determinan la cantidad del mRNA y la velocidad a la que ésta cambia en respuesta a los cambios ambientales. Para ello se proponen los siguientes objetivos específicos:

- Estudiar la validez de la asunción de un estado estacionario para la síntesis y degradación de los mRNAs durante el crecimiento exponencial de *S. cerevisiae*.
- Desarrollar un método para el cálculo de vidas medias de mRNAs a escala genómica basado en la hipótesis de estado estacionario y en la técnica del *Genomic Run-on* (GRO) y contrastarlo con otros disponibles en la bibliografía.
- Desarrollar un nuevo método para calcular las tasas de transcripción a escala genómica basado en la detección de polimerasas unidas a los genes que permita contrastar los datos obtenidos mediante GRO.
- Estudiar las posibles diferencias biológicas entre la cantidad de RNA polimerasas presentes en los genes y su actividad, para extraer conclusiones sobre los mecanismos de transcripción.
- Adaptar la técnica del *Genomic Run-on* al marcaje fluorescente para poder ser usada en chips de DNA de alta resolución que permitan un estudio más detallado de la transcripción.

## ***Materiales y Métodos***

# Materiales y métodos

## Técnicas microbiológicas

### Cepas de levadura

En este trabajo se han utilizado las siguientes cepas de *S. cerevisiae*:

Tabla 1. Cepas de *S. cerevisiae* usadas en este trabajo.

Cepa	Genotipo	Procedencia
BQS252	Mat a, <i>ura3-52</i> , procede de FY1679 (fondo genético S288c)	Susana Rodríguez Navarro, tesis doctoral.
Z676	Mat a, <i>his3, leu2-3, ura3-52, rpb1::HIS3</i> con plásmido RY2522 ( <i>rpb1-1<sup>ts</sup>, URA3, CEN, AMP</i> ) (fondo genético S288c)	(Thompson C.M. y Young, R.A. 1995)
ARG3	BY4741, Mat a; <i>RPB1:Myc<sub>18</sub>-TRP1<sub>K2</sub></i> (fondo genético S288c)	(Vanti M. <i>et al</i> 2009)
SCR101 pRAP1	Mat a, <i>ade2, his2, leu2, ura3, trp1, rap1:URA3:UAS<sub>GAL</sub>-rpl25-RAP1</i> con plásmido pRS415-RAP1 (fondo genético SCR101)	(Graham I.R. <i>et al</i> 1999)
SCR101 pRAP1ΔSil	Mat a, <i>ade2, his2, leu2, ura3, trp1, rap1:URA3:UAS<sub>GAL</sub>-rpl25-RAP1</i> con plásmido pRS415-RAP1ΔSil (fondo genético SCR101)	(Graham I.R. <i>et al</i> 1999)
Y01089	BY4741, Mat a, <i>his3, leu2, met15, ura3, tpk2::kanMX4</i> (fondo genético S288c)	EUROSCARF

### Medios y condiciones de cultivo

Para crecer todas las levaduras en general se usó el medio rico YPD. Este medio contiene extracto de levadura 1%, peptona bacteriológica 2% y glucosa 2%. Para los experimentos de cambio de fuente de carbono se usó YPGal (extracto de levadura 1%, peptona bacteriológica 2% y galactosa 2%). Todos los medios se esterilizaron autoclavándolos durante 20 minutos a 121°C.

Los cultivos se crecieron en matraces con una agitación de 190 rpm y a 28°C, excepto el caso de la cepa termosensible Z676 que se creció a 25°C.

## **Técnicas de Biología Molecular**

### **Extracción de ácidos nucleicos**

Para la extracción de ácidos nucleicos se rompieron las células en presencia de fenol y se recuperó el DNA o RNA de la fase acuosa. Para optimizar la recuperación de ácidos nucleicos se usó un Fenol a pH 4.5 en las extracciones de RNA y otro tamponado a pH 7.9 para las de DNA.

### **Extracción de RNA**

Para extraer RNA se partió de 50 mL de cultivo crecido a  $DO_{600}$  0.5. Las células se resuspendieron en 500  $\mu$ L de LETS (0.1M LiCl, 10 mM EDTA, 0.2% SDS y 10 mM Tris-HCl pH 7.4) y se añadieron a un tubo con 500  $\mu$ L de perlas de vidrio y 500  $\mu$ L de Fenol:Cloroformo:Alcohol Isoamílico (125:24:1). Las células se rompieron usando un agitador FastPrep (MP) con 4 pulsos de 30 s a 5.5 m/s. Se separaron las fases por centrifugación y se recuperó la fase acuosa. Posteriormente se añadió a esta fase 500  $\mu$ L más de Fenol:Cloroformo:Alcohol Isoamílico (125:24:1) y se volvió a recuperar la fase acuosa. A continuación, para eliminar los restos de fenol, se mezcló la fase acuosa con 500  $\mu$ L de Cloroformo:Alcohol Isoamílico (24:1) y se volvieron a separar las fases por centrifugación. Finalmente se precipitó la fase acuosa con 1 volumen de LiCl 5 M durante al menos 3 h a  $-20^{\circ}\text{C}$ . Para eliminar los restos de litio se realizó una segunda precipitación con 1/10 de acetato sódico 3 M y 2.5 volúmenes de etanol 96°. La cantidad de RNA se cuantificó espectrofotométricamente y se comprobó su integridad realizando una electroforesis en gel de TAE (1 mM EDTA, 40 mM Tris-Acético pH 8) agarosa al 1% o usando el sistema de electroforesis capilar Experion (Bio-Rad). En los casos en que resultó necesario eliminar los restos de DNA se realizó una incubación adicional del RNA con DNAsa I y posteriormente se eliminó la DNAsa mediante fenolización (proceso descrito en detalle en el apartado de marcaje de cDNA usando cebadores aleatorios).

### **Extracción de DNA.**

Para extraer DNA se partió también de 50 mL de cultivo crecido a  $DO_{600}$  0.5. Las células se resuspendieron en 500  $\mu$ L de 10'prep (2% TritonX-100, 1% SDS, 100 mM NaCl, 1 mM EDTA y 10 mM Tris-HCl pH 8) y se añadieron a un tubo con 500  $\mu$ L de perlas de vidrio y 500  $\mu$ L de Fenol:Cloroformo:Alcohol Isoamílico (25:24:1). Las células se



rompieron con un agitador FastPrep (MP) con 4 pulsos de 30 s a 5.5 m/s. Se separaron las fases por centrifugación y se recuperó la fase acuosa. Posteriormente se añadió a esta fase 500  $\mu\text{L}$  más de Fenol:Cloroformo:Alcohol Isoamílico (25:24:1) y se volvieron a separar las fases. A continuación, para eliminar los restos de fenol, se mezcló la fase acuosa con 500  $\mu\text{L}$  de Cloroformo:Alcohol Isoamílico (24:1) y se volvieron a separar las fases por centrifugación. Seguidamente se precipitó la fase acuosa con 1/10 de acetato sódico 3 M y 2 volúmenes de EtOH durante 3 h a  $-20^{\circ}\text{C}$ . El DNA se resuspendió en 400  $\mu\text{L}$  TE (1 mM EDTA, 10 mM Tris HCl pH 8) y se incubó durante 30 minutos a  $37^{\circ}\text{C}$  con 3  $\mu\text{L}$  RNAsa A libre de DNAsas (10 mg/mL, Roche). Posteriormente se añadieron 10  $\mu\text{L}$  Proteinasa K (20 mg/mL, Novagen) y se incubó durante 1 h a  $65^{\circ}\text{C}$ . Finalmente, se volvió a fenolizar el DNA y se precipitó con acetato sódico y EtOH para, a continuación, resuspenderlo otra vez en TE.

### **PCR**

Las condiciones habituales de amplificación de DNA mediante PCR consistieron en una desnaturalización inicial de 3 minutos a  $94^{\circ}\text{C}$  seguidas de 30 ciclos de amplificación (30 segundos a  $94^{\circ}\text{C}$ , 30 segundos a  $52^{\circ}\text{C}$  y 1 minuto a  $72^{\circ}\text{C}$ ) y una extensión final de 5 minutos a  $72^{\circ}\text{C}$ . La temperatura de apareamiento se modificó en función de la  $T_m$  de los cebadores empleados, y el tiempo de extensión se aumentó dependiendo del tamaño final del fragmento amplificado. Para las amplificaciones se usaron, indistintamente, los termocicladores *Mastercycler Personal* (Eppendorf) y *GeneAmp PCR System 9700* (Applied Biosystems).

Para cada reacción de 25  $\mu\text{L}$  totales se usó: 1  $\mu\text{L}$   $\text{MgCl}_2$  (50 mM), 2.5  $\mu\text{L}$  Tampón 10x, 0.5  $\mu\text{L}$  dNTPs (10 mM), 5  $\mu\text{L}$  del cebador directo y 5  $\mu\text{L}$  del reverso (ambos a 2.5  $\mu\text{M}$ ), 9.75  $\mu\text{L}$   $\text{H}_2\text{O}$ , 0.5  $\mu\text{L}$  molde (habitualmente DNA genómico a 50 ng/ $\mu\text{L}$ ) y 0.25  $\mu\text{L}$  Taq DNA pol (5 u/ $\mu\text{L}$ ) de Biotools.

### **PCR cuantitativa**

Todos los experimentos de qPCR se realizaron con el termociclador DNAEngine acoplado al detector Chromo4 (Bio-Rad) y se analizaron usando el programa Opticon Monitor (Bio-Rad). Para las amplificaciones se usó la mezcla de qPCR *Quantimix Easy SYG* con Rox (Biotools) basada en la detección del producto amplificado por *SYBR Green*.

El programa consistió en una desnaturalización inicial de 1 minuto a  $95^{\circ}\text{C}$  seguidos por un total de 45 ciclos consistentes en: 30 s

## Materiales y métodos

desnaturalizando a 95°C, 30 s de apareamiento de oligos a 50°C y 30 s de elongación a 72°C. En cada ciclo se midió la fluorescencia después de la elongación a 75°C, 80°C y 82°C, dependiendo de la temperatura de fusión del producto amplificado. Finalmente se realizó una curva de desnaturalización de los productos de PCR y se descartaron aquellos amplificados que produjeron más de un producto. Para cada reacción de PCR se utilizaron los datos de fluorescencia obtenidos a la mayor temperatura que no desnaturalizara el producto de PCR.

Para el análisis de los resultados se utilizó el método  $\Delta\Delta\text{CT}$  (Livak K.J. y Schmittgen, T.D. 2001), comparando los resultados del amplificado de la muestra inmunoprecipitada con el del extracto celular completo. Como normalizador se usó una región intergénica (Intg-V) (Komarnitsky P. *et al* 2000).

### **Northern blot**

Los RNAs se separaron por tamaños usando geles desnaturalizantes con 1% de agarosa, MOPS 1x (20 mM MOPS, 4 mM acetato sódico, 1 mM EDTA, pH 7) y 0.064% de formaldehído. En cada pocillo se cargaron entre 15 y 20  $\mu\text{g}$  de RNA resuspendidos en un volumen de 2  $\mu\text{L}$  totales. A esa cantidad de RNA se le añadió 9.75  $\mu\text{L}$  de una mezcla consistente en: 14  $\mu\text{L}$  MOPS 1x, 24.5  $\mu\text{L}$  formaldehído 36%, 70  $\mu\text{L}$  formamida, 24  $\mu\text{L}$  tampón de carga de electroforesis (30% glicerol y 0,25% azul de bromofenol) y 1  $\mu\text{L}$  BrEt 500 ng/ $\mu\text{L}$ . Posteriormente, se desnaturalizaron las muestras calentándolas durante 10 minutos a 65°C y enfriándolas en hielo. Una vez cargadas las muestras, se dejaron correr durante 2 a 3 horas a 50 V. Para cuantificar el rRNA y comprobar la carga de cada carrera se realizaron lecturas del gel con un lector *FLA3000 Phosphorimager* (FujiFilm) usando el láser de 473 nm y el filtro O580. Los geles se transfirieron a membranas de nylon *Hybond N<sup>+</sup>* (GE-healthcare) por capilaridad usando SSC 6x (0.9 M NaCl y 90 mM citrato sódico) durante toda la noche. Al día siguiente se fijaron las muestras a la membrana seca entrecruzándolas con 50 mJ de luz ultravioleta (GS Gene Linker de Bio-Rad).

Para detectar mRNAs concretos se usaron como sondas fragmentos de DNA obtenidos por PCR que cubrían toda la ORF. Estas sondas se marcaron con  $\alpha^{33}\text{P}$ -dCTP (3000 Ci/mmol y 10 mCi/mL) y el kit *Ready-to-Go* (-dCTP) (GE-healthcare).

Las membranas se prehibridaron durante 1 hora a 42°C y después se hibridaron con la sonda marcada durante 16 horas más. En ambos casos

se usó una solución de hibridación con 50% formamida desionizada, 5x SSPE (0.9 M NaCl, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 5 mM EDTA, pH 7.7), 10% sulfato de dextrano (500000 MW), 1% SDS, 5x *Denhardt's* (0.5% BSA, 0.5% Ficoll, 0.5% PVP) y 0.2 mg/mL DNA de esperma de salmón). Para eliminar las uniones no específicas de la sonda se hicieron 2 lavados de 10 minutos a 42°C con 2x SSPE y 0.1% SDS, y un lavado de 15 minutos a 65°C con 1x SSPE y 0.1% SDS. Finalmente se sellaron las membranas con film plástico alimentario (cloruro de polivinilideno), se expusieron en una pantalla IP (BAS-MP de FujiFilm) para obtener las imágenes con un lector *FLA3000 Phosphorimager* (FujiFilm).

### **Cálculo de vidas medias de mRNAs**

Los experimentos para calcular vidas medias individuales se realizaron parando la transcripción de las células y extrayendo RNA a diferentes tiempos tras la inhibición de la transcripción (Parker R. *et al* 1991). Posteriormente se determinó la variación en la cantidad de los mensajeros individuales por *northern blot* y se calculó su vida media.

Para inhibir la transcripción se usó tanto el antibiótico tiolutina (Tipper D.J. 1973) como la cepa Z676 (cuya RNA Pol II es termosensible (Thompson C.M. y Young, R.A. 1995)). En el caso de la tiolutina se añadió la droga a una concentración final de 3 µg/mL a un cultivo en crecimiento exponencial. En el caso de la cepa con la polimerasa termosensible se añadió al cultivo crecido a 25°C, YPD caliente a 65°C. De esta manera la temperatura del cultivo se incrementó instantáneamente a 37°C y se mantuvo a esa temperatura durante el resto del experimento. Una vez parada la transcripción se recogieron muestras a diferentes tiempos y se congelaron las células en N<sub>2</sub> líquido. Posteriormente, se extrajo en RNA y se analizó el resultado mediante *northern blot* usando sondas específicas para los mRNAs de interés.

Para normalizar los diferentes tiempos entre sí, se usó como control de carga el rRNA. Para realizar los cálculos se asumió que el rRNA es mucho más estable que los mRNAs y que su degradación durante el experimento es por tanto despreciable. De esta manera, y asumiendo una cinética de degradación de primer orden, se calculó la vida media de los mensajeros de interés. Concretamente, para realizar este cálculo se representó el log<sub>2</sub> de la concentración de mRNA en función del tiempo y se ajustaron los puntos a una recta. Finalmente la vida media se calculó directamente de esta recta, dado que corresponde al inverso de la pendiente cambiada de signo ( $t_{1/2}=1/-pte$ ).

## **Cuantificación de RNA poliadenilado**

Para cuantificar la cantidad de RNA poliadenilado en diferentes muestras de RNA total se depositaron sobre una membrana de nylon *Hybond N<sup>+</sup>* (GE-healthcare) diferentes concentraciones conocidas de las muestras usando el robot *BioGRID* (BioRobotics).

Para la detección de las colas poli (A) se marcó terminalmente una sonda poli (dT) con <sup>32</sup>P. Para ello, añadieron a 5 μL de poli (dT) (Sigma P6905), 5 μL tampón quinasa directo 10x (Roche), 2 μL polinucleotido quinasa (10 u/μL Roche), 33 μL de H<sub>2</sub>O y 5 μL γ<sup>32</sup>P-ATP (3000 Ci/mmol, 10 mCi/mL). Esta mezcla se incubó a 37°C durante 30 minutos y posteriormente se inactivó la quinasa con una incubación de 10 minutos a 70°C. Finalmente se purificó la sonda usando columnas de exclusión molecular (*ProbeQuant G-50 Micro Columns* de GE-healthcare) para eliminar la radiactividad no incorporada.

Para realizar la hibridación, inicialmente se prehibridó la membrana de nylon con 5 mL de solución de hibridación (SCC 5x, *Denhart's* 5x (0.5% BSA, 0.5% Ficoll, 0.5% PVP), SDS 0.5%, DNA esperma de salmón 100 μg/mL) a 40°C durante 1 hora. A continuación se reemplazó la solución de hibridación por una nueva y se añadió la sonda poli (dT) marcada radiactivamente. La hibridación se incubó a 40°C durante 16 horas en rotación. Para eliminar las señales no específicas se lavó la membrana durante 20 minutos con una solución de lavado I (2x SSC y 0.1% SDS) y dos veces durante 30 minutos con una solución de lavado II (0.2x SSC y 0.1% SDS). Finalmente se sellaron las membranas con film plástico alimentario (cloruro de polivinilideno) y se expusieron en una pantalla IP. Las imágenes fueron obtenidas usando un lector *Phosphorimager* (FujiFilm).

Para determinar el cambio en la proporción de RNA poliadenilado se calculó la relación entre la señal radiactiva y la cantidad de RNA depositada en cada punto. Posteriormente se refirieron todos los valores a la condición inicial para calcular la variación en la proporción de RNA poliadenilado en función del tiempo o el tratamiento.

## Oligonucleótidos utilizados en este trabajo

Tabla 2. Oligonucleótidos usados en este trabajo.

Nombre	Secuencia
Intg V-1	GGG TGT CAG AAT ATG GGG CGG TAG TA
Intg V-2	CAC CCC GAA GCT GCT TTC ACA ATA C
oJW102	GCG GTG ACC CGG GAG ATC TGA ATT C
oJW103	GAA TTC AGA TC
Oligo dT	TTT TTT TTT TTT TTT (ACG)(AGCT)
PMA12012	CTA TTA TTG ATG CTT TGA AGA CCT CCA G
PMA12290	TGC CCA AAA TAA TAG ACA TAC CCC ATA A
RPL25-1D	AGA TGA GGA GGC ATG GGT C
RPL25-1R	GGC AAG CCG AAG GTA GAA AG
RPL25-2D	CTA CCT TCG GCT TGC CTC AC
RPL25-2R	CCC GAT AAT TAC GTT GCT CAC C
RPL25-3D	TCC GAC GAC AAG AAT AGT GC
RPL25-3R	TTT CTT AGC GGC AGT AGC C
RPL25-4D	GCC GCT AAG AAA GCT GTC G
RPL25-4R	CTT GGT ACC GTT TGG TCT AAC C
RPL25-5D	AGG CCG TCA AGG AAT TAT ACG
RPL25-5R	TTT GTA AGG CAC AGG AAA CCC
RPS3-1D	AAA GGT CCA CGT CAG TTC CAC
RPS3-1R	CAG GAA AGG TAT GGA TGT TAC GG
RPS3-2D	TCC GTA ACA TCC ATA CCT TTC C
RPS3-2R	AAG ACA CCG TCA GCG ACT AGC
RPS3-3D	GCT GAC GGT GTC TTC TAC GC
RPS3-3R	TCT TCT GAT AGC CAA ACC GTT C
RPS3-4D	GCT GCT TAC GGT GTC GTC AG
RPS3-4R	GGT GAC AGC ATC TGG CAA AG
RPS3-5D	AGA ACT GGT CCA AAG GCT TTG
RPS3-5R	CAC TGC TTC AAC GTG CAC C
CYT1-1D	TGA GCC AAT CCC AAA TGT TTG
CYT1-1R	AAA GAA AGA TTG TGA AAC CGA TG
CYT1-2D	ATC GGT TTC ACA ATC TTT CTT TG
CYT1-2R	TCC TTT GAG CCC AAC GTT TAG
CYT1-3D	TCT AAA CGT TGG GCT CAA AGG
CYT1-3R	CAC CAA CCA AAG TTC TCC AAG C
CYT1-4D	GAG CTG CCA ATC AAG GTG C
CYT1-4R	TGT TCA GGT TCG GCA CAC C
CYT1-5D	GGG TGA AGA AGT TCA AAT GGG
CYT1-5R	TCG CTT GCA GCA GTA TCT CAG

## Técnicas genómicas

### Hibridación de Chips de DNA

En este trabajo se han utilizado los chips de DNA de *S. cerevisiae* realizados por el servicio de chips de DNA (SCSIE) de la Universitat de Valencia. Estos chips están realizados sobre una membrana de nylon de 7 x 11 cm y poseen 6144 puntos donde se distribuyen las sondas que cubren la zona codificante de la mayoría de los genes de levadura

## Materiales y métodos

(Alberola T.M. *et al* 2004). Concretamente se han usado chips de las versiones v3, v4, v5 y v6. En todas las aproximaciones genómicas se han usado las mismas condiciones de hibridación lavado y lectura.

La radiactividad incorporada en cada marcaje fue cuantificada mediante un contador de centelleo (Wallac 1409), usando 5 mL de solución de centelleo (*Cocktail 22 Normascint* (CO0135) de Scharlau) y un programa de conteo para  $^{33}\text{P}$  de 10 minutos. Previamente a la hibridación, las muestras de DNA fueron desnaturalizadas incubándolas 5 minutos a  $95^{\circ}\text{C}$  y enfriándolas en hielo.

Los chips se prehibridaron con 5 mL de solución de hibridación (SCC 5x, *Denhart's* 5x (0.5% BSA, 0.5% Ficoll, 0.5% PVP), SDS 0.5% y 100  $\mu\text{g}/\text{mL}$  DNA de esperma de salmón) a  $65^{\circ}\text{C}$  durante 1 hora. A continuación, se decantó la solución y se añadieron de 3 a 5 mL de solución de hibridación con la muestra marcada radiactivamente (de 3 a 5 millones de dpm por mL de solución de hibridación, dependiendo del tipo de muestra). Las hibridaciones se dejaron en una estufa con agitación rotatoria a  $65^{\circ}\text{C}$  durante 40-42 horas. Finalmente, para eliminar la señal inespecífica se realizó un lavado de 20 minutos con la solución de lavado I (2x SSC, 0.1% SDS) y dos de 30 minutos con la solución de lavado II (0.2x SSC, 0.1% SDS). Para evitar que los chips se secan se sellaron con film plástico alimentario (cloruro de polivinilideno) y se expusieron en pantallas IP Phosphorimager (BAS-MP de FujiFilm).

Para eliminar la señal radiactiva de los chips y poder reutilizarlos se eliminó la muestra hibridada con tres lavados de 15 minutos con una solución en ebullición con 0.1% de SDS y 5 mM tampón fosfato sódico pH 7.

## **Cálculo genómico de la cantidad de los mRNAs**

Para realizar un análisis de la cantidad de mRNAs a escala genómica se usaron dos aproximaciones diferentes. O un marcaje de los mRNAs poliadenilados usando un cebador oligo dT o un marcaje del RNA total usando hexámeros aleatorios.

## **Marcaje de cDNA usando oligo dT**

Para el marcaje de mRNAs poliadenilados se partió de 40-50  $\mu\text{g}$  de RNA total, al que se le añadieron 1  $\mu\text{L}$  de oligo dT (500  $\text{ng}/\mu\text{L}$ ) y 1  $\mu\text{L}$  de inhibidor de RNasas (*RNase OUT* de Invitrogen). Seguidamente se aumentó el volumen hasta 14.5  $\mu\text{L}$  y se calentó la muestra a  $70^{\circ}\text{C}$  durante 10 minutos, dejándola a continuación enfriar a temperatura ambiente.

Posteriormente se añadieron 6  $\mu\text{L}$  de 5x *First Strand Buffer*, 3  $\mu\text{L}$  de DTT 0.1 M, 1.5  $\mu\text{L}$  de dNTPs (16 mM de ACG y 0.1 mM de C), 1  $\mu\text{L}$  de Retrotranscriptasa SuperScript II (200 u/ $\mu\text{L}$  Invitrogen) y 4  $\mu\text{L}$  de  $\alpha^{33}\text{P}$ -dCTP (3000 Ci/mmol y 10 mCi/mL). Para realizar la retrotranscripción se incubó la muestra durante 1 hora a 42°C y se paró el marcaje añadiendo 1  $\mu\text{L}$  de EDTA 0.5 M. Para eliminar el exceso de nucleótidos no incorporados se aumentó el volumen de la muestra hasta 100  $\mu\text{L}$  y se purificó usando una columna de exclusión molecular (S-300 HR de GE-healthcare).

Se cuantificó la incorporación de la muestra y se usaron  $3 \cdot 10^6$  dpm/mL en la solución de hibridación.

### **Marcaje de cDNA usando cebadores aleatorios**

En el caso del marcaje de RNA total con cebadores aleatorios fue necesario eliminar previamente los restos de DNA contaminante mediante un tratamiento con DNasa I. Se partió de 50  $\mu\text{g}$  de RNA al que se le añadieron 2  $\mu\text{L}$  de 5x *First Strand Buffer*, 1  $\mu\text{L}$  de inhibidor de RNasas (RNase OUT de Invitrogen) y 0.6  $\mu\text{L}$  de DNasa I libre de RNasas (10 u/ $\mu\text{L}$  Roche), se aumentó el volumen hasta 20  $\mu\text{L}$  y se incubó durante 30 minutos a 37°C. Posteriormente se añadieron 0.6  $\mu\text{L}$  más de DNasa I libre de RNasas y se incubó durante 30 minutos más. Para eliminar la DNasa se extrajo el RNA con un primer pase de Fenol:Cloroformo:Alcohol Isoamílico (125:24:1) y un segundo pase de Cloroformo:Alcohol Isoamílico (24:1). Seguidamente se precipitaron las muestras durante 3 horas a -20°C el RNA con 1/10 de volumen de acetato sódico 3M y 2.5 volúmenes de EtOH 96° frío.

Para marcar la muestra, ya libre de DNA, se utilizó el mismo protocolo previamente descrito para el marcaje de cDNA. Pero sustituyendo el oligo dT por 1.2  $\mu\text{L}$  de hexámeros aleatorios (3  $\mu\text{g}/\mu\text{L}$  de Invitrogen) y purificando la muestra con una columna de exclusión molecular (*ProbeQuant G-50* de GE-healthcare).

Se cuantificó la incorporación de la muestra y se usaron  $5 \cdot 10^6$  dpm/mL en la solución de hibridación.

### **Marcaje de DNA genómico**

Para poder calcular valores absolutos de cantidad de mRNA o tasa de transcripción para los diferentes genes, es necesario realizar una normalización de la cantidad de sonda depositada en cada punto del chip usando un marcaje de DNA genómico. Para ello, se partió de 1  $\mu\text{g}$  de

## Materiales y métodos

DNA genómico disuelto en un volumen total de 46  $\mu\text{L}$  y se desnaturalizó incubándolo 5 minutos a 95°C y pasándolo a hielo. El marcaje de la muestra se realizó con 4  $\mu\text{L}$  de  $\alpha^{33}\text{P}$ -dCTP (3000 Ci/mmol y 10 mCi/mL) y usando el kit *Ready-to-Go* (-dCTP) (GE-healthcare). La reacción se incubó durante 2 horas a 37°C. Seguidamente se paró la reacción con 1  $\mu\text{L}$  de EDTA 0.5 M y se eliminaron los nucleótidos no incorporados con una columna de exclusión molecular (*ProbeQuant G-50* de GE-healthcare). Finalmente se cuantificó la incorporación, se desnaturalizó la muestra y se hibridó con los chips de DNA usando una concentración de  $5 \cdot 10^6$  dpm/mL de solución de hibridación.

## **RNA Pol II ChIP-on-Chip (RPCC)**

Este método se basa en ChIP-on-chip y permite detectar la densidad de polimerasas unidas a los genes combinando el método de inmunoprecipitación de cromatina y la hibridación de chips de DNA (fig 6).

Para cada experimento se tomaron 100 mL de células crecidas a  $\text{DO}_{600}$  0.5 y se entrecruzaron añadiendo formaldehído a una concentración final del 1% durante 15 minutos en agitación. A continuación se paró el entrecruzamiento con 5.4 mL de glicina 2.5 M y se pasaron las células a hielo. Finalmente se lavaron las células con 4 lavados de TBS frío (20 mM Tris-HCl pH 7.5 y 146 mM NaCl) y se guardaron a -20°C.

Para preparar la cromatina se resuspendieron las células en 300  $\mu\text{L}$  de Tampón de lisis (50 mM HEPES, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Desoxicolato sódico, 1 mM PMSF, 1mM Benzamidina e inhibidor de proteasas Complete<sup>TM</sup> (1 pastilla para 50 mL totales)) y añadiéndolas a un tubo cónico con 300  $\mu\text{L}$  de perlas de vidrio. A continuación, se rompieron las células a 4°C usando un Vortex2-Genie a máxima potencia durante 12 minutos. Para separar la cromatina de las perlas de vidrio se añadieron 300  $\mu\text{L}$  más de tampón de lisis y, perforando el tubo cónico con una aguja caliente, se pasó la cromatina a un tubo eppendorf. La cromatina se fragmentó hasta un tamaño adecuado para la inmunoprecipitación (alrededor de 400-600 pb) mediante sonicación usando un BioRuptor (Diagenode) con 5 pulsos de 30 segundos a alta potencia con una eficiencia de 0.5. Posteriormente se centrifugó la muestra para eliminar los restos celulares y se recuperó el sobrenadante. De ese sobrenadante se reservaron 10  $\mu\text{L}$  como control positivo (extracto celular completo).



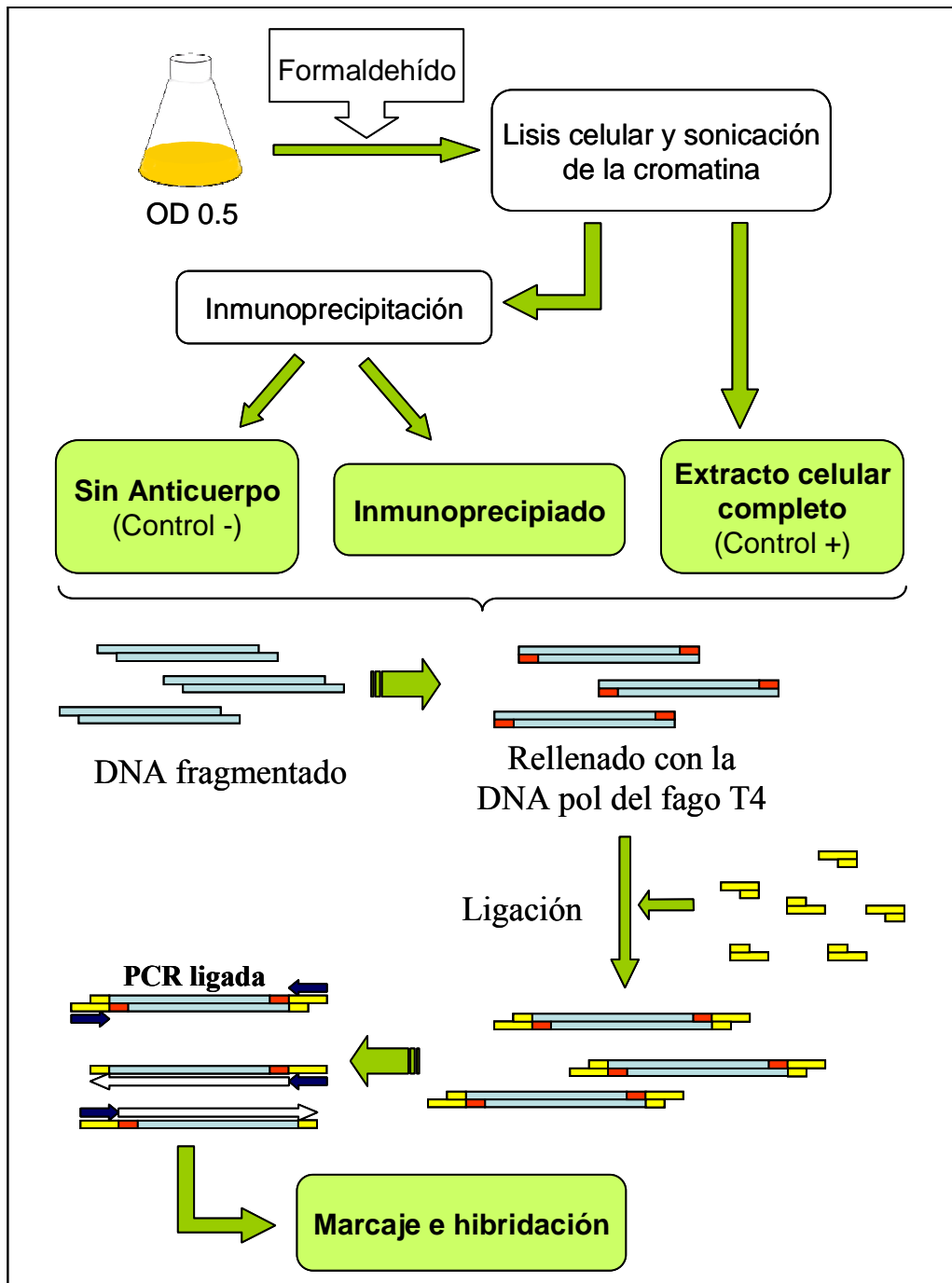


Figura 6. Diseño experimental del RPCC. Se añadió formaldehído a un cultivo para entrecruzar el DNA y las proteínas. Seguidamente se rompieron las células y se fragmentó la cromatina. Una parte de la cromatina total se conservó como control positivo (extracto celular completo), el resto se sometió a un proceso de inmunoprecipitación específico (IP) o un control negativo (Sin Anticuerpo). Posteriormente se amplificaron las muestras inmunoprecipitadas mediante PCR ligada y se hibridaron de forma sucesiva con los chips de DNA.

## Materiales y métodos

Para la inmunoprecipitación se usaron 50  $\mu\text{L}$  de Dynabeads por muestra. Dependiendo del anticuerpo utilizado en la inmunoprecipitación se usaron Dynabeads recubiertas con anticuerpos anti IgG de ratón (para el Ac 8WG16 y el anti-MYC), anti IgM de ratón (para el Ac H5) o con Proteína G (para los anticuerpos policlonales de conejo contra el CTD fosforilado en Ser5 o en Ser2). En todos los casos se lavaron dos veces las Dynabeads con 600  $\mu\text{L}$  de PBS/BSA (140 mM NaCl, 2.7 mM KCl, 10 mM  $\text{Na}_2\text{HPO}_4$ , 1.8 mM  $\text{KH}_2\text{PO}_4$ , 5 mg/mL BSA, pH 7.4) y se incubaron en un agitador orbital durante 16 horas a 4°C con el Anticuerpo de interés y 1  $\mu\text{L}$  de tRNA de levadura (10 mg/mL). Paralelamente se preparó un control negativo (Sin Anticuerpo) donde se resuspendieron las bolas magnéticas únicamente con PBS/BSA en vez del anticuerpo específico contra la proteína de interés. Después de la incubación se lavó el exceso de Ac no unido con 4 lavados de PBS/BSA y se resuspendieron las bolas magnéticas con 30  $\mu\text{L}$  de PBS/BSA. Seguidamente se añadió la cromatina sonicada a las esferas y se incubaron en el agitador orbital a 4°C durante 90-180 minutos. Para limpiar la muestras de la cromatina no unida específicamente se realizaron 2 lavados de 1 mL con tampón de lisis, 2 con tampón de lisis con 360 mM NaCl, 2 con tampón de lavado (10 mM Tris pH 8, 250 mM LiCl, 0.5% NP-40, 0.5% Desoxicolato sódico y 1 mM EDTA) y uno con TE (10 mM Tris HCl, 1 mM EDTA). Seguidamente, se eluyó la muestra inmunoprecipitada con 2 incubaciones consecutivas de 10 minutos con tampón de elución (50 mM Tris pH 8, 10 mM EDTA y 1 % SDS) a 65°C y en agitación. Finalmente se desentrecuzaron las muestras incubándolas durante 16 horas a 65°C y aumentando el volumen a 300  $\mu\text{L}$  con tampón de elución. Para eliminar las proteínas se incubaron las muestras durante 90 minutos a 37°C con 7.5  $\mu\text{L}$  Proteinasa K (20 mg/mL) y se aisló el DNA con una microcolumna de purificación de DNA (*QIAquick* de QIAGEN). Para comprobar la correcta inmunoprecipitación de la polimerasa se realizó una PCR con oligonucleótidos específicos para una zona intergénica (Intg V) y una zona dentro de un gen altamente expresado (*PMA1*).

Una vez realizada la inmunoprecipitación, para conseguir suficiente DNA para hibridar un chip de DNA, fue necesario amplificar el DNA inmunoprecipitado. Para ello se usó una amplificación por PCR ligada (Ren B. *et al* 2000). Inicialmente se rellenaron los fragmentos de DNA con la DNA pol del fago T4 en una incubación de 20 minutos a 12°C para generar extremos romos. Concretamente, a 100  $\mu\text{L}$  de muestra se le añadieron 0.6 unidades de polimerasa de fago T4 (Roche), 22  $\mu\text{L}$  de

tampón DNA pol 5x, 0.5 µL de BSA (10 mg/mL) y 1 µL dNTPs (10 mM). Posteriormente se fenolizó la muestra y se precipitó con acetato sódico y EtOH a -20°C. Una vez resuspendida la muestra en 25 µL, se le ligaron adaptadores unidireccionales. Para ello se incubó la muestra durante 16 h a 16°C con 0.5 unidades de ligasa de T4 (Invitrogen), 10 µL de tampón 5x de ligasa, 0.5 µL ATP (20 mM) y 6.7 µL de adaptador unidireccional LE59 (15 µM de oJW102 y oJW103, 250 mM Tris-HCl pH 8). A continuación se volvió a precipitar el DNA con acetato sódico y EtOH a -20°C. Finalmente, para amplificar la muestra, se realizó una PCR con cebadores que reconocen los adaptadores unidireccionales que se habían añadido. Para ello se añadieron a 32.15 µL de muestra con adaptadores ligados, 4 µL de tampón Taq DNA pol 10x, 1.6 µL MgCl<sub>2</sub> 50 mM, 1 µL dNTPs 10 mM y 1.25 µL oligo oJW102 40 µM. Se calentó la muestra durante 2 minutos a 55°C y a continuación se le añadieron 8 µL de agua, 1 µL tampón Taq DNA pol 10x, 0.4 µL MgCl<sub>2</sub> 50 mM y 1 µL Taq DNA pol (Biotools) a cada muestra. Para la PCR se usó el siguiente programa: 5 min a 72 °C, 2 min 95°C, 34 ciclos (30 s a 95°C, 30 s a 55°C y 2 min a 72°C) y una elongación final de 4 min a 72°C. El producto de PCR se precipitó con acetato sódico y etanol. Para comprobar la eficiencia y el tamaño de la amplificación se resuspendió en 50 µL de H<sub>2</sub>O y se corrieron 5 µL en una electroforesis en gel de agarosa al 1.2% en TBE 0.5X (1 mM EDTA, 45 mM Tris-Borato pH 8).

Para marcar las muestras se usó un único ciclo de PCR en presencia de <sup>33</sup>P-dCTP a partir de 5-15 µL del producto amplificado. A los 5-15 µL de DNA amplificado se le añadieron 5 µL tampón 10X, 2 µL de MgCl<sub>2</sub> 50 mM, 2 µL de dNTPmix (5 mM de AGT), 2.5 µL dCTP 0.5 mM, 1.25 oligo oJW102 40 µM, 5 µL de α<sup>33</sup>P-dCTP (3000 Ci/mmol y 10 mCi/mL), 1 µL DNA pol Taq (Biotools) y agua hasta 50 µL. Se incubó la muestra 5 min a 95°C, 5 min a 50°C y 30 min a 72°C. Seguidamente el DNA marcado se purificó en columna de exclusión molecular (*ProbeQuant G-50 Micro Columns* de GE-healthcare) para eliminar los nucleótidos no incorporados. Posteriormente, se cuantificó la incorporación, se desnaturalizó la muestra y se hibridó con los chips usando 5·10<sup>6</sup> dpm por mililitro de solución de hibridación.

### **Genomic run-on**

El *Genomic Run-on* (GRO) es una técnica que permite medir la densidad de polimerasas en transcripción activa basándose en la elongación de RNAs nacientes en presencia de un precursor radiactivo (García-

## Materiales y métodos

Martínez J. *et al* 2004). Además, combinando los datos de tasa de transcripción con la cuantificación de la cantidad de mRNAs, permite calcular datos indirectos de estabildades de mRNAs (fig 7).

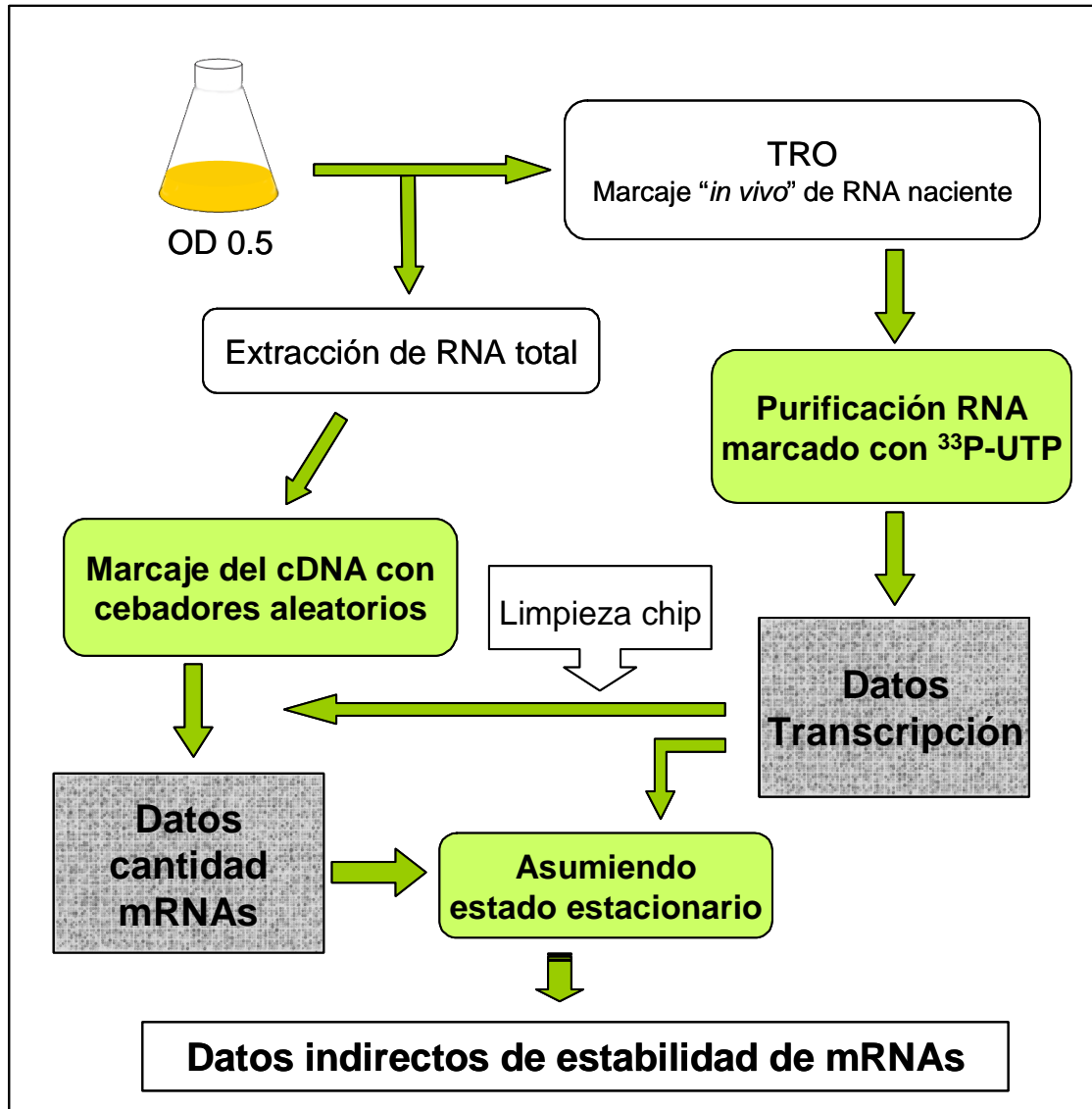


Figura 7. Diseño experimental del Genomic run-on (GRO). Partiendo de un mismo cultivo se separaron dos alícuotas. A una de ellas se la sometió a un marcaje de los RNAs nacentes usando  $\alpha^{33}\text{P}$ -UTP y se hibridó ese RNA marcado con un chip de DNA (datos de transcripción). A partir de la otra alícuota se obtuvo RNA total que se marcó e hibridó sobre el mismo chip (una vez eliminada la señal de transcripción) para obtener datos de la cantidad de mRNAs. La combinación de ambos datos, asumiendo un estado estacionario para la cantidad de mRNAs, permitió calcular la estabilidad para todos los mRNAs.

Para cada experimento se usaron 100 mL de células a una  $DO_{600}$  cercana a 0.5. Se recogieron las células por centrifugación y se lavaron con 5 mL sarkosil 0.5% frío. Se congelaron las células procedentes de 50 mL de células para el estudio de la cantidad de mRNAs y se continuó el GRO con la otra alícuota.

Se resuspendieron las células de la alícuota de GRO con 0.5% sarkosil y se incubaron durante 20 minutos en hielo para permeabilizarlas. Seguidamente se recogieron las células por centrifugación y se resuspendieron con 120  $\mu$ L de agua. A las células resuspendidas se le añadieron 120  $\mu$ L tampón de transcripción 2.5x (50 mM Tris-HCl pH 7.7; 500 mM KCl; 80 mM  $MgCl_2$ ), 16  $\mu$ L de NTPs (ATP, CTP y GTP 10 mM de cada), 6  $\mu$ L DTT 0.1 M y 16  $\mu$ L  $\alpha$ - $^{33}P$ -UTP (10 $\mu$ Ci/ $\mu$ L). Para realizar el pulso de transcripción se incubaron las células a 30°C durante 5 minutos. La transcripción se paró añadiendo 1 mL de agua a 4°C. Seguidamente se centrifugaron las células y se eliminó el sobrenadante con el exceso de radiactividad. Posteriormente se extrajo el RNA como ha sido descrito previamente, pero realizando únicamente una precipitación con 0.1 volúmenes de LiCl 5 M y 2.5 volúmenes de EtOH 96% durante 3 horas a -20°C. El RNA se resuspendió en 300  $\mu$ L de agua y se cuantificó la cantidad de RNA espectrofotométricamente. La incorporación radiactiva se cuantificó usando un contador de centelleo, y la muestra de RNA marcado se hibridó siguiendo el procedimiento general. El rendimiento del marcaje generalmente fue del orden de  $3 \cdot 10^7$  dpm/ $\mu$ g RNA total.

Para calcular la cantidad de mRNAs de cada una de las muestras se realizó un marcaje del RNA usando cebadores aleatorios siguiendo el protocolo que ha sido descrito previamente.

### **Genomic run-on fluorescente de alta densidad**

Esta técnica permite incorporar *in vivo* un nucleótido modificado al RNA naciente para permitir su posterior purificación. De esta manera es posible usar chips de DNA basados en tecnología fluorescente, lo que permite usar chips comerciales de alta densidad (p.ej. chips de embaldosado). En esta técnica se han usado dos aproximaciones diferentes usando dos nucleótidos modificados: el 4-tio-UTP (Trilink) y Biotina-11-UTP (Ambion). En ambas aproximaciones se han usado los chips *GeneChip® S. cerevisiae Tiling 1.0R Array* (Affymetrix) (Gresham D. *et al* 2006) que cubren una hebra del genoma a una resolución media de 5 pb.

## Materiales y métodos

### **GRO fluorescente usando 4-tio-UTP**

Para cada experimento se usaron 200 mL de células a una  $DO_{600}$  cercana a 0.5. Se recogieron las células por centrifugación en cuatro tubos de 50 mL y se lavaron con 5 mL sarkosil 0.5%. Cada alícuota se resuspendió en 1 mL de sarkosil 0.5% y fueron incubadas durante 2 minutos a temperatura ambiente. Seguidamente se recogieron las células por centrifugación y se resuspendieron en 116  $\mu$ L de agua. A las células se les añadieron 120  $\mu$ L tampón de transcripción 2.5x (50 mM Tris-HCl pH 7.7, 500 mM KCl, 80 mM  $MgCl_2$ ), 16  $\mu$ L de NTPs (ACG 10 mM de cada), 6  $\mu$ L DTT 0.1 M y 20  $\mu$ L 4-tio-UTP (0.5mM). Para realizar el pulso de transcripción se incubaron las células 30°C durante 5 minutos. La transcripción se paró añadiendo 1 mL de agua a 4°C. Seguidamente se centrifugaron las células y se eliminó el sobrenadante. Posteriormente se extrajo el RNA con fenol de la misma manera que en el experimento de GRO convencional (descrito en el apartado anterior).

Seguidamente se unió biotina al RNA total purificado a través del azufre del 4-tio-UTP. Para ello se añadieron 2  $\mu$ L de *EzLink Biotin-HPDP* (1 mg/mL de Pierce resuspendido en DMF) a cada 5  $\mu$ g de RNA total y se incubó durante 90 minutos a 22°C y en oscuridad con 10 mM Tris-HCl pH 7.4 y 1 mM EDTA en un volumen final de cinco veces la cantidad de Biotin-HPDP usado. El RNA se precipitó añadiendo 1/10 de NaCl 5 M y un volumen de isopropanol, y centrifugando a máxima velocidad durante 20 minutos. Seguidamente se lavó el RNA con etanol 70% y se resuspendió en agua.

Para enriquecer la muestra en RNA naciente se incubó el RNA total biotinilado en un volumen final de 500  $\mu$ L con 50  $\mu$ L de esferas magnéticas recubiertas de estreptoavidina (*RiboMinus* de Invitrogen) y se incubaron durante 10 minutos en un rotor a temperatura ambiente. Para eliminar el RNA no biotinilado se hicieron dos lavados a 65°C y tres a temperatura ambiente con 1 ml de una solución 1 M NaCl, 10 mM EDTA y 10 mM TrisHCl pH 7.3. Seguidamente, para liberar el RNA biotinilado, se rompió el puente disulfuro del RNA con la biotina mediante 2 incubaciones de 5 minutos con 200  $\mu$ L de DTT 0.1 M. Finalmente, se volvió a precipitar el RNA purificado con añadiendo 1/10 de NaCl 5 M y un volumen de isopropanol.

### **GRO fluorescente usando Biotina-11-UTP**

Se usó el mismo protocolo que para el caso anterior pero usando 1  $\mu$ L de Biotina-11-UTP (75mM) por reacción en vez del 4-tio-UTP. Sin

embargo, dado que la biotina puede arrastrar al RNA naciente a la fase orgánica durante una purificación con fenol, fue necesario realizar una extracción de RNA sin fenol. Para ello, se usó un kit de extracción sin fenol (*MasterPure Yeast RNA Purification Kit* de Epicentre) separando las células provenientes de 200 mL de cultivo a  $DO_{600}$  0.5 en 10 extracciones diferentes.

Una vez obtenido el RNA, y cuantificado espectrofotométricamente, resultó necesario realizar una fragmentación previa a la purificación del RNA biotinilado. Para ello, se dividió el RNA total en tubos eppendorfs con 200  $\mu$ g de RNA cada uno resuspendido en un volumen total de 320  $\mu$ L. A cada tubo se le añadieron 80  $\mu$ L de tampón de fragmentación (200 mM Tris-Acetato pH 8.1, 500 mM Acetato Potásico y 150 mM Acetato Sódico), y para fragmentar el RNA se incubaron durante 3 minutos a 94°C. El RNA fragmentado fue precipitado por centrifugación con 40  $\mu$ L acetato sódico 3 M y 400  $\mu$ L de isopropanol. Seguidamente se realizó un lavado con etanol 70% y se resuspendió el RNA fragmentado en agua.

Para purificar el RNA biotinilado se usaron bolas de agarosa recubiertas de avidina monomérica (Pierce). Se usó un tubo con 300  $\mu$ L de avidina monomérica por muestra. Para bloquear las bolas se le añadieron 200 mL de tampón D (2 mM TrisHCl pH 7.3 y 0.5 mM EDTA) y 1  $\mu$ L de DNA de salmón (10 mg/mL, fenolizado y sonicado) a cada tubo. Las bolas de agarosa se incubaron en un rotor a temperatura ambiente durante 15 minutos. Seguidamente, se pasaron a una columna vacía y se realizaron 6 lavados con 600  $\mu$ L de glicina 0.1 M pH 2 y 6 lavados con 600  $\mu$ L de tampón D. Una vez eliminado el exceso de DNA de salmón se resuspendieron las bolas en 300  $\mu$ L de tampón D y se pasaron a un eppendorf con el RNA total fragmentado (resuspendido en otros 300  $\mu$ L). Se incubó la muestra durante 20 minutos a 4°C en un rotor para permitir la unión del RNA biotinilado a las bolas. Para eliminar el RNA no biotinilado contaminante se pasaron las bolas a una columna y se realizaron 10 lavados con 600  $\mu$ L de tampón A (1 M NaCl, 0.1% NP-40, 5 mM EDTA y 20 mM Tris-HCl pH 7.3) y 8 lavados con 600  $\mu$ L de tampón D. Para liberar el RNA biotinilado se resuspendieron las bolas en 300  $\mu$ L de tampón D y se pasaron a un tubo eppendorf con 300  $\mu$ L de reactivo 2-ME (2 M Mercaptoetano, 50 mM MOPS, 1 M NaCl, 2.5 mM EDTA y 2 mM D-Biotina) donde se incubaron durante 5 minutos a 95°C. Para recuperar el RNA biotinilado se centrifugaron las bolas y se recuperó el sobrenadante. Este proceso se repitió una segunda vez. Finalmente, se añadieron 2  $\mu$ L de glucogeno para Biología Molecular (20

## Materiales y métodos

mg/mL de Roche) al RNA biotinilado y se precipitó con acetato sódico y etanol.

### **Comprobación de la incorporación de Biotina al RNA**

Para comprobar la incorporación de biotina al RNA se corrieron geles de agarosa al 1% en tampón TAE 1x (40 mM Tris-acetato, 1 mM EDTA, pH 8) y se transfirieron a una membrana de nylon Hybond N<sup>+</sup> (GE-healthcare) por capilaridad usando SSC 6x (0.9 M NaCl y 90 mM citrato sódico) durante toda la noche. Al día siguiente se fijaron las muestras a la membrana entrecruzándolo con 50 mJ de luz UV de 254 nm (GS Gene Linker de Bio-Rad). Las membranas se incubaron en agitación durante 20 minutos con tampón de bloqueo (125 mM NaCl, 9 mM Na<sub>2</sub>HPO<sub>4</sub>, 7 mM NaH<sub>2</sub>PO<sub>4</sub>, 10% SDS) y posteriormente se incubaron durante 5 minutos con *Immunopure Streptavidin-HRP* (Pierce) 1 µg/mL en tampón de bloqueo. Se hicieron 2 lavados de 10 minutos con tampón de bloqueo y 2 lavados de 5 minutos con tampón de lavado (100 mM TrisHCl, 100 mM NaCl, 10.5 mM MgCl<sub>2</sub>, pH 9.5). Posteriormente se incubó la membrana con ECL (GE-healthcare) durante 1 min y se reveló usando una pantalla fotográfica Super RX (Fuji).

### **Amplificación del RNA biotinilado**

Con el objeto de amplificar el RNA purificado, para tener suficiente cantidad para hibridar un chip primero fue necesario eliminar cualquier resto de DNA contaminante. Para ello se resuspendió el RNA purificado en 169 µL de agua y se le añadieron 20 µL de tampón 5x de DNasa, 1 µL inhibidor de RNasas ScriptGuard 40 u/µL, y 10 µL DNasa libre de RNasas 1 u/µL (todo ello del kit *MasterPure Yeast RNA Purification* de Epicentre). Las muestras se incubaron durante 30 minutos a 37°C, se inactivó la DNasa con una incubación de 15 minutos a 65°C y se eliminó siguiendo el protocolo del kit *MasterPure Yeast RNA Purification Kit* (Epicentre).

Además, para limpiar la muestra de RNA de cualquier contaminante que pudiera arrastrar, se usó el kit RNeasy (Qiagen) con las siguientes modificaciones. Se añadió al RNA resuspendido en 35 µL de agua 350 µL de tampón RLT y 600 µL de etanol 96°, se agitó y se pasó por una microcolumna de purificación de RNA. Seguidamente, se realizaron dos lavados con 500 µL de tampón RPE y se eluyó la muestra en 30 µL de agua. Finalmente se concentró la muestra hasta un volumen de 5 µL por evaporación.



Para confirmar la calidad y cantidad del RNA purificado se usó un sistema de electroforesis capilar (Experion RNA StdSens de Bio-Rad) y para amplificar la muestra se usó el protocolo estandar de Affymetrix *GeneChip® WT Double-Stranded Target Assay*.

### **Hibridación de los Chips de Affymetrix**

Las hibridaciones se llevaron a cabo en el servicio de análisis multigénico de la unidad central de investigaciones médicas (UCIM) de la Universidad de Valencia. Para todas las hibridaciones se siguió el protocolo estándar de Affymetrix *WT GeneChip® Double-Stranded Target Assay*, usando el *GeneChip® Hybridization, Wash and Stain Kit* (P/N 900720).

## **Técnicas Bioinformáticas**

### **Análisis de imagen y normalización**

#### **Análisis genérico para los *macroarrays* de Nylon**

Todas las imágenes de hibridaciones de los chips fueron obtenidas usando un lector *FLA3000 Phosphorimager* (FujiFilm) con 16 bits de profundidad y una resolución de 50  $\mu\text{m}$ . Las imágenes fueron analizadas usando el programa de análisis de imagen *ArrayVision* (Imaging Research). Se excluyeron del análisis todos los puntos cuya señal no fuera al menos 1.35 veces más intensa que el fondo. Posteriormente se calculó para cada punto el sARM (*Artifact Removed Median*), siendo este el valor de la mediana de la intensidad de los pixels cada punto al que se le han eliminado las posibles manchas artefactuales y se le ha restado la intensidad del fondo.

Todos los experimentos fueron realizados por triplicado barajando los chips entre réplicas. Para fusionar entre sí las diferentes replicas se usó el programa *ArrayStat* (Imaging Research). Se usó un modelo proporcional (dado que existe una relación entre la desviación estándar y la media) en escala logarítmica. Se normalizaron los datos a lo largo de las réplicas, asumiendo condiciones independientes y exigiendo un mínimo de 2 réplicas válidas (dentro de las 3 réplicas) para cada gen. Para calcular el error asociado a cada medida se usó una estimación del error conjunta para el ajuste de una curva. Finalmente, se eliminaron los valores anómalos.

## Materiales y métodos

Los datos obtenidos en este trabajo están depositados en las bases de datos *VYdBase* (<http://vydbase.uv.es/>) y GEO (Barrett T. *et al* 2007) (<http://www.ncbi.nlm.nih.gov/projects/geo/>). Las referencias son las siguientes:

- GSE11521 para el estudio cinético de la expresión génica durante el estado estacionario.
- GSE7261 para los cálculos de vida media usando 3 µg/mL de tiolulina. GSE8629 para la comparación entre el efecto de 3 y 10 µg/mL de tiolulina. Información adicional en el material suplementario de (Pelechano V. y Perez-Ortin, J.E. 2008).
- GSE14060 para los datos de RNA pol II ChIP on chip (RPCC) usando diferentes anticuerpos y en diferentes situaciones fisiológicas.
- GSE14077 y GSE14080 para los análisis de los mutantes *rap1-sil* y *tpk2* usando respectivamente RPCC o GRO.

### **Normalización de la señal de transcripción obtenida mediante GRO**

Se usó el mismo protocolo anteriormente descrito con las siguientes modificaciones: Inicialmente, se analizaron las réplicas de hibridaciones de GRO y de DNA genómico usando el programa *ArrayVision* como ha sido descrito. Posteriormente, se calculó la relación GRO/gDNA usando el programa *ArrayStat*. Como en el caso del GRO se incorpora UTP al RNA de una sola hebra, y en el del DNA genómico se incorpora dCTP en las dos hebras, resultó necesario realizar una corrección de la señal teniendo en cuenta la diferente proporción de nucleótidos para cada sonda.

Para comparar la transcripción entre diferentes puntos experimentales se trataron todas las muestras en paralelo, partiendo del mismo número de células y exponiendo los chips en la misma pantalla IP. La capacidad de hibridación de cada chip se normalizó usando un marcaje común de DNA genómico que se hibridó simultáneamente con todos los chips del estudio.

## **Normalización de la señal de transcripción obtenida mediante RPCC**

Para calcular el enriquecimiento en polimerasas se compararon los datos de la muestra inmunoprecipitada con el control positivo (extracto celular completo) usando el programa *ArrayStat*. No se normalizaron los datos respecto a los controles negativos debido a la alta variabilidad experimental que presentan las PCRs ligadas cuando se parte de tan poco material.

## **Análisis de los chips de embaldosado de Affymetrix**

Para analizar las imágenes de los chips *GeneChip® S. cerevisiae Tiling 1.0R Array* (Affymetrix) se utilizó el programa *Tiling Array Software* (Affymetrix). Se partió de los ficheros .cel fijándose un ancho de ventana de 30 nucleótidos, por lo que cada dato se obtiene a partir de 5-6 sondas diferentes. Para la identificación de transcritos se utilizaron los parámetros por defecto, que exigen una significación de al menos  $10^{-5}$  y fijan un mínimo de longitud de los fragmentos de 40 pares de bases con un intervalo máximo de 80 pb sin señal.

Para visualizar los resultados se utilizó el programa *Integrated Genome Browser* (Affymetrix).

Para analizar conjuntamente los resultados, así como para realizar comparaciones con otros disponibles en la bibliografía, se usó programación a medida usando el lenguaje R (<http://www.r-project.org/>).

## **Análisis bioinformático de los resultados.**

Para el análisis de los resultados e interpretación biológica de los resultados se han utilizado múltiples herramientas y bases de datos. Entre estas herramientas se pueden destacar las siguientes:

Bases de datos de para obtener información biológica:

- Base de datos SGD. "*Saccharomyces Genome Database*" (<http://www.yeastgenome.org/>).
- El programa para búsqueda bibliográfica de funciones *IHOP. Information Hyperlinked over Proteins* (<http://www.pdg.cnb.uam.es/UniPub/iHOP/>) (Hoffmann R. y Valencia, A. 2005).

## Materiales y métodos

- La herramienta de minería de información bibliográfica *Textpresso for yeast* (<http://textpresso.yeastgenome.org/>) (Muller H.M. *et al* 2004).

### Programas de análisis estadístico de datos:

- La suite para el análisis de datos online GEPAS. *Gene Expression Pattern Analysis Suite* (<http://gepas.bioinfo.cipf.es/>) (Montaner D. *et al* 2006).
- El programa para realizar cálculos masivos de diferencias por test-t: *T-profiler* (<http://www.t-profiler.org/>) (Boorsma A. *et al* 2005)
- El programa para análisis de series temporales STEM. *Short Time-series Expression Miner* (<http://www.cs.cmu.edu/~jernst/stem/>) (Ernst J. y Bar-Joseph, Z. 2006)

### Programas para la anotación funcional:

- La suite para la anotación genómica de resultados genómicos *Babelomics* (<http://babelomics.bioinfo.cipf.es/>) (Al-Shahrour F. *et al* 2006).
- La herramienta para el cálculo de enriquecimiento en categorías funcionales: *FuncAssociate*. (<http://llama.med.harvard.edu/cgi/func/funcassociate>) (Berriz G.F. *et al* 2003).
- La herramienta para el cálculo de enriquecimiento en categorías funcionales: *Funcspec*. (<http://funcspec.med.utoronto.ca/>) (Robinson M.D. *et al* 2002)

Además de estos servicios on-line se utilizaron las herramientas informáticas Microsoft Office Excel, Matlab (MathWorks) y el lenguaje estadístico R (<http://www.r-project.org/>).

## **Resultados y discusión**

## Resultados y discusión

En este trabajo se ha pretendido realizar un estudio exhaustivo sobre el recambio de los mRNAs a escala genómica en *S. cerevisiae*.

Inicialmente se ha tratado de confirmar la asunción de estado estacionario para la expresión génica en condiciones de crecimiento exponencial. De ser correcta esta asunción, siempre que las condiciones fisiológicas no sufran cambios bruscos, las tasas de transcripción se compensan con las tasas de degradación para mantener una concentración de mRNAs constante y, por lo tanto, será válido calcular indirectamente valores de estabilidad a partir de datos de tasa de transcripción y de cantidad de mRNA.

Una vez comprobada la validez del estado estacionario para la expresión génica en las condiciones estudiadas, se ha realizado un cálculo alternativo de estabilidades de mRNAs usando un método directo. Concretamente se ha realizado una parada general de la transcripción usando tiolutina y se ha cuantificado la desaparición de los mRNAs. Los datos obtenidos por cada procedimiento, directo e indirecto, se han comparado.

Por otra parte, se ha desarrollado un método para medir la tasa de transcripción a escala genómica basado en la inmunoprecipitación de cromatina (RPCC). Este método ha permitido descubrir y corregir alguno de los sesgos técnicos presentes en el GRO contribuyendo a generar unos datos genómicos de transcripción en levadura más robustos y fiables. Además, la comparación de estas dos técnicas también ha permitido descubrir diferencias biológicas en la forma que tiene la célula de transcribir los genes pertenecientes a diferentes grupos funcionales. Se ha profundizado en el estudio de la regulación de estas diferencias y se ha descrito un posible mecanismo de acción para el caso de los genes de las proteínas ribosómicas.

Finalmente se ha puesto a punto una variante del GRO basada en la purificación selectiva de mRNAs nacientes. Esta alternativa permitirá en un futuro el estudio de la transcripción a una mayor resolución usando chips de DNA de embaldosado u otras técnicas genómicas de nueva generación.

## **1. Estudio del estado estacionario de la expresión génica**

La utilización de datos de transcripción y de cantidad de mRNA para el cálculo indirecto de las estabilidades de los mRNAs requiere asumir que estos dos procesos están equilibrados para mantener una concentración de mRNA constante en el tiempo. Si asumimos un volumen promedio constante para la célula de levadura, se puede hablar de cantidad de mRNA como equivalente a la concentración de mRNA ( $[mRNA]$ ). Este tipo de cálculo permite una determinación instantánea de las estabilidades de mRNA utilizando los resultados del GRO sin necesidad de estudiar la desaparición de los mRNA durante un tiempo prolongado. Sin embargo, es importante confirmar la validez de esta asunción. Por lo tanto, decidimos realizar un estudio detallado de la validez de esta asunción durante la condición más habitual de trabajo con levadura.

### **1.1 Estado estacionario de la expresión génica durante el crecimiento exponencial**

La condición más habitual de trabajo en levadura, y por tanto para la que se poseen más datos, es un cultivo asincrónico de levaduras en crecimiento exponencial en medio rico (YPD) en un matraz con agitación a 28°C. Ésa es la condición óptima para el crecimiento donde la levadura crece a máxima velocidad y de forma continuada y estable durante un periodo prolongado de tiempo.

En esta situación, donde las células permanecen en una situación de crecimiento óptimo durante varias generaciones y están completamente adaptadas a las condiciones del medio que mantiene concentraciones no limitantes de todos los requerimientos nutricionales, es lógico asumir que no hay variaciones importantes para la cantidad de mRNAs al nivel poblacional. Como la cantidad de mRNA es el producto directo de dos procesos opuestos, la síntesis de mRNA (transcripción) y su degradación, si no hay variaciones en la cantidad de RNA con el tiempo es por que las tasas de estos dos procesos son iguales. Esa situación, donde la cantidad de los mRNAs permanece estable es lo que podemos denominar estado estacionario para la expresión génica. Además, si se cumple esta hipótesis, se puede calcular para un instante determinado una variable (ej. la tasa de degradación), conociendo las otras dos variables (ej. cantidad de mRNA y tasa de transcripción). Este cálculo se puede realizar fácilmente si se considera que la degradación de los mRNAs es un proceso de primer orden que sólo depende de la cantidad de mRNAs presente y de la constante de

## Resultados y discusión

degradación ( $k_d$ ) [4], y a la transcripción como un proceso de orden cero que sólo depende de la tasa de transcripción (TR) [5]. Durante el estado estacionario, por definición, se pueden igualar estas dos tasas relacionando las tres magnitudes [6].

$$[4] \quad d[\text{mRNA}]/dt = k_d \cdot [\text{mRNA}]$$

$$[5] \quad d[\text{mRNA}]/dt = \text{TR}$$

$$[6] \quad \text{TR} = k_d \cdot [\text{mRNA}]$$

Para comprobar si es correcto asumir una situación de estado estacionario respecto a la síntesis y degradación de mRNA en condiciones de crecimiento exponencial en YPD y, por lo tanto, si se pueden usar datos de tasa de transcripción y cantidad de mRNA obtenidos mediante experimentos de GRO para calcular indirectamente las estabilidades de los mRNAs, se decidió realizar un estudio detallado de la variación de la cantidad de mRNA durante el crecimiento exponencial.

### **1.2 Reanálisis de datos genómicos de variación de mRNAs durante el crecimiento exponencial**

En la bibliografía no existen muchos estudios que analicen la expresión génica global durante un intervalo amplio de tiempo del crecimiento exponencial. Sólo el estudio de Radonjic M. et al. 2005 contiene datos de tiempos diversos durante este periodo de crecimiento celular. En ese trabajo se estudia con detalle un ciclo de 9 días desde la dilución de un cultivo en fase estacionaria hasta que el cultivo vuelve a alcanzar la saturación con un total de 39 puntos experimentales diferentes, situados mayoritariamente en las fases iniciales de experimento (fig 8) (Radonjic M. *et al* 2005).

En el trabajo citado, las células fueron crecidas durante 9 días, hasta una  $\text{DO}_{600}$  de 12 y, en el tiempo 0 del experimento, se diluyeron a  $\text{DO}_{600}$  0.5 en medio YPD fresco. Los autores de este estudio delimitan la fase de crecimiento exponencial durante el intervalo entre 4 y 9 horas, cuando las células ya han superado la fase de latencia posterior a la dilución, pero antes de que la velocidad de crecimiento se vea afectada por el consumo de nutrientes del medio de cultivo. Esta fase comprende 9 puntos experimentales (3.9 h, 4 h, 4.25 h, 4.75 h, 5.25 h, 5.75 h, 6.5 h, 7.25 h y 9 h) marcados en azul en la figura 8.



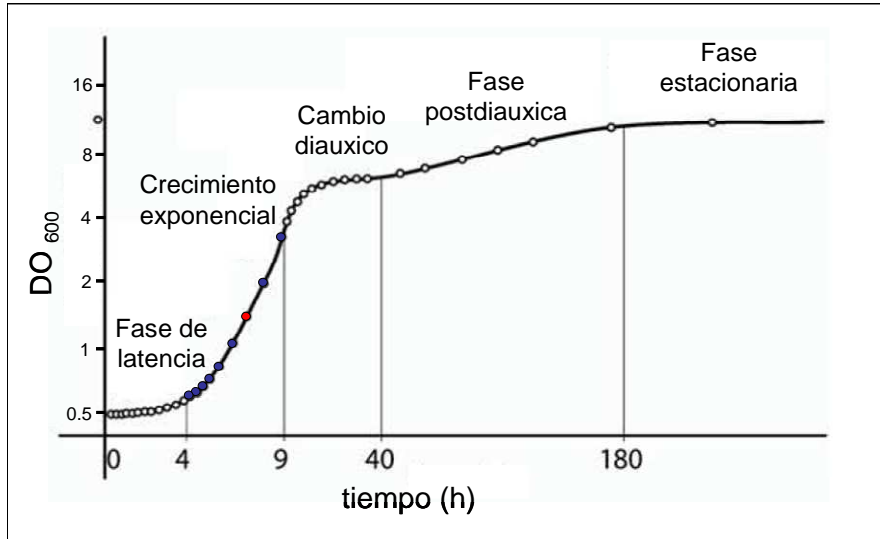


Figura 8. Cambio de la densidad óptica de un cultivo desde su dilución en medio fresco hasta la fase estacionaria. Cada círculo representa un tiempo para el que se disponen de datos genómicos de cambio en la cantidad de mRNA. Todos los cambios de cantidad de mRNA se calculan frente al tiempo 6.5 h representado por un círculo rojo. El resto de puntos reanalizados se representan en azul. Figura adaptada de Radonjic M. et al. (2005).

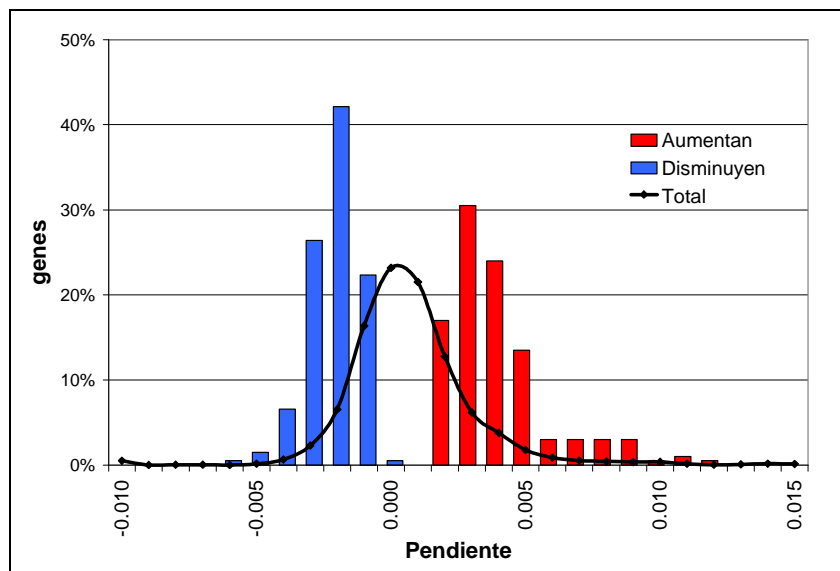


Figura 9. Cambio en la cantidad de mRNA durante el intervalo entre 3.9 y 9 horas. Se representan los histogramas de las pendientes del total de los genes (línea negra) y de los genes con pendientes significativamente negativas (en azul) o positivas (en rojo), que corresponden a un total de 197 y 200 genes respectivamente. Las pendientes están calculadas usando los datos de cantidad de mRNA en escala natural frente al tiempo en minutos.

## Resultados y discusión

Se calculó para el intervalo temporal de 3.9 a 9 h si existen genes con una pendiente significativamente diferente de 0 para la cantidad de mRNA usando la herramienta de análisis diferencial de la suite GEPAS (Montaner D. *et al* 2006). Si se fija un umbral  $FDR < 0.01$ , sólo se encuentran 197 genes con pendiente significativamente negativa y 200 genes con pendiente positiva, de un total de 4404 genes para los que se dispone de datos (fig 9).

Tabla 3. Grupos funcionales de genes con pendientes significativamente diferentes de 0 para el intervalo entre 3.9 y 9 horas. Se representan únicamente los grupos funcionales más informativos. Los valores han sido calculados usando el programa *Funcassociate*, que se basa en un test exacto de Fisher de una cola para el cálculo de los p-valores. El p-valor ajustado corresponde a la iteración de 1000 hipótesis nulas (Berriz G.F. *et al* 2003).

<b><u>Grupos funcionales con incremento significativo en la cantidad de mRNA</u></b>				
<u>nº genes</u>	<u>total</u>	<u>p-valor</u>	<u>p-adj</u>	<u>Ontología Génica</u>
153	2449	$3.50 \cdot 10^{-10}$	<0.001	citoplasma
10	28	$1.90 \cdot 10^{-7}$	<0.001	fosforilación oxidativa
60	710	$3.80 \cdot 10^{-7}$	<0.001	mitocondria
7	18	$7.60 \cdot 10^{-6}$	0.003	síntesis de ATP acoplada al transporte electrónico
8	31	$5.10 \cdot 10^{-5}$	0.014	acumulo cortical de actina
9	41	$6.90 \cdot 10^{-5}$	0.027	26S Proteasoma ( <i>sensu Eukaryota</i> )
<b><u>Grupos funcionales con disminución significativa en la cantidad de mRNA</u></b>				
<u>nº genes</u>	<u>total</u>	<u>p-valor</u>	<u>p-adj</u>	<u>Ontología Génica</u>
52	179	$3.90 \cdot 10^{-30}$	<0.001	nucleolo
51	198	$9.10 \cdot 10^{-27}$	<0.001	biogénesis y ensamblaje de ribosomas
37	132	$9.60 \cdot 10^{-21}$	<0.001	Procesado del rRNA
102	1389	$1.50 \cdot 10^{-9}$	<0.001	núcleo
11	55	$2.30 \cdot 10^{-5}$	0.007	actividad metiltransferasa
12	67	$3.20 \cdot 10^{-5}$	0.009	modificación de RNA
7	23	$4.30 \cdot 10^{-5}$	0.015	actividad RNA helicasa dependiente de ATP

Entre los genes que aumentan la cantidad de mRNA durante este periodo se pueden destacar las categorías funcionales relacionadas con la

## Resultados y discusión

fosforilación oxidativa y la mitocondria. Por otra parte, entre los genes que disminuyen su cantidad de mRNA se encuentran básicamente los relacionados con la biogénesis de ribosomas (tabla 3). Estos cambios que se observan son coherentes con lo esperado durante la fase final del crecimiento exponencial, dado que las células tras esta fase ralentizan su crecimiento, disminuyendo por tanto la producción de ribosomas. Además, las células cambian a un metabolismo energético con una respiración más activa, lo que justifica el incremento de mRNAs de genes relacionados con la mitocondria.

Sin embargo, si se repite el mismo análisis pero usando sólo los 7 primeros puntos (de 3.9 h a 6.5 h) no se detecta ningún gen con una pendiente significativamente diferente de 0. Este resultado indica que la cantidad de mRNA permanece estable durante gran parte del crecimiento exponencial para casi todos los genes hasta al menos dos horas antes de que se pueda detectar una disminución en la velocidad de crecimiento.

### **1.3 Análisis detallado de los cambios de cantidad de mRNA y tasa de transcripción durante el crecimiento exponencial.**

Una vez determinado que no existen cambios importantes en la cantidad de mRNA durante el crecimiento exponencial, se decidió realizar un estudio más detallado durante una fase inicial de esta fase de crecimiento. Para ello, se seleccionó una zona inicial del crecimiento exponencial con una densidad celular adecuada para la realización de los experimentos, pero bastante alejada del cambio diaúxico, para que éste no afecte a los resultados. En este estudio detallado se midieron simultáneamente tanto las variaciones en cantidad de mRNA como las variaciones en las tasas de transcripción mediante GRO, para poder determinar a su vez si existen variaciones de la estabilidad de los mRNAs en estas condiciones. Se partió de un cultivo en crecimiento exponencial inoculado al menos 16 horas antes, y se empezó el experimento cuando las células alcanzaron una  $DO_{600}$  de 0.36. Se tomaron un total de 5 alícuotas del mismo cultivo separadas cada una 10 minutos ( $DO_{600}$  de 0.36, 0.38, 0.41, 0.44 y 0.47). Todos los experimentos fueron realizados por triplicado e igualando lo máximo posible las densidades ópticas iniciales entre las diferentes réplicas experimentales.

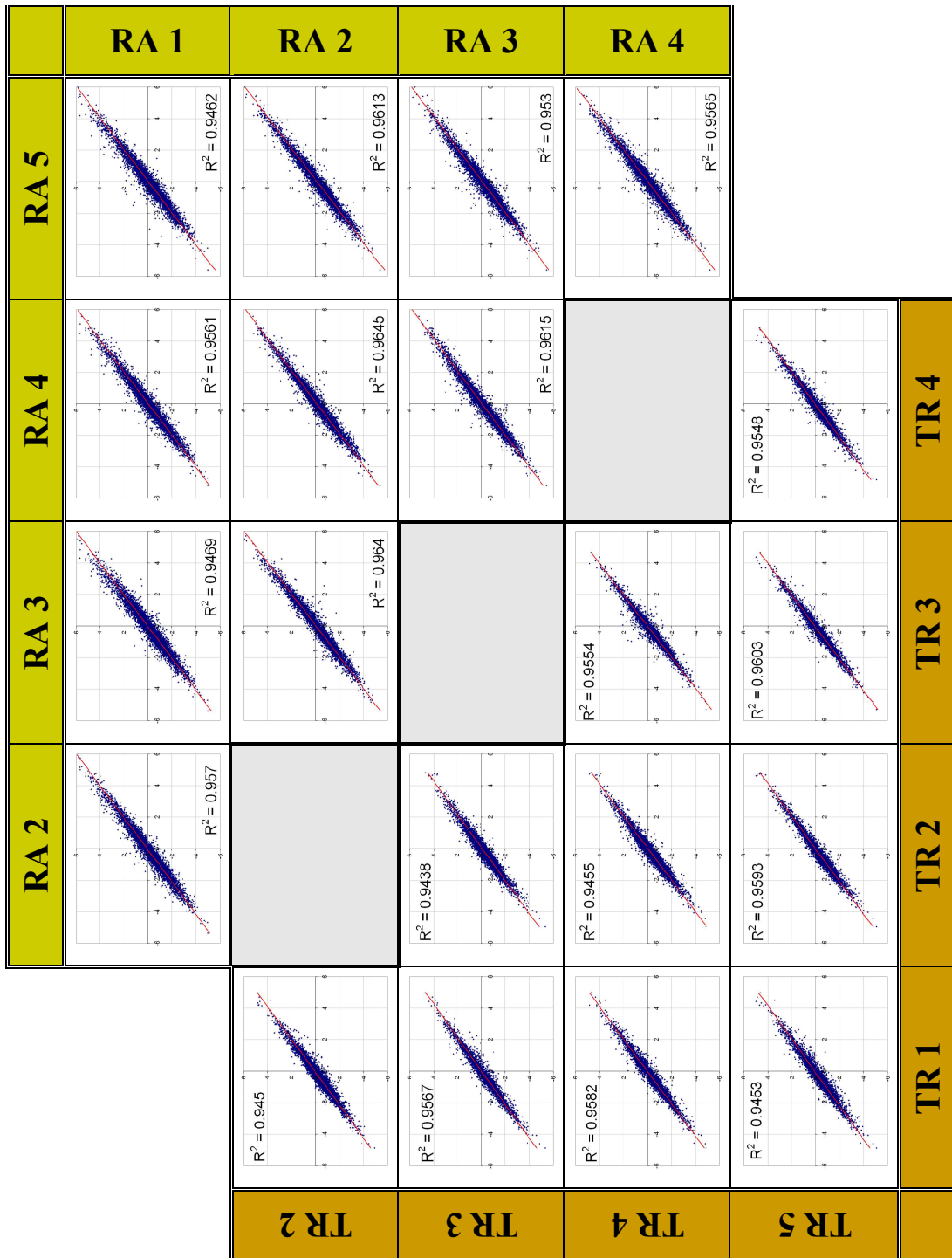


Figura 10. Correlaciones de cantidad de mRNA y tasas de transcripción entre diferentes momentos durante el crecimiento exponencial. Los cinco puntos experimentales están separados 10 minutos entre ellos, y corresponden respectivamente a  $DO_{600}$  de 0.36, 0.38, 0.41, 0.44 y 0.47. Cada punto experimental es el resultado del promedio de tres réplicas biológicas independientes. Tanto las tasas de transcripción (TR) como la cantidad de mRNAs (RA) están representadas en escala logarítmica ( $\log_2$ ) y centradas en 0.

## Resultados y discusión

Como era esperable, existe una elevada correlación tanto entre las tasas de transcripción como en la cantidad de mensajeros entre los cinco puntos experimentales. Las variabilidades de los datos de transcripción medidos por GRO y la cantidad de mRNA son similares, siendo el promedio las correlaciones entre réplicas biológicas de 0.934 para GRO y 0.936 para cantidad de mRNA. Por otra parte, la correlación entre diferentes tiempos del mismo experimento es de 0.965 para GRO y de 0.964 para el mRNA. Estos resultados indican que la variabilidad entre tiempos experimentales es menor que la producida por las réplicas biológicas, y por lo tanto apoyan la hipótesis del estado estacionario. Una vez fusionadas las réplicas entre sí, los coeficientes de correlación de Pearson al cuadrado entre los tiempos experimentales son, en todos los casos, del orden de 0.95, lo que indica que el 95% de la variabilidad para la medida de TR o RA en un tiempo concreto, se puede explicar directamente por la variabilidad observada en los otros tiempos experimentales (fig 10).

Una vez que se determinaron las tasas de transcripción y las cantidades de mRNAs para diferentes tiempos durante el crecimiento exponencial, y que se comprobó que no existen grandes variaciones entre ellas, se realizó un análisis para estudiar si estos valores presentan algún tipo de variación significativa con respecto al tiempo o a la densidad óptica del cultivo. Para ello se volvió a utilizar la herramienta de análisis de expresión génica diferencial de la suite GEPAS (Montaner D. *et al* 2006). En este caso no se encontró ningún gen con variaciones estadísticamente significativas. Por lo que todos estos resultados confirman que la asunción general de que existe un estado estacionario para la producción y degradación de mRNA durante el crecimiento exponencial es correcta. Y, por lo tanto, es correcto aplicar la ecuación [6] para el cálculo indirecto de la  $k_d$ .

### **1.4 Desviaciones del estado estacionario para la expresión génica**

Como se ha descrito anteriormente, si se realiza un análisis de expresión génica diferencial analizando individualmente los genes, no se detectan cambios significativos. Sin embargo, si en vez de analizar los genes individualmente, se usan herramientas que analizan sesgos para grupos de genes relacionados funcionalmente, sí que se observan algunos cambios. Usando la herramienta de anotación funcional Fatiscan (Al-Shahrour F. *et al* 2006) se pueden detectar sesgos significativos en las pendientes de los valores de cantidad de mRNA y de tasas de transcripción (tabla 4). El resultado más significativo es que se da una disminución tanto en la transcripción, como en la cantidad de mRNA, de los genes relacionados

con los ribosomas. Estos cambios concuerdan con los observados entre puntos experimentales más separados entre si en los datos de Radonjic M. *et al.* (2005).

Tabla 4. Resumen de las clases funcionales para las que se puede detectar un cambio en las tasas de transcripción o en la cantidad de mRNA respecto al tiempo durante el crecimiento exponencial. Los valores de significación corresponden a los p-valores ajustados por test múltiple obtenidos mediante Fatiscan. Las significaciones se calculan usando tests exactos de Fisher a lo largo de diferentes particiones de los datos con el objetivo de detectar asimetrías en su distribución (Al-Shahrour F. *et al* 2006). El p-valor ajustado se calcula utilizando una corrección por FDR (*False Discovery Rate*). Sólo se representan aquellas categorías con un p valor ajustado menor de 0.01.

<u>Grupos funcionales cuya transcripción aumenta</u>	
<u>p-adj</u>	<u>Grupo funcional</u>
0.004	Fase del ciclo celular
<u>Grupos funcionales cuya transcripción disminuye</u>	
<u>p-adj</u>	<u>Grupo funcional</u>
$2.34 \cdot 10^{-15}$	Componente estructural de ribosoma
$8.76 \cdot 10^{-15}$	Traducción
$6.7 \cdot 10^{-5}$	Biogénesis y ensamblaje de ribosomas
$3.19 \cdot 10^{-5}$	Respuesta a proteínas desplegadas
0.0012	Unión a RNA
0.0012	Unión a proteínas desplegadas
<u>Grupos funcionales cuya cantidad de mRNA disminuye</u>	
<u>p-adj</u>	<u>Grupo funcional</u>
$2.31 \cdot 10^{-28}$	Traducción
0.007	Gluconeogénesis

Usando intervalos cortos de cuatro puntos y analizando los sesgos mediante Fatiscan también se pueden encontrar resultados similares en los datos de Radonjic M. *et al.* (2005). Estas diferencias son especialmente importantes en el intervalo inicial y en el final, mientras que en la zona intermedia del crecimiento exponencial se detectan menos variaciones respecto al estado estacionario (tabla 5).

## Resultados y discusión

Tabla 5. Grupos funcionales para los que se detectan cambios en la cantidad de mRNA en el estudio de Radonjic M. et al. (2005). Los valores de significación corresponden a los p-valores ajustados por test múltiple obtenidos mediante Fatiscan (Al-Shahrour F. et al 2006). Se representan en rojo aquellos grupos con pendientes significativamente positivas (**incrementa la cantidad de mRNA**) y en azul los grupos con pendientes negativas (**disminuye la cantidad de mRNA**). Sólo se representan los grupos funcionales más informativos.

Grupo funcional	3.9 – 4.75 h	4 – 5.25 h	4.25 – 5.75 h	4.75 – 6.5 h	5.25 – 7.25 h	5.75 – 9 h
Citoplasma	4.06·10 <sup>-8</sup>	-	-	5.56·10 <sup>-3</sup>	4.26·10 <sup>-16</sup>	2.90·10 <sup>-7</sup>
Mitocondria	-	-	-	-	1.24·10 <sup>-17</sup>	2.08·10 <sup>-5</sup>
Actividad oxidoreductasa	3.72·10 <sup>-3</sup>	-	-	-	5.07·10 <sup>-5</sup>	1.82·10 <sup>-7</sup>
Procesos del metabolismo del alcohol	-	-	-	-	3.29·10 <sup>-5</sup>	4.13·10 <sup>-5</sup>
Respuesta a proteínas desplegadas	-	-	-	-	1.87·10 <sup>-11</sup>	1.06·10 <sup>-3</sup>
Respuesta al estrés	3.06·10 <sup>-3</sup>	-	-	-	1.20·10 <sup>-5</sup>	8.17·10 <sup>-3</sup>
Integral de membrana	6.92·10 <sup>-3</sup>	4.07·10 <sup>-6</sup>	9.44·10 <sup>-5</sup>	-	7.58·10 <sup>-5</sup>	-
Ensamblaje y biogénesis de ribosomas	6.04·10 <sup>-5</sup>	-	-	-	6.32·10 <sup>-19</sup>	-
Componente estructural del ribosoma	-	-	-	-	4.85·10 <sup>-13</sup>	-
Traducción	-	-	-	-	7.22·10 <sup>-8</sup>	8.16·10 <sup>-3</sup>
Núcleo	1.30·10 <sup>-6</sup>	-	-	-	-	-

La mayoría de las desviaciones respecto al estado estacionario que se observan en la tabla 4 son comunes tanto para las tasas de transcripción como para la cantidad de mRNA. Sin embargo existen ciertas discrepancias que deben ser explicadas por variaciones en la estabilidad de los mRNAs. Para calcular la influencia en la cantidad de mRNA de los cambios en la estabilidad se calculó la cantidad de mRNA teórica que debería tener cada gen usando únicamente los datos de variación de tasa de transcripción y asumiendo una estabilidad constante a lo largo del experimento. De esta forma, la diferencia entre la cantidad de mRNA teórica asumiendo una estabilidad constante y la cantidad real, se deberá



exclusivamente a variaciones en la estabilidad de los mRNAs durante el crecimiento exponencial (fig 11).

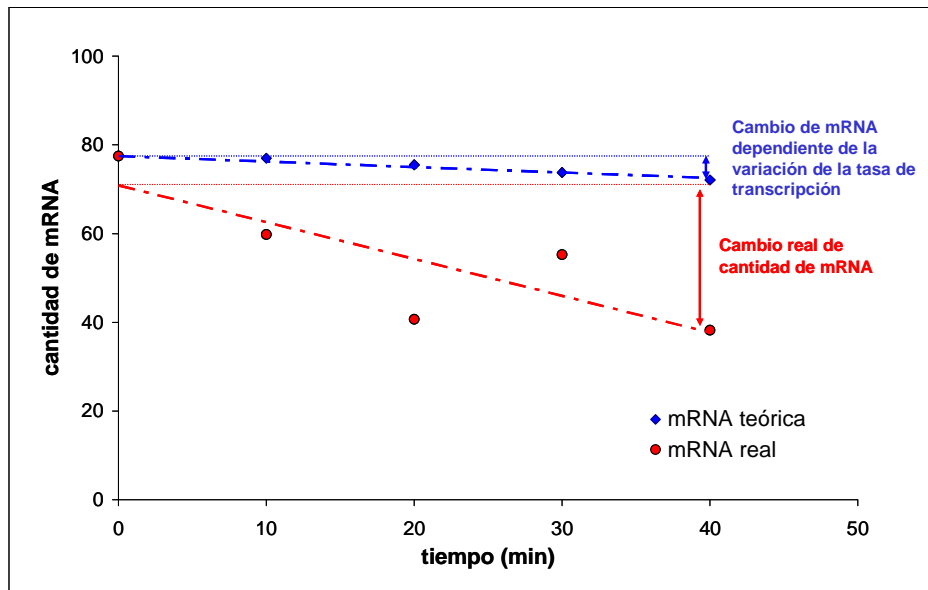


Figura 11. Cálculo de la dependencia de la variación de mRNA respecto a la transcripción y la estabilidad. Los puntos **rojos** representan datos reales de cantidad de mRNA para un gen concreto, mientras que los **azules** representan la variación teórica que presentaría ese gen si la estabilidad del mRNA fuera constante y las variaciones sólo se debieran a cambios en la tasa de transcripción. La diferencia entre el cambio real de la cantidad de mRNA y el cambio dependiente de transcripción es el cambio dependiente de variaciones en la estabilidad e los mRNAs.

Se calculó qué porcentaje del cambio total en la cantidad de mRNA para cada gen se debe a cambios en la estabilidad, y se analizó si existen diferencias significativas entre grupos funcionales. Para ello se ordenó jerárquicamente el porcentaje de variación debido a la estabilidad para cada gen y se analizaron mediante la herramienta Fatiscan. Los resultados indican principalmente una desestabilización de mRNAs relacionados con los ribosomas y una estabilización de genes relacionados con la mitocondria (tabla 6). Además, estos cambios van en el mismo sentido (cambio homodireccional), incremento o descenso, que los observados previamente en la cantidad de mRNA.

## Resultados y discusión

Tabla 6. Grupos funcionales con cambios significativos para la estabilidad de los mRNAs. Entre paréntesis se muestra el p-valor ajustado por test múltiple calculado mediante Fatican (Al-Shahrour F. *et al* 2006) a partir de la ordenación jerárquica del porcentaje de variación de los mRNAs que se debe a cambios en la estabilidad. Sólo se muestran aquellas categorías con un p-adj<0.01.

<u>Grupos funcionales cuyos mRNAs se estabilizan</u>	
<u>p-adj</u>	<u>Ontología Génica</u>
9·10 <sup>-5</sup>	Procesos del metabolismo de biopolímeros
0.004	Mitocondria
0.009	Regulación de procesos celulares

<u>Grupos funcionales cuyos mRNAs se desestabilizan</u>	
<u>p-adj</u>	<u>Ontología Génica</u>
0.004	Citosol
0.004	Subunidad grande del ribosoma
0.01	Proceso de biosíntesis celular

Una vez calculada la variación de los tres parámetros que afectan a la cantidad de mRNA durante el crecimiento exponencial (transcripción, cantidad de mRNA y estabilidad), se seleccionaron aquellos grupos que presentan variaciones significativas en alguno de los parámetros para estudiarlos en detalle. En la fig 12 se representan la aportación de los cambios en transcripción o en estabilidad al cambio total en la cantidad de mRNA. Para estos grupos funcionales, las variaciones en la cantidad de mRNA se deben normalmente al efecto conjunto de cambios en la tasa de transcripción y en la estabilidad que actúan de forma coordinada (cambios homodireccionales). Sin embargo, también se pueden encontrar casos donde es más importante uno de los dos factores, como el caso de los genes relacionados con la gluconeogénesis donde la mayoría de la disminución de la cantidad de mRNA se debe a una disminución de la vida media. O incluso algunos casos donde las dos magnitudes parecen comportarse de manera opuesta como el caso de los genes de biogénesis y ensamblaje de ribosomas o de unión a proteínas desplegadas (cambios heterodireccionales).

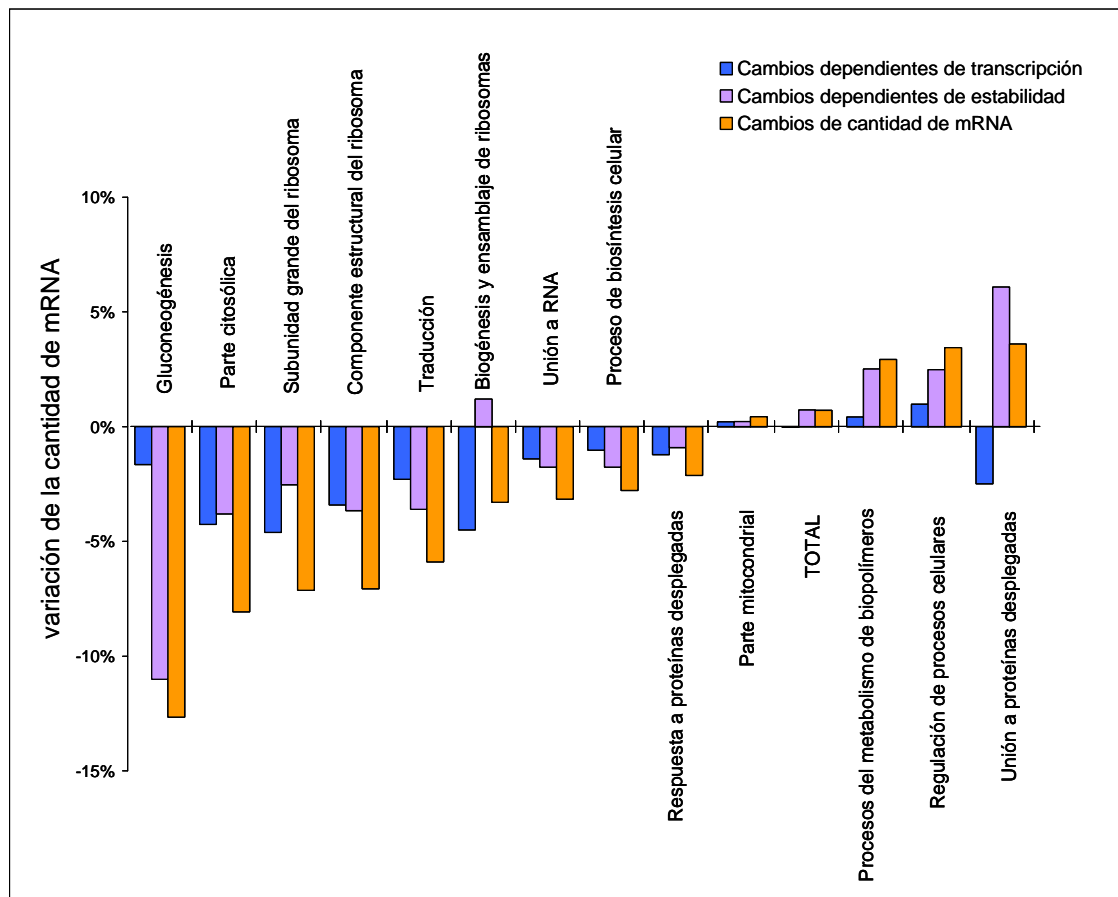


Figura 12. Contribución de la transcripción y de la estabilidad a la variación de la cantidad de mRNA. Se representan las medianas para la variación global en la cantidad de mRNA de unos pocos grupos funcionales para los que se detectan desviaciones significativas del estado estacionario para la expresión génica.

## 1.5 Asunción del estado estacionario para la expresión génica.

Se puede, pues, asumir que para la mayor parte de los genes la cantidad de mRNA y la tasa de transcripción son constantes durante el crecimiento exponencial. Y por lo tanto, también es correcto asumir también que no hay variaciones importantes de la estabilidad de los mRNAs. En este trabajo se ha estudiado con detalle una situación de crecimiento exponencial bastante alejada temporalmente del cambio diáuxico. Sin embargo, teniendo en cuenta el análisis de los datos de Radonjic M. *et al.* (2005), esta situación de expresión génica estacionaria se puede extrapolar a otros momentos del crecimiento exponencial, hasta al menos, un par de horas antes de que las células empiecen a disminuir su velocidad de crecimiento.

## Resultados y discusión

El hecho de que sea correcto realizar una asunción de estadio estacionario para la expresión génica, permite usar los datos de tasa de transcripción y de cantidad de mRNA para calcular una vida media indirecta durante el crecimiento exponencial. Esto confirma la validez del uso del GRO para calcular las estabildades de los mRNAs durante el crecimiento exponencial así como en otras situaciones donde se pueda asumir el estado estacionario para la expresión génica (García-Martínez J. *et al* 2004).

Sin embargo, aunque ningún gen considerado individualmente se aparte del estado estacionario, se pueden detectar algunos grupos funcionales que presentan variaciones ligeras pero significativas respecto al estado estacionario. Y estas variaciones son consecuencia de cambios tanto en las tasas de transcripción como en la estabilidad de los mensajeros. Lo que apoya la importancia de la regulación postraduccional de la estabilidad de los mRNAs como un mecanismo fundamental de la regulación génica (Keene J.D. y Lager, P.J. 2005; Pérez-Ortín J.E. *et al* 2007). Además, esas variaciones parecen anticipar cambios importantes en el metabolismo celular que serán más importantes a medida que progrese el crecimiento celular. Estos cambios empiezan a preparar a la célula para una menor velocidad de crecimiento y un metabolismo menos fermentativo y más respiratorio. Sin embargo, es importante destacar que estas variaciones, pese a existir, son muy ligeras, y sólo se pueden detectar cuando se estudian conjuntamente grupos de genes relacionados.

También es interesante señalar que, para la mayor parte de los genes y grupos funcionales que presentan cambios en la cantidad de mRNA, los cambios de tasa de transcripción van en el mismo sentido. Además estos cambios de transcripción suelen ir acompañados de cambios de estabilidad del mRNA que refuerzan su efecto. Sólo en dos grupos se observan cambios de transcripción y estabilidad de mRNA en sentido opuesto. El grupo de “unión a proteínas desplegadas” exhibe un cambio de transcripción negativo, mientras que el cambio de la cantidad de mRNA es positivo gracias a un gran incremento de la estabilidad. Por el contrario, el grupo de “biogénesis de ribosomas” presenta un cambio mayor de transcripción que de mRNA para compensar una pérdida de estabilidad de los mRNAs. Estos resultados apoyan la idea de que en circunstancias de estabilidad ambiental las variaciones de cantidad de RNA se consiguen mediante cambios tanto en la tasa de transcripción como en la estabilidad de los mRNAs que actúan en el mismo sentido (homodireccionales). Sólo en casos excepcionales los efectos son opuestos. Es importante destacar que, en otras situaciones, cuando se producen cambios bruscos en la situación fisiológica se pueden encontrar más frecuentemente grupos de

genes que presenten comportamientos opuestos para la transcripción y la degradación de mRNAs en un momento determinado. Esto ha sido descrito recientemente usando también el método GRO durante el estrés oxidativo (Molina-Navarro M.M. *et al* 2008) y el estrés osmótico (Lorena Romero-Santacreu y Paula Alepuz, comunicación personal).

### **1.6 Discrepancias entre cálculos directos e indirectos para el estudio del recambio de los mRNAs**

Una vez comprobada la validez de la asunción del estado estacionario para la expresión génica durante el crecimiento exponencial, se puede tratar de relacionar entre sí los datos de transcripción, cantidad y estabilidad de los mRNA mediante la ecuación [6] ( $TR = k_d \cdot [mRNA]$ ). De esta forma se pueden calcular, por ejemplo, datos de TR a partir de la cantidad de mRNA y la vida media para compararlos con los datos directos de transcripción. La similitud entre los dos tipos de cálculos, experimentales o directos frente a los matemáticos o indirectos, nos darán idea de la bondad de ambas mediciones. Si ambos tipos de datos ofrecen resultados similares se reforzará la validez de los dos tipos de medidas. Si por el contrario los resultados son muy diferentes eso será indicativo de que uno o los dos tipos de los resultados tendrán errores o sesgos que hay que corregir. En este último caso el estudio de las diferencias entre las medidas directas e indirectas puede permitir descubrir grupos de genes para los que se obtengan resultados significativamente diferentes y tratar de encontrar el motivo de esta discrepancia.

Para tratar de descubrir si existen diferencias entre estos dos tipos de medidas se realizó una comparación de los datos directos e indirectos de tasas de transcripción. Otra posibilidad sería obtener datos indirectos de estabilidades de mRNA a partir de datos de TR y cantidad de mRNA y compararlos con datos de directos estabilidades. Sin embargo esta posibilidad, y los problemas asociados a estos cálculos, serán discutidos con más detalle en el segundo capítulo de este trabajo.

Para realizar las comparaciones de tasa de transcripción se utilizaron datos robustos de TR (calculados en el apartado 3.1.2 de este trabajo) y un conjunto de diferentes de datos indirectos de transcripción. Para obtener estas medidas indirectas de transcripción se combinaron los datos de cantidad de mRNA calculados a partir de los resultados de PCR competitiva y RNA-seq (Miura F. *et al* 2008; Nagalakshmi U. *et al* 2008) con 5 cálculos diferentes de estabilidades de mRNA usando el mutante termosensible *rpb1-1<sup>ts</sup>* (Grigull J. *et al* 2004; Wang Y. *et al* 2002) o drogas

## Resultados y discusión

inhibidoras de la transcripción como la tiolutina y la fenantrolina (Grigull J. *et al* 2004). De esta forma se obtuvieron 10 tablas de datos diferentes de TR indirecta usando la ecuación [6] ( $TR = k_d \cdot [mRNA]$ ).

Una vez obtenidos datos más robustos para la transcripción calculados de forma directa o indirecta, se compararon entre sí. Aunque la correlación entre los diferentes datos es relativamente alta (tabla 7) existen importantes diferencias. Para tratar de descubrir si existe algún grupo de genes para el que estas diferencias se especialmente importantes, se realizó un análisis funcional de las diferencias. Sin embargo, para evitar posibles artefactos debidos a los diferentes métodos de medida de estabildades de mRNA sólo se eligieron grupos de genes que fueran sistemáticamente diferentes entre las medidas directas e indirectas de TR para todos los cálculos indirectos diferentes, y que, por lo tanto, no dependieran exclusivamente de unos datos de estabilidad de mRNAs concretos.

Tabla 7. Correlaciones entre diferentes cálculos indirectos de TR respecto a datos. Los valores representan los coeficientes de correlación no paramétricos de Spearman. Todos los valores son estadísticamente diferentes de 0 ( $\text{sig} < 10^{-100}$ ). Los datos utilizados para realizar los cálculos indirectos proceden de cálculos de vida media (Grigull J. *et al* 2004; Wang Y. *et al* 2002) y cantidades de mRNA (Miura F. *et al* 2008; Nagalakshmi U. *et al* 2008). Todos estos datos se comparan con los datos robustos de GRO calculados en el apartado 3.1.2 de este trabajo.

Correlación respecto a GRO	[mRNA] PCR competitiva (Miura <i>et al</i> 2008)	[mRNA] RNA-Seq (Nagalakshmi <i>et al</i> 2008)
rpb1-1 con hexameros (Wang)	0.551	0.567
rpb1-1 con oligo (dT) (Wang)	0.534	0.557
rpb1-1(Grigull)	0.490	0.521
Tiolutina (Grigull)	0.519	0.526
Fenantrolina (Grigull)	0.492	0.516

Dado que en una situación de crecimiento exponencial es correcto asumir un estado estacionario (como se ha demostrado en este capítulo), las discrepancias entre las medidas directas e indirectas se deberán a asunciones incorrectas realizadas durante el cálculo de algunas de las magnitudes implicadas o bien en la relación entre ellas.

Entre las posibles asunciones incorrectas que podrían explicar las discrepancias entre medidas directas e indirectas podemos destacar:

- Que no sea correcta alguna de las asunciones realizadas para calcular las tasas de transcripción. Así, por ejemplo para transformar los datos de densidad de polimerasas activas a TR es necesario asumir que la velocidad de la RNA polimerasa II es constante y común para todos los genes. Esta asunción es obviamente una simplificación en la que los genes con diferente velocidad para la RNA pol II tendrán también valores diferentes de TR calculada directamente, pero no para la TR calculada indirectamente.
- Que no sea correcta alguna de las asunciones realizadas para calcular la estabilidad de los mRNAs. Por ejemplo en el caso de genes cuya estabilidad sea afectada de forma diferencial por las condiciones de estrés necesarias para producir la parada transcripcional y para los que, consecuentemente, se calcule una vida media más larga o más corta de la real. En este caso la TR calculada indirectamente será también diferentes de la calculada directamente.
- Otra posibilidad sería que exista cualquier asunción errónea en la relación entre las tres magnitudes (TR, RA y estabilidad de mRNAs). Así, por ejemplo, la transcripción *nuclear* que se mide mediante GRO (que hace referencia la cantidad de mRNAs nacientes) no corresponde exactamente a la velocidad de aparición de nuevos transcritos maduros en el citoplasma. Estas dos medidas sólo serán iguales en el caso que la eficiencia de la maduración y exporte de los mRNAs sea el mismo para todos los genes. Sin embargo, todos los eventos de transcripción no dan lugar a un mRNA maduro, sino que también se producen transcripciones de RNAs cortos crípticos (Struhl K. 2007; Xu Z. *et al* 2009) que darán señal de transcripción por métodos directos pero en realidad no producirán un incremento en la cantidad de mRNAs maduros.

Entre los grupos de genes con valores más altos para la TR calculada de forma indirecta respecto a la calculada de forma directa se pueden encontrar principalmente genes relacionados con la glicosilación de proteínas (mínimo p-adj  $10^{-2}$ ), retículo endoplasmático (mínimo p-adj  $10^{-5}$ ) y proteínas intrínsecas de membrana (mínimo p-adj  $10^{-2}$ ). Estas diferencias podrían explicarse por ser genes con RNA pol II más veloces, por tener mRNAs que se desestabilizan especialmente en condiciones de inhibición transcripcional o por ser genes que sufren menos transcripciones no productivas que el global de la población.

## Resultados y discusión

Por el contrario entre los grupos de genes con valores más altos para la TR calculada de forma indirecta respecto a la calculada de forma directa se pueden encontrar genes relacionados con el ciclo celular (mínimo p-adj  $10^{-3}$ ) y la regulación de la transcripción (mínimo p-adj  $10^{-3}$ ). Estas diferencias podrían explicarse por ser genes con RNA pol II menos veloces, por tener mRNAs que se estabilizan especialmente en condiciones de inhibición transcripcional o ser genes que sufren mayor número de transcripciones no productivas que el global de la población.

Sin embargo, estos resultados son sólo indicios que pueden dar pistas sobre genes con comportamientos diferenciales para los procesos de transcripción o degradación de los mRNAs.



## **2 Cálculo Genómico de estabilidades de mRNA**

Uno de los objetivos de este trabajo es realizar un cálculo de las estabilidades de los mRNAs de *S. cerevisiae* a escala genómica. Al inicio de este trabajo sólo existía un estudio genómico de estabilidades de mRNAs en el que usaban un mutante termosensible de la polimerasa para parar la transcripción (Wang Y. *et al* 2002). Sin embargo, durante la realización de este trabajo han surgido más estudios genómicos de estabilidades de mRNAs (Grigull J. *et al* 2004; Guan Q. *et al* 2006). El realizar un cálculo alternativo de estabilidades de mRNAs con nuestros macrochips permitirá estudiar la fiabilidad y validez de este tipo de medidas genómicas y compararlas también con los datos de estabilidad calculados indirectamente mediante GRO (García-Martínez J. *et al* 2004) obtenidos en el apartado anterior.

Para realizar este cálculo alternativo de la estabilidad de los mRNAs a escala genómica se decidió usar tiolutina, una droga inhibidora de la transcripción. La tiolutina es un antibiótico producido por *Streptomyces luteoreticuli* (Celmer W.D. y Solomons I.A. 1955) que es capaz de inhibir la transcripción de todas las RNA polimerasas de *S. cerevisiae* (Jiménez A. *et al* 1973). La inhibición de la transcripción se produce principalmente en la iniciación, dado que la droga interacciona principalmente con la RNA pol libre y no con la RNA pol que ya ha iniciado la transcripción (Tipper D.J. 1973), aunque otros autores han descrito que también es capaz de inhibir la elongación de la transcripción en *E. coli* (Khachatourians G.G. y Tipper, D.J. 1974). Además, esta droga ha sido ampliamente usada para el cálculo de estabilidades de mRNA en levadura (Grigull J. *et al* 2004; Guan Q. *et al* 2006; Herrick D. *et al* 1990; Michan C. *et al* 2005; Parker R. *et al* 1991).

### **2.1 Medida de vidas medias individuales con tiolutina**

Debido a la discrepancia entre las concentraciones de tiolutina usadas en diferentes estudios, se decidió buscar que concentración de tiolutina era la óptima para la medida de estabilidades de mRNAs en nuestras condiciones experimentales. Este estudio inicial se realizó calculando las estabilidades para unos pocos mRNAs individuales. Para ello, se añadió la droga a un cultivo de levaduras en crecimiento exponencial y se recogieron alícuotas de células a diferentes tiempos tras la inhibición transcripcional. Seguidamente, se extrajo el RNA y se cuantificó la desaparición de los mRNAs de interés mediante *northern blot*. Para poder normalizar las diferentes muestras de RNA entre sí y poder calcular la vida media de los

## Resultados y discusión

mRNAs, se asumió que la concentración del rRNA permanece constante a lo largo del experimento, dado que este RNA es mucho más estable que los mRNAs. De esta manera es posible cuantificar la desaparición de los mRNAs y por tanto calcular su estabilidad (fig 13).

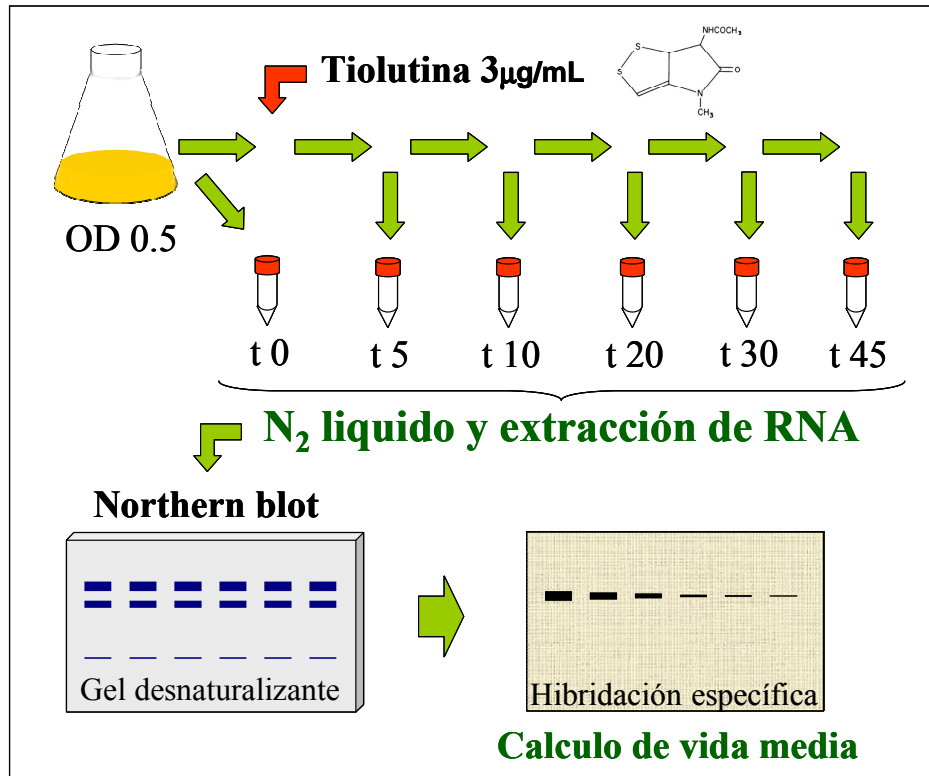


Figura 13. Diseño experimental para el cálculo de vidas medias usando tiolutina. Se partió de un cultivo de levaduras en crecimiento exponencial al que se le añadieron 3 µg/mL de tiolutina. Posteriormente se recogieron células a diferentes tiempos tras la adición de la droga, se congelaron las células en N<sub>2</sub> líquido y se extrajo el RNA. Posteriormente se realiza un *northern blot* hibridando la membrana con una sonda específica para el mRNA de interés y se calcula su velocidad de desaparición.

La concentración de tiolutina más habitualmente usada para el cálculo de estabilidades de mRNAs es de 3 µg/mL (Grigull J. *et al* 2004; Herrick D. *et al* 1990) pero también se ha descrito el uso de concentraciones más altas (Guan Q. *et al* 2006; Michan C. *et al* 2005). En las condiciones experimentales de este estudio se comprobó que a partir de 3 µg/mL de tiolutina se produce una parada de varias horas en el crecimiento celular. Para confirmar si esta concentración de tiolutina era la óptima para producir una parada en la transcripción y poder calcular así la estabilidad de los mRNAs, se realizaron diferentes cinéticas de cálculo de vidas medias a concentraciones crecientes del inhibidor. Se usaron concentraciones de tiolutina entre 3 y 30 µg/mL, y se calculó la vida media

para los mRNAs de *ACT1*, para el que está descrita una vida media de entre 31-71 minutos, y *RPL24*, que tiene una vida media descrita de entre 11-24 min (Herrick D. *et al* 1990; Santiago T.C. *et al* 1986; Wang Y. *et al* 2002). Usando 3  $\mu\text{g/mL}$  de tiolutina se calculó una vida media de 50 minutos para el mRNA de *ACT1* y de 20.3 minutos para el de *RPL24*, siendo estas vidas medias coherentes con las presentes en la bibliografía. Por el contrario, usando concentraciones superiores de tiolutina la vida media calculada se incrementaba significativamente.

Si, como es esperable, la parada de la transcripción por parte de la tiolutina fuera menos eficaz a concentraciones más bajas de la droga, un incremento en la concentración de la droga conduciría, en todo caso, a una parada más eficaz de la transcripción y, por lo tanto, un incremento en la concentración de tiolutina usada produciría un cálculo de vidas medias más cortas, dado que se reduciría la transcripción residual que podría quedar a concentraciones bajas de droga. Sin embargo, contrariamente a lo esperado, el incremento de la concentración de tiolutina usada para parar la transcripción conduce a la aparición de vidas medias aparentemente más largas (fig 14). La forma más fácil para explicar este fenómeno es que la tiolutina, además de su efecto inhibitorio sobre la transcripción, tiene también un efecto sobre el proceso celular opuesto: degradación de los mRNAs.

## Resultados y discusión

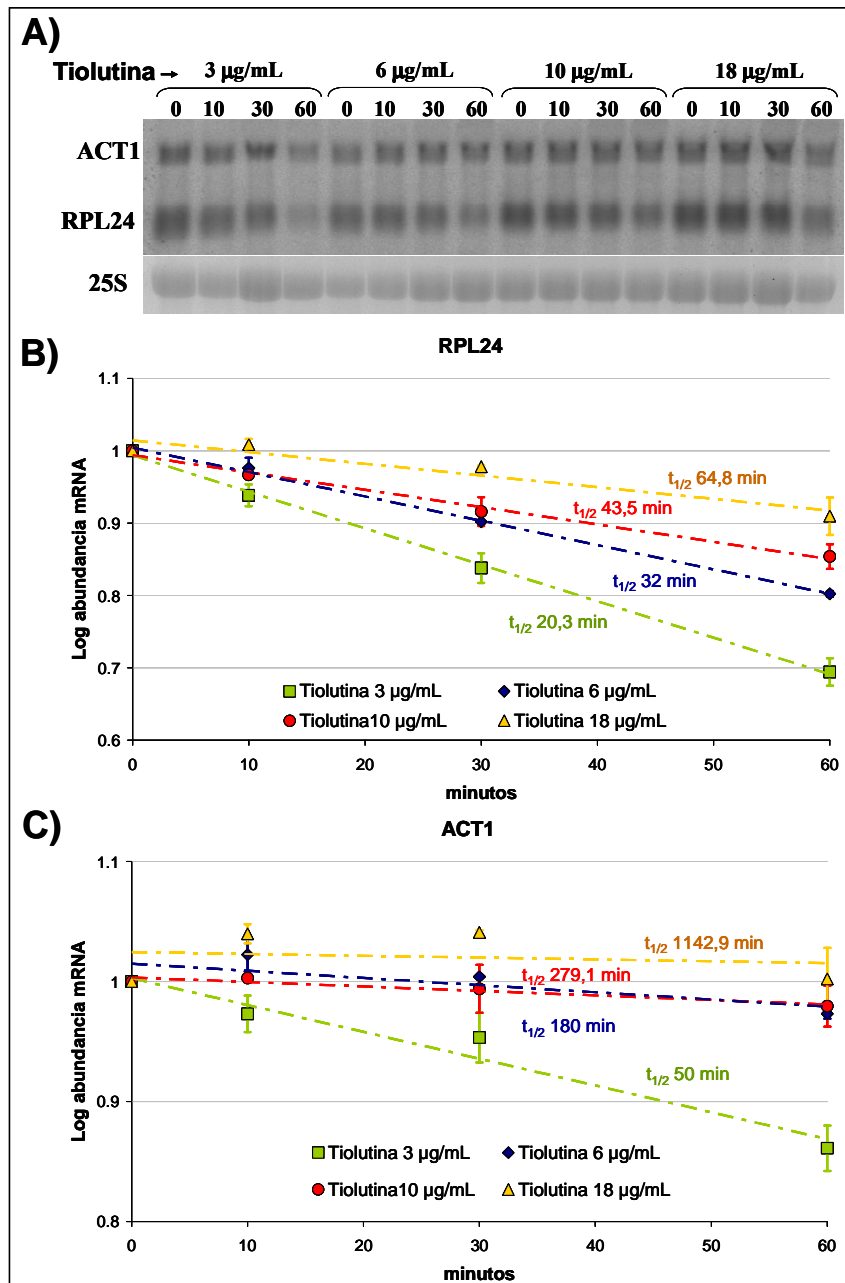


Figura 14. Cálculo de vidas medias de mRNAs individuales usando concentraciones crecientes de tiolutina. (A) Northern blot de muestras obtenidas a diferentes tiempos tras la adición de concentraciones crecientes de tiolutina hibridado con sondas para los mensajeros de *ACT1* y *RPL24*. (B) Calculo de la vida media aparente del mRNA de *RPL24* para las diferentes concentraciones de tiolutina usadas en (A). (C) Calculo de la vida media aparente del mRNA de *ACT1* para las diferentes concentraciones de tiolutina usadas en (A). Los experimentos se realizaron por duplicado y las barras de error representan la desviación estándar entre réplicas. Todas las muestras han sido normalizadas usando la señal de 25S rRNA teñido con bromuro de etídio.

## 2.2 Comprobación del efecto inhibitorio de la tiolutina sobre la degradación de los mRNAs usando mutante *rpb1-1*

Para comprobar la hipótesis de que la tiolutina, además de su efecto inhibitorio sobre la transcripción, presenta un efecto inhibitorio sobre la degradación de los mRNAs, se diseñó un experimento en el que se produce una parada general de la transcripción de forma independiente al uso de la tiolutina. Una vez producida la parada de la transcripción, la adición de la tiolutina permite estudiar si ésta tiene algún efecto adicional sobre la degradación del mRNA.

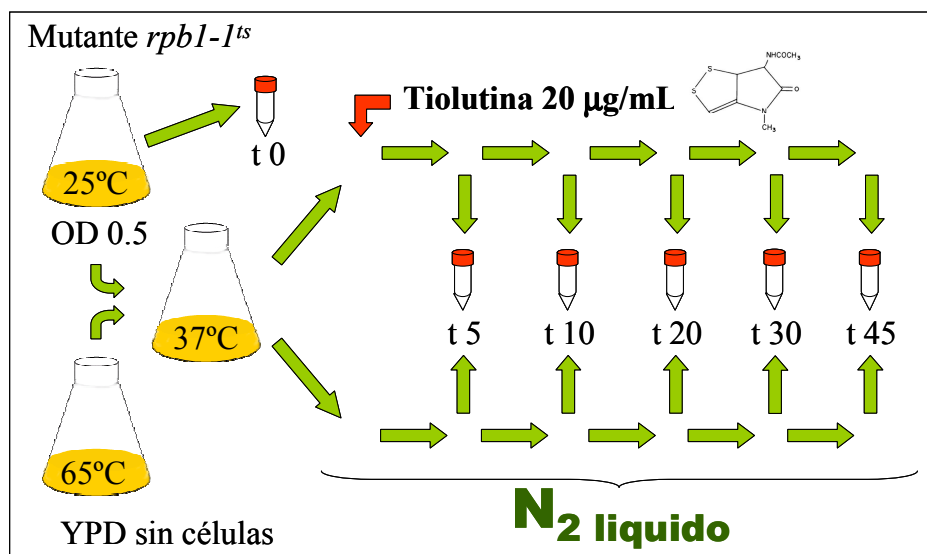


Figura 15. Diseño experimental del estudio del efecto inhibitorio de la tiolutina sobre la degradación de los mRNAs. Se partió de un cultivo de levaduras en crecimiento exponencial con una RNA pol II termosensible y se paró la transcripción mediante un choque térmico. A una alícuota se le añadió tiolutina hasta 20 µg/mL para observar si esta tiene un efecto sobre la cantidad de mRNA independiente a la parada de transcripción. Posteriormente se recogieron células a diferentes tiempos congelándolas en N<sub>2</sub> líquido y se extrajo el RNA para ser analizado mediante *northern blot*.

Para realizar este experimento se usó una cepa (Z676) que posee una RNA pol II termosensible (*rpb1-1<sup>ts</sup>*) y que también es habitualmente usada para el cálculo de vidas medias de mRNAs (Grigull J. *et al* 2004; Parker R. *et al* 1991; Wang Y. *et al* 2002). Se crecieron las células a temperatura permisiva (25°C) y posteriormente se aumentó la temperatura del cultivo rápidamente a 37°C mediante la adición de medio caliente (fig 15). De esta forma, es el choque térmico el que detiene la transcripción y permite, recogiendo células a diferentes tiempos, calcular la vida media de los

## Resultados y discusión

mRNAs de la misma manera que se hacía en el caso de la tiolutina. Una vez producida una parada en la transcripción de forma independiente a la tiolutina, se separó el cultivo en dos alícuotas y a una de ellas se le añadió tiolutina hasta una elevada concentración (20  $\mu\text{g}/\text{mL}$ ) para maximizar los efectos secundarios de la droga. Posteriormente, se mantuvieron las dos alícuotas a 37°C y se recogieron células a diferentes tiempos para calcular las vidas medias de los mRNAs en presencia o ausencia de una concentración elevada de tiolutina (fig 15).

Usando exclusivamente la parada de transcripción producida por el choque térmico en una cepa *rpb1-1<sup>ts</sup>* se calculó una vida media de 31.1 minutos para el mRNA de *ACT1* y de 14.3 minutos para el de *RPL24*. Sin embargo, cuando además del choque térmico se añade tiolutina a 20  $\mu\text{g}/\text{mL}$  se produce una gran estabilización de estos mRNAs (fig 16). Como la transcripción está inhibida debido al uso de la polimerasa termosensible en ambos casos, es la tiolutina la responsable de estabilizar los mRNAs, y por lo tanto posee un efecto inhibitorio sobre la degradación de los mRNAs de forma independiente a la inhibición general de la transcripción.

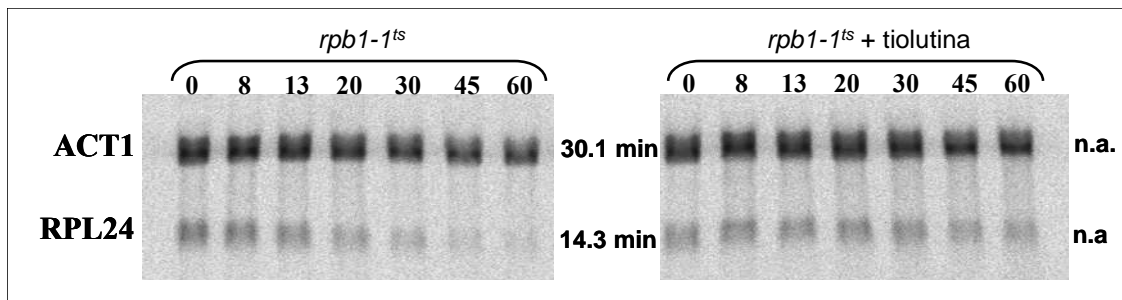


Figura 16. Cálculo de vidas medias usando un mutante termosensible de la RNA pol II en presencia y ausencia de tiolutina. Northern blot de muestras a diferentes tiempos tras un choque térmico a 37°C en una cepa *rpb1-1<sup>ts</sup>* con o sin la adición de tiolutina 20  $\mu\text{g}/\text{mL}$ . La vida media calculada se muestra a la derecha de cada serie. n.a. indica que la cantidad de mRNA aumenta durante el experimento, y por lo tanto el cálculo de una vida media no es aplicable. Todas las muestras han sido normalizadas usando la señal de 25S rRNA teñido con bromuro de etidio.

### 2.3 Estudio del efecto de la tiolutina sobre la desadenilación de los mRNAs

Una vez demostrado que la tiolutina posee un efecto inhibitorio sobre la degradación de los mRNAs, se decidió tratar de averiguar a qué nivel se produce esta inhibición. El primer paso en la vía de degradación de los

mRNAs es el recorte de la longitud de de la cola poli (A). Además, la desadenilación es un paso limitante para la degradación y es común para las diferentes rutas de degradación, ya que, tras el recorte de la cola poli (A), los mRNAs pueden seguir su degradación por diferentes vías (Parker R. y Song,H. 2004). Para estudiar si el efecto de la tiolutina sobre la inhibición de la degradación de mRNAs es anterior o posterior a este proceso, se realizó una cuantificación de la proporción de RNA poli (A) global a diferentes tiempos tras la adición de la droga.

Para ello, se realizó un *dot-blot* depositando cantidades de RNA total conocidas extraídas de células a diferentes tiempos tras la adición de concentraciones crecientes de tiolutina y se hibridó con una sonda poli (dT) marcada radiactivamente. De esta manera, la intensidad de cada punto es proporcional a la cantidad de RNA poliadenilado y, dado que la cantidad total de RNA depositado es conocida, es posible calcular la proporción de RNA poliadenilado presente en cada muestra.

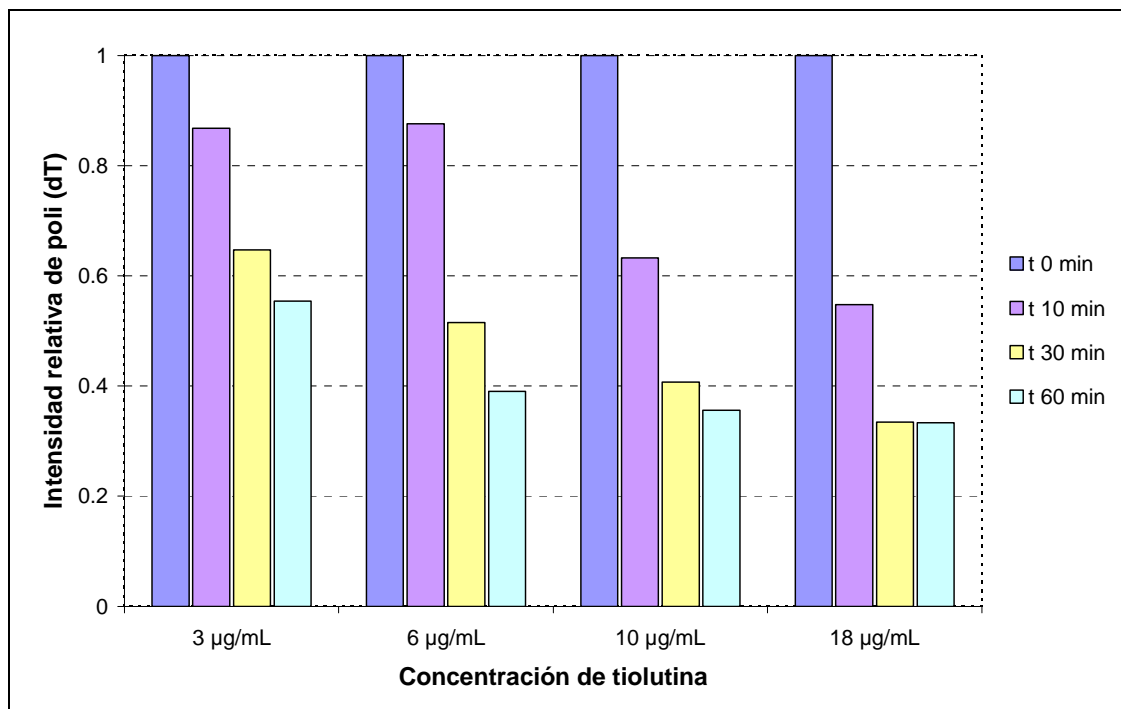


Figura 17. Estudio del efecto de la tiolutina sobre la desadenilación de los mRNAs. Cuantificación de la cantidad de poli (A) total después de la adición de diferentes concentraciones de tiolutina. Los datos corresponden a la intensidad de los puntos de un dot blot realizado con RNA total extraído tras la adición de tiolutina e hibridado con un oligo poli (dT) marcado terminalmente con  $^{32}\text{P}$ . Las intensidades de los tiempos 0 (antes de la adición de la tiolutina) se han tomado arbitrariamente como uno.

## Resultados y discusión

Cuando se añade tiolutina se puede observar una disminución global de la población de RNAs poliadenilados que concuerda con una inhibición de la transcripción (fig 17). Como ha sido previamente descrito (Herrick D. *et al* 1990; Minvielle-Sebastia L. *et al* 1991), la disminución de señal de RNA poliadenilado tras la inhibición de la transcripción por tiolutina se debe tanto a un recorte en la longitud global de las colas poli (A) como de la cantidad de RNAs poliadenilados. Sin embargo, el incremento en la concentración de la tiolutina usada no conduce a una estabilización del RNA poliadenilado. De hecho, se produce una degradación sensiblemente más rápida a concentraciones superiores. Este resultado sugiere que a concentraciones superiores la inhibición de la transcripción es más eficiente. Por lo tanto, la estabilización de los mRNAs que se observan mediante *northern* blot con concentraciones altas de tiolutina se debe producir afectando a la estabilidad de los mRNAs en una etapa posterior al recorte de la cola poli (A).

Cabe destacar, que la cantidad de RNA poliadenilado parece estabilizarse en un valor mínimo. Esto puede deberse a que, durante el proceso inicial de desadenilación, no se produce una eliminación completa de la cola poli (A) y, por lo tanto, los mRNAs en degradación presentan colas cortas que podrían seguir dando alguna señal. Otra posibilidad no excluyente con la anterior es que un pequeño porcentaje del RNA ribosómico se poliadenila (Kuai L. *et al* 2004). El pequeño porcentaje de poliadenilación de rRNA (entre el 0.1 y el 0.01%) puede ser suficiente para producir una señal significativa, dado que el rRNA representa el 80% del RNA total (frente al 5% que representa el mRNA) y es muy estable (Warner J.R. 1999).

### **2.4 Efecto diferencial de la estabilización de los mRNAs a escala genómica**

Una vez determinado que el efecto inhibitorio de la tiolutina sobre la degradación de los mRNAs se produce en un paso posterior al recorte general de la cola poli (A) se estudió si la inhibición de la degradación es específica de algunos mRNAs concretos (como *ACT1* y *RPL24*) o es un efecto común para todos los mRNAs de levadura. Además, un estudio genómico de estas características resulta interesante porque puede aportar pistas sobre que rutas de degradación usan preferencialmente los diferentes mRNAs tras el paso inicial de desadenilación (Parker R. y Song, H. 2004).

Para realizar este estudio se usaron dos concentraciones diferentes de tiolutina (3 y 10  $\mu\text{g/mL}$ ). Se tomaron muestras de RNA de células en



crecimiento exponencial y a los 30 minutos después de la adición de estas dos concentraciones de tiolutina. Se marcaron los RNA copiándolos a cDNA en presencia de  $^{33}\text{P}$ -dCTP usando como cebador un oligo dT<sub>15</sub>VN para generar una copia por molécula de mRNA de forma independiente a la longitud de la cola poli (A) y finalmente se hibridaron las muestras con chips de DNA. Una vez determinada la cantidad de mRNAs en cada punto se normalizó la variación en la cantidad de mRNAs usando como referencia la estabilidad calculada para *ACT1* mediante *northern blot*, dado que este gen permanece en las dos distribuciones en una posición similar (fig 18). De esta manera, y asumiendo una cinética de primer orden, se puede calcular una constante de degradación aparente para todos los genes usando la cantidad de mRNA a tiempo 0 y 30 minutos después de la adición de la tiolutina.

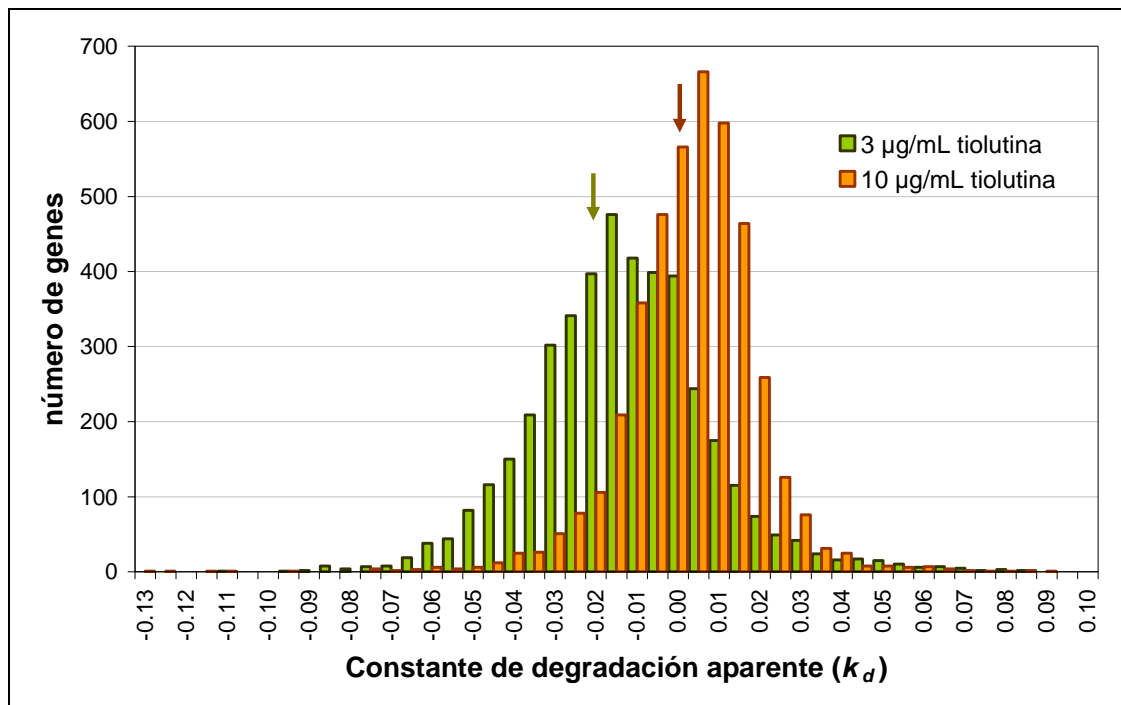


Figura 18. Análisis genómico del efecto de la tiolutina sobre la degradación de los mRNAs. Histograma de las constantes de degradación aparentes [ $k_d = (-1/t_{1/2})$ ] usando diferentes concentraciones de tiolutina. Para normalizar los resultados se utilizó la vida media de *ACT1* calculada por *northern blot*. Este gen permanece aproximadamente en la misma posición relativa en las dos distribuciones (representado por una flecha). Los experimentos se realizaron por duplicado y tuvieron un coeficiente de correlación de Pearson entre las réplicas igual o superior a 0.9.

## Resultados y discusión

En la figura 18 se puede observar que cuando se usa una concentración más alta de tiolutina (10  $\mu\text{g}/\text{mL}$ ) se produce una estabilización general de todos los mRNAs respecto a la vida media calculada con 3  $\mu\text{g}/\text{mL}$ . Con el aumento de la concentración de tiolutina, la distribución de constantes de degradación se desplaza hacia valores positivos (asociados con la estabilización de mRNAs). De hecho, de los 4222 genes analizados 3533 presentan una vida media aparentemente más larga a 10  $\mu\text{g}/\text{mL}$  de tiolutina que a 3  $\mu\text{g}/\text{mL}$ .

Una vez confirmado que la estabilización de los mRNAs es un fenómeno global, que afecta a la mayoría de la población de mRNAs, se decidió estudiar si algunos mRNAs se ven especialmente afectados por la tiolutina. Para ello se realizó un análisis funcional para determinar si existen grupos funcionales que resultan más afectados por el efecto inhibitorio de la tiolutina sobre la degradación de los mRNAs (vidas medias aparentemente más altas). Para ello se compararon las señales de mRNA obtenidas mediante chips de DNA a los 30 minutos después de la incubación con 3 o con 10  $\mu\text{g}/\text{mL}$  de tiolutina. Los genes cuyos mRNAs son relativamente más abundantes en las muestras incubadas con 10  $\mu\text{g}/\text{mL}$  de tiolutina serán aquellos que han sufrido una mayor estabilización debido a la elevada concentración de tiolutina. Por el contrario, los genes cuyos mRNAs son relativamente más abundantes en las muestras incubadas con 3  $\mu\text{g}/\text{mL}$  serán aquellos en los que la tiolutina tiene menos efecto como inhibidor de la degradación de los mRNAs. Así se podrá determinar si todos los genes son igualmente afectados por el efecto estabilizador de la tiolutina o existen diferencias funcionales entre ellos.

Entre los mRNAs que presentan una vida media aparentemente más corta se encuentran los genes relacionados con el metabolismo de sulfuros (p-adj  $9.02 \cdot 10^{-5}$ ), el ribosoma mitocondrial (p-adj  $7.03 \cdot 10^{-9}$ ), la mitocondria (p-adj  $4.28 \cdot 10^{-15}$ ), el transporte de hierro (p-adj 0.008), el transporte de zinc (p-adj 0.0008) o el citoesqueleto de actina (p-adj 0.0005). La identidad de varias de estas categorías funcionales sugiere que este resultado puede deberse a que a bajas concentraciones de tiolutina se induzca su transcripción como respuesta al estrés producido por la droga (Grigull J. *et al* 2004), y que esto no ocurra cuando el bloqueo de la transcripción sea mayor a elevadas concentraciones de la misma. Otra posibilidad no excluyente es que sean genes cuyos mRNAs se degradan principalmente por rutas que no son inhibidas por la tiolutina. Por otra parte, entre los genes que presentan una vida media aparente más larga podemos destacar los genes de las proteínas ribosómicas (p-adj  $1.12 \cdot 10^{-25}$ ) y relacionados con la síntesis de los ribosomas (p-adj  $2.92 \cdot 10^{-25}$ ). Estos genes serán aquellos

cuyos mRNAs se degradan principalmente por las rutas que son más inhibidas por la tiolutina.

## **2.5 Modelo hipotético de inhibición de la degradación de mRNAs por la tiolutina**

La degradación de los mRNAs en eucariotas comienza habitualmente por un recorte de la cola poli (A) (Parker R. y Song,H. 2004). Sin embargo, nuestros resultados muestran que la inhibición de la degradación por parte de la tiolutina se produce principalmente en un punto posterior a la desadenilación, aunque no se puede descartar que el mecanismo de estabilización sea diferente para diferentes mRNAs. Por otra parte, está descrito que la adición de tiolutina produce un estrés a las células (Grigull J. *et al* 2004) y es conocido que una situación de estrés puede producir una estabilización de los mRNAs (Bond U. 2006). Sin embargo, ese tipo de estabilización se produce inhibiendo el paso inicial del recorte de la cola poli (A) (Hilgers V. *et al* 2006) y por lo tanto la estabilización que encontramos no parece deberse a un efecto directo del estrés producido por la tiolutina.

## Resultados y discusión

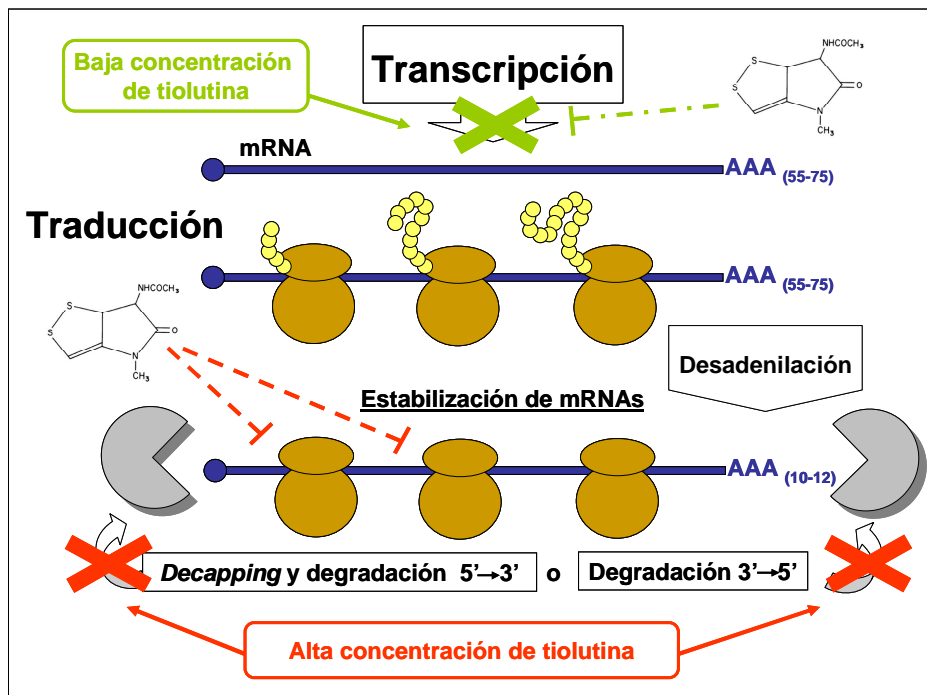


Figura 19. Modelo hipotético de la estabilización de los mRNAs por parte de la tiolutina. A bajas concentraciones, la tiolutina inhibiría principalmente la transcripción, pero a concentraciones superiores se produciría también un bloqueo de la traducción atrapando a los mRNAs en polisomas. De esta manera los mRNAs no serían accesibles a la maquinaria de degradación estabilizándose, pero si podrían sufrir el recorte de la cola poli (A).

Sin embargo, la tiolutina también posee efectos inhibidores de la traducción a altas concentraciones (Jiménez A. *et al* 1973) y está descrito que la inhibición de la traducción por parte de otras drogas, como la cicloheximida produce un bloqueo de la degradación de los mRNAs al nivel de la eliminación de la caperuza, pero sin afectar al recorte de la cola poli (A) (Beelman C.A. y Parker,R. 1994). En el caso de la cicloheximida, los mRNAs quedan secuestrados por los polisomas y no son accesibles al resto de maquinaria de degradación. Estos hechos sugieren que la inhibición de la degradación de los mRNAs por parte la tiolutina podría estar relacionada con su actividad de inhibidor traduccional a elevadas concentraciones, de forma que los mRNAs fueran retenidos en los polisomas de forma análoga a lo que ocurre con la cicloheximida (fig 19).

### 2.6 Uso de la tiolutina como inhibidor transcripcional

El uso de la tiolutina en levadura para inhibir la transcripción de todas las polimerasas seguido con la determinación de la cantidad de mRNAs a diferentes tiempos tras la adición de la droga es un método muy sencillo

para determinar vidas medias de mRNAs (Parker R. *et al* 1991). Para estos estudios se usa habitualmente tiolutina en concentraciones que van desde 3 a 25  $\mu\text{g}/\text{mL}$  (Adams C.C. y Gross, D.S. 1991; Grigull J. *et al* 2004; Guan Q. *et al* 2006; Herrick D. *et al* 1990; Michan C. *et al* 2005; Minvielle-Sebastia L. *et al* 1991; Parker R. *et al* 1991), y se asume que cualquier concentración por encima de 3  $\mu\text{g}/\text{mL}$  tiene básicamente el mismo efecto (Herrick D. *et al* 1990). Sin embargo, en este estudio se ha demostrado que concentraciones superiores a 3  $\mu\text{g}/\text{mL}$  de tiolutina pueden conducir al cálculo de vidas medias aparentemente más largas. Anteriormente ya se habían descrito problemas con el cálculo de vidas medias usando tiolutina, aunque no se había explicado la causa (Brendolise C. *et al* 2002).

Este trabajo sugiere que este fenómeno se debe a que a concentraciones superiores la tiolutina produce también una inhibición de la degradación de los mRNAs. Además, este bloqueo no es una consecuencia directa del propio bloqueo de la transcripción, sino que afecta por una vía independiente de la transcripción al proceso de la degradación de mRNAs. Por lo tanto, en el cálculo de estabilidades de mRNAs, bajas concentraciones de tiolutina pueden no acabar de bloquear eficientemente la transcripción, mientras que concentraciones altas pueden bloquear la degradación de los mRNAs, produciendo una curva de vidas medias aparentes en función de la concentración de tiolutina usada (fig 20).

## Resultados y discusión

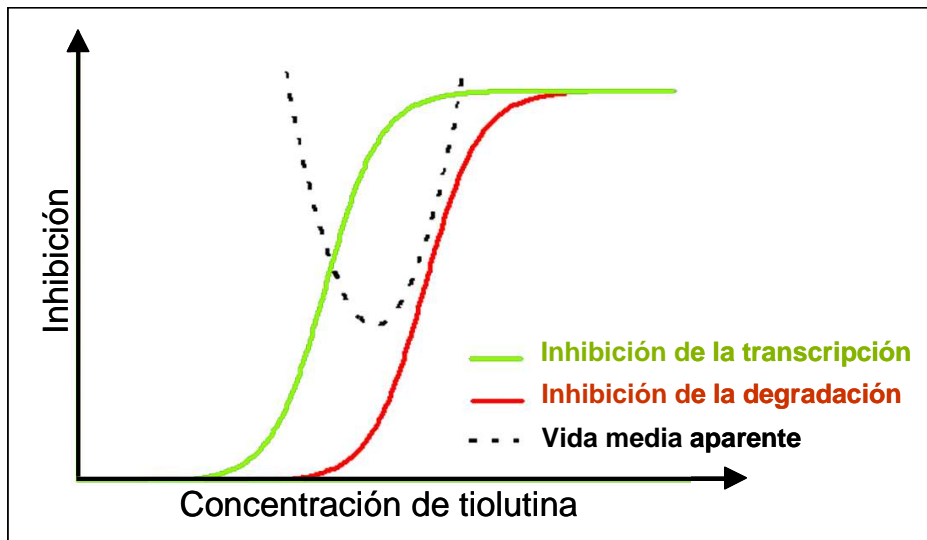


Figura 20. Modelo de acción de la tiolutina en células de levadura. A) La inhibición de la transcripción (**verde**) y la degradación (**rojo**) de los mRNAs produce un efecto complejo en la medida de las vidas medias. La línea punteada representa los efectos añadidos de las dos inhibiciones sobre la vida media aparente calculada.

Otros autores han realizado experimentos similares a los realizados en este trabajo para estudiar los posibles efectos secundarios de la tiolutina añadiéndola a un cultivo cuya transcripción habían parado previamente usando una polimerasa termosensible y un choque térmico (Guan Q. *et al* 2006). Sin embargo, usando una concentración de 10  $\mu\text{g}/\text{mL}$  de tiolutina no observan diferencias significativas en las vidas medias calculadas entre las alícuotas con y sin tiolutina. Contrariamente, usando esas concentraciones en nuestro laboratorio ya se observa un efecto inhibitorio sobre la degradación de mRNAs. La explicación más fácil para estas discrepancias es que son debidas a la diferente composición de los lotes de tiolutina que hasta hace poco tiempo no era un producto comercial. Guan *et al.*(2006) describen que necesitan 10  $\mu\text{g}/\text{mL}$  de tiolutina para medir las vidas medias en un fondo genético S288c (como el que usamos nosotros) y 25  $\mu\text{g}/\text{mL}$  para un fondo W303 (2.5 veces más). Mientras que en nuestro laboratorio se determinó que es necesario una concentración de 3  $\mu\text{g}/\text{mL}$  para un fondo S288c y de 8  $\mu\text{g}/\text{mL}$  para un fondo W303 (2.67 veces más) (comunicación personal de Lorena Romero-Santacreu y Paula Alepuz). Es decir, en ambos casos se mantiene la relación entre las concentraciones necesaria para inhibir la transcripción entre las dos cepas y lo que parece cambiar es la actividad biológica de los diferentes lotes de tiolutina. Para comprobar esa variabilidad en actividad biológica entre los diferentes lotes de tiolutina, se comprobó mediante espectrometría de masas la composición de dos lotes diferentes de tiolutina, y en uno de ellos se detecto hasta un 20% de contaminantes que no estaban presentes en el otro

lote. Esto confirma la variabilidad en la composición de diferentes lotes de tiolutina y explica el uso de concentraciones más altas de la misma por otros autores (Guan Q. *et al* 2006; Michan C. *et al* 2005).

Por todo esto, el uso de la tiolutina como inhibidor transcripcional, además del problema ya conocido de la inducción de genes de respuesta a estrés (Adams C.C. y Gross, D.S. 1991; Grigull J. *et al* 2004) presenta un problema debido a su acción inhibitoria sobre la degradación de los mRNAs. Así, si la concentración usada es demasiado baja, no se bloqueará eficientemente la transcripción y si es demasiado alta se producirá una estabilización de los mRNAs. Por lo tanto solo habrá un intervalo de concentraciones donde se pueda usar la tiolutina para calcular vidas medias de mRNAs, y en cualquier caso las vidas medias aparentes calculadas usando tiolutina serán ligeramente superiores a las reales (ya sea debido a una ineficiente parada de la transcripción o a una ligera inhibición de la degradación). Esto hace que sea necesario realizar un estudio previo de las concentraciones de tiolutina a usar en un determinado experimento, para asegurar que se trabaja en las concentraciones óptimas para el cálculo de vidas medias, dado que esta depende tanto de las condiciones experimentales como de la cepa utilizada.

### **2.7 Medida de vidas medias genómicas**

Una vez determinada que entre las concentraciones utilizadas, la mejor concentración de tiolutina para la medida de vidas medias en estas condiciones experimentales era de 3  $\mu\text{g/mL}$ , se amplió el cálculo de vidas medias de mRNAs a escala genómica.

Para ello, se realizó un protocolo análogo al usado para el cálculo de estabilidades para mRNAs individuales. Se añadió tiolutina a una concentración de 3  $\mu\text{g/mL}$  a un cultivo en crecimiento exponencial, se recogieron las células a diferentes tiempos tras la inhibición de la transcripción (0, 10, 20, 30 y 45 minutos) y se extrajo el RNA. Posteriormente, en vez de usar los RNAs para un análisis mediante *northern blot*, se marcaron e hibridaron en macrochips de DNA (fig 21).

## Resultados y discusión

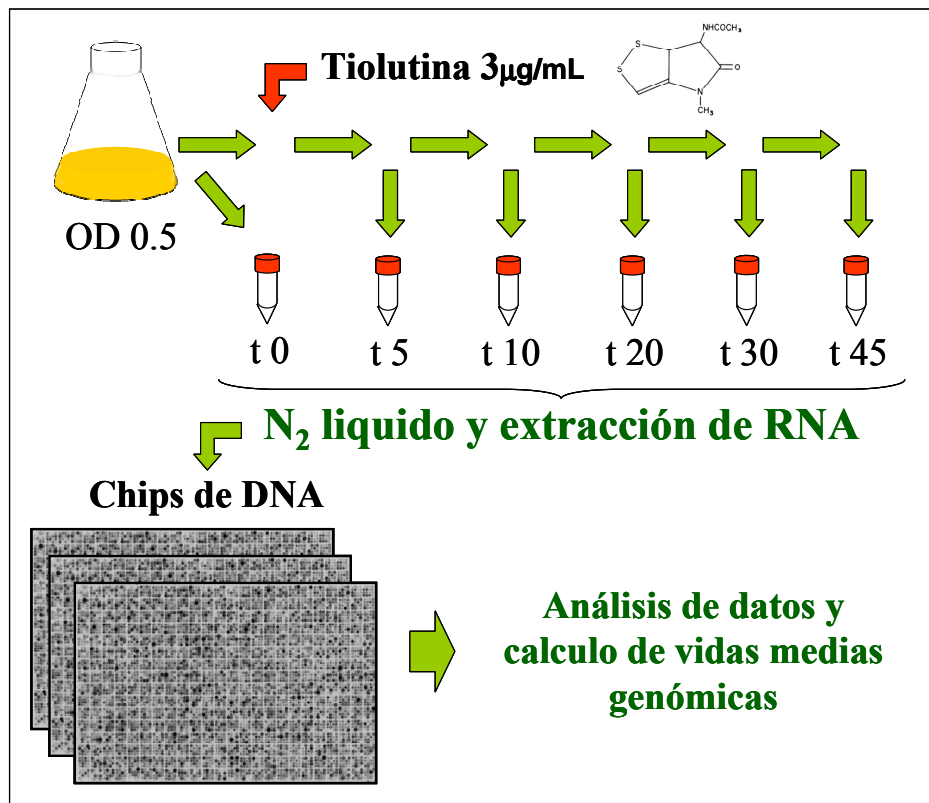


Figura 21. Diseño experimental del cálculo de vidas medias a escala genómica. Se partió de un cultivo de levaduras en crecimiento exponencial al que se le añadieron 3 µg/mL de tiolutina. Posteriormente se recogieron células a diferentes tiempos tras la adición de la droga, se congelaron las células en N<sub>2</sub> líquido y se extrajo el RNA. Seguidamente se marcaron las muestras y se hibridaron en chips de DNA. Posteriormente se analizaron los datos para calcular las vidas medias para los mRNAs de todos los genes.

Con las señales obtenidas de estos chips de DNA se puede calcular la variación en la cantidad de mensajeros tras la inhibición de la transcripción. Pero, para calcular las estabildades de los mRNAs, es necesario normalizar entre sí los resultados obtenidos a diferentes tiempos tras la parada transcripcional. Para ello, se asumió que el global de los mRNAs desaparecen con una cinética similar al global de RNA poliadenilado calculado mediante un *dot blot*.

Usando esta normalización, se calcularon datos de vida media para un total de 4700 genes, de los cuales 512 presentaban valores negativos de vida media (incrementa la cantidad de mRNA). Los genes para los que se calculan vidas medias negativas son aquellos cuya transcripción se ve inducida por el efecto de la droga y, por lo tanto, no es posible calcular sus vidas medias utilizando esta aproximación. Estos genes están relacionados con la respuesta a estrés y ya estaba descrito que ni el uso de drogas



inhibidoras de transcripción ni el mutante termosensible permiten realizar un cálculo adecuado de la estabilidad de sus mRNAs (Adams C.C. y Gross, D.S. 1991; Grigull J. *et al* 2004; Wang Y. *et al* 2002). Entre estos genes se pueden destacar los relacionados con el metabolismo de carbohidratos (p-adj <0.001), asimilación de sulfatos (p-adj 0.001) o el plegado de proteínas (p-adj 0.003).

### **2.8 Comparación de estabildades de mRNAs medidas con métodos genómicos diversos.**

Una vez obtenidos datos de vidas medias por el método directo con tiolutina y macrochips de DNA, se decidió compararlos con otros estudios genómicos para tratar de extraer algunas conclusiones sobre este tipo de estudios. En general, las correlaciones entre diferentes cálculos de vidas medias genómicas son bastante bajas (tabla 8). Las comparaciones entre los cálculos de vidas medias directa (basadas en la parada de la transcripción y la posterior medida de la desaparición de los mRNAs) son significativas en todos los casos, pero la misma técnica empleada en diferentes laboratorios sólo presenta correlaciones entorno al 0.4 (tanto en el caso de la parada por tiolutina como en el uso del mutante termosensible). También se aprecia mejor correlación entre los datos procedentes del mismo laboratorio, aun usando técnicas diferentes, que entre laboratorios diferentes, lo que indica que existe una gran variabilidad experimental inherente a estas metodologías. Por otro lado, la comparación de las medidas directas con las calculadas de manera indirecta asumiendo un estado estacionario entre la transcripción y la degradación de los mRNAs presenta en todos los casos una correlación nula o negativa. Previamente ya se había descrito esta nula correlación (Pérez-Ortín J.E. 2007) usando los resultados de un único experimento (García-Martínez J. *et al* 2004). Aquí se han usado los valores promedio de 8 experimentos diferentes de GRO, realizando o no una corrección por la longitud del gen (detallada en el apartado 3.1 de este trabajo), y se sigue observando un resultado similar. Este hecho parece deberse a un artefacto matemático. Este artefacto se debe a que las tres magnitudes (TR, [mRNA] y la  $k_d$ ) presentan una correlación positiva entre si. Cuando se trata de obtener un listado de  $k_d$  indirectas dividiendo TR/[mRNA], la correlación de estas dos magnitudes con la  $k_d$  hace que se anulen entre si. Por lo tanto los coeficientes de correlación no son una buena forma para comparar datos directos e indirectos de estabildades de mRNAs.

## Resultados y discusión

Tabla 8. Correlaciones entre diferentes cálculos de estabildades de mRNAs a escala genómica. Los valores representan los coeficientes de correlación no paramétricos de Spearman. Todos los valores son estadísticamente diferentes de 0 ( $\text{sig} < 10^{-5}$ ) excepto los resaltados en **gris**. Los datos utilizados para las comparaciones proceden de cálculos alternativos usando un mutante termosensible de la polimerasa, tiolutina como inhibidor transcripcional (Grigull J. *et al* 2004; Wang Y. *et al* 2002) o el cálculo indirecto de estabildades utilizando un conjunto de datos de GRO antes o después de realizar una corrección por longitud (capítulo 3.1 de este trabajo).

Correlación	Indirecta corregida (GRO)	Indirecta (GRO)	rpb1-1 (Grigull)	rpb1-1 (Wang)	Tiolutina (Grigull)
Tiolutina (este estudio)	<b>-0.034</b>	<b>-0.038</b>	0.189	0.172	<b>0.421</b>
Tiolutina (Grigull)	-0.071	-0.073	0.33	0.23	
rpb1-1 (Wang)	-0.124	-0.127	<b>0.365</b>		
rpb1-1 (Grigull)	<b>-0.06</b>	<b>-0.056</b>			

Las bajas correlaciones existentes entre los diferentes métodos genómicos para el cálculo de estabildades de mRNAs pone de manifiesto la dificultad para obtener medidas precisas de estos parámetros y la gran variabilidad que presentan. Además, el hecho de requerir diversas medidas durante un tiempo prolongado dificulta la comparación de las diferentes aproximaciones y aumenta la variabilidad de los resultados. Por otra parte, es necesario recordar que los datos de vida media calculados por diferentes autores para genes individuales también presentan variaciones importantes (Herrick D. *et al* 1990; Santiago T.C. *et al* 1986), por lo que no es sorprendente que estas variaciones sean incluso mayores cuando se realizan estudios a escala genómica.

Si se comparan las vidas medias de unos pocos grupos funcionales entre los diferentes experimentos (fig 22) se observa que hay importantes sesgos dependiendo del método usado para parar la transcripción. Sin embargo, estas diferencias son repetitivas entre experimentos independientes, de forma que la ordenación de estabildad de mRNAs entre los grupos funcionales diferentes es similar cuando se usa la misma técnica. Además de estas diferencias debidas a la técnica usada para calcular las vidas medias, también existen diferencias de escala entre los diferentes experimentos, de forma que el global de vidas medias puede ser superior o inferior dependiendo de las diferentes normalizaciones y asunciones aplicadas.

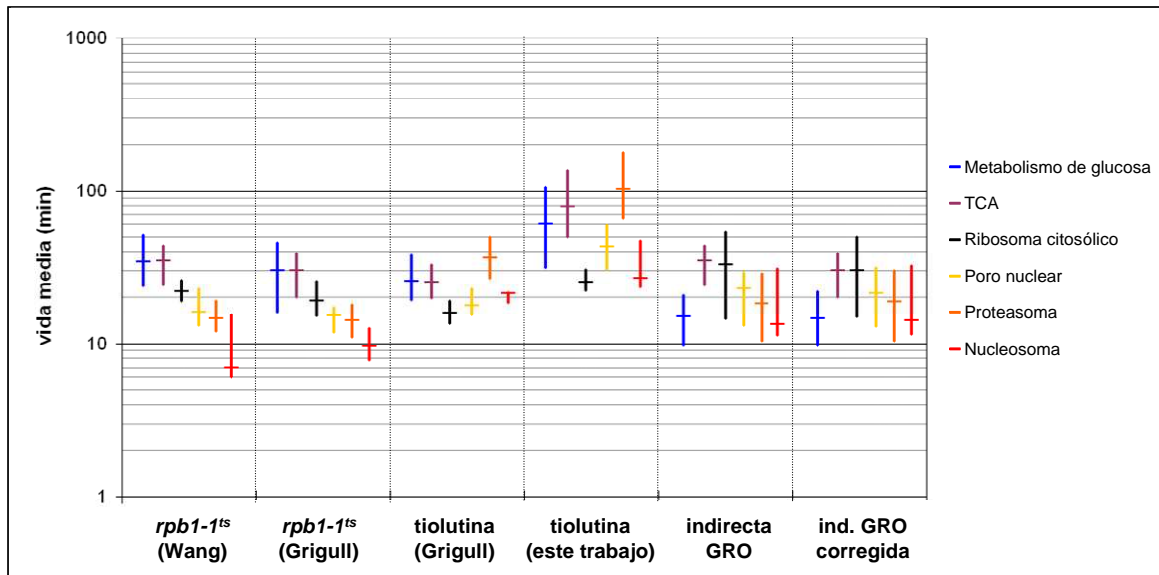


Figura 22. Comparación de vidas medias para algunos grupos funcionales entre diferentes experimentos. Los valores mostrados representan la distancia intercuartílica de las vidas medias de esas grupos funcionales, mientras que las señales horizontales representan las medianas para estos grupos. Se representan los valores calculados usando el mutante *rpb1-1<sup>ts</sup>* (Wang Y. *et al* 2002), usando es mismo mutante, tiolutina (Grigull J. *et al* 2004), los resultados de este estudio y los datos indirectos obtenido mediante un compendio de experimentos de GRO antes y después de realizar una corrección por la longitud del gen (descrito en el capítulo 3.1 de este trabajo).

Estos resultados confirman la gran variabilidad existente entre los diferentes datos de vidas medias genómicas de mRNAs y la dificultad para realizar comparaciones entre ellos, incluso usando la misma técnica.

### **3. Desarrollo de nuevas técnicas genómicas para el estudio de la transcripción**

Una vez realizado un estudio sobre el estado estacionario para la expresión génica y sobre la medida genómica de estabilidades de mRNAs, se decidió desarrollar nuevas técnicas que permitieran un mejor estudio de la transcripción. Para ello se desarrolló una técnica para la medida genómica de tasas de transcripción basada en la inmunoprecipitación de cromatina, el RPCC. Esta técnica ha permitido tanto realizar una corrección de los datos de transcripción obtenidos mediante GRO como descubrir diferencias biológicas entre la presencia y actividad de la RNA pol II para diferentes grupos de genes.

Por otra parte, también se ha puesto a punto un protocolo experimental para aislar mRNAs nacientes e hibridar chips de alta resolución basados en detección fluorescente.

#### **3.1. Cálculo de la transcripción basado en ChIP-on-chip**

##### **3.1.1 RPCC: RNA Pol ChIP-on-chip**

Existen algunas técnicas que permiten realizar una estimación de la transcripción o del RNA de nueva síntesis (Cleary M.D. *et al* 2005; Fan J. *et al* 2002; Ohtsu M. *et al* 2008), pero la única técnica que permite actualmente una cuantificación de las tasas de transcripción a escala genómica es el *Genomic Run-on* (García-Martínez J. *et al* 2004). Esta técnica se basa en la incorporación *in vivo* por parte de las RNA polimerasas activas de un nucleótido marcado radiactivamente. Dado que la tasa de transcripción calculada está basada únicamente en el *run-on* nuclear, los sesgos que éste pueda tener no son fácilmente identificables. Para tratar de solucionar este problema, resulta útil desarrollar otras técnicas para medir la tasa de transcripción de manera independiente. De esta forma, comparando los resultados obtenidos mediante dos aproximaciones experimentales diferentes para estudiar el mismo proceso (transcripción), se puede contrastar la validez de los resultados genómicos de tasas de transcripción, así como tratar de descubrir si existen sesgos biológicos o experimentales asociados a cada una de las técnicas.

Para obtener una medida de la tasa de transcripción para todos los genes de manera alternativa al GRO hemos desarrollado el RNA pol ChIP-on-chip (RPCC). Esta técnica esta basada en la inmunoprecipitación de cromatina

seguida de un análisis con chips de DNA (ChIP-on-chip) (Iyer V.R. *et al* 2001; Ren B. *et al* 2000; Rosaleny L.E. *et al* 2007). En el RPCC se usan anticuerpos que reconocen la RNA pol II para determinar la densidad de polimerasas unidas a cada gen (Sandoval J. *et al* 2004) y posteriormente se realizan una serie de normalizaciones para transformar los datos de unión de polimerasa a tasas de transcripción.

### **Diseño experimental del RPCC**

El diseño experimental del RPCC es básicamente el mismo que el del ChIP-on-chip tradicional, siendo la principal diferencia que, en vez de determinar la posición más o menos cualitativa a lo largo del genoma, se pretende determinar cuantitativamente la densidad de polimerasas para todos los genes.

La primera etapa del proceso consistió en entrecruzar las células usando formaldehído para producir uniones covalentes DNA-Proteína y Proteína-Proteína que fijen las polimerasas a los fragmentos de DNA a los que estaban asociadas en ese momento. Posteriormente, se extrajo la cromatina y se realiza una inmunoprecipitación específica de la RNA pol II. De esta manera se consiguió una muestra de DNA enriquecida en aquellos fragmentos donde estaba unida la polimerasa (IP). Este enriquecimiento será mayor cuando más frecuente sea la presencia de la polimerasa en esos fragmentos. Además de esta muestra de DNA inmunoprecipitada, se obtuvieron otras dos muestras: un control positivo (extracto celular completo) para normalizar los datos del inmunoprecipitado y un control negativo (una inmunoprecipitación realizada en paralelo sin el anticuerpo contra la RNA pol II) para descartar artefactos durante el proceso de inmunoprecipitación.

Como la inmunoprecipitación no permite obtener suficiente cantidad de DNA para poder hibridar directamente un chip de DNA, fué necesario realizar una amplificación de las muestras obtenidas. Para ello se utilizó una amplificación por PCR-ligada (Ren B. *et al* 2000), en la que se liga a las moléculas de DNA inmunoprecipitadas unos adaptadores unidireccionales comunes que son usados posteriormente para una amplificación por PCR de toda la población de fragmentos de DNA. Para minimizar posibles sesgos que se pudieran dar durante el proceso de la amplificación, se sometieron todas las muestras (control e inmunoprecipitadas) al mismo proceso de amplificación. Finalmente se marcó el DNA resultante con un único ciclo de PCR en presencia de <sup>33</sup>P-

## Resultados y discusión

dCTP y se hibridaron las muestras con macrochips de DNA (Alberola T.M. *et al* 2004).

### **Cálculo de la tasa de transcripción**

Para poder realizar un cálculo de tasas de transcripción a partir de los datos de unión relativa de polimerasas a los diferentes fragmentos de DNA obtenidos por ChIP-on-chip, fue necesario realizar una serie de normalizaciones.

Las muestras inmunoprecipitadas (IP) están enriquecidas en fragmentos de DNA que tenían polimerasas unidas, y por tanto contienen información sobre la densidad de polimerasa presente en esos fragmentos. Sin embargo, no se pueden usar directamente estos datos, dado que el valor que se puede calcular para cada gen se ve influido también por otros factores, como el contenido en GC de la muestra (se usó un marcaje de DNA de doble cadena con <sup>33</sup>P-dCTP) o el número y longitud de las sondas depositadas en el chip. Para obtener un valor proporcional a la densidad de polimerasas, es necesario normalizar los resultados de la muestra inmunoprecipitada con los provenientes de la hibridación del control positivo (extracto celular completo), que representa a la cromatina total sin ningún enriquecimiento y ha sido amplificada y marcada por el mismo procedimiento. Por lo tanto, el resultado de dividir la señal de la muestra inmunoprecipitada por el DNA genómico es directamente proporcional a la densidad de polimerasa en cada sonda. Además, como el chip de DNA usado en este estudio posee sondas para casi todos los genes de levadura (Alberola T.M. *et al* 2004), con este método se puede calcular la densidad de polimerasa relativa para casi todos los genes.

Una vez calculada la densidad relativa de polimerasas presentes en cada gen por ChIP-on-chip, para transformar estos valores en tasas de transcripción es necesario asumir que la cantidad de polimerasa presente en los genes es directamente proporcional a la tasa de transcripción y que la velocidad de transcripción a lo largo de los genes es constante. Esta asunción no es exclusiva del RPCC, dado que en el GRO se obtienen datos de densidades de polimerasas activas y también es necesario asumir que las velocidades de transcripción a lo largo del gen son constantes para transformar estos datos en tasas de transcripción. La principal diferencia entre GRO y RPCC es que en el GRO se obtienen densidades de polimerasas elongantes activas, mientras que en el RPCC se obtienen densidades de polimerasas físicamente unidas a los genes pero no necesariamente en elongación activa. En general se puede asumir que la

presencia de polimerasa en los genes está directamente relacionada con la transcripción de los mismos. Pero esta asunción, aunque puede ser válida de forma general, no se cumple para todos los genes. De hecho, se han encontrado genes con polimerasas reclutadas en sus promotores de forma independiente a su transcripción (Margaritis T. y Holstege, F.C. 2008; Price D.H. 2008; Radonjic M. *et al* 2005; Tamkun J.W. 2007). En el apartado 3.1.3 de este trabajo nos centraremos en el estudio de genes para los que no es válido realizar esta asunción y en los cuales existen diferencias entre RPCC y GRO.

Otra de las limitaciones del ChIP-on-chip “clásico” es que sólo permite calcular la densidad relativa de proteína unida a los diferentes genes respecto al promedio de la población del experimento, pero no permite calcular valores absolutos de presencia de proteína, y por tanto tampoco de transcripción. Sin embargo, asumiendo que no hay presencia de RNA pol II en los locus del rDNA se pueden obtener valores de transcripción comparables entre diferentes experimentos. Los rDNAs son lugares de alta transcripción por parte de la RNA pol I y III, y aunque en algunos casos excepcionales se puede dar actividad de la RNA pol II en esas regiones (Vu L. *et al* 1999) en general se asume que son regiones libres de RNA pol II (Steinmetz E.J. *et al* 2006). Por lo tanto, si consideramos que las señales de los 32 puntos con sondas para el rDNA presentes en el chip no se deben a ningún enriquecimiento específico por la presencia de polimerasas, sino que son debidos al ruido de “fondo” del experimento (la señal mínima debido a un arrastre inespecífico de cromatina durante el proceso de la inmunoprecipitación), se puede fijar que la transcripción medida por RPCC para esos puntos es siempre cero. Consecuentemente se pueden usar estos puntos para determinar el ruido de fondo y normalizar la transcripción del resto de genes entre diferentes experimentos y poder compararlos.

## Resultados y discusión

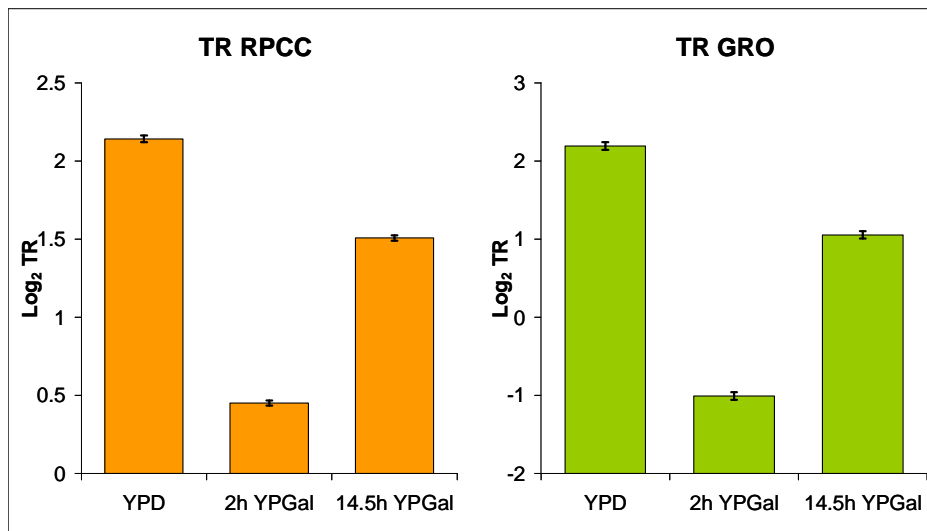


Figura 23. Promedio de tasa de transcripción para todos los genes en diferentes situaciones fisiológicas usando RPCC y GRO. En los dos casos representamos el logaritmo de las tasas de transcripción en unidades arbitrarias, lo que nos permiten comparar entre diferentes situaciones fisiológicas dentro de la misma técnica. En ambos casos se usan datos de células en crecimiento exponencial en glucosa (YPD), a las dos horas tras el cambio a galactosa (2 h YPGal) y en crecimiento exponencial en YPGal a las 14.5 h. En el caso del RPCC se ha usado el Ac 8WG16. Las barras de error representan el intervalo de confianza para la media para una significación del 0.05.

Para estudiar la validez de esta normalización, se realizaron experimentos de RPCC en diferentes situaciones fisiológicas para las que habían datos de transcripción medidos por GRO (García-Martínez J. *et al* 2004). Se usaron 3 condiciones fisiológicas diferentes: crecimiento exponencial en YPD, a las dos horas de cambiar las células a un medio con galactosa como fuente de carbono (YPGal) y cuando las células vuelven a estar en crecimiento exponencial en galactosa (14.5 horas tras el cambio a YPGal). Usando los valores del rDNA para normalizar los diferentes experimentos de RPCC entre sí, se observa que las variaciones en tasas de transcripción calculadas por RPCC son consistentes con las calculadas por GRO para las mismas condiciones (fig 23).

Se puede observar que hay una elevada tasa de transcripción promedio en condiciones de crecimiento exponencial en glucosa, pero que esta disminuye drásticamente a las dos horas de cambiar las células a un cultivo con galactosa. Esta disminución en la transcripción ya se observó mediante GRO (García-Martínez J. *et al* 2004), y es debida a la completa parada en el crecimiento celular que se produce cuando la célula se está adaptando al uso de la nueva fuente de carbono. Posteriormente, en crecimiento exponencial en galactosa, aumenta la tasa de transcripción, pero no llega a



los niveles observados en crecimiento exponencial en glucosa. La coincidencia de los perfiles observados es un dato a favor de la validez del RPCC como técnica para medida de las tasas de transcripción. Sin embargo, se puede observar que el GRO tiene un mayor rango dinámico que el RPCC, dado que la transcripción medida por GRO en glucosa es aproximadamente 8 veces superior a la de 2 horas en galactosa ( $3 \log_2$ ), mientras que en caso del RPCC sólo es unas 4 veces superior ( $2 \log_2$ ) (fig 23). Esto se debe, probablemente, a que en el RPCC los valores de transcripción provienen de un enriquecimiento selectivo de la muestra que convierte a los datos de RPCC en datos más ruidosos y dificulta la medida de valores bajos de transcripción. Por otra parte, el GRO tiene menos ruido de fondo, dado que se evalúa directamente el efecto de la transcripción sin someter a las muestras a ningún tipo de enriquecimiento selectivo.

### **Comparación de cálculos genómicos de la tasa de transcripción**

Para comprobar la validez del RPCC para medir las tasas de transcripción se compararon los resultados de experimentos realizados por triplicado usando los dos anticuerpos diferentes con los datos obtenidos mediante GRO en García-Martínez J. *et al.* (2004). Como era esperable, existe una correlación positiva entre las dos técnicas de medida de transcripción en cuanto a tendencias globales (fig 24). Concretamente, se puede calcular una correlación de Spearman entre el RPCC realizado con el anticuerpo 8WG16 y el GRO de 0.47, y entre el RPCC realizado con Rpb1-MYC y el GRO de 0.38 (mientras que la correlación entre los dos RPCC es de 0.56). Sin embargo, hay que tener en cuenta que a pesar de estas correlaciones globales existen diferencias significativas en cuanto a genes individuales.

Comparando las diferencias existentes entre las dos técnicas se puede intentar descubrir si existe algún sesgo específico en alguna de ellas de forma independiente al proceso de la transcripción.

## Resultados y discusión

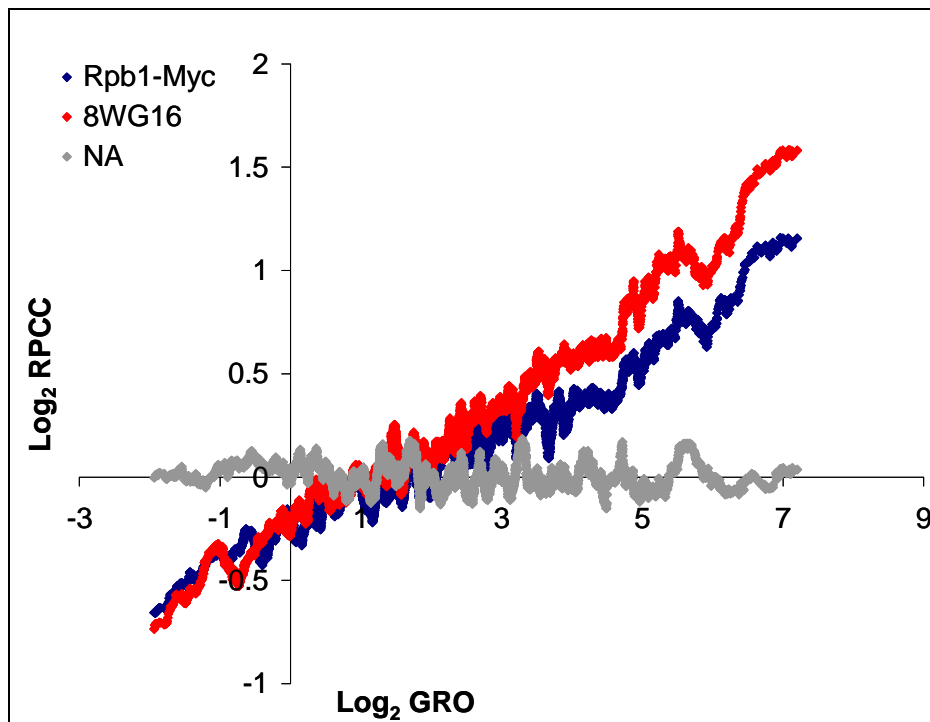


Figura 24. Comparación de datos de RPCC y GRO en crecimiento exponencial en YPD. RPCC usando la subunidad grande de la RNA pol II etiquetada (**Rpb1-MYC**) en azul, usando un anticuerpo específico contra el dominio CTD hipofosforilado (**8WG16**) en rojo o un control negativo sin usar anticuerpo (**No Anticuerpo**) en gris. Se representan las tendencias de los datos usando las medias de ventanas deslizantes de 100 genes a lo largo de toda la distribución.

Algunas de estas diferencias corresponden a grupos funcionales específicos, y se deben a que el RPCC mide la presencia física de polimerasa y el GRO su actividad elongante. Estas diferencias con fundamento biológico las estudiaremos en detalle en el apartado 3.1.3 de este trabajo. Pero además de estas diferencias, también se han encontrado algunas otras diferencias no explicables biológicamente y que, probablemente, se deban a sesgos dependientes del método experimental. Estos sesgos dependientes de técnica no son importantes cuando se comparan diferentes situaciones usando la misma técnica, dado que el sesgo afectaría a las diferentes medidas, pero sí son importantes cuando se trata de obtener valores absolutos de transcripción. Por otra parte, una vez identificado un posible sesgo que diferencie los datos de transcripción calculados por RPCC y GRO, hay que determinar que técnica es la que ofrece datos menos sesgados para poder realizar una corrección de los datos de transcripción.

El principal sesgo que se ha podido detectar comparando las dos técnicas es un efecto dependiente de la longitud de la sonda depositada en el chip, y ese efecto parece ser debido a sesgos intrínsecos en el GRO. Cuando se comparan los valores de RPCC con GRO en función de la longitud de la sonda (fig 25) se puede observar que ambas técnicas detectan una bajada de la transcripción a medida que aumenta la longitud del gen. Esta bajada es de origen biológico (los genes más transcritos son generalmente más cortos) y también se encuentra en datos de cantidades de mRNA. La bajada es ligera en el caso del RPCC y mucho más acusada en el GRO. Además, en el caso del GRO se puede observar un aumento brusco de la tasa media de transcripción de los genes cuando superan las 3000 pb. Este aumento no se observa en el RPCC. Como el chip de DNA que se ha utilizado en este estudio (Alberola T.M. *et al* 2004) tiene la particularidad de que para todos los genes de más de 3 kb solo está depositada una sonda que corresponde a la kilobase 3' final, en lugar de la longitud completa de la ORF como en el resto de los genes, el efecto observado debe ser causado por el cambio en la longitud de la sonda. De hecho, el valor de la curva de la media de transcripción mediada por GRO aumenta en los genes mayores de 3 kb aproximadamente hasta situarse aproximadamente en el valor de transcripción de los genes de 1 kb.

Para identificar el origen de este efecto es necesario centrar la atención en las diferencias técnicas entre el RPCC y el GRO. Por una parte los dos han sido realizados usando el mismo tipo de chip, por lo que las diferencias observadas deben proceder de diferencias en el proceso de marcaje o en el de normalización de los datos. En el caso del RPCC, los valores de transcripción se obtienen a partir de la relación entre la señal de muestra inmunoprecipitada (IP) y del control positivo (DNA genómico obtenido del extracto celular completo). En la figura 26 se muestra una suavización de las intensidades de estas dos señales en función de la longitud del gen en la que se puede observar que no existe ninguna diferencia importante entre las señales de IP y DNA genómico (líneas roja y verde). Por lo tanto, la señal de RPCC que se obtiene dividiendo estas dos señales tampoco presentará un sesgo dependiente de la longitud del gen. Este resultado era esperable porque en ambos casos se usa PCR ligada para amplificar las muestras, cualquier sesgo que pudiera existir en el marcaje se eliminaría.

## Resultados y discusión

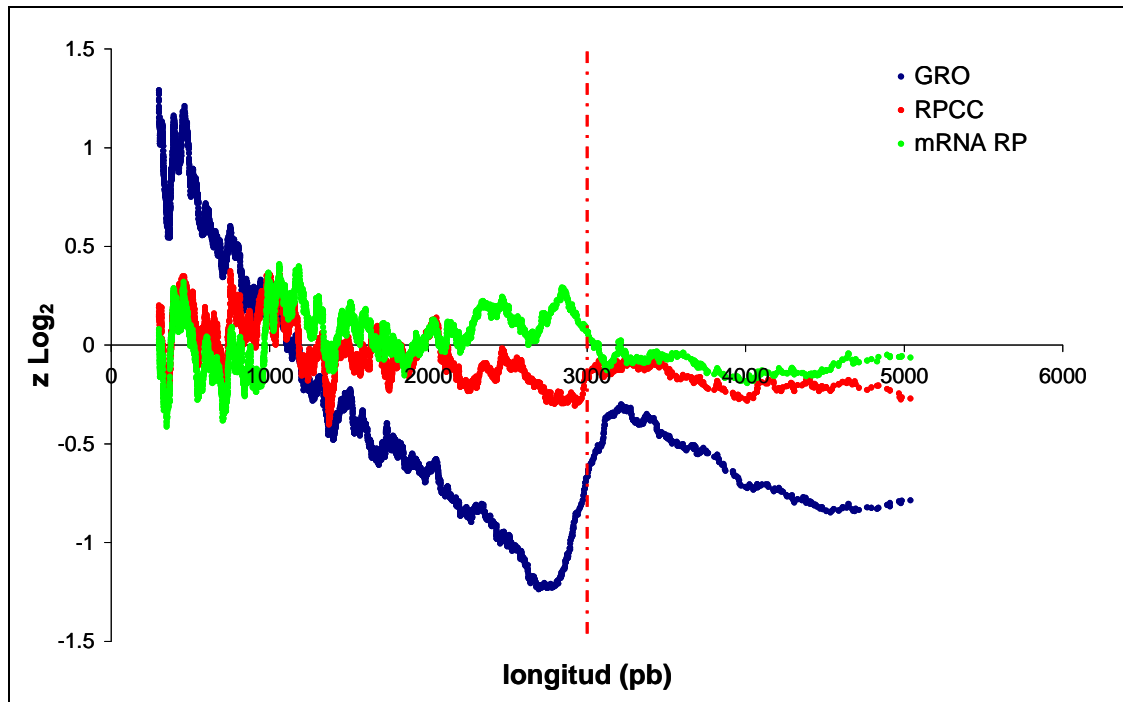


Figura 25. Sesgo de los datos de TR de GRO en función de la longitud del gen. Se puede observar que la curva de GRO decae con mayor pendiente y que los genes con más de 3000 pb (en los cuales sólo se encuentra depositado en el chip la kilobase final) presentan valores superiores de transcripción medida por GRO (en azul). Este fenómeno no se observa ni en el RPCC (en rojo) ni en la cantidad de mRNA (en verde) marcado con cebadores aleatorios (*Random Primers*). En todos los casos se muestra una suavización usando la media de una ventana deslizante de 100 genes. Los datos de RPCC provienen de la media de 9 experimentos diferentes, los de GRO de 14 experimentos diferentes y los de mRNA RP de 15 experimentos diferentes.

Por el contrario, la tasa de transcripción por GRO se obtiene dividiendo la señal de un marcaje *in vivo* por parte de la RNA pol II y la de un DNA genómico marcado con cebadores aleatorios. Por lo tanto el sesgo podría deberse al diferente procedimiento de marcaje entre esas dos técnicas. En la figura 26 se puede observar que la curva de GRO (azul) presenta valores más altos que su control de DNA genómico (gris) especialmente para los genes de más de 3 kb (todos los cuales poseen sondas de la misma longitud que cubren exclusivamente su kilobase final). Concretamente, esta anomalía de la curva de GRO sugiere que el efecto dependiente de la longitud del gen podría estar relacionado con un sesgo en la marca del GRO hacia los extremos 3' de los genes. Este fenómeno puede deberse a que el marcaje de GRO, al producirse por la propia polimerasa es intrínsecamente un marcaje direccional, puesto que la marca se introduce

por el movimiento de síntesis de la RNA pol en sentido 5'→3' desde el punto en que se encuentra la polimerasa. De esta forma, la marca en los RNAs elongantes tiende a ser mayor hacia 3'. Al dividir por la señal inicial del DNA genómico se produce una normalización por la longitud del gen dando densidades de polimerasas promediadas para toda la ORF. Pero si sólo existe una sonda de 1 kb de la región 3' del gen se incrementa la densidad aparente de RNA pol medidas.

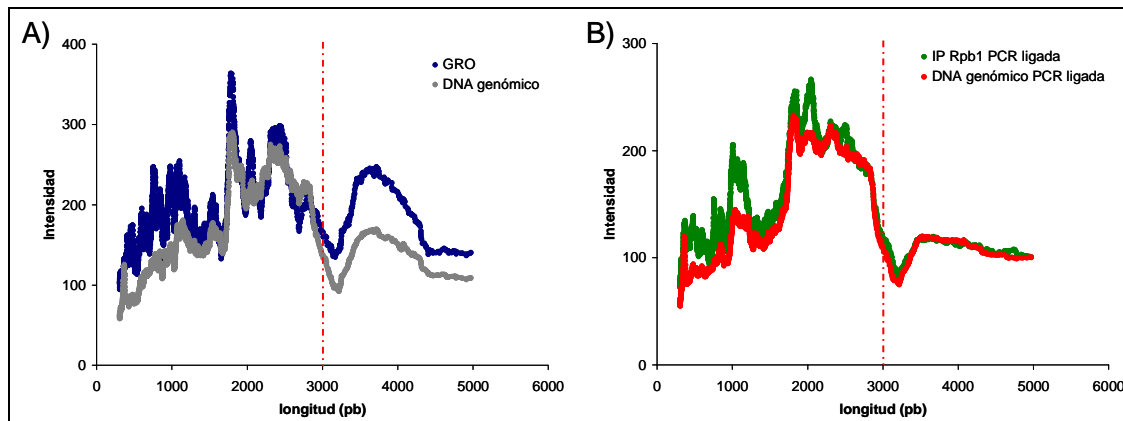


Figura 26. Intensidades por separado para el GRO y el RPCC. A) Separación de las señales del GRO entre la señal directa de la hibridaciones de “run-on” (**GRO**) y del DNA genómico que se usa para normalizar las señales (**DNA genómico**). B) Separación de las señales del RPCC entre la señal de la muestra inmunoprecipitada (**IP**) y del extracto celular completo (**DNA genómico**). Se muestran la suavización de las señales usando ventanas deslizantes de 100 genes en función de la longitud de las ORFs a partir de experimentos realizados por triplicado.

Para comprobar si el efecto que observamos se puede explicar por un marcaje preferente de las zonas 3' de los genes realizamos el mismo análisis comparando marcajes para los que se conoce que existe un sesgo en el marcaje. Comparamos mRNAs marcados usando cebadores aleatorios (que producen una marca homogénea a lo largo de todo el transcrito) o usando como cebador un oligo dT (que al marcar los transcritos siempre desde el extremo 3' produce una distribución de la marca sesgada hacia esa zona). Para normalizar los datos usamos en ambos casos un marcaje de DNA genómico tal como se hace en el RPCC. Al representar los valores en función de la longitud de los genes se puede observar que en caso del mRNA marcado con oligo dT se produce una pendiente acusada y un cambio brusco en la zona de 3000 pb, muy similar

## Resultados y discusión

al que se observa para el GRO (fig 27). Por el contrario en el caso del mRNA marcado con cebadores aleatorios no se observa este efecto.

La explicación más probable, pues, para la mayor pendiente y el efecto en las 3 kb es un sesgo en el marcaje por *run-on* hacia la zona 3' de los genes. Y por lo tanto este es un posible sesgo a corregir en los datos de transcripción obtenidos mediante *run-on*. Sin embargo hay que remarcar que este sesgo sólo tiene importancia cuando se quieren obtener valores absolutos de transcripción, no cuando se evalúan cambios en la transcripción medida por la misma técnica.

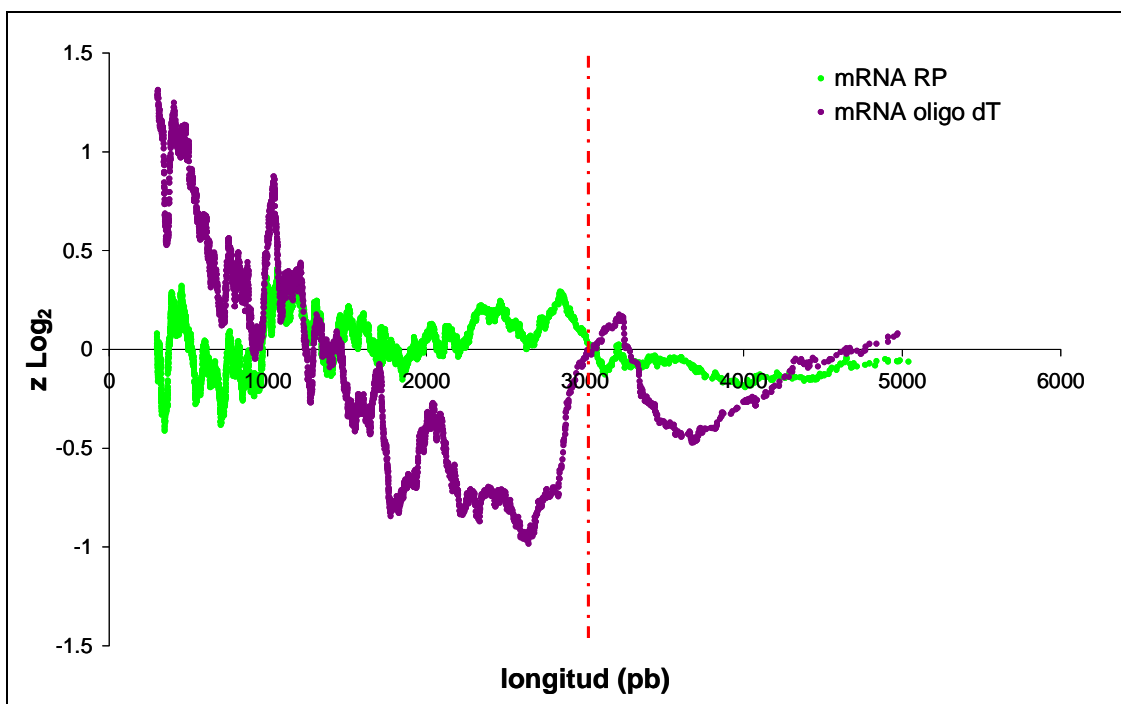


Figura 27. Sesgos en los marcajes de mRNAs en función de la longitud del gen usando diferentes cebadores para la transcripción reversa. Los datos de cDNA usando cebadores aleatorios (**mRNA RP**) provienen de 15 marcajes diferentes y los de cDNA marcados con oligo dT (**mRNA oligo dT**) provienen de 3 hibridaciones diferentes. Se representa media de una ventana deslizante de 100 genes.

Otro factor que puede tener efecto en el cálculo de las tasas de transcripción es el solapamiento de fragmentos marcados entre sondas de genes adyacentes. En el caso del RPCC, como se marcan fragmentos de DNA de un determinado intervalo de tamaños, la señal de polimerasa detectada en un determinado gen puede estar contaminada por la señal de las polimerasas de un gen adyacente y muy próximo. Este efecto se puede minimizar reduciendo el tamaño de la cromatina sonicada, pero al

disminuir el tamaño de la cromatina también se disminuye la eficiencia de la inmunoprecipitación. Por otra parte, en el caso del GRO, la transcripción detectada en un gen también puede estar influenciada por polimerasas de genes vecinos. Pero debido a la direccionalidad del marcaje del GRO, solo influirán los genes cercanos cuya transcripción esté orientada hacia la región codificante del gen. Sin embargo, al contrario que en el caso de la longitud del gen, en este caso resulta muy difícil averiguar que parte de la señal de un gen puede ser debida a otros genes muy próximos. Y, como el efecto de genes muy próximos afecta tanto el RPCC como el GRO, resulta muy difícil cuantificar el posible sesgo que este fenómeno introduce en las medidas de transcripción. De hecho, analizando la influencia que tienen genes altamente transcritos en la transcripción calculada para sus genes vecinos, no se pudieron detectar diferencias importantes que fueran dependientes de la orientación de los genes comparando los datos de GRO y RPCC.

### **Uso de diferentes anticuerpos para el RPCC**

Una de las ventajas del RPCC es que permitiría, usando diferentes anticuerpos, evaluar la tasa de transcripción por separado para las tres RNA polimerasas. Asimismo, usando anticuerpos contra diferentes formas de RNA pol II se puede estudiar en detalle el proceso de la transcripción.

Las moléculas de la RNA pol II, a lo largo de la transcripción, sufren cambios de fosforilación en el CTD, el dominio carboxiterminal de Rpb1p, que está formado por una serie de repeticiones de un heptapéptido y cuyo estado de fosforilación está relacionado con las diferentes etapas de la transcripción sirviendo de lugar de unión para múltiples proteínas (Komarnitsky P. *et al* 2000). En este estudio se han usado cuatro anticuerpos diferentes para inmunoprecipitar la RNA polimerasa II.

La primera opción ha sido usar una cepa (ARG3) que tiene la subunidad mayor de la RNA pol II etiquetada con 18 repeticiones del epítipo MYC en N-terminal (RPB1-Myc<sub>18</sub>). De esta manera, usando un anticuerpo anti-MYC se puede detectar todas las polimerasas independientemente del estado de fosforilación de su CTD. Otra opción que permite detectar la mayoría de polimerasas ha sido usar el anticuerpo 8WG16 (Jones J.C. *et al* 2004) que reconoce preferentemente el CTD no fosforilado. Este estado del CTD de la RNA pol II está presente preferentemente en polimerasas que se encuentran en la región del promotor y aun no han iniciado transcripción activa. Sin embargo, el anticuerpo 8WG16 también es capaz de unirse a polimerasas a lo largo del gen incluso aunque el CTD este

## Resultados y discusión

parcialmente fosforilado (Komarnitsky P. *et al* 2000). Por otra parte, también se han usado en este estudio anticuerpos contra fosforilaciones específicas del CTD. Concretamente un anticuerpo contra la fosforilación en Serina 5 del CTD da la RNA pol II que reconoce las polimerasas que han iniciado la transcripción y están empezando la elongación, y el anticuerpo H5 que reconoce el CTD fosforilado en Serina 2 que se asocia con la elongación y terminación de la transcripción.

La medida de RPCC usando la RNA pol II etiquetada (RPB1-MYC<sub>18</sub>) o el anticuerpo específico contra el dominio CTD hipofosforilado (8WG16) de células en crecimiento exponencial producen resultados muy similares (fig 28). Si se realiza un análisis convencional para detectar uniones significativamente diferentes entre los dos experimentos mediante test-t y fijando un 5% de falsos positivos (corrección FDR) sólo se detectan 20 genes significativamente enriquecidos y 31 empobrecidos en polimerasa con el CTD hipofosforilado. El análisis posterior de esos genes no conduce a ninguna categoría funcional estadísticamente sobrerrepresentada. Sin embargo, si se realiza un análisis de los sesgos de las distribuciones de los diferentes grupos funcionales (Al-Shahrour F. *et al* 2006) sí que se pueden detectar algunas diferencias, como un enriquecimiento en polimerasas hipofosforiladas en los genes RP (fig 28). Sin embargo, el enriquecimiento general en polimerasa hipofosforiladas que se pueden detectar para los genes muy transcritos (línea roja) no se debe únicamente a los genes RP. El estudio en detalle de estas diferencias está englobado dentro del apartado 3.1.3 de este trabajo.



Tabla 9. Correlaciones entre datos de transcripción medidos por RPCC o GRO. Se muestran las correlaciones de Pearson para tres experimentos de RPCC usando anticuerpos diferentes y para datos de GRO previamente publicados (García-Martínez J. *et al* 2004). Todos los datos provienen de experimentos realizados por triplicado a células en crecimiento exponencial en YPD.

Correlación	RPCC Rpb1-MYC	RPCC 8WG16	RPCC Ser5P
GRO	0.422	0.509	0.428
RPCC Ser5P	0.59	0.723	
RPCC 8WG16	0.651		
Correlaciones entre replicas usando la misma técnica			
GRO	RPCC Rpb1-MYC	RPCC 8WG16	RPCC Ser5P
0.934	0.705	0.751	0.716

## Resultados y discusión

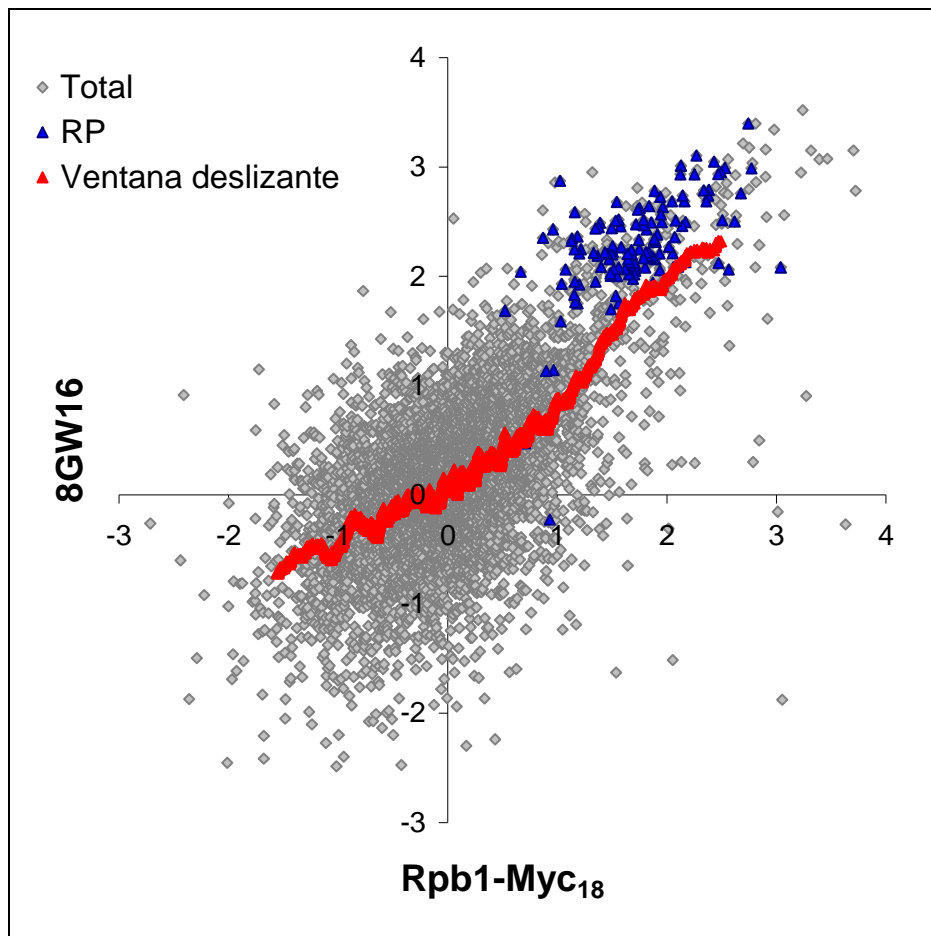


Figura 28. Comparación de RPCC usando polimerasa total (Rpb1-MYC) o un anticuerpo contra el CTD hipofosforilado (8WG16). En rojo se muestra una suavización usando la media de una ventana deslizante de 100 genes y en azul los genes de las proteínas ribosomales (RP) que son especialmente ricos en polimerasa hipofosforilada. Todos los datos representan el enriquecimiento por IP en  $\log_2$ .

El anticuerpo contra la forma fosforilada del CTD de la RNA pol II en Serina 5 produce unos resultados de RPCC muy similares a los obtenidos con los otros anticuerpos (tabla 9). Además, también se puede encontrar un enriquecimiento en los genes RP similar al encontrado en el caso del 8WG16 (apartado 3.1.3 de este trabajo). Por otra parte hemos observado que el anticuerpo H5 (frente al CTD fosforilado en Ser2) no produce un enriquecimiento suficientemente alto mediante inmunoprecipitación para obtener unos valores fiables de transcripción a escala genómica. Por lo tanto no se muestran datos referentes a él.

## **RPCC: ventajas e inconvenientes**

El RPCC presenta una serie de ventajas e inconvenientes como metodología alternativa al GRO para el cálculo de tasas de transcripción a nivel genómico.

Entre las ventajas podemos destacar que sería directamente aplicable al estudio de la transcripción en organismos superiores sin la necesidad de aislar núcleos al contrario de lo que ocurre en el caso del *run-on* (Hirayoshi K. y Lis, J.T. 1999). Además, al basarse en una técnica tan ampliamente utilizada como el ChIP-on-chip es fácilmente aplicable por diferentes laboratorios. Otra de las ventajas es que se pueden usar diferentes Ac para detectar la transcripción por separado de las tres RNA pol y de las diferentes formas fosforiladas del CTD de la RNA pol II.

Por el contrario, también presenta una serie de limitaciones. La principal limitación es que detecta la presencia física de la polimerasa y no su actividad. Y aunque estos dos factores suelen estar muy ligados, no siempre ocurre así (como discutiremos en el apartado 3.1.3 de este trabajo). Además, al basarse en un proceso de enriquecimiento selectivo de la muestra seguido de amplificación de la misma, da resultados con más ruido de fondo y más variabilidad que el GRO. Por otra parte, usando el GRO se pueden detectar diferencias más grandes entre los genes con mayor y menor TR (mayor rango dinámico) mientras que en RPCC las señales entre los diferentes genes son más similares. Esto se debe a que las señales en el GRO marcan directamente las diferencias de transcripción, mientras que en el RPCC estas diferencias están atenuadas por la necesidad de realizar un enriquecimiento selectivo de la muestra. Una tercera limitación es que el ChIP-on-chip sólo sitúa la polimerasa en una región del genoma de forma independiente de su orientación, mientras que el GRO produce RNAs marcados que si se hibridaran con un chip con sondas específicas de hebra permitiría determinar el sentido de la transcripción.

### **3.1.2. Correcciones de las tasas de transcripción del GRO usando los datos de RPCC**

Una vez determinado que existe un sesgo dependiente del tamaño de la sonda depositada en el chip, y que éste se debe a una peculiaridad técnica del GRO, se decidió realizar una corrección de los datos de TR obtenidos mediante GRO usando los de RPCC.

Para tratar de obtener unos valores de TR lo más representativos posible se generó un lote de datos con diferentes experimentos de GRO realizados en

## Resultados y discusión

nuestro laboratorio. Concretamente se usaron 8 grupos de resultados, realizados cada uno por triplicado, de GRO de cultivos en crecimiento exponencial en YPD realizados tanto en la cepa BQS252, (comunicación personal de Oreto Antúnez y José E Pérez Ortín) y este mismo trabajo, como de la W303 (comunicaciones personales de Lorena Romero, Paula Alepuz y Enrique Herrero). Una vez recopilados los resultados de GRO se generó un único grupo de datos usando el programa Arraystat y exigiendo para cada gen un mínimo de 5 réplicas válidas de las 8, con lo que sólo 4670 genes que superaron este filtro. Para poder comparar los datos de GRO y RPCC se realizó una estandarización de ambos, restándole a cada punto el promedio de la población y dividiéndolo por la desviación estándar. Una vez estandarizados los datos se calculó la suavización independientemente para las ORFs menores y mayores de 3000 pares de bases de los datos usando el algoritmo *lowess* y se corrigieron los resultados de GRO para que coincidieran con el perfil de la media de 3 grupos de experimentos diferentes de RPCC (fig 29). La correlación entre el conjunto de resultados de GRO y el de RPCC antes de la corrección por la longitud de la sonda era de 0.69 (Pearson), y ésta mejora a 0.7 tras la corrección. Posteriormente se deshizo la estandarización de los datos devolviéndole a la población su desviación estándar.

Finalmente, para transformar los valores de TR de unidades arbitrarias a moléculas por minuto, se normalizaron los datos asumiendo un estado estacionario para la transcripción y degradación de mRNA. Concretamente se utilizaron los datos de cantidad de mRNA calculados a partir de PCR competitiva y RNA-Seq (Miura F. *et al* 2008; Nagalakshmi U. *et al* 2008) y la estabilidad calculados en Wang Y. *et al.* (2002) para generar un conjunto alternativo de TR indirectas. Seguidamente, se calculó un factor de corrección igualando el valor promedio de TR de los 2000 genes con mayor transcripción entre GRO y los cálculos indirectos, y se aplicó este factor a los datos de GRO en unidades arbitrarias para transformarlos en datos en unidades reales.

De esta manera se han obtenido datos robustos de TR para 4670 genes en los que se ha corregido el efecto de la longitud de la sonda. Aunque existen grupos funcionales con cambios significativos tras la corrección (tabla 10) estos cambios son pequeños. Por ejemplo en el caso de los genes RP, que presentan la diferencia más significativa, la corrección supone sólo una disminución del 10% de la TR. Además estas diferencias parecen centrarse en aquellos grupos de genes que presentan mayoritariamente un tamaño similar y que, consecuentemente, son afectados de forma conjunta por la corrección aplicada. Por ejemplo, los genes RP tienen un tamaño promedio

de 744 nucleótidos con una desviación estandar de sólo 270 nucleótidos, comparado con el total del genoma que tiene un promedio de 1338 nucleótidos con una desviación estandar de 1100.

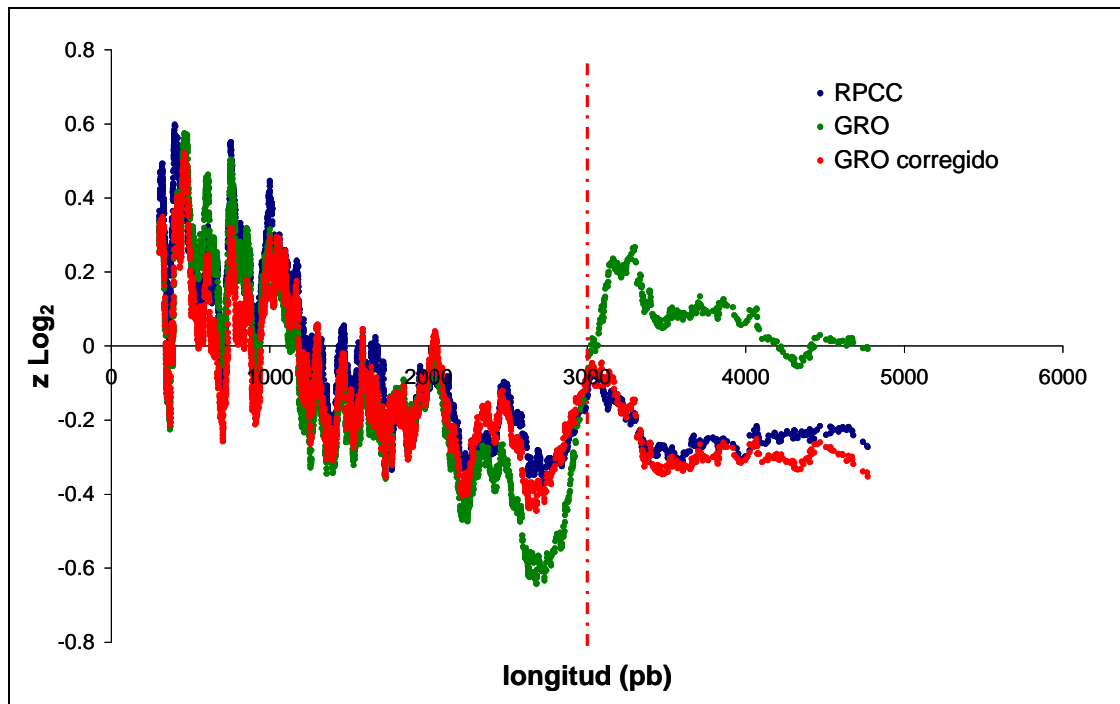


Figura 29. Corrección de la señal de GRO dependiente de la longitud de la sonda usando los datos de RPCC. Se muestran las suavizaciones por ventanas deslizantes de 100 genes de los valores de GRO antes (**verde**) y después (**rojo**) de la normalización por *lowess* respecto a los valores de RPCC (**azul**).

Tabla 10. Anotación funcional de las diferencias introducidas por la corrección por la longitud de los genes. Sólo se muestran las categorías funcionales más representativas con un p-valor ajustado menor de  $10^{-4}$ .

<u>Grupos funcionales cuya TR aumenta por la corrección</u>		<u>Grupos funcionales cuya TR disminuye por la corrección</u>	
<u>p-adj</u>	<u>Grupo funcional</u>	<u>p-adj</u>	<u>Grupo funcional</u>
$1.28 \cdot 10^{-7}$	Unión a DNA	$8.8 \cdot 10^{-27}$	Componente estructural del ribosoma
$3.47 \cdot 10^{-5}$	Parte mitocondrial	$1.36 \cdot 10^{-10}$	Unión a ATP
		$3.3 \cdot 10^{-7}$	Complejo central del Proteasoma
		$3.81 \cdot 10^{-5}$	Gema celular
		$6.77 \cdot 10^{-5}$	Secreción

Estos datos robustos de transcripción se pueden expresar también en unidades reales utilizando datos absolutos de cantidad (en

## Resultados y discusión

moléculas/célula) y estabilidad de los mRNAs (en minutos<sup>-1</sup>). Los datos de transcripción presentan una distribución log-normal, en el que el 90% de los genes tienen transcripciones entre 0.017 y 0.91 moléculas por minuto, y con una TR mediana de 0.1 mol/min (fig 30). Hay que aclarar que estos cálculos se basan en la asunción de que los valores de cantidad de moléculas de mRNA son del orden de 26000 moléculas por célula en condiciones de crecimiento exponencial. Estos valores han sido calculados a partir de los datos de secuenciación masiva de mRNAs (Nagalakshmi U. *et al* 2008) transformados a unidades reales usando los datos de cantidad de mRNA obtenidos mediante PCR competitiva a escala genómica (Miura F. *et al* 2008). Recientemente se han publicado también otros trabajos que apoyan que el valor de mRNAs por célula podría ser de entre 20000 y 60000 mRNAs/célula (Miura F. *et al* 2008; Zenklusen D. *et al* 2008). Sin embargo, todos estos datos son superiores a la estimación clásica de 15000 mRNAs por célula de Hereford L.M. y Rosbash M. (1977) que han sido ya usados en cálculos anteriores. Por lo tanto, asumiendo que las vidas medias son correctas, las TR obtenidas con estos nuevos datos serían también superiores a las calculadas anteriormente (Bon M. *et al* 2006; García-Martínez J. *et al* 2004).

Una vez calculada la tasa de transcripción en moléculas/minuto se pueden transformar estos datos a densidades de polimerasas. Si se asume una velocidad de transcripción constante para la RNA pol II de 25 nt/segundo (o 1.5 kb/min) (Edwards A.M. *et al* 1991) se puede calcular que un mRNA que se produzca a una tasa de una molécula por minuto tendría una densidad de 0.66 polimerasas por kb (fig 30). El hecho de asumir una velocidad constante para todos los genes es obviamente una simplificación que debe promediar el tiempo necesario para las diferentes etapas de la transcripción y asumir que esta velocidad es la misma para todos los genes. De hecho existen algunos estudios recientes que sugieren que diferentes genes podrían tener diferentes velocidades de transcripción. Así, por ejemplo, se postula que los genes más activamente transcritos tendrían RNA pol II capaces de avanzar a mayor velocidad (Zenklusen D. *et al* 2008). Por lo tanto, en esos casos, la densidad de RNA pol II que se calcule asumiendo una velocidad de 25 nt/segundo en realidad será diferente de la real.

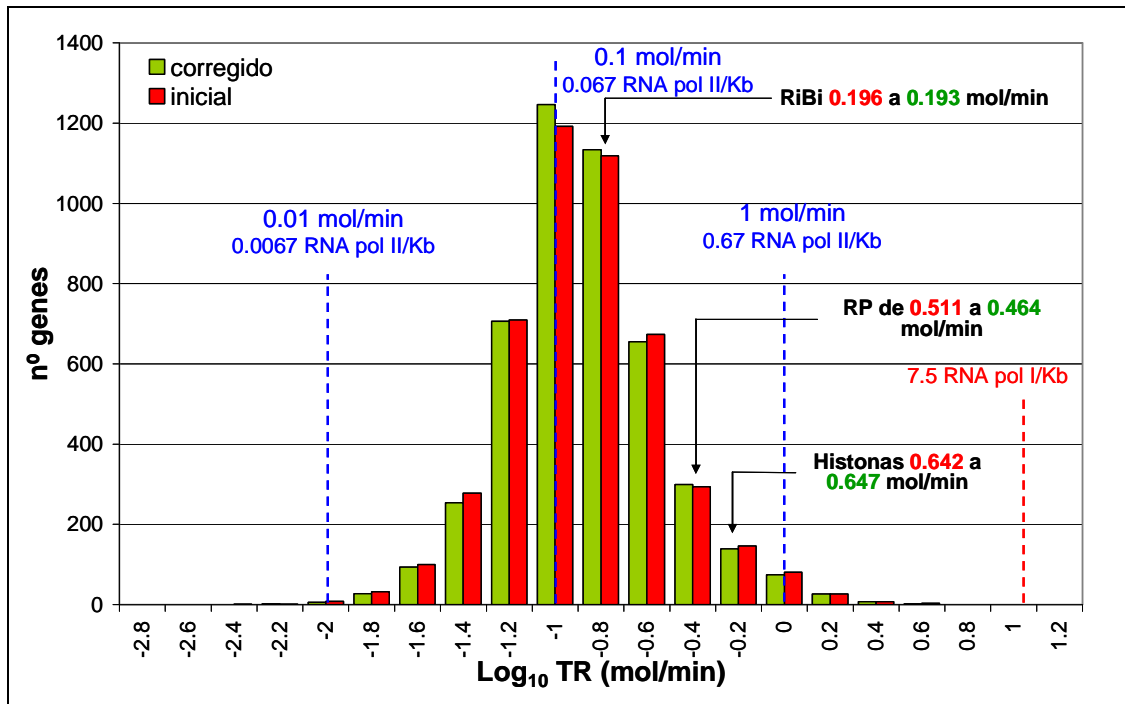


Figura 30. Datos de transcripción antes y después de aplicar una corrección por la longitud de los genes. Se muestran los datos de transcripción obtenidos a partir de un lote de 8 experimentos diferentes de GRO realizados por triplicado antes y después de aplicar la corrección por longitud del gen. Se resaltan la mediana de la transcripción para algunos grupos funcionales antes (**verde**) y después (**rojo**) de la corrección. En líneas punteadas se resaltan las tasas de transcripción calculadas para la RNA pol II (**azul**) y con una línea discontinua **roja** se muestra como referencia la densidad de RNA pol I en el rDNA (French S.L. *et al* 2003). La densidad Los valores se muestran en escala logarítmica.

Sin embargo, usando estas simplificaciones se puede calcular la densidad de RNA pol II de los 4670 genes para los que se disponen datos. En este caso, la densidad de polimerasas para los genes en un cultivo asincrónico en crecimiento exponencial tendría una mediana de 0.067 RNA pol II/kb lo que, teniendo en cuenta la longitud de los genes, corresponde a 0.082 RNA pol II/ORF. Esto implica que en un momento determinado la mayoría de los genes de una célula no tendría ninguna polimerasa activa en su ORF. De hecho, con estos datos, sólo el 0.005% de los genes tienen como promedio 1 o más polimerasas por ORF. Asumiendo que los 4670 datos de transcripción son representativos de los 5795 genes non-dubious

## Resultados y discusión

de levadura, se puede calcular que cada minuto en la célula habría alrededor de 856 eventos de transcripción productiva, lo que haría un total de unos 77000 eventos de transcripción productiva por ciclo celular. Como la cantidad de mRNA es de aproximadamente 26000 moléculas por célula esto implica que durante un ciclo celular hay un recambio en el que se producen y degradan el triple de la cantidad de moléculas presentes en un momento determinado. Estos resultados de recambio de mRNA que inicialmente pueden parecer sorprendentemente bajos, no lo son tanto. De hecho otros autores realizaron cálculos en este mismo sentido usando datos previos de cantidad y estabilidad de mRNAs concluyendo que muchos genes sólo son transcritos una única vez por ciclo celular y que por tanto para muchos genes las células hijas utilizan sólo mRNAs provenientes de las células madres (Bon M. *et al* 2006). Por otra parte, de estos resultados también se puede concluir que sólo el 9% de los eventos de transcripción de la RNA pol II son consecuencia de la transcripción de los genes RP, muy por debajo del 50% que se había estimado previamente (Warner J.R. 1999).

Si nos centramos en el caso extremo de las histonas, que son el grupo de genes más activamente transcritos cuya transcripción supone el 1% del total, podemos calcular una transcripción promedio de 0.64 mol/min que, asumiendo una velocidad de 25 nt/s, corresponde a una densidad de 0.43 RNA pol II/kb. Sin embargo, estos datos están obtenidos utilizando células de un cultivo asincrónico en el que sólo las células en fase S están en realidad transcribiendo los genes de las histonas. Como la proporción de células en fase S corresponde aproximadamente a un 27% del total (Niemisto A. *et al* 2007) la densidad real de polimerasas en las células con transcripción activa de histonas sería de 1.58 RNA pol II/kb (o 2.38 mol/min). Esta densidad, como era esperable, es inferior a la densidad de 7.5 moléculas de RNA pol I/kb que se puede detectar en los rDNAs en transcripción activa (French S.L. *et al* 2003). Si usamos estos datos de velocidad de transcripción de la RNA pol II en el contexto de la organización de la cromatina, una velocidad máxima de la RNA pol II de alrededor de 2.38 moléculas por minuto implica que cada 25 segundos una polimerasa con toda su maquinaria asociada pasa por un punto concreto de un gen de histonas en máxima transcripción. Si suponemos que en ese punto existe un nucleosoma posicionado, y que la polimerasa tarda unos 6 segundos en atravesar esa región, el nucleosoma sólo podrá estar posicionado durante unos 19 segundos hasta que llegara la siguiente polimerasa. Estos resultados implican que, incluso en los genes más



activamente transcritos, los nucleosomas pueden estar presentes un 75% del tiempo.

A partir de estos datos también se puede concluir que la situación en que dos polimerasas están transcribiendo el mismo gen es extremadamente improbable, y por lo tanto el hecho que se ayuden mutuamente en el caso de genes de alta transcripción para aumentar su velocidad (Zenklusen D. *et al* 2008) no puede ser mediado por la interacción directa de las RNA pol II, sino por las modificaciones que la transcripción tenga en otros factores, como la estructura de la cromatina.

### **3.1.3 Comparación entre diferentes aproximaciones genómicas para el estudio de la transcripción**

Una vez puesto a punto el RPCC, que permite calcular la densidad de polimerasa asociada físicamente a los genes, se decidió realizar una comparación detallada de los resultados de este método con los obtenidos mediante GRO, en el que sólo se obtienen señales de las polimerasas que son capaces de elongar en condiciones de *run-on*. Por lo tanto, las diferencias entre estas técnicas pueden revelar la diferencia biológica que existe entre la mera presencia de las polimerasas unidas a los diferentes genes y su actividad elongante.

#### **Estudio funcional de las diferencias entre GRO y RPCC**

En el apartado anterior de este trabajo se han estudiado los posibles sesgos técnicos asociados a las técnicas genómicas de medida de la transcripción y se ha realizado una corrección de los datos de GRO en función de la longitud de los genes. Sin embargo, incluso después de esta corrección siguen existiendo diferencias entre el RPCC y el GRO. Para determinar si estas diferencias tienen significado biológico se realizó un estudio de anotación funcional de las diferencias. Además, como el GRO mide las RNA pol en transcripción activa y el RPCC mide la presencia de la polimerasa en los genes, la comparación entre ambos aporta información sobre la cantidad de polimerasas que están presentes en los genes respecto a las que están en transcripción activa en los mismos.

Para poder comparar estos resultados entre sí se estandarizaron los resultados (restándole a cada valor individual la media de la población y dividiéndolo por la desviación estándar) y se realizó una anotación funcional de las diferencias usando la herramienta bioinformática Fatican (Al-Shahrour F. *et al* 2006). Estos análisis se realizaron usando datos de RPCC obtenidos con tres anticuerpos diferentes: contra la polimerasa etiquetada (Rpb1-MYC), el CTD hipofosforilado (8WG16) y el CTD fosforilado en Serina5.

Como ya se ha explicado anteriormente, se puede observar que existe una buena correlación general entre los datos de GRO y los datos de RPCC usando los tres anticuerpos diferentes (fig 24). Sin embargo, en los tres casos se puede detectar que los genes relacionados con los ribosomas y la membrana están enriquecidos en presencia de polimerasa medida por RPCC frente a la tasa de transcripción medida por GRO (relación RPCC/GRO alta). Por el contrario, los genes relacionados con el núcleo o

la biogénesis de ribosomas presentan menos polimerasa que la esperable por su señal de GRO. Estas diferencias son especialmente importantes cuando se usan los datos de RPCC realizado con el anticuerpo Ser 5 y con el 8WG16 (tabla 11).

Tabla 11. Grupos funcionales cuya relación entre RPCC y GRO es significativamente diferente. Se representa los grupos funcionales más representativos para las diferencias entre datos de RPCC usando tres anticuerpos diferentes y los datos de GRO corregidos en el apartado anterior.

<p><b><u>Grupos funcionales enriquecidos en RNA pol II total respecto a la Transcripción</u></b>  <b><u>Relación (Rpb1-MYC/GRO) alta</u></b></p>		<p><b><u>Grupos funcionales empobrecidos en RNA pol II total respecto a la Transcripción</u></b>  <b><u>Relación (Rpb1-MYC/GRO) baja</u></b></p>	
<u>p-adj</u>	<u>Grupo funcional</u>	<u>p-adj</u>	<u>Grupo funcional</u>
$2.07 \cdot 10^{-10}$	Membrana	$3.38 \cdot 10^{-13}$	Núcleo
$3.15 \cdot 10^{-4}$	Pared celular	$3.95 \cdot 10^{-9}$	Orgánulo unido a membrana
		$3.95 \cdot 10^{-7}$	Procesos del metabolismo del RNA
		$3.44 \cdot 10^{-6}$	Transcripción
		$2.01 \cdot 10^{-4}$	Biogénesis y ensamblaje de complejos ribonucleoproteicos
<p><b><u>Grupos funcionales enriquecidos en CTD hipofosforilado respecto a la Transcripción</u></b>  <b><u>Relación (8WG16/GRO) alta</u></b></p>		<p><b><u>Grupos funcionales empobrecidos en CTD hipofosforilado respecto a la Transcripción</u></b>  <b><u>Relación (8WG16/GRO) baja</u></b></p>	
<u>p-adj</u>	<u>Grupo funcional</u>	<u>p-adj</u>	<u>Grupo funcional</u>
$3.24 \cdot 10^{-10}$	Membrana	$5.52 \cdot 10^{-14}$	Núcleo
$2.15 \cdot 10^{-5}$	Componente citoplásmico	$3.48 \cdot 10^{-7}$	Transcripción
$3.29 \cdot 10^{-4}$	Componente estructural del Ribosoma	$7.51 \cdot 10^{-7}$	Orgánulo unido a membrana
		$9.62 \cdot 10^{-4}$	Biogénesis y ensamblaje de complejos ribonucleoproteicos
<p><b><u>Grupos funcionales enriquecidos en CTD fosforilado en Ser5 respecto a la Transcripción</u></b>  <b><u>Relación (Ser5P/GRO) alta</u></b></p>		<p><b><u>Grupos funcionales empobrecidos en CTD fosforilado en Ser5 respecto a la Transcripción</u></b>  <b><u>Relación (Ser5P/GRO) baja</u></b></p>	
<u>p-adj</u>	<u>Grupo funcional</u>	<u>p-adj</u>	<u>Grupo funcional</u>
$2.88 \cdot 10^{-9}$	Componente citoplásmico	$4.88 \cdot 10^{-6}$	Orgánulo unido a membrana
$6.69 \cdot 10^{-8}$	Ribosoma	$8.16 \cdot 10^{-4}$	complejo del Proteasoma
$1.11 \cdot 10^{-7}$	Componente estructural del Ribosoma		
$6.24 \cdot 10^{-7}$	Integral de membrana		

Estas diferencias también se pueden observar si se representan directamente los datos de RPCC frente a los datos de GRO (figuras 31 y 32).

## Resultados y discusión

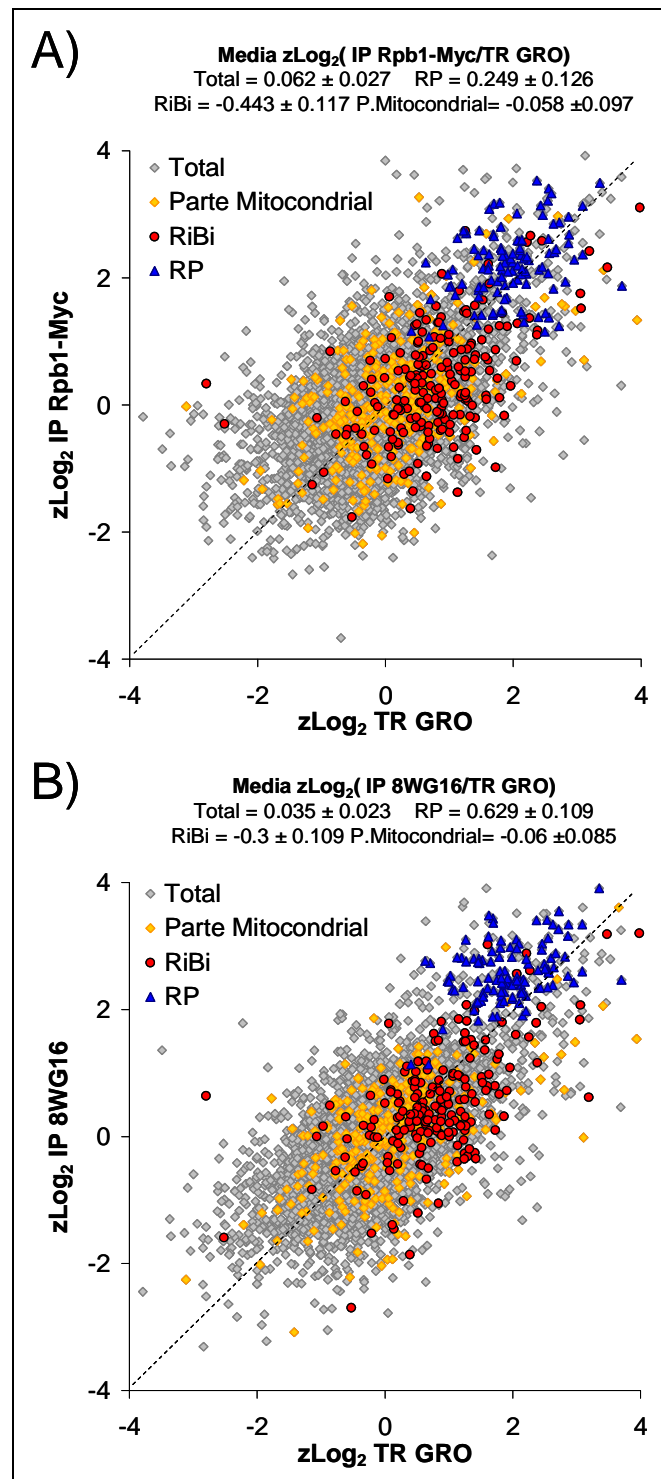


Figura 31. Comparación entre los datos de transcripción medidos por GRO y RPCC (I). Se representan los valores de RPCC usando un anticuerpo contra la polimerasa total etiquetada (Rpb1-MYC) (A) y un Ac contra el CTD hipofosforilado (8WG16) (B). Se resaltan los datos para algunos grupos funcionales y se indica el intervalo de confianza al 95% para la media de la relación RPCC/GRO de estos mismos grupos. Todos los datos se representan estandarizados (*z-scores*) y en escala logarítmica.

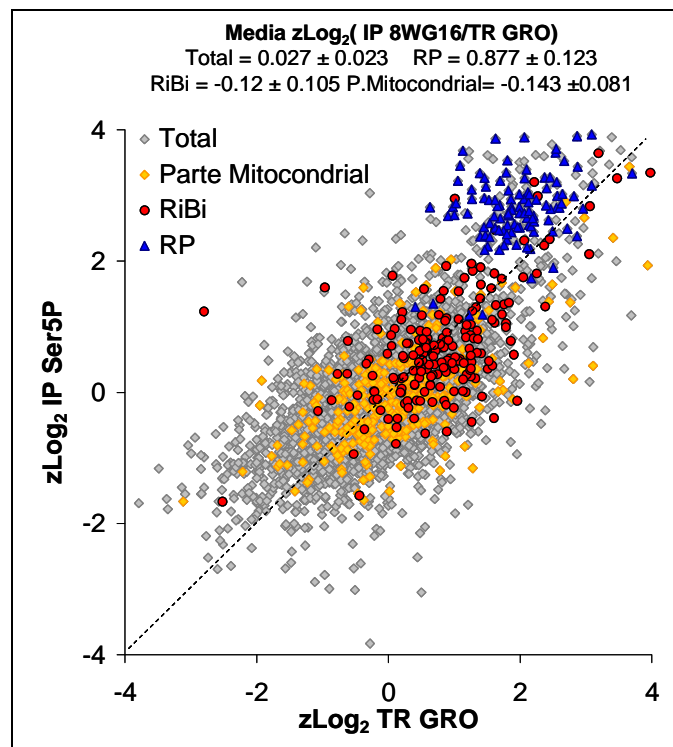


Figura 32 Comparación entre los datos de transcripción medidos por GRO y RPCC (II). Se representan los valores de RPCC usando un anticuerpo contra el CTD fosforilado en Ser2. Se resaltan los datos para algunos grupos funcionales y se indica el intervalo de confianza al 95% para la media de la relación RPCC/GRO de estos mismos grupos. Todos los datos se representan estandarizados (z-scores) y en escala logarítmica.

De las diferencias que se encontraron, una de las más interesantes es el enriquecimiento en maquinaria transcripcional de los genes relacionados con los ribosomas (RP). Además, estos genes también presentan diferencias cuando se comparan entre sí los datos de RPCC usando diferentes anticuerpos. Concretamente, estos genes están especialmente enriquecidos en polimerasa con el CTD hipofosforilado respecto a la polimerasa total (tabla 12 y fig 33).

## Resultados y discusión

Tabla 12. Grupos funcionales con diferencias para el estado de fosforilación del CTD de la RNA pol II. Se representan los grupos funcionales con diferencias entre la IP usando anticuerpos específicos contra estados de fosforilación del CTD (8WG16 y Ser5P) y IP de la polimerasa total (Rpb1-MYC). Sólo se muestran aquellos grupos más representativos.

<b><u>Grupos funcionales enriquecidos en CTD hipofosforilado</u></b> <b><u>Relación (8WG16/Rpb1-MYC) alta</u></b>		<b><u>Grupos funcionales empobrecidos en CTD hipofosforilado</u></b> <b><u>Relación (8WG16/Rpb1-MYC) baja</u></b>	
<u>p-adj</u>	<u>Grupo funcional</u>	<u>p-adj</u>	<u>Grupo funcional</u>
4.76 · 10 <sup>-13</sup>	Componente estructural de ribosoma	0.0180	Interacción de comportamiento entre organismos
3.18 · 10 <sup>-8</sup>	Subunidad grande del ribosoma citosólico	0.0180	Comportamiento reproductivo
9.4 · 10 <sup>-7</sup>	Biosíntesis de proteínas	0.0256	Respuesta a feromonas durante la conjugación
0.0003	Subunidad pequeña del ribosoma citosólico	0.0285	Integral de membrana
0.0043	Unión a ácido nucleico		
0.0166	Ensamblaje de subunidades ribosómicas		
0.0301	Nucleolo		
<b><u>Grupos funcionales enriquecidos en CTD fosforilado en Ser5</u></b> <b><u>Relación (Ser5P/Rpb1-MYC) alta</u></b>		<b><u>Grupos funcionales empobrecidos en CTD fosforilado en Ser5</u></b> <b><u>Relación (Ser5P/Rpb1-MYC) baja</u></b>	
<u>p-adj</u>	<u>Grupo funcional</u>	<u>p-adj</u>	<u>Grupo funcional</u>
4.76 · 10 <sup>-6</sup>	Unión a ácidos nucleicos	0.0006	Pared celular con quitina y β-glucano
0.0014	Componente nuclear	0.0023	Pared celular
0.0026	Lumen de orgánulos	0.0039	Componente citosólico
0.0182	Ensamblaje de complejos macromoleculares	0.0136	Actividad oxidoreductasa
0.0239	Biogénesis y ensamblaje del complejo ribonucleoproteico	0.0146	Peroxisoma
0.0239	Proceso metabólico de ácidos nucleicos	0.0312	Procesos metabólicos de aminoácidos y derivados
0.0485	Proceso metabólico de biopolímeros		

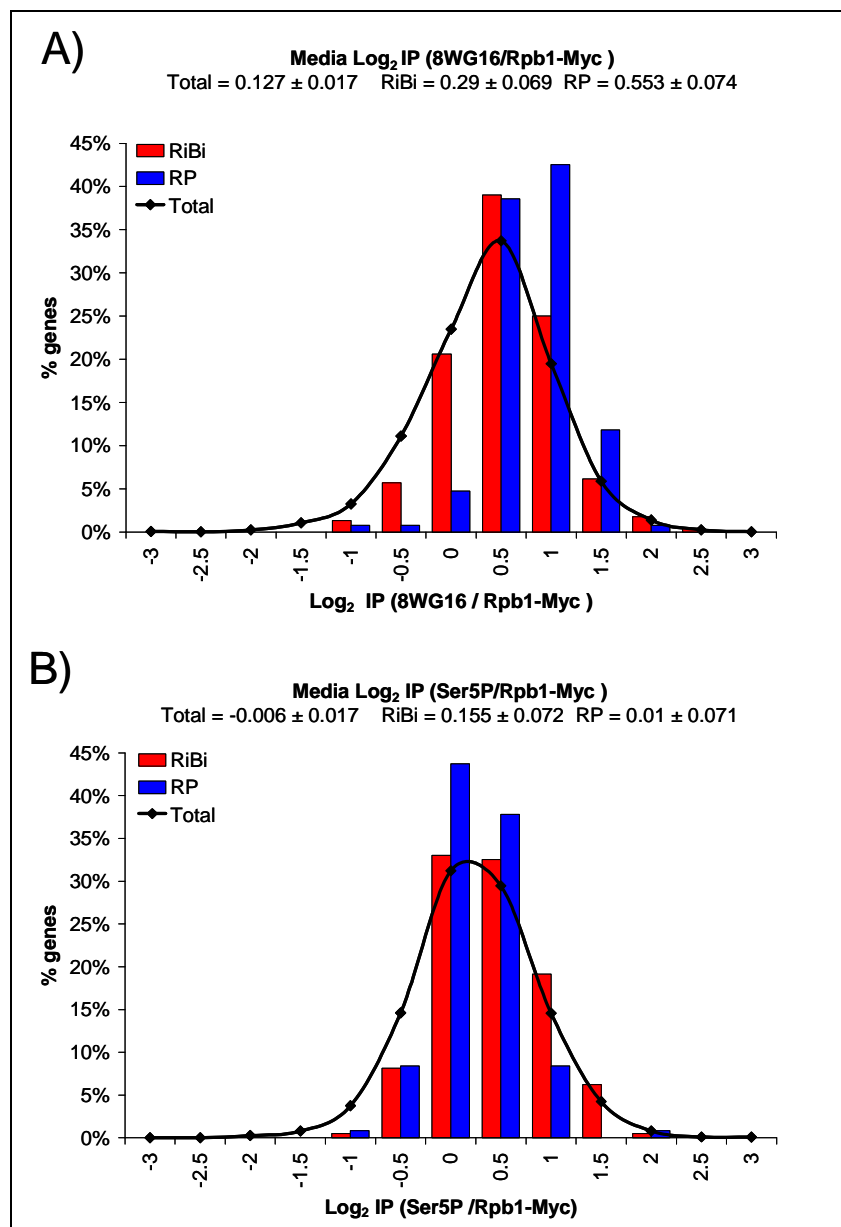


Figura 33. Comparación del estado de fosforilación del CTD respecto a la polimerasa total. A) Comparación del RPCC usando un anticuerpo contra el CTD hipofosforilado (8WG16) frente a la polimerasa total etiquetada (Rpb1-MYC). B) Comparación usando un anticuerpo contra el CTD fosforilado en Serina 5 (Ser5P) frente a la polimerasa total etiquetada (Rpb1-MYC). En ambos casos se representan los genes **RP** y **RiBi** en barras y el global de la población con una línea.

Por lo tanto parece que algunos genes enriquecidos en polimerasa lo están especialmente en polimerasa con el CTD hipofosforilado.

## Resultados y discusión

### **Comparación RPCC/GRO en diferentes condiciones de crecimiento**

Para confirmar que las diferencias entre RPCC y GRO reflejan en realidad diferencias biológicas en la forma de transcribir los genes, y no son el resultado de sesgos intrínsecos de las técnicas, se decidió ampliar el estudio a diferentes condiciones de crecimiento celular. De esta manera, si se diera una variación en la relaciones RPCC/GRO entre diferentes condiciones experimentales, ésta no se deberá a artefactos en los métodos de medida (dado que son los mismos en todos los casos), sino que será consecuencia de un cambio en la forma de transcribir los genes por parte de las células.

Para realizar este experimento se partió de células crecidas en medio rico con glucosa como fuente de carbono (YPD), que son las condiciones estándar de crecimiento en laboratorio, y se cambiaron a un medio rico con galactosa como fuente de carbono (YPGal). Este cambio de fuente de carbono hace que las células reorganicen todo su metabolismo para adaptarse a las nuevas condiciones de crecimiento. Además, como las diferencias encontradas entre las dos técnicas se centran en los genes de las proteínas ribosómicas, que son un grupo cuya regulación está muy relacionada con los nutrientes del medio y la velocidad de crecimiento celular (Jorgensen P. *et al* 2004; Regenber B. *et al* 2006) este tipo de experimento puede reflejar bien las variaciones que experimenten estos genes.

Se analizaron 3 puntos experimentales:

- Crecimiento exponencial en YPD, que es la condición óptima de crecimiento.
- A las dos horas después del cambio de glucosa a galactosa, donde las células han detenido su crecimiento y se están adaptando a la nueva situación.
- Crecimiento exponencial en galactosa, donde las células se han adaptado a la fuente de carbono y crecen activamente (tras 14.5 horas en YPGal).

Con el fin de resaltar las posibles diferencias se usó el anticuerpo 8WG16 para los experimentos de RPCC, dado que que permite detectar mayores diferencias entre GRO y RPCC.



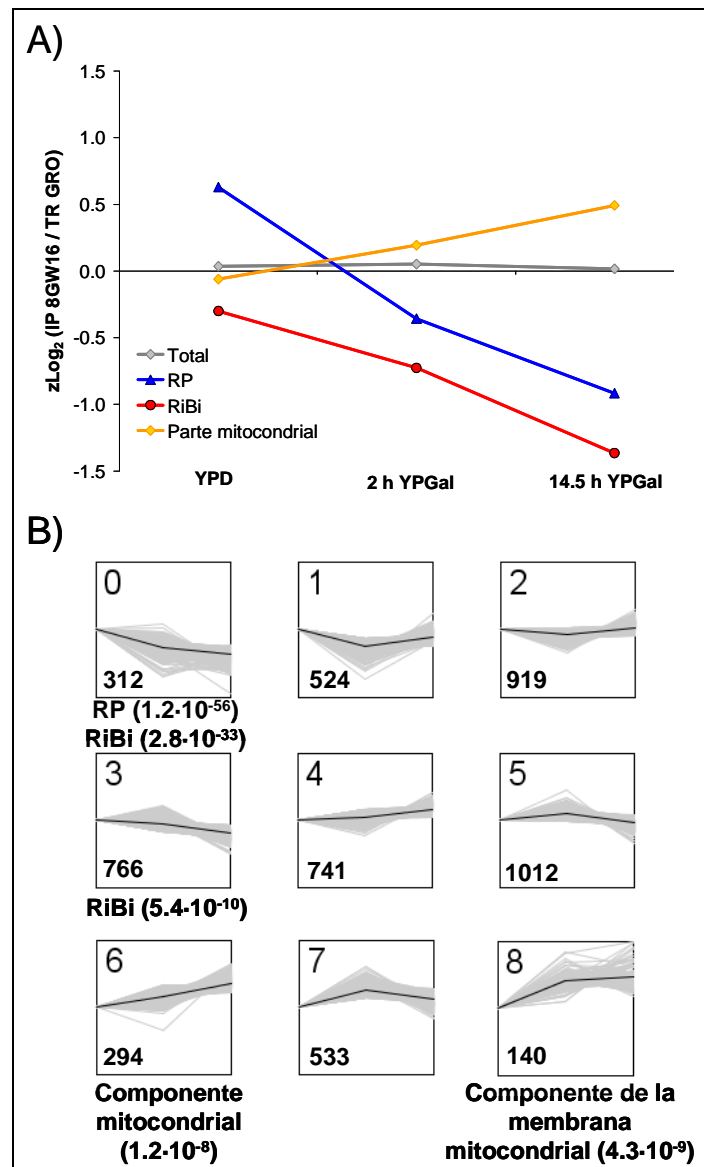


Figura 34. Relación RPCC/GRO a lo largo de la transición glucosa galactosa. A) valores medios de la relación RPCC/GRO, usando el Ac 8WG16, de diferentes grupos funcionales a lo largo de la transición de glucosa galactosa a galactosa. Los valores de GRO han sido corregidos usando los datos calculados en el apartado anterior. B) análisis de conglomerados de las relaciones RPCC/GRO en crecimiento exponencial en glucosa (YPD), a las 2 horas tras el cambio a galactosa (sin crecimiento celular) y en crecimiento exponencial en galactosa (14.5 horas). Los genes se han agrupado en 9 conglomerados mediante el algoritmo k-medias representando las diferencias respecto al crecimiento en YPD. Se muestra tanto el número de genes que forman cada conglomerado como la significación estadística (p-valor) de algunos grupos funcionales.

## Resultados y discusión

Este experimento permitió detectar cambios de la relación RPCC/GRO entre las diferentes condiciones fisiológicas. De hecho, el exceso en maquinaria transcripcional (relación RPCC/GRO alta) detectado en los genes RP durante el crecimiento en YPD, desaparece a las dos horas del cambio a YPGal y se invierte cuando las células crecen exponencialmente en galactosa (fig 34). Los genes RiBi, a pesar de que parten de valores bajos de RPCC/GRO, disminuyen su relación RPCC/GRO de forma similar a los RP. Por lo que concluimos que existe un fenómeno biológico regulable tras estas diferencias, y que éste se ve afectado por el cambio de fuente de carbono. Para confirmar que estas variaciones son significativas se realizó un análisis de conglomerados usando el algoritmo k-medias y el programa STEM (Ernst J. y Bar-Joseph, Z. 2006). Se agruparon los perfiles de RPCC/GRO para las diferentes condiciones fisiológicas en 9 conglomerados y se estudió si existía enriquecimiento significativo para algún grupo funcional (fig 34).

Tabla 13. Anotación funcional de los conglomerados para el cambio de la relación RPCC/GRO en la transición YPD-YPGal. Sólo se muestran las categorías funcionales con p-valor ajustado menor de 0.001 más representativas. Todos los cálculos han sido realizados con el programa STEM (Ernst J. y Bar-Joseph,Z. 2006)

<b>Enriquecimiento significativo en los diferentes conglomerados</b>			
conglomerado	p-valor	p-adj	Categorías funcionales
0	$1.2 \cdot 10^{-56}$	<0.001	Ribosoma citosólico ( <i>sensu Eukaryota</i> )
	$3.5 \cdot 10^{-39}$	<0.001	Componente estructural del ribosoma
	$2.8 \cdot 10^{-33}$	<0.001	Biogénesis y ensamblaje de ribosomas
	$1.3 \cdot 10^{-30}$	<0.001	Traducción
	$1.7 \cdot 10^{-9}$	<0.001	Procesado de rRNA
3	$1.2 \cdot 10^{-10}$	<0.001	Procesos del metabolismo del rRNA
	$5.4 \cdot 10^{-10}$	<0.001	Biogénesis y ensamblaje de complejos ribonucleoprotéicos
	$2.2 \cdot 10^{-9}$	<0.001	Biogénesis y ensamblaje de ribosomas
6	$1.2 \cdot 10^{-8}$	<0.001	Componente mitocondrial
	$2.9 \cdot 10^{-7}$	<0.001	Fosforilación oxidativa
	$5.7 \cdot 10^{-7}$	<0.001	Mitocondria
	$4 \cdot 10^{-6}$	<0.001	Síntesis de ATP acoplada a transporte de protones
8	$1.5 \cdot 10^{-15}$	<0.001	Generación de metabolitos precursores y energía
	$6.4 \cdot 10^{-10}$	<0.001	Procesos del metabolismo de carbohidratos
	$4.3 \cdot 10^{-9}$	<0.001	Componente de la membrana mitocondrial
	$1.9 \cdot 10^{-8}$	<0.001	Fosforilación oxidativa
	$7.1 \cdot 10^{-7}$	<0.001	Síntesis de ATP acoplada a transporte de electrones

De esta forma se confirmó la existencia de diferencias significativas en la relación RPCC/GRO entre las diferentes condiciones fisiológicas. Entre estas diferencias se puede destacar que los genes de proteínas ribosómicas (RP) y de Biogénesis de Ribosomas (RiBi) disminuyen su relación RPCC/GRO al cambiar de crecimiento en glucosa a galactosa. Por el contrario los genes relacionados con la mitocondria y la respiración oxidativa presentan un comportamiento opuesto con un valor RPCC/GRO elevado en crecimiento en galactosa y bajo en glucosa (tabla 13).

Hay que tener en cuenta que todas las variaciones de la relación RPCC/GRO hacen referencia a variaciones respecto al comportamiento promedio de la población. Estas variaciones se expresan respecto al global

## Resultados y discusión

del genoma porque ni el RPCC ni el GRO disponen de referencias internas que permitan expresar sus resultados directamente en valores absolutos. Por lo tanto, al comparar valores de RPCC y GRO para las mismas condiciones se hizo referencia al comportamiento general de la población. Un valor elevado en la relación RPCC/GRO en un gen, informa que éste presenta más maquinaria transcripcional asociada (unida físicamente y detectada por RPCC) de la que cabría esperar por la TR medida por GRO. Por el contrario, un valor bajo en la relación RPCC/GRO hacer referencia a genes que presentan menos maquinaria no activa transcripcionalmente que el promedio de la población. Sin embargo, es importante remarcar que la maquinaria físicamente unida a los genes (detectada por RPCC) siempre será mayor o igual que la que es transcripcionalmente activa (detectable por GRO), y cuando hablamos de genes con más polimerasa detectada por GRO que por RPCC (relación RPCC/GRO baja) siempre es en referencia al promedio del genoma.

A pesar de que no existen controles internos que permitan comparar los valores absolutos de RPCC y GRO entre sí, sí se pueden comparar los resultados de diferentes experimentos usando la misma técnica (discutido en el apartado 3.1.1 de este trabajo). De esta forma se puede estudiar como varían por separado los valores de RPCC y GRO para las diferentes condiciones y observar como los genes relacionados con el ribosoma y la mitocondria se apartan del comportamiento global de la población (fig 35). Así, se puede ver que, para los genes RP, la presencia de polimerasa medida por RPCC en glucosa respecto a la media de la población es mayor que la diferencia medida por GRO. Y por contra, en galactosa se puede observar un comportamiento opuesto (el valor de RPCC es más similar a la media, mientras que el de GRO es mucho más grande).

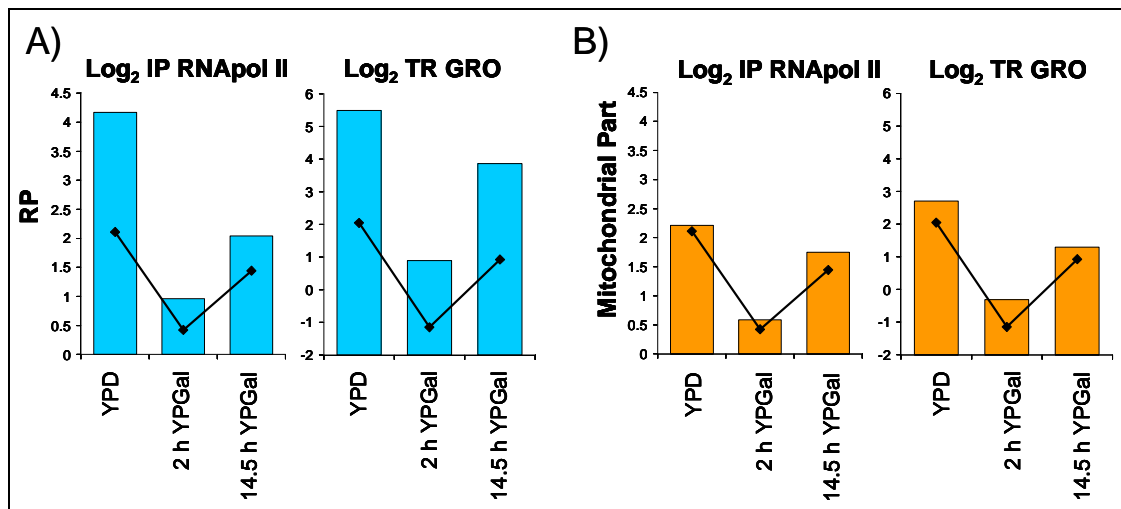


Figura 35. Datos de transcripción calculados por RPCC y GRO para los grupos de proteínas ribosomales y genes de respiración en diferentes situaciones fisiológicas. Datos medidos por RPCC usando el anticuerpo 8WG16 y medidos por GRO. Se representan los datos para los genes de las proteínas ribosomales, **RP**(A) y para los genes de **respiración** (B). En los dos casos se representa con una línea negra el valor medio de transcripción para el global del genoma y mediante barras el valor medio para el grupo especificado. Las escalas se representan en unidades arbitrarias en escala logarítmica que sólo permite comparación directa entre los experimentos realizados con la misma técnica.

Estas diferencias también se pueden observar si se comparan las relaciones RPCC/GRO para todos los genes en las diferentes condiciones de crecimiento (fig 36 y 37).

## Resultados y discusión

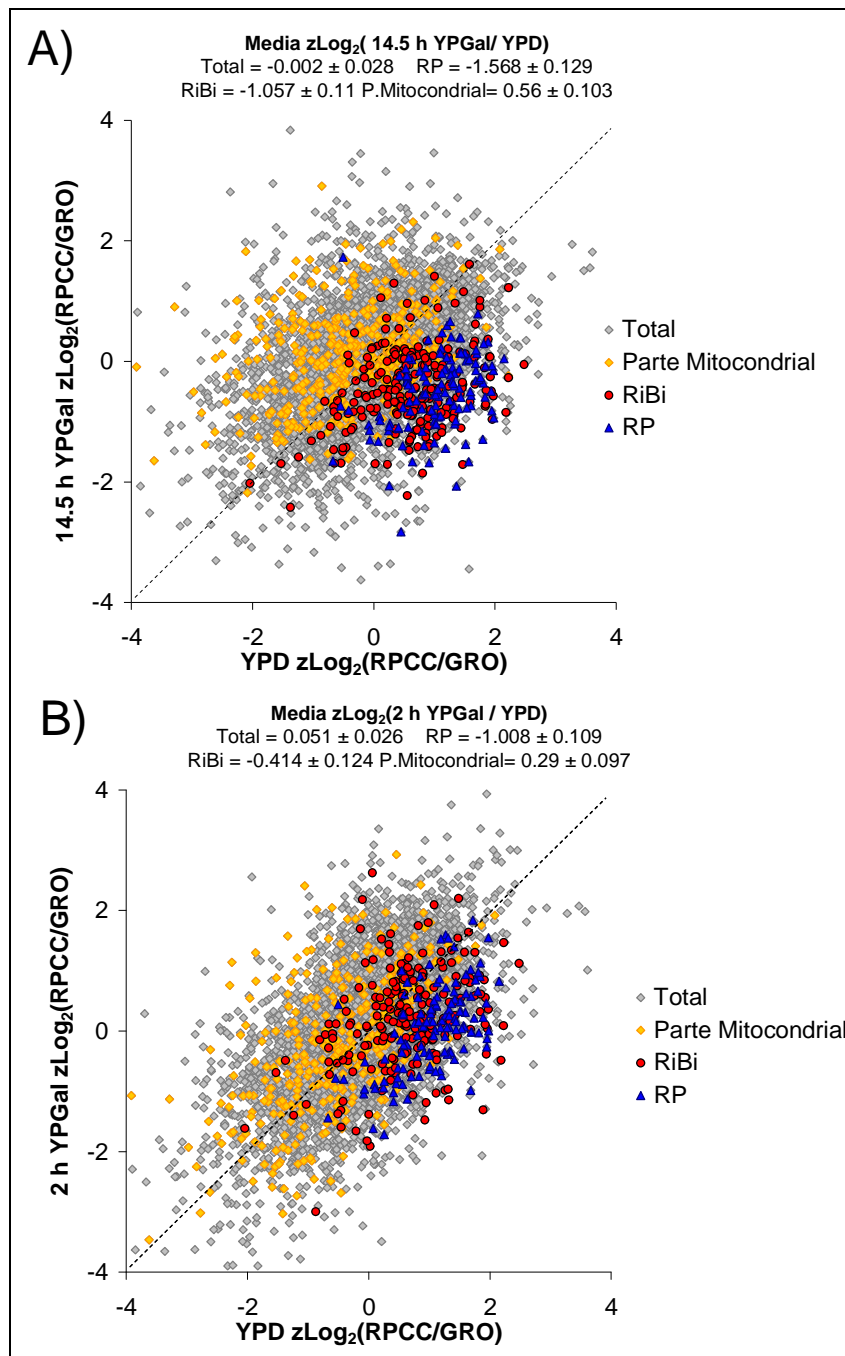


Figura 36. Relación RPCC/GRO de diferentes situaciones fisiológicas respecto a crecimiento en YPD. Comparación de los datos de: A) células paradas tras 2 h en galactosa (2 h YPGal) con células creciendo en glucosa (YPD). B) células creciendo en galactosa (14.5 h YPGal) con células creciendo en glucosa (YPD). Todos los valores hacen referencia a la relación RPCC/GRO en  $\log_2$  y normalizada por z-scores.

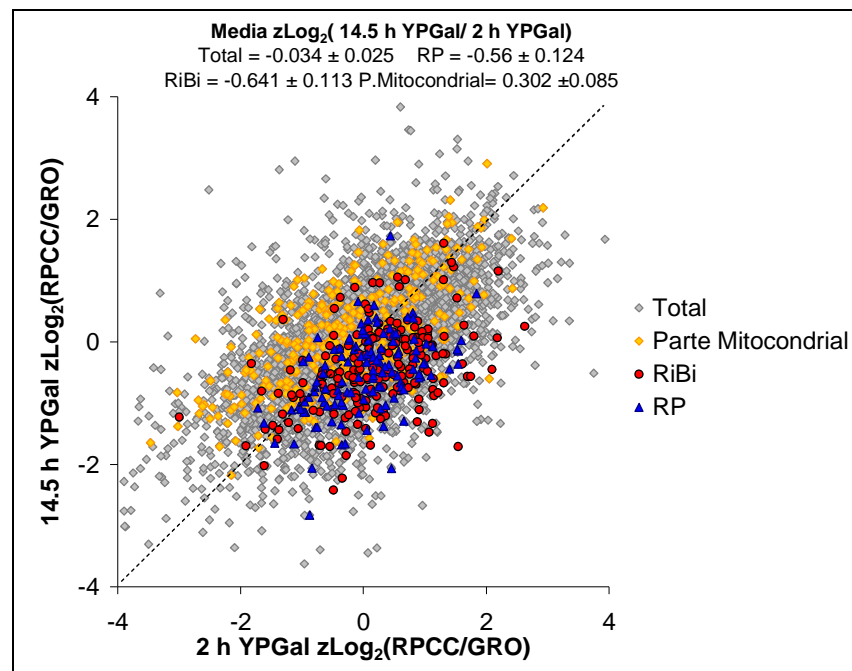


Figura 37. Relación RPCC/GRO entre 2 y 14.5 h en presencia de galactosa. Se comparan los datos de células creciendo en galactosa (14.5 h YPGal) con células paradas tras el cambio de fuente de carbono (2.5 h YPGal). Todos los valores hacen referencia a la relación RPCC/GRO en  $\log_2$  y normalizada por z-scores.

Si se analizan por separado las tasas de transcripción aparentes calculadas por cada método (fig 38), se puede observar que la nube de puntos en el caso de los RPCC es mucho más dispersa que en el caso del GRO. Esto se puede deber a que el RPCC, al basarse en un proceso de enriquecimiento selectivo, es una técnica con mayor variabilidad y más sujeta a introducir ruido que el GRO.

Por otra parte, observando el comportamiento de las categorías funcionales, se puede ver que la mayor parte de la variación de RPCC/GRO se debe al cambio en la polimerasa unida detectada por RPCC. De esta forma, los genes RP y RiBi presentan valores bajos de unión de polimerasa relativa (RPCC) en YPGal, mientras que las variaciones en la tasa de transcripción respecto al promedio de la población medidas por GRO apenas disminuyen. Por el contrario, los genes relacionados con la mitocondria presentan un importante incremento de unión de polimerasa (RPCC) en YPGal respecto al promedio de la población, mientras que apenas cambian en los experimentos de GRO.

## Resultados y discusión

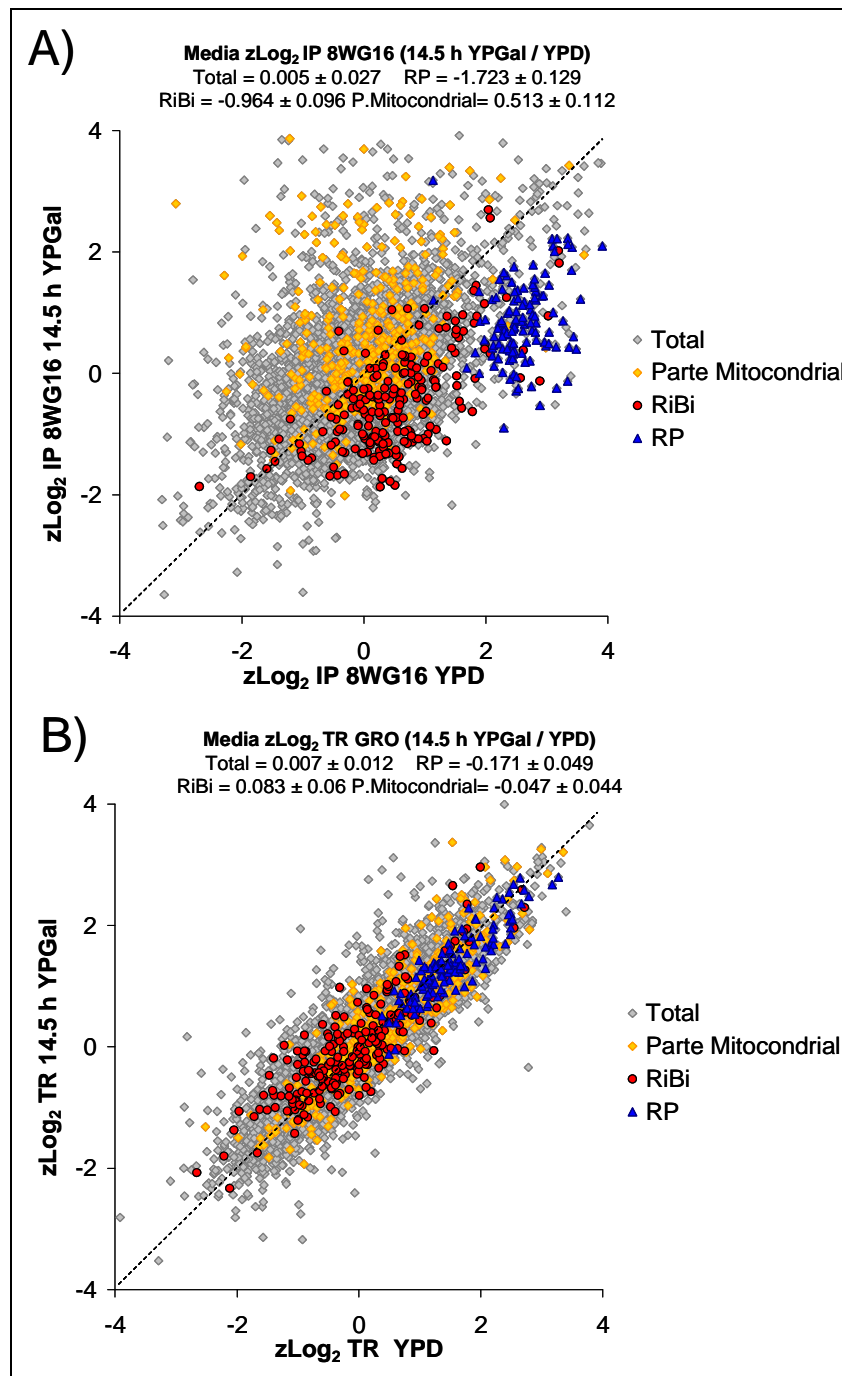


Figura 38. Comparación de tasas de transcripción para crecimiento exponencial en galactosa o glucosa. A) RPCC realizado con el Ac 8WG16. B) Datos de GRO para ambas condiciones. En todos los casos se han resaltado los grupos funcionales analizados con detalle en este estudio.

El estudio de las relaciones entre RPCC y GRO ha permitido, pues, descubrir que existen diferencias entre la cantidad de RNA pol II transcripcionalmente activa presente en los genes y la cantidad de RNA pol



total unida a los mismos. Además, estos *excesos* o *defectos* de maquinaria transcripcional (siempre respecto al global del genoma) son específicos de determinados grupos funcionales y se comportan de manera dinámica respondiendo a cambios fisiológicos de las células. Todo esto apunta a que la proporción de polimerasas unida no activas transcripcionalmente, es un parámetro regulado por la célula.

### **Regulación de la acumulación de maquinaria transcripcional en los genes RP**

El grupo funcional que presenta variaciones más claras para la acumulación de maquinaria transcripcional durante el cambio de fuente de carbono es el de los genes de las proteínas ribosómicas (RP). Estos genes presentan una transcripción elevada que a su vez está muy regulada y vinculada con la velocidad de crecimiento celular (García-Martínez J. *et al* 2007; García-Martínez J. *et al* 2004; Jorgensen P. *et al* 2004; Warner J.R. 1999). Por lo tanto, es un grupo de genes interesante para estudiar en detalle el fenómeno de la acumulación de maquinaria transcripcional.

Inicialmente, se podría pensar en asociar el enriquecimiento en maquinaria transcripcional no activa con una situación de baja transcripción, como en caso que las polimerasas estuvieran prerreclutadas en los promotores de los genes pero no iniciaran transcripción activa. Sin embargo, el exceso en RNA pol II en los genes de las proteínas ribosómicas ocurre cuando estos genes están transcribiéndose a máxima velocidad en condiciones de crecimiento óptimas (YPD). Por otra parte, los genes respiratorios acumulan maquinaria inactiva durante el crecimiento en galactosa, donde su transcripción también es más elevada. Es decir, esos dos grupos de genes acumulan maquinaria transcripcional no activa en condiciones de máxima transcripción y con una regulación opuesta en cuanto a la fuente de carbono. Sin embargo, el exceso de maquinaria transcripcional tampoco es un fenómeno directamente relacionado con una alta de tasa de transcripción. De hecho, los genes ribosómicas durante el crecimiento en galactosa siguen transcribiéndose a una velocidad mayor que los genes respiratorios y, sin embargo, presentan menor presencia de maquinaria transcripcionalmente no activa que el global de la población. Además, a las dos horas tras el cambio de las células a galactosa, cuando el crecimiento celular está detenido y la transcripción es más baja, el enriquecimiento en maquinaria transcripcional de los genes respiratorios y de las proteínas ribosómicas es similar al del global del genoma. Por lo tanto deben existir factores específicos independientes de la tasa de transcripción que permitan regular esta acumulación de polimerasa.

## Resultados y discusión

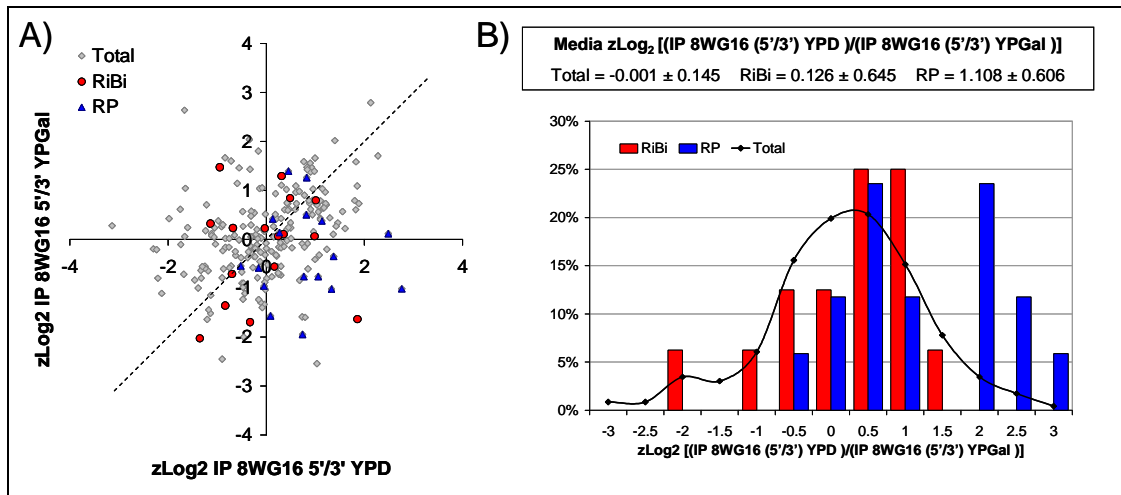


Figura 39. Sesgo en la distribución de polimerasas a lo largo de los genes en crecimiento exponencial en glucosa o galactosa. A) Relación 5'/3' de la presencia de polimerasa en crecimiento exponencial en glucosa (YPD) y en galactosa (YPGal) detectada con el Ac 8WG16. B) Histograma del cambio de la distribución de polimerasa entre las dos condiciones de crecimiento.

Como los chips de DNA que hemos usado en este estudio cubren las ORF completas (Alberola T.M. *et al* 2004), y no informan sobre la distribución de las moléculas de RNA pol dentro de los genes, realizamos experimentos de RPCC usando unos chips que contienen sondas para la región 5' y 3' de 377 genes (diseñados en colaboración con A. Rodríguez y S. Chávez). El análisis del resultado de este experimento (fig 39) muestra que los genes RP durante el crecimiento en YPD presentan una relación 5'/3' para la presencia de polimerasa más alta que el resto de genes estudiados. Esto indica que los genes RP presentan un sesgo en la distribución de polimerasa, acumulándose ésta preferencialmente en la región inicial de los genes. Además esta acumulación que se observa durante el crecimiento en glucosa desaparece cuando las células se crecen en galactosa. Por lo tanto, la acumulación de maquinaria transcripcional no activa de los genes RP durante el crecimiento en glucosa ocurre preferentemente en la región 5' de las ORFs.

Una vez determinada que la acumulación de polimerasa se produce preferentemente en la región 5' de los genes, se decidió delimitar con mayor resolución dónde se produce esta acumulación. Para ello se realizaron ensayos de inmunoprecipitación de cromatina con dos genes de proteínas ribosómicas y se analizó mediante PCR cuantitativa la distribución de las polimerasas a lo largo de los genes (fig 40).

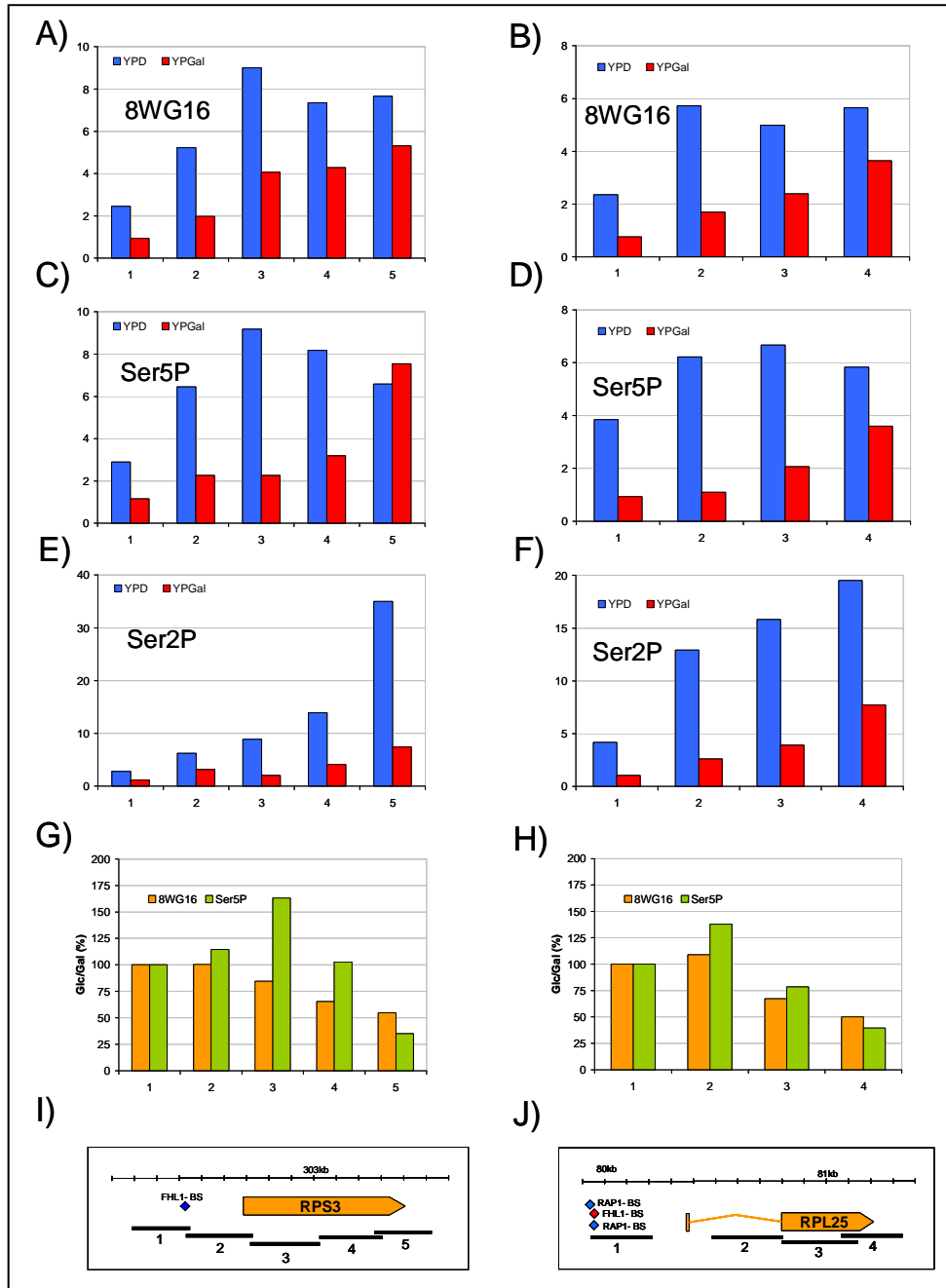


Figura 40. Distribución de la RNA pol II en los genes *RPS3* y *RPL25* durante el crecimiento exponencial en YPD y en YPGal. Se muestran los resultados de inmunoprecipitaciones de la RNA pol usando diferentes Ac y analizados mediante qPCR. Las señales están corregidas respecto al extracto celular completo y normalizadas respecto a una región intergénica. Se analizó el gen *RPS3* (A, C, E, G y I) y el gen *RPL25* (B, D, F, H y J). En G y H se muestra la relación entre la presencia de polimerasa en glucosa respecto a galactosa normalizando los valores respecto a la primera sonda. En I y J se muestra el mapa de las regiones analizadas. Se ha utilizado tanto el Ac comercial 8WG16, como dos Ac policlonales contra las fosforilaciones del CTD en Ser2 y 5 (cedidos por el Dr. D. Bentley).

## Resultados y discusión

En el caso de la inmunoprecipitación usando el Ac 8WG16, que reconoce preferentemente el CTD hipofosforilado, se produce una acumulación de polimerasa hacia la región 5' de los genes durante el crecimiento en glucosa. Esto también ocurre cuando se estudia el cambio de distribución de la polimerasa con el CTD fosforilado en Ser5, pero en este caso la acumulación se da de forma más importante en el cuerpo de los genes. Sin embargo, en galactosa se detecta siempre una mayor presencia de RNA pol en la región 3' de los genes. Por último, en el caso de la polimerasa con el CTD fosforilado en Ser2 se da tanto en crecimiento en YPD como en YPGal una acumulación en la región 3' de los genes y la única diferencia parece ser la cantidad de polimerasa detectada en ambos casos que es coherente con las tasas de transcripción correspondiente a cada condición de crecimiento. Hay que remarcar que, en este caso, se utilizó un Ac contra el CTD fosforilado en la Ser2 diferente del H5 y que permite una mayor inmunoprecipitación (cedido por el Dr. D. Bentley).

Estos resultados se ven más claramente si se normalizan las inmunoprecipitaciones igualando la primera sonda en YPD y en YPGal, y se estudia la variación de las señales lo largo de los genes. De esta manera se puede observar que, independientemente de la tasas de transcripción, en glucosa existe una acumulación de polimerasas hacia la zona 5' de los genes (fig 40 G y H). Este comportamiento se puede observar tanto con el anticuerpo contra el CTD hipofosforilado como contra el fosforilado en Ser5. Aunque en el caso de este último se produce una acumulación en la zona central del gen disminuyendo hacia 5'. Por otra parte, en el caso del anticuerpo contra el CTD fosforilado en Ser2 no tiene sentido realizar este análisis porque apenas se detecta señal en la primera sonda.

Estos resultados refuerzan la idea de que la acumulación de polimerasas transcripcionalmente no activas se produce principalmente en la región 5' de los genes. Pero esta acumulación no se produce sólo en el promotor, como podría pensarse inicialmente, sino que esta afectando a polimerasas que ya han iniciado su transcripción y tienen el CTD fosforilado en la Ser 5.

Por otra parte, *CYT1*, un gen relacionado con la mitocondria, presenta patrón diferente de acumulación de polimerasa al de los genes RP (fig 41). Concretamente, se puede observar que presenta más RNA pol durante el crecimiento en YPGal, y que su distribución es bastante homogénea a lo largo de toda la ORF. Por lo tanto, la distribución de RNA pol no activas en la región 5' de la ORF es característica de los genes RP.

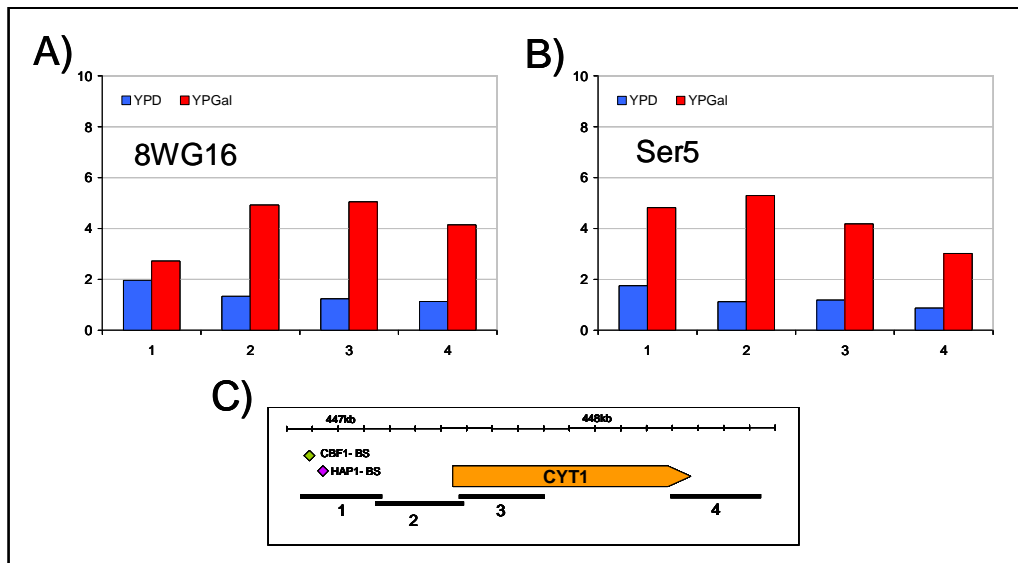


Figura 41. Distribución de la RNA pol II en el gene *CYT1* durante el crecimiento exponencial en YPD y en YPGal. Se muestran los resultados de inmunoprecipitaciones de la RNA pol usando diferentes Ac y analizados mediante qPCR. Las señales están corregidas respecto al extracto celular completo y normalizadas respecto a una región intergénica.

## Efecto de la ruta PKA en la acumulación de maquinaria transcripcional en los genes RP durante el crecimiento en glucosa.

Una vez caracterizada la acumulación de RNA polimerasa en los genes ribosómicos en condiciones de máxima transcripción en glucosa resulta interesante descubrir que mecanismo es el responsable de dicha acumulación.

Los genes RP están altamente regulados por las condiciones de crecimiento, y las dos principales rutas de regulación que actúan sobre ellos son la ruta TOR (“*target of Rapamycin*”) y la ruta PKA (“*Protein Kinase A*”). En condiciones en las que existe crecimiento la ruta TOR está constitutivamente activa y actúa sobre los genes RP, RiBi y respiratorios. Por el contrario, la ruta PKA está activa en condiciones de crecimiento en glucosa, pero no en crecimiento en galactosa (Howard S.C. *et al* 2006). Además, la ruta PKA posee un efecto positivo sobre la transcripción de los genes RP y RiBi, y un efecto negativo sobre los genes de respiración (Chen J.C. y Powers, T. 2006). El hecho que la ruta PKA cambie su actividad durante la transición de crecimiento en glucosa a galactosa, y que tenga efectos opuestos sobre la transcripción de los genes RP y los de

## Resultados y discusión

respiración, la hacen una buena candidata para regular la acumulación de maquinaria transcripcional.

Una vez seleccionada la ruta PKA como posible punto de regulación de la acumulación de polimerasas se buscó qué componente de la ruta podría tener más influencia en este fenómeno. La proteína quinasa dependiente de cAMP (PKA) en levadura presenta tres isoformas de la subunidad catalítica (Tpk1p, Tpk2p y Tpk3p), que en condiciones de represión (cuando no hay cAMP) están secuestradas por la subunidad reguladora Bcy1p permaneciendo así inactivas. Aunque las tres isoformas de Tpk son redundantes por lo que respecta a la viabilidad celular, cada una de ellas parece tener funciones específicas. De hecho, ha sido descrito que Tpk2p se encuentra físicamente unida a los promotores de los genes de las proteínas ribosómicas (Pokholok D.K. *et al* 2006).

Para tratar de descubrir si la ausencia de la subunidad Tpk2p produce algún efecto específico en la acumulación de RNA pol II en estos genes se realizaron experimentos de RPCC y GRO con esta cepa. La cepa mutante  $\Delta tpk2$  es viable y presenta un crecimiento en glucosa muy similar al de la cepa referencia, por lo tanto no son esperables diferencias debidas a la velocidad de crecimiento de la cepa ni a una falta de actividad PKA general.

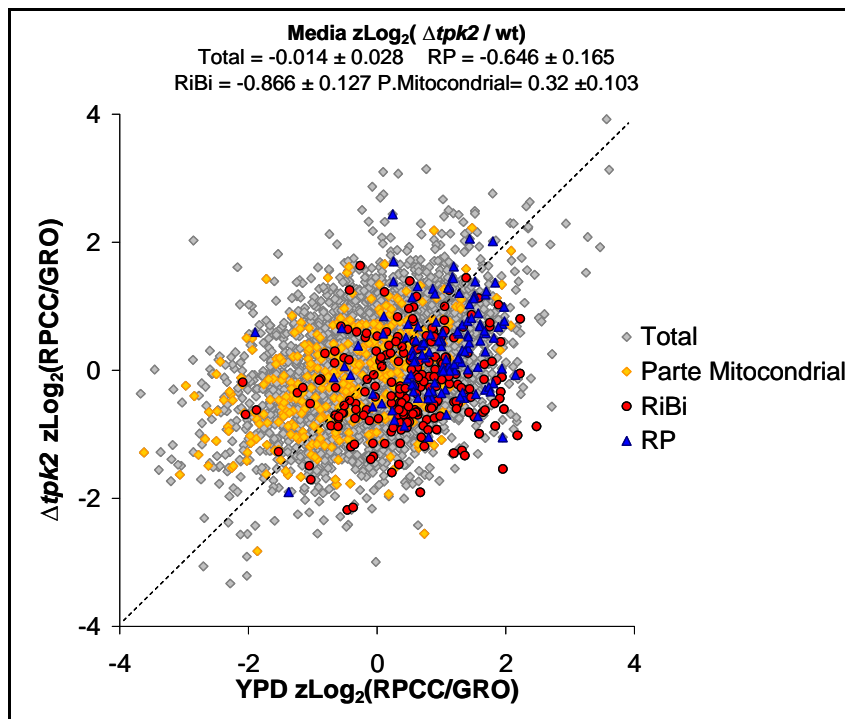


Figura 42. Cambio en la relación RPCC/GRO para el mutante  $\Delta tpk2$  durante el crecimiento en glucosa. Los valores hacen referencia a la relación RPCC/GRO en  $\text{Log}_2$  y estandarizados usando z-scores.

Nuestros resultados muestran (fig 42) que el mutante  $\Delta tpk2$  presenta tanto un aumento de acumulación de maquinaria no activamente elongante en los genes de respiración, como un descenso en la acumulación de maquinaria en los genes RP y RiBi. Es decir, el mutante  $\Delta tpk2$  en glucosa presenta un comportamiento muy similar al que tiene la cepa de referencia durante el crecimiento en galactosa (fig 36A). Lo que sugiere que este comportamiento puede ser debido a la actividad de Tpk2p durante el crecimiento en glucosa.

Por otro lado, los cambios en la relación RPCC/GRO se deben principalmente a cambios en la transcripción medida por GRO sin que varíe mucho la presencia de RNA pol II unida (fig 43). Es importante destacar que el hecho de que aumente la transcripción medida por GRO, sin afectar a la presencia de polimerasa puede parecer inicialmente contrario a los resultados obtenidos en el crecimiento en galactosa. De hecho, en la transición del crecimiento de YPD a YPGal la mayor parte de las diferencias en la relación RPCC/GRO parecen deberse a cambios en presencia de polimerasas (RPCC) (fig 38). Esto puede deberse a que en el cambio de fuente de carbono se está afectando la regulación de la transcripción a múltiples niveles, desde el reclutamiento de polimerasas hasta su acumulación. Sin embargo en este mutante, se está eliminando específicamente la acumulación de polimerasas sin afectar a la regulación de estos genes por parte de otros muchos factores transcripcionales que pueden responder a su vez al cambio de fuente de carbono.

Estos resultados sugieren que la ruta PKA controla la variación de la acumulación de RNA pol II durante el crecimiento en glucosa de una manera general. De hecho, el mutante  $\Delta tpk2$  presenta un comportamiento respecto a los cocientes RPCC/GRO similar al de una cepa normal creciendo en galactosa.

## Resultados y discusión

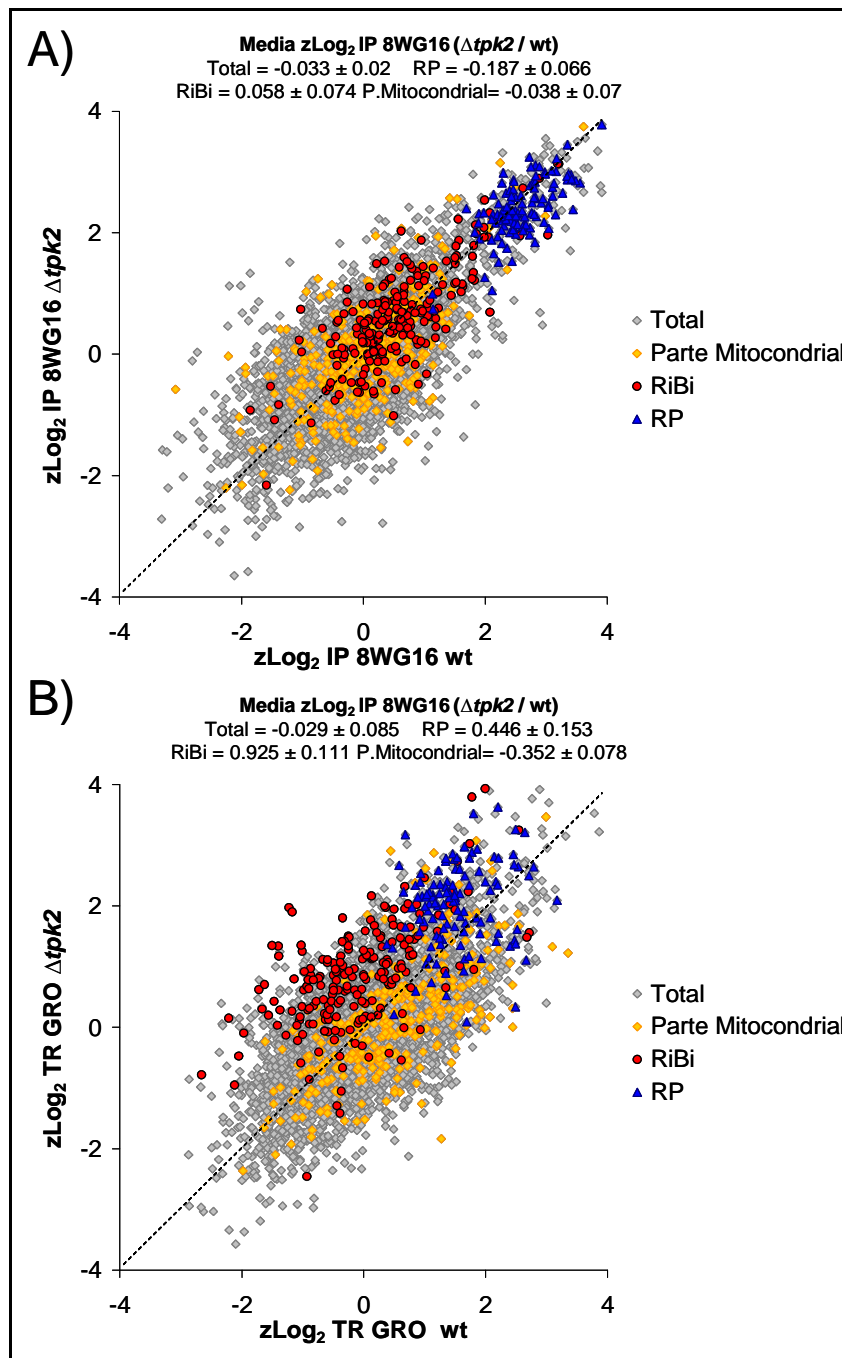


Figura 43. Comparación de RPCC y GRO entre el mutante  $\Delta\text{tpk2}$  y una cepa control. A) RPCC usando el anticuerpo 8WG16. B) Comparación mediante GRO. Todos los valores están estandarizados por z-scores y representados en  $\text{Log}_2$ .



## **Efecto de dominio silenciador de Rap1p en la acumulación de maquinaria transcripcional en los genes RP**

Una vez determinada la ruta responsable de la acumulación de RNA pol durante el crecimiento en presencia de glucosa, se trató de profundizar en el mecanismo que produce esta acumulación en los genes de las proteínas ribosómicas.

El mecanismo de transcripción de los genes ribosómicas está ampliamente estudiado, y se conocen múltiples factores transcripcionales implicados en él (Garbett K.A. *et al* 2007; Rudra D. *et al* 2005; Rudra D. *et al* 2007; Wade J.T. *et al* 2004). De existir un factor transcripcional que fuera el responsable del acúmulo de polimerasas inactivas, éste debía ser uno que estuviera unido en condiciones de alta transcripción, como suelen estarlo los activadores transcripcionales, pero que pudiera producir directa o indirectamente efectos negativos en la tasa de transcripción. De hecho, Rap1p (*Repressor Activator Protein* 1), que es uno de los factores transcripcionales más importantes en la regulación de los genes ribosómicas, cumple estas características. Rap1p por una parte es una proteína implicada en el reclutamiento de la maquinaria de silenciamiento transcripcional a los telómeros y por otra parte es un factor transcripcional necesario para la activación de genes glucolíticos y de proteínas ribosómicas. Se han determinado dominios diferenciados dentro de la proteína necesarios para cada una de estas funciones (Graham I.R. *et al* 1999; Idrissi F.Z. *et al* 2001; Pina B. *et al* 2003). Estos datos, junto con que la ruta PKA también controla la actividad de Rap1p (Klein C. y Struhl, K. 1994), hacen que esta proteína sea un buen candidato para regular la acumulación de maquinaria específicamente en los genes RP.

Para investigar la hipótesis de que el dominio silenciador de Rap1p esté implicado en la acumulación de polimerasas se realizaron experimentos de RPCC y GRO comparando los resultados de una cepa a que le falta el dominio silenciador de Rap1p (*rap1 $\Delta$ sil*) con una cepa con la misma construcción pero con la proteína Rap1 completa (Graham I.R. *et al* 1999). Estos experimentos se realizaron en crecimiento exponencial en glucosa dado que es la situación donde se observa mayor acumulación de maquinaria transcripcional inactiva en los genes RP. Los resultados muestran que en el mutante *rap1 $\Delta$ sil* disminuye de forma importante el enriquecimiento de maquinaria inactiva en los genes de las proteínas ribosómicas (fig 44).

## Resultados y discusión

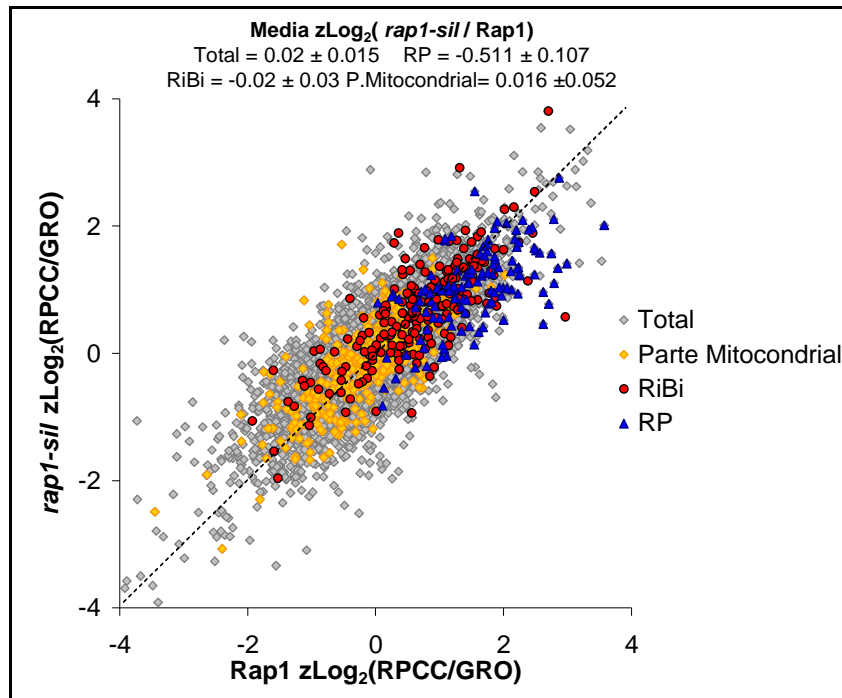


Figura 44. Cambio en la relación RPCC/GRO para una cepa sin el dominio silenciador de Rap1p. Los valores hacen referencia a la relación RPCC/GRO en  $\text{Log}_2$  y estandarizados usando z-scores. La cepa control (Rap1) es una cepa idéntica a la que lleva la delección parcial *rap1-sil* en un plásmido pero con la copia completa de Rap1.

Además, de igual manera que ocurría en el mutante  $\Delta\text{tpk2}$ , esta diferencia se debe principalmente a un aumento en la transcripción de los genes RP (medida por GRO) que no va acompañado de un aumento de presencia de polimerasa (RPCC) (fig 45). Por lo tanto, las polimerasas transcripcionalmente no activas que se acumulan en estos genes en el tipo silvestre durante el crecimiento en glucosa, en el mutante sin el dominio silenciador de Rap1p, sí son capaces de producir transcripción activa. Sin embargo, el mutante *rap1 $\Delta$ sil* sólo produce una variación significativa para la relación RPCC/GRO para el caso de los genes RP, sin afectar apreciablemente ni a los genes RiBi ni a los de respiración. Estos resultados confirman que el dominio silenciador de Rap1p está implicado específicamente en mantener el exceso de maquinaria en los genes de las proteínas ribosómicas durante el crecimiento exponencial en YPD.

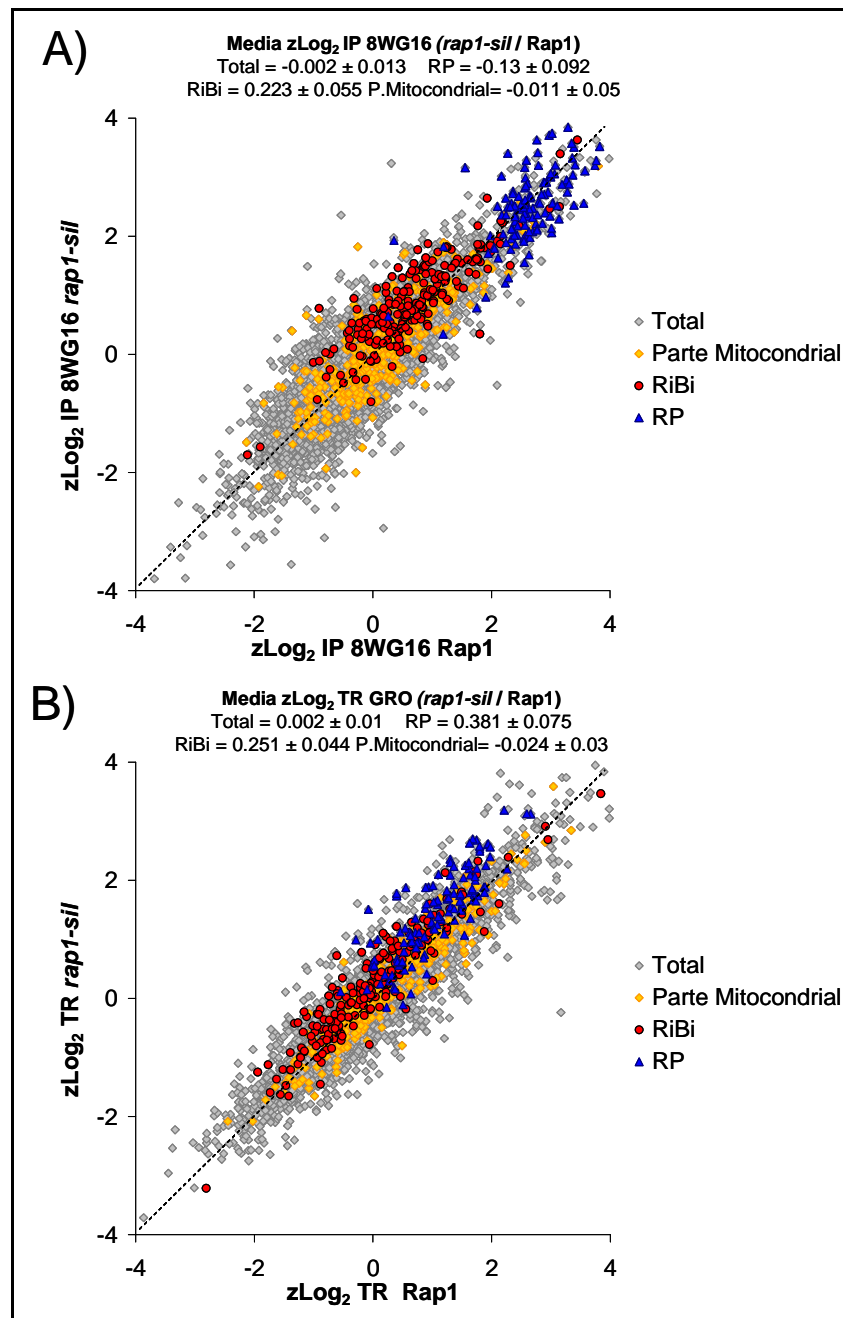


Figura 45. Comparación de RPCC y GRO entre el mutante *rap1Δsil* y una cepa control. A) RPCC usando el anticuerpo 8WG16. B) Comparación mediante GRO. Todos los valores están estandarizados por z-scores y representados en Log<sub>2</sub>.

## Resultados y discusión

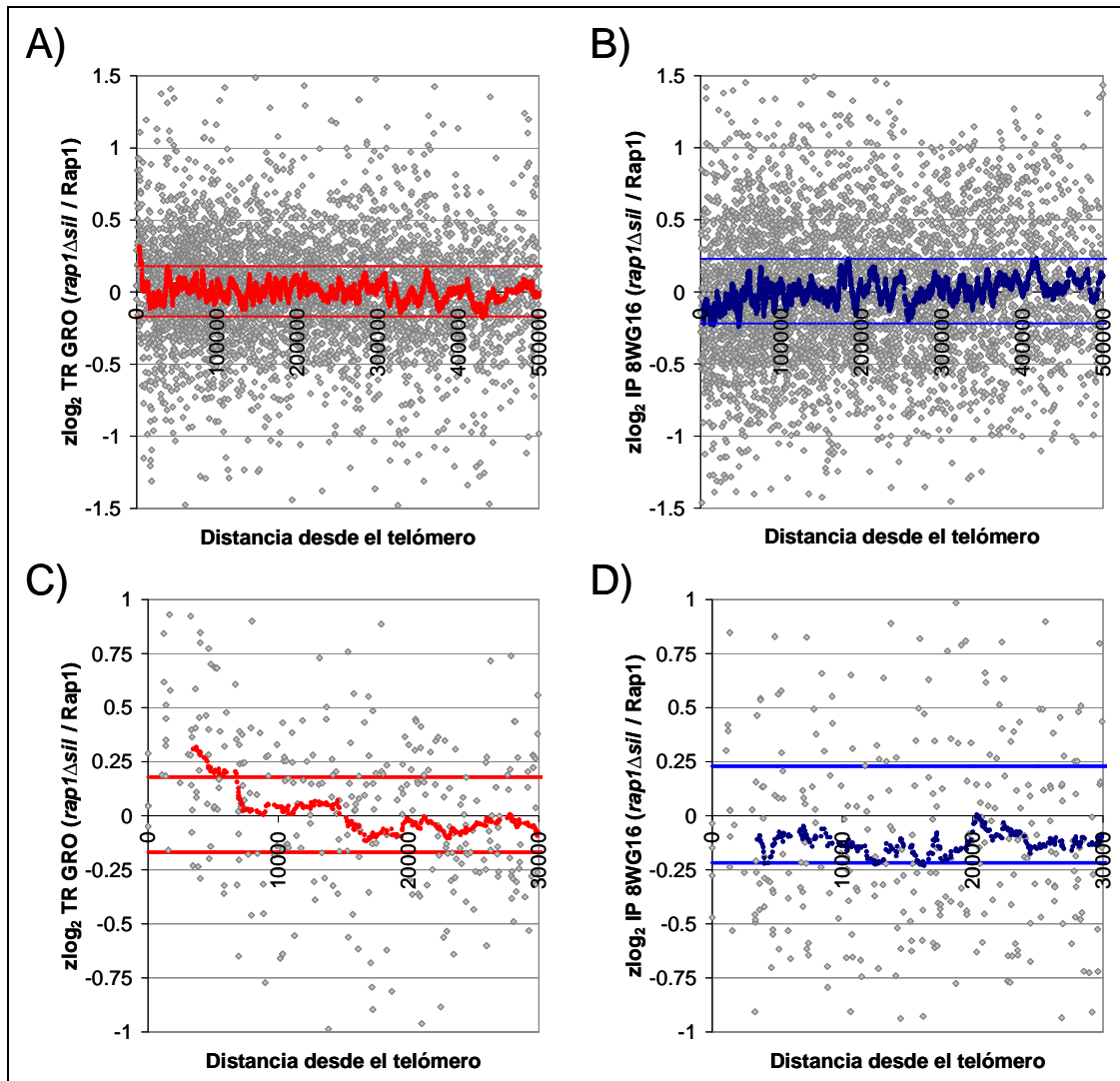


Figura 46. Cambios en la tasa de transcripción y en la presencia de RNA pol II en el mutante  $rap1\Delta sil$  respecto a la distancia al telómero. Cambio en la tasa de transcripción medida por GRO (A y C) y por RPCC (B y D) respecto al telómero. Los puntos grises representan los valores individuales, mientras que los rojos y azules representan la media de una ventana deslizante de 50 genes para los datos de **GRO** y **RPCC** respectivamente. C y D representan una ampliación de las últimas 30 kb del mapa ofrecido en A y B (500 kb) Las barras horizontales representan bandas de control para la media (Shewhart) con un coeficiente de confianza del 0.999.

Para estudiar si el efecto que tiene el dominio silenciador de Rap1p sobre la transcripción de los genes RP presenta alguna similitud con el efecto que tiene este dominio sobre las regiones silenciadas, se analizó la variación de RPCC y GRO que tiene el mutante  $rap1\Delta sil$  para las regiones subtelméricas. Si se ordenan todos los genes respecto a su distancia de los telómeros se puede ver que en el mutante  $rap1\Delta sil$  los genes cercanos a los telómeros tienen valores más altos de TR medida por GRO. Sin embargo la

presencia de polimerasa en esas regiones no varía significativamente en el mutante frente a la cepa control (fig 46). Hay que señalar que, a pesar de ser regiones silenciadas, la RNA pol esta presente en ellas, aunque inactiva (Pirrotta V. y Gross, D.S. 2005; Sekinger E.A. y Gross, D.S. 2001). Por lo tanto, la desrepresión que sufren en el mutante *rap1 $\Delta$ sil* los genes subteloméricos silenciados, es similar al cambio en la relación RPCC/GRO que se observa en el caso de los genes RP: un aumento de la transcripción sin aumento del número de polimerasa

Estos resultados, junto con los presentados en el apartado anterior, sugieren que la ruta PKA controla la variación de la acumulación de RNA pol II a un nivel superior que Rap1p. De hecho, en el mutante *rap1 $\Delta$ sil* se afecta específicamente a la acumulación de RNA pol II en los genes de las proteínas ribosómicas sin afectar a la relación RPCC/GRO de los otros grupos que presentan variaciones de esta relación en el cambio de glucosa a galactosa, como los genes de respiración o de biogénesis de ribosomas. Por lo tanto, aunque el dominio silenciador de Rap1p sea el responsable de la acumulación de polimerasas en los genes RP, deben existir otros mecanismos diferentes que controlen este fenómeno para otros grupos funcionales de genes.

### **Modelo de regulación génica mediante la maquinaria de silenciamiento génico**

Los resultados anteriores apoyan un modelo en el que el dominio silenciador de la proteína Rap1 tiene un efecto negativo sobre la transcripción de los genes RP. De hecho, en la bibliografía existen datos que apoyan estas conclusiones. Así, en experimentos de *northern* blot realizados con la cepa *rap1 $\Delta$ sil*, donde se ha eliminado de la proteína Rap1 el dominio de silenciamiento, se han detectados incrementos en la cantidad de mRNAs de genes de proteínas ribosómicas (Graham I.R. *et al* 1999), resultados que concuerdan con el aumento de tasa de transcripción determinado en este trabajo. Por otra parte, el dominio C terminal de Rap1p, donde se sitúa este dominio de silenciamiento, es también necesario para poder reprimir eficientemente la transcripción de los genes RP en determinadas condiciones (Mizuta K. *et al* 1998).

Al ser el dominio silenciador de Rap1p el responsable de la acumulación de polimerasas en los genes RP durante crecimiento en YPD, y estar este dominio vinculado con la función de silenciamiento génico de Rap1p, es lógico preguntarse si el resto de la maquinaria de silenciamiento está implicada en esta acumulación.

## Resultados y discusión

Aunque tradicionalmente la presencia de maquinaria de silenciamiento siempre se asocia con una disminución drástica de la transcripción, la unión de las proteínas Sir no impide la unión física de la RNA pol II a los genes silenciados (Pirrotta V. y Gross, D.S. 2005; Sekinger E.A. y Gross, D.S. 2001). Para determinar si la maquinaria de silenciamiento está unida a los genes de las proteínas ribosómicas en crecimiento en glucosa se reanalizaron experimentos genómicos de localización genómica de Rap1p y de las proteínas Sir (Lieb J.D. *et al* 2001). Sin embargo, en vez de usar exclusivamente los valores más altos de unión de estas proteínas como habían hecho los autores del estudio, se buscaron grupos de genes significativamente enriquecidos para la presencia de estas proteínas usando el programa FatiScan (Al-Shahrour F. *et al* 2006). De esta manera se detectó que los genes de las proteínas ribosómicas están enriquecidos tanto en Sir2p, como en Sir3p y Sir4p, con p-valores ajustados de  $2.64 \cdot 10^{-32}$ ,  $7.96 \cdot 10^{-15}$  y  $1.74 \cdot 10^{-05}$  respectivamente, mientras que no se detectó ningún grupo funcional significativamente enriquecido en el control negativo. Estos resultados apoyan un modelo en el que la maquinaria de silenciamiento está actuando, al menos parcialmente, sobre la transcripción de los genes de las proteínas ribosómicas cuando éstos se transcriben a alta velocidad. Pero estas proteínas, en vez de ejercer una represión total de la transcripción, ejercen una represión parcial de la tasa de transcripción durante la elongación transcripcional.

Por otra parte, el exceso de polimerasas corresponde a polimerasas pausadas durante la elongación y que no son capaces de transcribir en condiciones de *run-on*. Por lo que es lógico pensar que se trate de RNA polimerasas bloqueadas que necesiten de la acción de TFIIS para retroceder y poder reanudar la transcripción (Adelman K. *et al* 2005). Este sería un fenómeno similar al de la parada próxima al promotor descrito recientemente en eucariotas superiores (Core L.J. y Lis, J.T. 2008; Koch F. *et al* 2008; Price D.H. 2008; Tamkun J.W. 2007), pero que no se limitaría exclusivamente a la zona inicial del gen, sino que se podría producir durante toda la elongación. Además, esta diferencia en el lugar de parada de las polimerasas (PrPPS, *Promoter Proximal Pausing Site*) podría ser debido a la diferente organización de la cromatina en el inicio de los genes en levadura con respecto a los eucariotas superiores. Concretamente en *Drosophila* el primer nucleosoma dentro de la zona transcrita está a unos 62 nucleótidos del lugar de inicio de la transcripción (justo donde podría producir el PrPPS), mientras que en levadura el primer nucleosoma solapa con el lugar de inicio de la transcripción (Mavrich T.N. *et al* 2008b). Por lo tanto las pausas de transcripción deberían producirse en nucleosomas

posteriores (y relacionados con la presencia de proteínas Sir), dado que es necesario que se desplace el nucleosoma inicial para comenzar la transcripción.

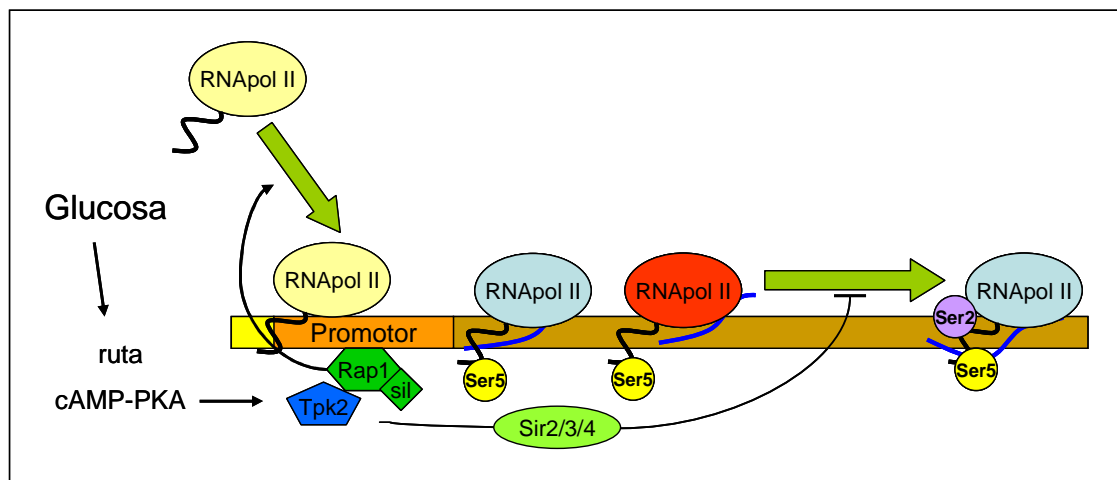


Figura 47. Modelo de regulación durante la elongación de los genes RP. Las RNA pol son reclutadas al promotor. Durante el crecimiento en glucosa la ruta cAMP-PKA, vía Tpk2p y el dominio silenciado de Rap1p, provoca la acumulación de polimerasas elongantes no activas. Esta acumulación de polimerasas permitiría regular la transcripción de los genes RP en una etapa posterior al inicio de la transcripción.

Una regulación de la transcripción durante la elongación podría aportar a la célula diferentes ventajas.

Una de las ventajas que podría ofrecer a la célula limitar la transcripción máxima de las proteínas ribosómicas en un punto posterior al reclutamiento de la polimerasa sería que permitiría un ajuste más rápido y fino del producto final, sin necesidad de esperar al siguiente ciclo de transcripción. De esta forma sería posible incluso una regulación cotranscripcional de la transcripción durante la elongación, en la que la maduración y el exporte de los mRNAs pudieran afectar a su propia transcripción antes de que esta finalice.

Otra ventaja sería que un control a nivel de la elongación permitiría asegurar un flujo constante de polimerasa a lo largo del gen, de forma que se limitaría el efecto de los “fogonazos” transcripcionales que ocurren cuando los promotores quedan accesibles a la maquinaria transcripcional. De esta manera se conseguiría una producción más continua y controlada de estos mRNAs (con un menor *ruido* transcripcional). Esto es especialmente importante en el caso de los genes de las proteínas ribosómicas, dado que su transcripción está altamente regulada para producir una determinada cantidad de cada una de las proteínas que

## Resultados y discusión

formaran parte del ribosoma y el exceso una de ellas puede ser tóxica para la célula (Sun X.X. *et al* 2008). De hecho, se ha demostrado que los genes de las proteínas ribosómicas son de hecho uno de los grupos funcionales con menor ruido transcripcional (Newman J.R. *et al* 2006).

Otro aspecto que hay que tener en cuenta por lo que respecta a la regulación de la transcripción de los genes ribosómicas es que existe una regulación espacial de la localización subnuclear donde se produce la transcripción. De hecho, se ha postulado que el crecimiento en presencia de glucosa puede organizar una maquinaria de transcripción estable en la zona perinuclear siendo Rap1p uno de los principales factores implicados (Santangelo G.M. 2006). Si se reanalizan los datos de asociación genómica de proteínas del poro nuclear durante el crecimiento en glucosa y en galactosa (Casolari J.M. *et al* 2004), se pueden detectar cambios significativos para la asociación de los genes RPs a diferentes proteínas relacionadas con el poro nuclear y la maquinaria de transporte. Concretamente, durante el crecimiento en YPD aumenta la asociación de los genes RP a Cse1p (p-adj  $4.35 \cdot 10^{-18}$ ), que está relacionada con una transcripción activa. Por otra parte, durante el crecimiento en galactosa aumenta la asociación de los genes RP con Prp20 (p-adj  $7.02 \cdot 10^{-11}$ ), que se relaciona con una disminución de la transcripción.

La regulación de la transcripción durante la elongación es un fenómeno complejo que requerirá estudios más detallados para tratar de descubrir sus vinculaciones con otros procesos biológicos. Además, hay que tener en cuenta que, aunque en este estudio nos hemos centrado en el mecanismo de regulación para los genes RP, la regulación de la acumulación de maquinaria transcripcional parece ser un fenómeno que afecta a otros grupos funcionales. De hecho, en este estudio donde sólo hemos estudiado la variación en función de la fuente de carbono, se han encontrado diferencias para otros grupos funcionales como los genes RiBi y los de respiración que deben tener sus propios mecanismos de regulación que estarán controlados por Tpk2 y la ruta PKA.



### **3.2. GRO fluorescente de alta densidad**

Una de las principales limitaciones del actual método de GRO es la necesidad de usar nucleótidos radiactivos durante la elongación de los RNAs nacientes para su posterior detección con chips de DNA. Esto representa una limitación principalmente por dos factores. Uno radica en la incomodidad asociada al manejo de sustancias radiactivas y la necesidad de utilizar instalaciones adecuadas. El otro factor es que los chips de DNA basados en un marcaje radiactivo no pueden alcanzar el nivel de miniaturización que se puede alcanzar en plataformas basadas en fluorescencia. Esta última es la principal limitación, dado que sólo permite el uso del GRO con chips de DNA con relativamente pocas sondas realizados sobre nylon y excluyendo la gran mayoría de chips de DNA comerciales. Esto no supone ninguna limitación cuando se realizan estudios con resolución de ORFs en levadura porque, debido al pequeño tamaño del genoma de la levadura, se pueden utilizar chips de DNA de nylon con sondas para la gran mayoría de los genes (Alberola T.M. *et al* 2004). Sin embargo, debido a la aparición de los chips de DNA de embaldosado que poseen sondas solapantes que cubren la mayor parte del genoma (Mockler T.C. *et al* 2005), la adaptación del GRO a un marcaje fluorescente permitiría aumentar significativamente la resolución de los estudios de transcripción.

Los chips de DNA de embaldosado se caracterizan por tener una altísima densidad de puntos de forma que las sondas cubren todo genoma independientemente de la anotación del mismo. Aunque existen diferentes tecnologías que permiten fabricar chips de embaldosado, todas están basadas en plataformas comerciales que realizan una detección fluorescente de la muestra. Por lo tanto, un GRO adaptado para un marcaje fluorescente, combinado con el uso de chips de embaldosado, aportaría información sobre la transcripción a una resolución de unas pocas bases, permitiendo así el estudio en detalle de la transcripción dentro de los genes.

En este trabajo hemos desarrollado un protocolo que permite realizar un GRO fluorescente basado en la incorporación de biotina para su uso posterior con un chip de embaldosado de Affymetrix (*GeneChip® S. cerevisiae Tiling 1.0R Array*). Como este protocolo permite aislar moléculas de RNA nacientes podría ser usado con cualquier otro tipo de chip comercial u otra técnica, tal como la secuenciación masiva paralela. Sin embargo, esta nueva metodología aun esta en una fase inicial de

## Resultados y discusión

desarrollo y requerirá una mayor optimación para poder ser aplicada con todo su potencial.

### **Diseño experimental**

El protocolo que se ha empleado se basa en la incorporación *in vivo* de un nucleótido (UTP) modificado para marcar el RNA naciente. Se ha buscado un nucleótido que pueda ser incorporado eficientemente por parte de las RNA polimerasas de levadura y que además permita la posterior purificación del RNA naciente. Una vez obtenida una muestra enriquecida en RNA naciente, y dependiendo de la modificación que éste tenga, se ha ensayado tanto la hibridación directa de la muestra como la amplificación y marcaje de la misma previamente a su hibridación en un chip de DNA.

En este trabajo se han incorporado directa o indirectamente moléculas de biotina al RNA naciente para poder usar la interacción biotina estreptoavidina tanto para purificar el RNA naciente como detectarlo en el chip de DNA de Affymetix.

### **Incorporación *in vivo* de diferentes nucleótidos modificados.**

En nuestro laboratorio se había ensayado previamente a este trabajo un GRO incorporando al RNA naciente de aminoalil-UTP para someter posteriormente al RNA modificado a un proceso de esterificación con fluoróforo con un grupo reactivo NSH (N-hidroxisuccinimida). Sin embargo, a pesar que el UTP modificado parece incorporarse al RNA, existe una reacción inespecífica del éster de fluoróforo con el RNA no marcado, lo que imposibilita el uso de este nucleótido (comunicación personal de J García-Martínez y JE Pérez-Ortín).

Una vez descartado el uso del aminoalil-UTP, y debido a la aparición en la bibliografía de un trabajo en el que se utiliza tiouracilo para incorporarlo al RNA *in vivo* en el protozoo *Toxoplasma gondii* para realizar posteriormente una purificación del RNA modificado (Cleary M.D. *et al* 2005), se decidió utilizar 4-tio-UTP en un experimento de *run-on* en levadura. Para ello se realizó un experimento de GRO pero sustituyendo el <sup>33</sup>P-UTP por diferentes concentraciones de 4-tio-UTP. Posteriormente se trató el RNA con un reactivo que une una molécula de biotina a los átomos de azufre mediante un puente disulfuro (EZ-Link Biotin-HPDP de Pierce). De esta manera, sólo aquellos RNAs nacientes que han incorporado 4-tio-UTP serán etiquetados con biotina. Finalmente, para comprobar la unión específica se realizó un gel de RNA total y se transfirió a una membrana de nylon para detectar la biotina incorporada mediante una reacción

quimioluminiscente catalizada por peroxidada unida a estreptoavidina. Se comprobó que sólo las muestras provenientes de un experimento de *run-on* en presencia de 4-tio-UTP daban señal quimioluminiscente. Además, esta señal es dependiente de la unión de la biotina al RNA mediante puentes disulfuro, dado que si se incuban las muestras con un agente reductor como el DTT se disminuye la señal al liberarse las moléculas de biotina del RNA modificado (fig 48).

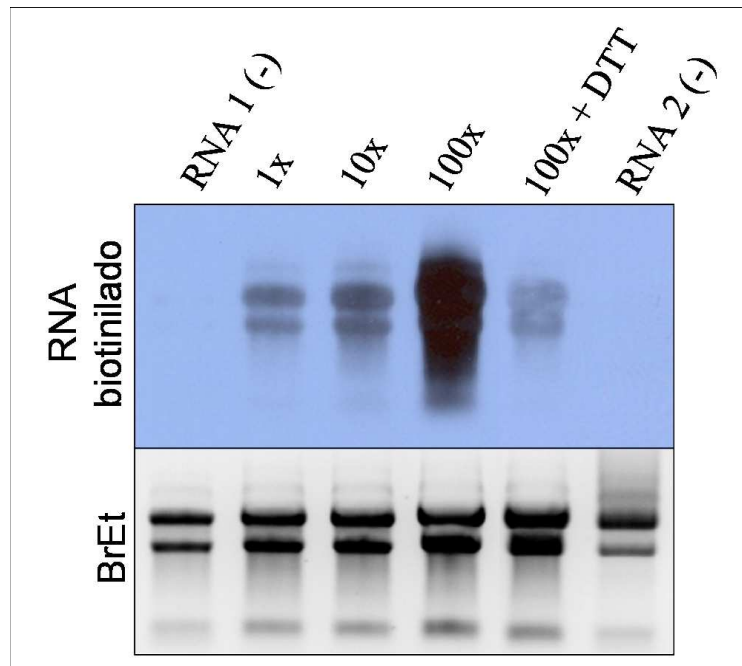


Figura 48. Incorporación de 4-tio-UTP mediante *run-on*. Se muestra la señal del RNA total obtenida mediante tinción con bromuro de etidio y la señal específica debida a la biotina incorporada (mediante unión de estreptoavidina-peroxidasa e incubación con un reactivo quimioluminiscente). Las primeras 4 carreras representan RNA procedente de experimentos de *run-on* con diferentes concentraciones de 4-tio-UTP, correspondiendo la 1X a un experimento realizado con 0.33  $\mu$ M de 4-tio-UTP. La quinta carrera corresponde a la misma muestra 100x, pero incubada con DTT para romper el puente disulfuro que une la biotina con el RNA. La sexta muestra (RNA 2 (-)) corresponde a un control negativo de RNA de levadura total al que no se ha sometido a ninguna incubación con reactivo biotinilante.

Los resultados anteriores parecen indicar que existe una correcta incorporación de 4-tio-UTP dependiente de la concentración de 4-tio-UTP usado. Sin embargo, si se analizan estos resultados con más atención, se puede ver que la mayoría de señal de RNA biotinilado corresponde a RNA ribosómico maduro (18S y 25S), que es el mayoritario en el RNA total, pero no debe serlo en la fracción de RNA naciente. Por el contrario, el resultado esperable para una incorporación transcripcional durante el *run-*

## Resultados y discusión

on debería ser un rastro difuso formado por RNAs de diferentes tamaños, que es lo que se observa cuando se analiza la incorporación de  $^{33}\text{P}$ -UTP en un GRO convencional, y no las bandas definidas correspondientes a los rRNAs maduros. Por lo tanto, existe una incorporación no transcripcional del 4-tio-UTP, al menos en el caso de los rRNAs. Esta incorporación no transcripcional que se detecta en levadura podría darse también en otros organismos, pero hasta el momento no ha sido descrita por otros autores (Cleary M.D. *et al* 2005; Dolken L. *et al* 2008). Esto puede ser debido a que en esos estudios no se realizan *run-on*, sino que se deja a las células crecer durante un tiempo prolongado en presencia de la base modificada y por lo tanto es esperable que las moléculas modificadas mayoritarias correspondan a los rRNAs maduros. Sin embargo por la coherencia de los resultados obtenidos en estos estudios, la incorporación no transcripcional, de producirse sería una incorporación adicional a la producida transcripcionalmente por las RNA polimerasas. En cualquier caso, este problema no tiene una solución simple. Por ello, y dado que el objetivo final es obtener RNA biotinilado se descartó el uso del 4-tio-UTP y se decidió utilizar directamente un nucleótido biotinilado. Concretamente se usó biotina-11-UTP, que es el nucleótido usado rutinariamente en los marcajes de RNA *in vitro* de Affymetrix.

En el caso de la biotina-11-UTP se realizó el mismo protocolo que en el 4-tio-UTP para incorporarlo al RNA naciente mediante *run-on*. Sin embargo, este nucleótido presenta unas dificultades adicionales. Uno de los inconvenientes es que no es posible utilizar una purificación estándar de RNA con fenol, dado que la biotina arrastra a las moléculas biotiniladas a la fase orgánica. Por lo tanto, para extraer el RNA biotinilado se usó un sistema de extracción alternativo al fenol basado en la digestión por proteinasa K y la precipitación selectiva del RNA.

Una vez purificado el RNA biotinilado se utilizó el mismo sistema de detección quimioluminiscente que se usó en el caso del 4-tio-UTP. Sin embargo, en este caso no se observa una incorporación mayoritaria del nucleótido a los rRNAs, sino que aparece una mezcla de RNAs biotinilados de diferentes tamaños como cabría esperar de una incorporación por *run-on* (fig 49). Además, en esta figura se pueden observar unas zonas con menos señal debidas a la deformación en el gel producidas por las moléculas mayoritarias del rRNA maduro, y que es lo contrario de lo que se observaba anteriormente con el 4-tio-UTP.

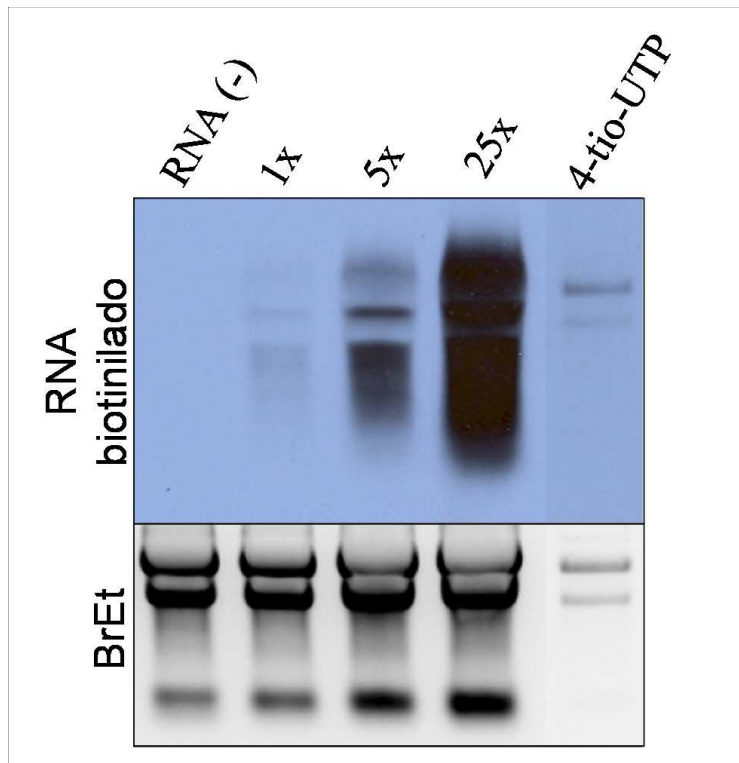


Figura 49. Incorporación de biotina-11-UTP mediante *run-on*. Se muestra la señal del RNA total obtenida mediante tinción con bromuro de etidio y la señal específica debida a la biotina incorporada. La primera carrera corresponde a un control negativo donde se usa un RNA al que no se le ha añadido biotina-11-UTP. Las siguientes tres carreras corresponden a RNAs obtenidos tras *run-on* en presencia de diferentes concentraciones de nucleótido modificado, correspondiendo la concentración 1x a 3.3  $\mu\text{M}$  de biotina-11.UTP. La última carrera corresponde a un RNA al que se ha añadido tiouracilo mediante *run-on* y posteriormente se le ha adicionado químicamente la biotina.

## Purificación de RNA naciente

Una vez demostrado que sí que se produce una incorporación de biotina-11-UTP al RNA en condiciones de *run-on* se trató de obtener una muestra enriquecida en RNA biotinilado.

Inicialmente se utilizaron esferas magnéticas cubiertas de estreptoavidina que permiten separar muy eficientemente los RNAs biotinilados del resto de moléculas de la población. Sin embargo, el problema de este método es que la unión estreptoavidina-avidina es tan fuerte ( $K_d = 10^{-15}\text{M}$ ) que la mayoría del RNA biotinilado queda unido irreversiblemente a las esferas y no es posible liberarlo. Para solucionar este problema se usaron bolas de agarosa recubiertas de avidina monomérica, la cual posee una afinidad por la biotina mucho menor ( $K_d = 10^{-8}\text{M}$ ) y permite recuperar el RNA naciente con mayor eficiencia. Además, para aumentar la resolución y facilitar la

## Resultados y discusión

interacción del RNA con la avidina se realizó una fragmentación suave del RNA total. Este procedimiento permite recuperar una cantidad de RNA biotinilado de alrededor de 200 ng a partir de un total de 2 mg de RNA total inicial. Finalmente, para comprobar la pureza del RNA biotinilado aislado se realizó una curva patrón usando un RNA sintetizado *in vitro* al que se le incorporó exclusivamente biotina-11-UTP en vez de UTP (fig 50). Se fabricó un *dot blot* sobre una membrana de nylon donde se depositaron concentraciones decrecientes y conocidas tanto de un RNA naciente biotinilado aislado después de un *run-on* como de un RNA biotinilado *in vitro* y de un control negativo de RNA sin biotina. Seguidamente se detectó la presencia de biotina mediante el sistema basado en estreptavidina-peroxidasa usado anteriormente y se calculó la señal quimioluminiscente de cada punto. Posteriormente se seleccionó un punto donde la señal quimioluminiscente mantuviera un relación lineal con la concentración de RNA y comparando la señal quimioluminiscente de la muestra problema con el RNA biotinilado *in vitro* se estimó que alrededor del 80% del RNA purificado corresponde a RNA biotinilado.

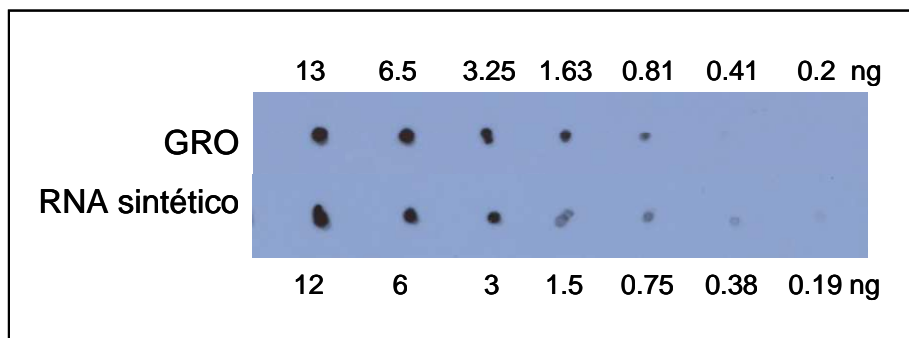


Figura 50. Comprobación de la pureza del RNA biotinilado *in vivo*. Comparación de la señal quimioluminiscente del RNA purificado tras un GRO con biotina con un patrón de RNA biotinilado generado *in vitro*.

Inicialmente se hibridó directamente la muestra de RNA naciente biotinilado *in vivo* y purificado con los chips de Affimetrix. Este procedimiento es el más directo, dado que sólo los RNAs que hayan incorporado biotina-11-UTP durante el *run-on* darán señal detectable y no puede haber señal procedente de contaminación por parte de moléculas de RNA maduras. Sin embargo, la hibridación directa del RNA purificado da señales muy débiles cercanas al límite de detección del sistema, de forma que sólo es posible detectar unos pocos genes muy transcritos. Paralelamente se realizó una hibridación del RNA total obtenido en un *run-on* con biotina pero sin realizar ninguna purificación previa del RNA biotinilado, sin embargo en este caso las señales fueron aun más bajas.

Para solucionar este problema de sensibilidad se realizó una amplificación del RNA aislado usando una amplificación basada en transcripción *in vitro*. Se eligió este protocolo de amplificación porque es un paso estándar en los experimentos con chips de Affymetrix y produce señales adecuadas para la lectura de los chips. Este paso adicional de amplificación permite obtener señales adecuadas durante la hibridación de los chips (fig 51). Sin embargo, introduce el riesgo de que parte de la señal obtenida pueda deberse a moléculas de RNA maduro contaminantes aisladas durante el proceso de purificación del RNA naciente.

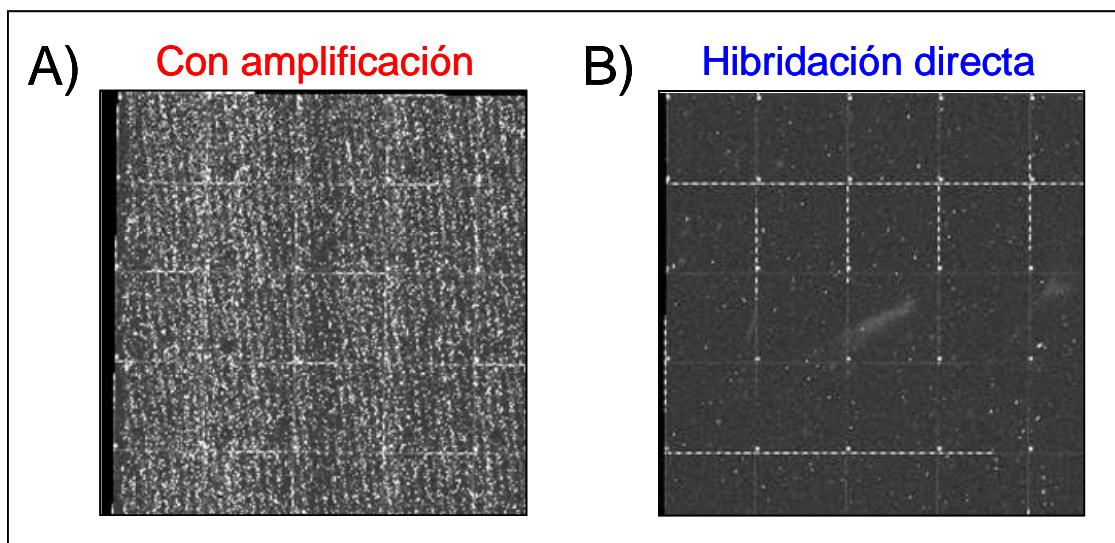


Figura 51. Chips de embaldosado hibridados con el RNA biotinilado *in vivo* o con una amplificación de ese mismo RNA biotinilado *in vitro*. En el chip con hibridación directa sólo se observan unos pocos puntos con señal apreciable a parte de los puntos donde hibrida el oligo control. Por otra parte, en el chip donde se ha hibridado el RNA amplificado se detectan señales importantes para la mayoría de regiones. Se muestran dos campos de la imagen procesada (.cel) de chips de Affymetrix (*GeneChip® S. cerevisiae Tiling 1.0R Array*).

### Resultados preliminares

Los resultados preliminares de esta técnica muestran un mapa de la transcripción en alta resolución para todo el genoma de *S. cerevisiae*. Las regiones que presentan una mayor transcripción se pueden detectar con la hibridación directa del RNA naciente biotinilado *in vivo*. De esta manera se pueden detectar genes de la RNA pol II muy transcritos y otras regiones transcritas como rRNAs, tRNAs o transposones (fig 52). Sin embargo, para obtener señales adecuadas para todo el genoma resulta necesario recurrir a la amplificación del RNA naciente aislado.

## Resultados y discusión

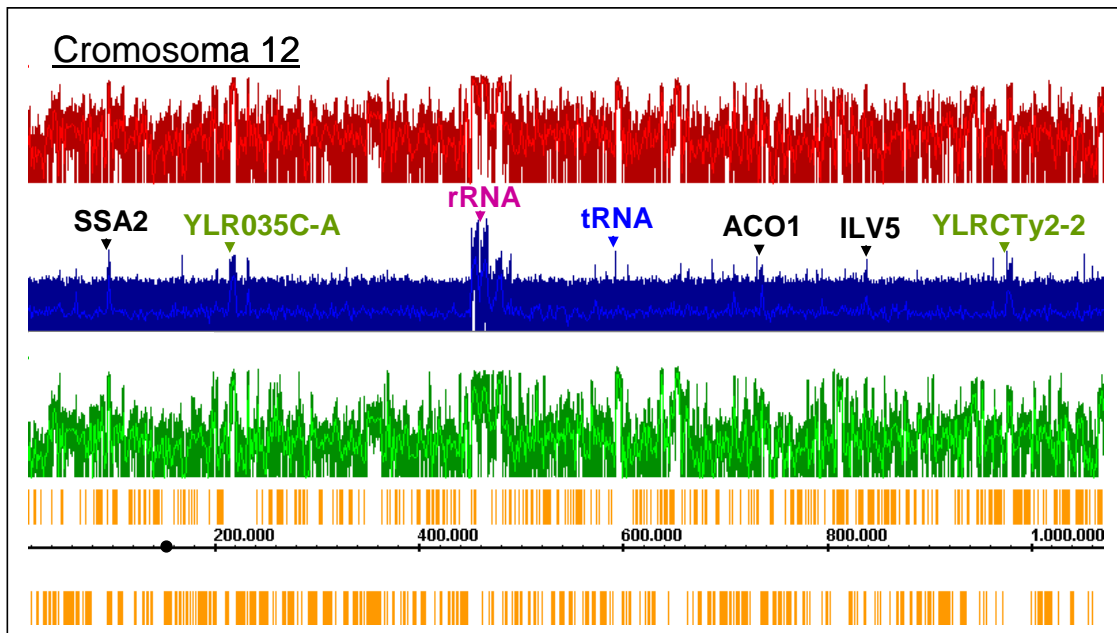


Figura 52. Mapa de transcripción para el cromosoma XII. Se muestra la señal de transcripción para una hibridación de RNA naciente amplificado *in vitro* (rojo), la hibridación directa del RNA biotinilado *in vivo* (azul) y un control de RNA maduro amplificado (verde). En la hibridación directa se pueden detectar señales tanto para genes de la RNA pol II como para rRNAs, tRNAs y transposones. Se muestran los máximos y mínimos resaltándose la media de las señales. En todos los casos se ha filtrado el 10% de las señales más bajas y se ha realizado la representación usando el programa *Integrated Genome Browser* (Affymetrix).

Si se observan a mayor resolución estos resultados se puede determinar la extensión de las regiones transcritas. Sin embargo, sólo se pueden observar mesetas claras para la mayoría de los genes en el caso de la hibridación de la muestra amplificada (fig 53). Como era esperable, la región transcrita es en general más grande que las ORFs anotadas (zonas codificantes). Otra diferencia entre la hibridación de la muestra directa y la de la muestra amplificada es que, como el chip de DNA utilizado sólo tiene sondas para una de las hebras, sólo permite detectar los transcritos codificados en la hebra Crick. Por el contrario, en la muestra amplificada se ha producido DNA de doble cadena para todos los transcritos y por lo tanto permite detectar los transcritos de las dos hebras, aunque no permite diferenciarlos.



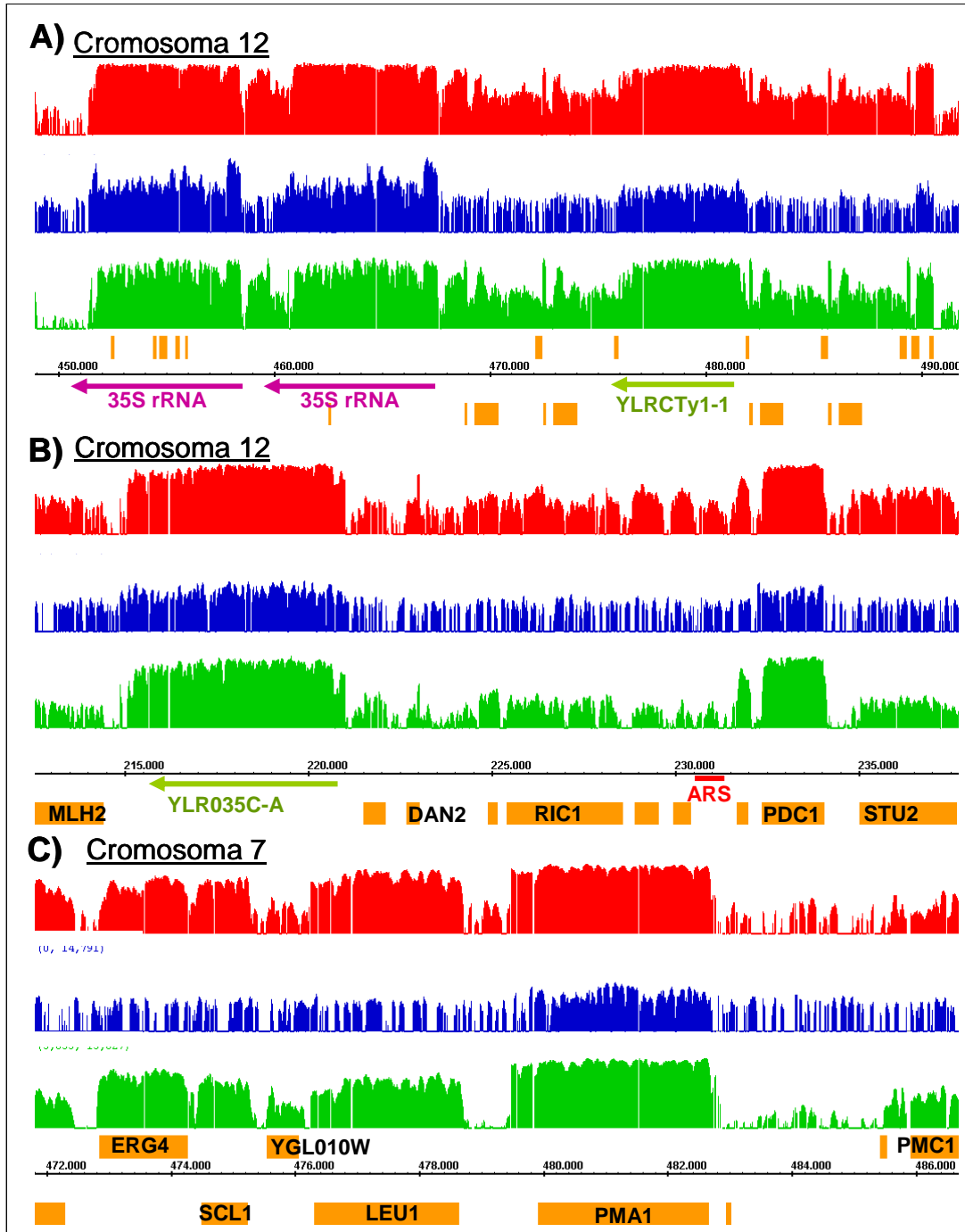


Figura 53. Mapa de transcripción detallado de algunas zonas del genoma. Se muestra la señal de transcripción para una hibridación de RNA naciente amplificado *in vitro* (rojo), la hibridación directa del RNA biotinilado *in vivo* (azul) y un control de RNA maduro amplificado (verde). A) Región del cromosoma XII donde se encuentran situadas las diferentes copias de los rRNAs. Sólo están representadas dos de las repeticiones. B) Zona del cromosoma XII donde se puede detectar la transcripción de un gen de un retrotransposón y del gen *PDC1* C) Región del cromosoma VII donde se puede ver la transcripción de diferentes genes por parte de la RNA pol II. Sólo el gen *PMA1* es detectable con la hibridación del RNA naciente biotinilado.

## Resultados y discusión

Por otro lado en la amplificación de la muestra, resulta imprescindible confirmar que los resultados obtenidos se deben en realidad a RNA naciente y no son consecuencia de RNA maduro contaminante. Además, el hecho de que no existan mapas alternativos de transcripción a esa resolución con los que compararlos, dificulta la confirmación de la validez de estos resultados. Por el contrario, en la bibliografía sí que existen datos de mRNAs maduros de alta resolución realizados con chips de embaldosado (David L. *et al* 2006) o mediante secuenciación masiva (Nagalakshmi U. *et al* 2008) con los que se puede tratar de comparar los resultados del GRO, así como datos de localización genómica de la RNA pol II con alta resolución (Steinmetz E.J. *et al* 2006). La principal dificultad reside en encontrar qué características particulares deben tener los datos de RNA nacientes frente a los RNA maduros para poder diferenciarlos. Algunas de estas diferencias podrían ser las siguientes:

- Los datos de GRO deberían presentar señales que se extendieran una mayor longitud en las zonas 3' de los genes, más allá del sitio de corte y poliadenilación, dado que, aunque estas regiones son transcritas, no están presentes en los transcritos maduros.
- Los datos de GRO deberían dar una mayor señal en la zona correspondiente a los intrones. Esto se debe a que en el caso de los mRNA maduros los intrones habrán sido procesados, mientras que en el caso de los mRNAs nacientes sólo se habrán eliminado los intrones en la medida que esta eliminación sea co-transcripcional.
- Los datos de GRO también deberían facilitar la detección de transcripción fuera de zonas codificantes de RNAs de vida media muy corta que son rápidamente degradados (y por lo tanto difíciles de detectar en análisis convencionales de mRNAs maduros), así como de transcripción antisentido (David L. *et al* 2006; Koch F. *et al* 2008; Struhl K. 2007). No es fácil predecir, sin embargo, donde se van a detectar estos transcritos.

En la figura 54 se representan los resultados para algunos genes con intrones. En ella se pueden ver algunos genes que presentan una clara disminución de la señal dentro de los intrones (como *ACT1*), pero también otros genes donde se encuentra una señal importante dentro de la zona del intrón (como el caso *RPS13* o *RPL30*). Además, estos resultados aparentemente contradictorios por lo que respecta a la señal dentro de los intrones, son similares a los descritos anteriormente para RNAs maduros (David L. *et al* 2006).

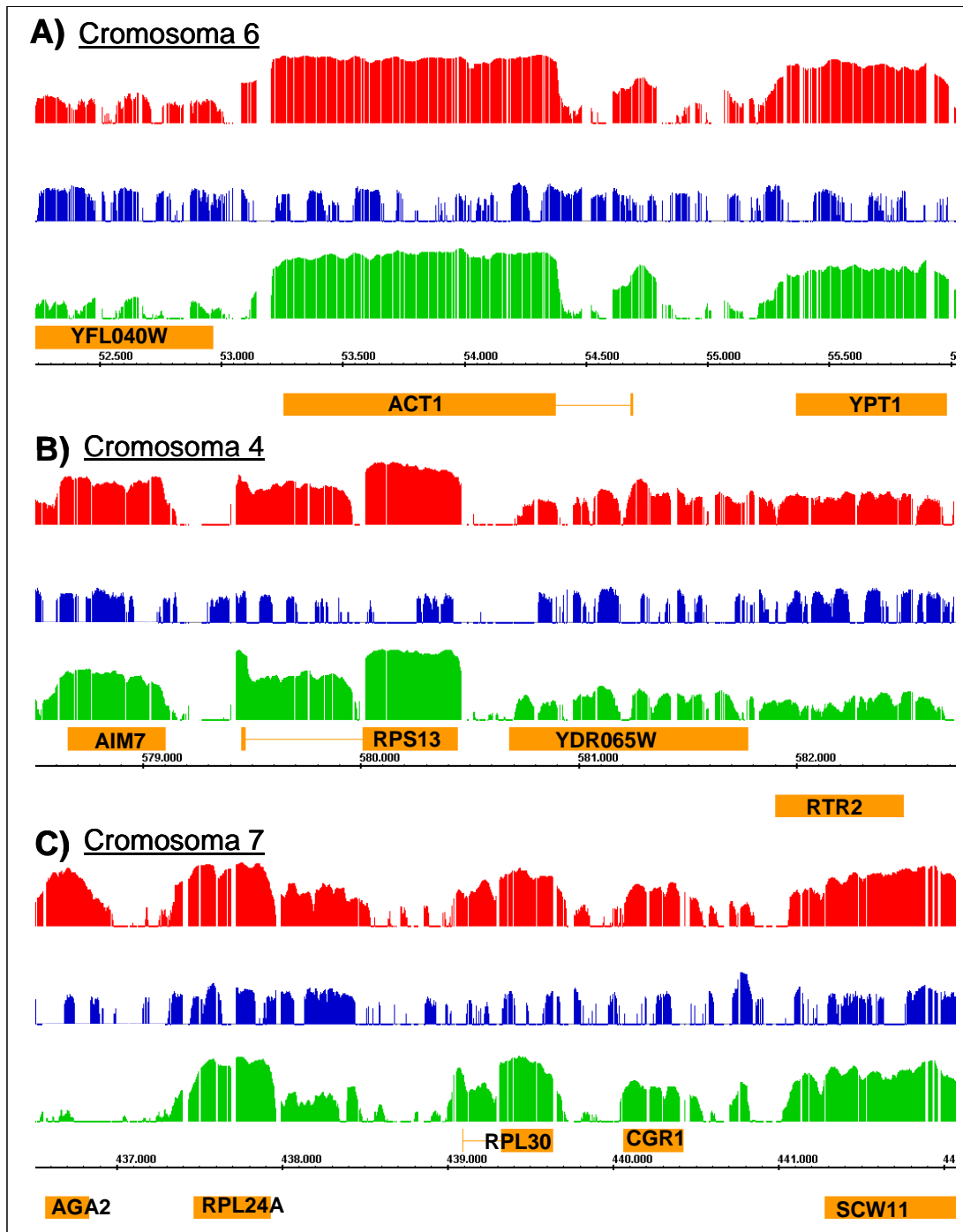


Figura 54. Mapa de transcripción detallado de algunos genes con intrones. Se muestra la señal de transcripción para el una hibridación de RNA naciente amplificado *in vitro* (rojo), la hibridación directa del RNA biotinilado *in vivo* (azul) y un control de RNA maduro amplificado (verde). Transcripción en detalle para diversos genes con intrón, como *ACT1* (A), *RPS13* (B) y *RPL30* (C).

### **Perspectivas de aplicación del GRO fluorescente**

Actualmente el GRO fluorescente aún presenta una serie de inconvenientes que deben ser solucionados antes de que esta técnica pueda ser utilizada de forma más productiva.

Dado que la hibridación directa del RNA marcado *in vivo* da señales muy pobres y sólo es útil para genes muy transcritos parece necesario hacer una amplificación *in vitro*. El principal inconveniente de esta amplificación, como se ha dicho anteriormente, es asegurar que las señales que se detectan sean debidas exclusivamente a la hibridación de RNA naciente y no procedan de una contaminación con RNA maduro. Para ello será necesario realizar diferentes controles. Uno de ellos puede ser un *run-on* en presencia de concentraciones de  $\alpha$ -amanitina que inhiban la transcripción de la RNA pol II pero no la de la RNA pol III y I. De esta manera debería poder detectarse sólo transcripción activa para genes transcritos por la RNA pol III, como *SCR1* o los tRNAs, y no para los genes transcritos por la RNA pol II. Por el contrario, si existiera una contaminación con RNAs maduros si que se observarían señales importantes dentro de los genes transcritos por la RNA pol II.

Otra posible mejora del método sería sustituir el uso de biotina-11-UTP por otro nucleótido que permita la posterior purificación del RNA naciente. Esto es especialmente importante debido al elevado coste de este nucleótido y a la elevada concentración a la que deben usarse durante el *run-on*. Además, la biotina-11-UTP presenta una baja eficiencia de incorporación *in vitro* por parte de la RNA pol del fago T7 y por lo tanto es esperable que la incorporación por parte de las RNA pol de levadura sea también baja. Entre los nucleótidos que pueden sustituir a la biotina-11-UTP se puede destacar el Br-UTP, con el que recientemente se han obtenido resultados prometedores (Ohtsu M. *et al* 2008), o incluso el 5-metil-CTP. Sin embargo, será necesario comprobar que estos nucleótidos no se incorporen de manera no transcripcional a las moléculas de RNA maduro como ocurría en el caso del 4-tio-UTP. Por otro lado, debido a la reciente aparición de técnicas de secuenciación masiva, puede ser interesante realizar estudios de secuenciación masiva de los RNAs nacientes de forma análoga a lo que se conoce como RNA-seq (Nagalakshmi U. *et al* 2008) en vez de utilizar chips de DNA de embaldosado. Esta última técnica permitiría tener datos con una resolución de un solo nucleótido, así como poder estudiar de forma independiente cada molécula de transcrito naciente aislado.

El desarrollo del GRO-fluorescente presentará importantes mejoras respecto al ChIP-on-chip de la polimerasa que ya se realiza usando chips de embaldosado (Brodsky A.S. *et al* 2005; Steinmetz E.J. *et al* 2006). La principal mejora es que marcará la posición de polimerasas transcripcionalmente activas, a diferencia del ChIP-on-chip que detecta todas las polimerasas unidas físicamente a la región. Algo de especial interés a la vista de los resultados presentados en el apartado 3.1.3 de esta tesis. Además, y dado que el marcaje del *run-on* es transcripcional, marcará la orientación de la polimerasa detectada, facilitando así la detección de transcritos antisentido de vida media corta y también la transcripción activa fuera de zonas codificantes. Sin embargo, el mayor potencial de esta tecnología reside en que permitirá estudiar como ocurre la transcripción dentro de los genes. Esto es especialmente importante en levadura porque, dado que existen mapas concretos de cómo se distribuyen los nucleosomas a escala genómica (Lee W. *et al* 2007; Mavrich T.N. *et al* 2008a), se puede tratar de correlacionar las variaciones en la transcripción con la estructura de la cromatina.

## **Conclusiones**

# Conclusiones

1. Estudio del estado estacionario para la expresión génica.
  - 1.1. Las tasas de transcripción y degradación de los mRNAs están equilibradas durante el crecimiento exponencial en YPD para la gran mayoría de los genes.
  - 1.2. Existen pequeñas desviaciones del estado estacionario para algunos grupos de genes en el mismo sentido que los cambios que experimentará la célula durante las fases posteriores del crecimiento.
2. Estudio genómico de la estabilidad de mRNAs.
  - 2.1. La tiolutina a altas concentraciones tiene un efecto de inhibición sobre la degradación de mRNAs que afecta a la medida de sus estabildades.
  - 2.2. Existen diferencias significativas entre los cálculos genómicos de estabildades de mRNAs en levadura calculados mediante la parada transcripcional y aplicación del estado estacionario. Estas diferencias se deben fundamentalmente a sesgos propios de cada método y al error introducido durante el procesado matemático de los datos.
3. Desarrollo de nuevas tecnologías para el estudio de la transcripción.
  - 3.1. Se ha desarrollado un método para medir tasas de transcripción basado en la inmunoprecipitación de la RNA pol II denominado RPCC y se han usado sus resultados para corregir los datos de tasas de transcripción obtenidos por GRO.
  - 3.2. Comparando los resultados de RPCC y GRO se han encontrado acumulaciones específicas de polimerasas no activas en diferentes grupos funcionales de genes que son reguladas por la célula y dependen de la situación fisiológica.
  - 3.3. Se ha determinado que la ruta PKA, en concreto la isozima Tpk2p, es importante para esta regulación.
  - 3.4. Se ha determinado que, para el caso de los genes de las proteínas ribosómicas, el dominio silenciador de Rap1p es el responsable de esta acumulación.

3.5. Se ha puesto a punto una versión inicial de un protocolo de GRO fluorescente para el estudio de la transcripción usando chips de embaldosado de alta resolución.



## **Bibliografía**

## Bibliografía

1. Adams,C.C. y Gross,D.S. (1991). The yeast heat shock response is induced by conversion of cells to spheroplasts and by potent transcriptional inhibitors. *J.Bacteriol.* 23(173):7429-7435.
2. Adelman,K., Marr,M.T., Werner,J., Saunders,A., Ni,Z., Andrulic,E.D. y Lis,J.T. (2005). Efficient release from promoter-proximal stall sites requires transcript cleavage factor TFIIS. *Mol.Cell.* 1(17):103-112.
3. Al-Shahrour,F., Minguéz,P., Tarraga,J., Montaner,D., Alloza,E., Vaquerizas,J.M., Conde,L., Blaschke,C., Vera,J. y Dopazo,J. (2006). BABELOMICS: a systems biology perspective in the functional annotation of genome-scale experiments. *Nucleic Acids Res. Web Server issue(34):W472-W476.*
4. Alberola,T.M., Garcia-Martinez,J., Antunez,O., Viladevall,L., Barcelo,A., Arino,J. y Perez-Ortin,J.E. (2004). A new set of DNA macrochips for the yeast *Saccharomyces cerevisiae*: features and uses. *Int.Microbiol.* 3(7):199-206.
5. Anderson,P. y Kedersha,N. (2006). RNA granules. *J.Cell Biol.* 6(172):803-808.
6. Barrett,T., Troup,D.B., Wilhite,S.E., Ledoux,P., Rudnev,D., Evangelista,C., Kim,I.F., Soboleva,A., Tomashevsky,M. y Edgar,R. (2007). NCBI GEO: mining tens of millions of expression profiles--database and tools update. *Nucleic Acids Res. Database issue(35):D760-D765.*
7. Beelman,C.A. y Parker,R. (1994). Differential effects of translational inhibition in cis and in trans on the decay of the unstable yeast MFA2 mRNA. *J.Biol.Chem.* 13(269):9687-9692.
8. Belli,G., Gari,E., Piedrafita,L., Aldea,M. y Herrero,E. (1998). An activator/repressor dual system allows tight tetracycline-regulated gene expression in budding yeast. *Nucleic Acids Res.* 4(26):942-947.

9. Benard,L. (2004). Inhibition of 5' to 3' mRNA degradation under stress conditions in *Saccharomyces cerevisiae*: from GCN4 to MET16. *RNA*. 3(10):458-468.
10. Bentley,D.L. y Groudine,M. (1986). A block to elongation is largely responsible for decreased transcription of c-myc in differentiated HL60 cells. *Nature*. 6071(321):702-706.
11. Bernstein,B.E., Mikkelsen,T.S., Xie,X., Kamal,M., Huebert,D.J., Cuff,J., Fry,B., Meissner,A., Wernig,M., Plath,K., Jaenisch,R., Wagschal,A., Feil,R., Schreiber,S.L. y Lander,E.S. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*. 2(125):315-326.
12. Berriz,G.F., King,O.D., Bryant,B., Sander,C. y Roth,F.P. (2003). Characterizing gene sets with FuncAssociate. *Bioinformatics*. 18(19):2502-2504.
13. Bon,M., McGowan,S.J. y Cook,P.R. (2006). Many expressed genes in bacteria and yeast are transcribed only once per cell cycle. *FASEB J*. 10(20):1721-1723.
14. Bond,U. (2006). Stressed out! Effects of environmental stress on mRNA metabolism. *FEMS Yeast Res*. 2(6):160-170.
15. Boorsma,A., Foat,B.C., Vis,D., Klis,F. y Bussemaker,H.J. (2005). T-profiler: scoring the activity of predefined groups of genes using gene expression data. *Nucleic Acids Res. Web Server issue*(33):W592-W595.
16. Bracken,A.P., Dietrich,N., Pasini,D., Hansen,K.H. y Helin,K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev*. 9(20):1123-1136.
17. Brendolise,C., Rouillard,J.M., Dufour,M.E. y Lacroute,F. (2002). Expression analysis of RNA14, a gene involved in mRNA 3' end maturation in yeast: characterization of the rna14-5 mutant strain. *Mol.Genet.Genomics* 4(267):515-525.
18. Brodsky,A.S., Meyer,C.A., Swinburne,I.A., Hall,G., Keenan,B.J., Liu,X.S., Fox,E.A. y Silver,P.A. (2005). Genomic mapping of

## Bibliografía

- RNA polymerase II reveals sites of co-transcriptional regulation in human cells. *Genome Biol.* 8(6):R64-
19. Cao,D. y Parker,R. (2001). Computational modeling of eukaryotic mRNA turnover. *RNA.* 9(7):1192-1212.
  20. Casolari,J.M., Brown,C.R., Komili,S., West,J., Hieronymus,H. y Silver,P.A. (2004). Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell.* 4(117):427-439.
  21. Celmer,W.D. y Solomons I.A. (1955). The structures of thiolutin and aureothricin, antibiotics containing a unique pyrrolidinonodithiole nucleus. *Journal of the American Chemical Society*(77):2861-2865.
  22. Chen,J.C. y Powers,T. (2006). Coordinate regulation of multiple and distinct biosynthetic pathways by TOR and PKA kinases in *S. cerevisiae*. *Curr.Genet.* 5(49):281-293.
  23. Chubb,J.R., Trcek,T., Shenoy,S.M. y Singer,R.H. (2006). Transcriptional pulsing of a developmental gene. *Curr.Biol.* 10(16):1018-1025.
  24. Cleary,M.D., Meiering,C.D., Jan,E., Guymon,R. y Boothroyd,J.C. (2005). Biosynthetic labeling of RNA with uracil phosphoribosyltransferase allows cell-specific microarray analysis of mRNA synthesis and decay. *Nat.Biotechnol.* 2(23):232-237.
  25. Core,L.J. y Lis,J.T. (2008). Transcription regulation through promoter-proximal pausing of RNA polymerase II. *Science.* 5871(319):1791-1792.
  26. Core,L.J., Waterfall,J.J. y Lis,J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science.* 5909(%19;322):1845-1848.
  27. David,L., Huber,W., Granovskaia,M., Toedling,J., Palm,C.J., Bofkin,L., Jones,T., Davis,R.W. y Steinmetz,L.M. (2006). A high-resolution map of transcription in the yeast genome. *Proc.Natl.Acad.Sci.U.S.A.* 14(103):5320-5325.

28. Dekker,J. (2008). Gene regulation in the third dimension. *Science*. 5871(319):1793-1794.
29. Dolken,L., Ruzsics,Z., Radle,B., Friedel,C.C., Zimmer,R., Mages,J., Hoffmann,R., Dickinson,P., Forster,T., Ghazal,P. y Koszinowski,U.H. (2008). High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. *RNA*. 9(14):1959-1972.
30. Edwards,A.M., Kane,C.M., Young,R.A. y Kornberg,R.D. (1991). Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter in vitro. *J.Biol.Chem.* 1(266):71-75.
31. Ernst,J. y Bar-Joseph,Z. (2006). STEM: a tool for the analysis of short time series gene expression data. *BMC.Bioinformatics*.(7:191.):191-
32. Eulalio,A., Behm-Ansmant,I. y Izaurralde,E. (2007). P bodies: at the crossroads of post-transcriptional pathways. *Nat.Rev.Mol.Cell Biol.* 1(8):9-22.
33. Fan,J., Yang,X., Wang,W., Wood,W.H., III, Becker,K.G. y Gorospe,M. (2002). Global analysis of stress-regulated mRNA turnover by using cDNA arrays. *Proc.Natl.Acad.Sci.U.S.A.* 16(99):10611-10616.
34. French,S.L., Osheim,Y.N., Cioci,F., Nomura,M. y Beyer,A.L. (2003). In exponentially growing *Saccharomyces cerevisiae* cells, rRNA synthesis is determined by the summed RNA polymerase I loading rate rather than by the number of active genes. *Mol.Cell Biol.* 5(23):1558-1568.
35. Gaillard,H. y Aguilera,A. (2008). A novel class of mRNA-containing cytoplasmic granules are produced in response to UV-irradiation. *Mol.Biol.Cell.* 11(19):4980-4992.
36. Gao,L. y Gross,D.S. (2008). Sir2 Silences Gene Transcription by Targeting the Transition Between RNA Polymerase II Initiation and Elongation. *Mol.Cell Biol*.(.):

## Bibliografía

37. Gao,L. y Gross,D.S. (2006). Using genomics and proteomics to investigate mechanisms of transcriptional silencing in *Saccharomyces cerevisiae*. *Brief.Funct.Genomic.Proteomic.* 4(5):280-288.
38. Garbett,K.A., Tripathi,M.K., Cencki,B., Layer,J.H. y Weil,P.A. (2007). Yeast TFIID serves as a coactivator for Rap1p by direct protein-protein interaction. *Mol.Cell Biol.* 1(27):297-311.
39. García-Martínez,J., Aranda,A. y Pérez-Ortín,J.E. (2004). Genomic run-on evaluates transcription rates for all yeast genes and identifies gene regulatory mechanisms. *Mol.Cell.* 2(15):303-313.
40. García-Martínez,J., González-Candelas,F. y Pérez-Ortín,J.E. (2007). Common gene expression strategies revealed by genome-wide analysis in yeast. *Genome Biol.* 10(8):R222-
41. Gari,E., Piedrafita,L., Aldea,M. y Herrero,E. (1997). A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in *Saccharomyces cerevisiae*. *Yeast.* 9(13):837-848.
42. Gasch,A.P., Spellman,P.T., Kao,C.M., Carmel-Harel,O., Eisen,M.B., Storz,G., Botstein,D. y Brown,P.O. (2000). Genomic expression programs in the response of yeast cells to environmental changes. *Mol.Biol.Cell.* 12(11):4241-4257.
43. Goffeau,A., Barrell,B.G., Bussey,H., Davis,R.W., Dujon,B., Feldmann,H., Galibert,F., Hoheisel,J.D., Jacq,C., Johnston,M., Louis,E.J., Mewes,H.W., Murakami,Y., Philippsen,P., Tettelin,H. y Oliver,S.G. (1996). Life with 6000 genes. *Science* 5287(274):546, 563-546, 567.
44. Graham,I.R., Haw,R.A., Spink,K.G., Halden,K.A. y Chambers,A. (1999). In vivo analysis of functional regions within yeast Rap1p. *Mol.Cell Biol.* 11(19):7481-7490.
45. Greenbaum,D., Colangelo,C., Williams,K. y Gerstein,M. (2003). Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol.* 9(4):117-

46. Gresham,D., Ruderfer,D.M., Pratt,S.C., Schacherer,J., Dunham,M.J., Botstein,D. y Kruglyak,L. (2006). Genome-wide detection of polymorphisms at nucleotide resolution with a single DNA microarray. *Science*. 5769(311):1932-1936.
47. Grigull,J., Mnaimneh,S., Pootoolal,J., Robinson,M.D. y Hughes,T.R. (2004). Genome-wide analysis of mRNA stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors. *Mol.Cell Biol*. 12(24):5534-5547.
48. Gross,D.S. (2001). Sir proteins as transcriptional silencers. *Trends Biochem.Sci*. 11(26):685-686.
49. Guan,Q., Zheng,W., Tang,S., Liu,X., Zinkel,R.A., Tsui,K.W., Yandell,B.S. y Culbertson,M.R. (2006). Impact of nonsense-mediated mRNA decay on the global expression profile of budding yeast. *PLoS.Genet*. 11(2):e203-
50. Hampsey,M. (1998). Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiol.Mol.Biol.Rev*. 2(62):465-503.
51. He,F., Li,X., Spatrick,P., Casillo,R., Dong,S. y Jacobson,A. (2003). Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5' to 3' mRNA decay pathways in yeast. *Mol.Cell*. 6(12):1439-1452.
52. Herrick,D., Parker,R. y Jacobson,A. (1990). Identification and comparison of stable and unstable mRNAs in *Saccharomyces cerevisiae*. *Mol.Cell Biol*. 5(10):2269-2284.
53. Hilgers,V., Teixeira,D. y Parker,R. (2006). Translation-independent inhibition of mRNA deadenylation during stress in *Saccharomyces cerevisiae*. *RNA*. 10(12):1835-1845.
54. Hirayoshi,K. y Lis,J.T. (1999). Nuclear run-on assays: assessing transcription by measuring density of engaged RNA polymerases. *Methods Enzymol*.(304):351-362.
55. Hirose,Y. y Ohkuma,Y. (2007). Phosphorylation of the C-terminal domain of RNA polymerase II plays central roles in the

## Bibliografía

- integrated events of eucaryotic gene expression. *J.Biochem.* 5(141):601-608.
56. Hofer,T. y Rasch,M.J. (2005). On the kinetic design of transcription. *Genome Inform.* 1(16):73-82.
  57. Hoffmann,R. y Valencia,A. (2005). Implementing the iHOP concept for navigation of biomedical literature. *Bioinformatics.*(21 Suppl 2:ii252-8.):ii252-ii258.
  58. Howard,S.C., Deminoff,S.J. y Herman,P.K. (2006). Increased phosphoglucomutase activity suppresses the galactose growth defect associated with elevated levels of Ras signaling in *S. cerevisiae*. *Curr.Genet.* 1(49):1-6.
  59. Idrissi,F.Z., Garcia-Reyero,N., Fernandez-Larrea,J.B. y Pina,B. (2001). Alternative mechanisms of transcriptional activation by Rap1p. *J.Biol.Chem.* 28(276):26090-26098.
  60. Iyer,V.R., Horak,C.E., Scafe,C.S., Botstein,D., Snyder,M. y Brown,P.O. (2001). Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* 6819(409):533-538.
  61. Jiménez,A., Tipper,D.J. y Davies,J. (1973). Mode of action of thiolutin, an inhibitor of macromolecular synthesis in *Saccharomyces cerevisiae*. *Antimicrob.Agents Chemother.* 6(3):729-738.
  62. Jones,J.C., Phatnani,H.P., Haystead,T.A., MacDonald,J.A., Alam,S.M. y Greenleaf,A.L. (2004). C-terminal repeat domain kinase I phosphorylates Ser2 and Ser5 of RNA polymerase II C-terminal domain repeats. *J.Biol.Chem.* 24(279):24957-24964.
  63. Jorgensen,P., Rupes,I., Sharom,J.R., Schneper,L., Broach,J.R. y Tyers,M. (2004). A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size. *Genes Dev.* 20(18):2491-2505.
  64. Keene,J.D. y Lager,P.J. (2005). Post-transcriptional operons and regulons co-ordinating gene expression. *Chromosome.Res.* 3(13):327-337.



65. Khachatourians,G.G. y Tipper,D.J. (1974). In vivo effect of thiolutin on cell growth and macromolecular synthesis in *Escherichia coli*. *Antimicrob.Agents Chemother.* 3(6):304-310.
66. Klein,C. y Struhl,K. (1994). Protein kinase A mediates growth-regulated expression of yeast ribosomal protein genes by modulating RAP1 transcriptional activity. *Mol.Cell Biol.* 3(14):1920-1928.
67. Koch,F., Jourquin,F., Ferrier,P. y Andrau,J.C. (2008). Genome-wide RNA polymerase II: not genes only! *Trends Biochem.Sci.* 6(33):265-273.
68. Komarnitsky,P., Cho,E.J. y Buratowski,S. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.* 19(14):2452-2460.
69. Komili,S. y Silver,P.A. (2008). Coupling and coordination in gene expression processes: a systems biology view. *Nat.Rev.Genet.* 1(9):38-48.
70. Kuai,L., Fang,F., Butler,J.S. y Sherman,F. (2004). Polyadenylation of rRNA in *Saccharomyces cerevisiae*. *Proc.Natl.Acad.Sci.U.S.A.* 23(101):8581-8586.
71. Kuo,M.H. y Allis,C.D. (1999). In vivo cross-linking and immunoprecipitation for studying dynamic Protein:DNA associations in a chromatin environment. *Methods* 3(19):425-433.
72. Lee,T.I., Jenner,R.G., Boyer,L.A., Guenther,M.G., Levine,S.S., Kumar,R.M., Chevalier,B., Johnstone,S.E., Cole,M.F., Isono,K., Koseki,H., Fuchikami,T., Abe,K., Murray,H.L., Zucker,J.P., Yuan,B., Bell,G.W., Herbolsheimer,E., Hannett,N.M., Sun,K., Odom,D.T., Otte,A.P., Volkert,T.L., Bartel,D.P., Melton,D.A., Gifford,D.K., Jaenisch,R. y Young,R.A. (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell.* 2(125):301-313.
73. Lee,W., Tillo,D., Bray,N., Morse,R.H., Davis,R.W., Hughes,T.R. y Nislow,C. (2007). A high-resolution atlas of nucleosome occupancy in yeast. *Nat.Genet.* 10(39):1235-1244.

## Bibliografía

74. Lieb, J.D., Liu, X., Botstein, D. y Brown, P.O. (2001). Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. *Nat. Genet.* 4(28):327-334.
75. Livak, K.J. y Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* 4(25):402-408.
76. Lu, P., Vogel, C., Wang, R., Yao, X. y Marcotte, E.M. (2007). Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat. Biotechnol.* 1(25):117-124.
77. Maheshri, N. y O'Shea, E.K. (2007). Living with noisy genes: how cells function reliably with inherent variability in gene expression. *Annu. Rev. Biophys. Biomol. Struct.* (36:413-34.):413-434.
78. Marenduzzo, D., Faro-Trindade, I. y Cook, P.R. (2007). What are the molecular ties that maintain genomic loops? *Trends Genet.* 3(23):126-133.
79. Margaritis, T. y Holstege, F.C. (2008). Poised RNA polymerase II gives pause for thought. *Cell.* 4(133):581-584.
80. Mata, J., Marguerat, S. y Bahler, J. (2005). Post-transcriptional control of gene expression: a genome-wide perspective. *Trends Biochem. Sci.* 9(30):506-514.
81. Mavrich, T.N., Ioshikhes, I.P., Venters, B.J., Jiang, C., Tomsho, L.P., Qi, J., Schuster, S.C., Albert, I. y Pugh, B.F. (2008a). A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res.* 7(18):1073-1083.
82. Mavrich, T.N., Jiang, C., Ioshikhes, I.P., Li, X., Venters, B.J., Zanton, S.J., Tomsho, L.P., Qi, J., Glaser, R.L., Schuster, S.C., Gilmour, D.S., Albert, I. y Pugh, B.F. (2008b). Nucleosome organization in the *Drosophila* genome. *Nature.* 7193(453):358-362.

83. Meyer,S., Temme,C. y Wahle,E. (2004). Messenger RNA turnover in eukaryotes: pathways and enzymes. *Crit Rev.Biochem.Mol.Biol.* 4(39):197-216.
84. Michan,C., Monje-Casas,F. y Pueyo,C. (2005). Transcript copy number of genes for DNA repair and translesion synthesis in yeast: contribution of transcription rate and mRNA stability to the steady-state level of each mRNA along with growth in glucose-fermentative medium. *DNA Repair (Amst)* 4(4):469-478.
85. Minvielle-Sebastia,L., Winsor,B., Bonneaud,N. y Lacroute,F. (1991). Mutations in the yeast RNA14 and RNA15 genes result in an abnormal mRNA decay rate; sequence analysis reveals an RNA-binding domain in the RNA15 protein. *Mol.Cell Biol.* 6(11):3075-3087.
86. Miura,F., Kawaguchi,N., Yoshida,M., Uematsu,C., Kito,K., Sakaki,Y. y Ito,T. (2008). Absolute quantification of the budding yeast transcriptome by means of competitive PCR between genomic and complementary DNAs. *BMC.Genomics.* 1(9):574-
87. Miyoshi,K., Shirai,C. y Mizuta,K. (2003). Transcription of genes encoding trans-acting factors required for rRNA maturation/ribosomal subunit assembly is coordinately regulated with ribosomal protein genes and involves Rap1 in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 7(31):1969-1973.
88. Mizuta,K., Tsujii,R., Warner,J.R. y Nishiyama,M. (1998). The C-terminal silencing domain of Rap1p is essential for the repression of ribosomal protein genes in response to a defect in the secretory pathway. *Nucleic Acids Res.* 4(26):1063-1069.
89. Mockler,T.C., Chan,S., Sundaresan,A., Chen,H., Jacobsen,S.E. y Ecker,J.R. (2005). Applications of DNA tiling arrays for whole-genome analysis. *Genomics.* 1(85):1-15.
90. Molina-Navarro,M.M., Castells-Roca,L., Belli,G., Garcia-Martinez,J., Marin-Navarro,J., Moreno,J., Perez-Ortin,J.E. y Herrero,E. (2008). Comprehensive transcriptional analysis of the oxidative response in yeast. *J.Biol.Chem.* 26(283):17908-17918.

## Bibliografía

91. Montaner,D., Tarraga,J., Huerta-Cepas,J., Burguet,J., Vaquerizas,J.M., Conde,L., Minguez,P., Vera,J., Mukherjee,S., Valls,J., Pujana,M.A., Alloza,E., Herrero,J., Al-Shahrour,F. y Dopazo,J. (2006). Next station in microarray data analysis: GEPAS. *Nucleic Acids Res. Web Server issue*(34):W486-W491.
92. Muller,H.M., Kenny,E.E. y Sternberg,P.W. (2004). Textpresso: an ontology-based information retrieval and extraction system for biological literature. *PLoS.Biol.* 11(2):e309-
93. Muse,G.W., Gilchrist,D.A., Nechaev,S., Shah,R., Parker,J.S., Grissom,S.F., Zeitlinger,J. y Adelman,K. (2007). RNA polymerase is poised for activation across the genome. *Nat.Genet.* 12(39):1507-1511.
94. Nagalakshmi,U., Wang,Z., Waern,K., Shou,C., Raha,D., Gerstein,M. y Snyder,M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science.* 5881(320):1344-1349.
95. Nam,H.G. y Fried,H.M. (1986). Effects of progressive depletion of TCM1 or CYH2 mRNA on *Saccharomyces cerevisiae* ribosomal protein accumulation. *Mol.Cell Biol.* 5(6):1535-1544.
96. Newman,J.R., Ghaemmaghami,S., Ihmels,J., Breslow,D.K., Noble,M., DeRisi,J.L. y Weissman,J.S. (2006). Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature.* 7095(441):840-846.
97. Niemisto,A., Nykter,M., Aho,T., Jalovaara,H., Marjanen,K., Ahdesmaki,M., Ruusuvoori,P., Tiainen,M., Linne,M.L. y Yli-Harja,O. (2007). Computational methods for estimation of cell cycle phase distributions of yeast cells. *EURASIP.J.Bioinform.Syst.Biol.*46150-
98. Nonet,M., Scafe,C., Sexton,J. y Young,R. (1987). Eucaryotic RNA polymerase conditional mutant that rapidly ceases mRNA synthesis. *Mol.Cell Biol.* 5(7):1602-1611.
99. Ohtsu,M., Kawate,M., Fukuoka,M., Gunji,W., Hanaoka,F., Utsugi,T., Onoda,F. y Murakami,Y. (2008). Novel DNA Microarray System for Analysis of Nascent mRNAs. *DNA Res.*

100. Parker,R., Herrick,D., Peltz,S.W. y Jacobson,A. (1991). Measurement of mRNA decay rates in *Saccharomyces cerevisiae*. *Methods Enzymol.*(194):415-423.
101. Parker,R. y Sheth,U. (2007). P bodies and the control of mRNA translation and degradation. *Mol.Cell.* 5(25):635-646.
102. Parker,R. y Song,H. (2004). The enzymes and control of eukaryotic mRNA turnover. *Nat.Struct.Mol.Biol.* 2(11):121-127.
103. Pelechano,V. y Perez-Ortin,J.E. (2008). The transcriptional inhibitor thiolutin blocks mRNA degradation in yeast. *Yeast.* 2(25):85-92.
104. Pérez-Ortín,J.E. (2007). Genomics of mRNA turnover. *Brief.Funct.Genomic.Proteomic.* 4(6):282-291.
105. Pérez-Ortín,J.E., Alepuz,P.M. y Moreno,J. (2007). Genomics and gene transcription kinetics in yeast. *Trends Genet.* 5(23):250-257.
106. Pina,B., Fernandez-Larrea,J., Garcia-Reyero,N. y Idrissi,F.Z. (2003). The different (sur)faces of Rap1p. *Mol.Genet.Genomics.* 6(268):791-798.
107. Pirrotta,V. y Gross,D.S. (2005). Epigenetic silencing mechanisms in budding yeast and fruit fly: different paths, same destinations. *Mol.Cell.* 4(18):395-398.
108. Pokholok,D.K., Zeitlinger,J., Hannett,N.M., Reynolds,D.B. y Young,R.A. (2006). Activated signal transduction kinases frequently occupy target genes. *Science.* 5786(313):533-536.
109. Pradet-Balade,B., Boulme,F., Beug,H., Mullner,E.W. y Garcia-Sanz,J.A. (2001). Translation control: bridging the gap between genomics and proteomics? *Trends Biochem.Sci.* 4(26):225-229.
110. Price,D.H. (2008). Poised polymerases: on your mark...get set...go! *Mol.Cell.* 1(30):7-10.
111. Puig,S., Askeland,E. y Thiele,D.J. (2005). Coordinated remodeling of cellular metabolism during iron deficiency through targeted mRNA degradation. *Cell.* 1(120):99-110.

## Bibliografía

112. Radonjic,M., Andrau,J.C., Lijnzaad,P., Kemmeren,P., Kockelkorn,T.T., van,L.D., van Berkum,N.L. y Holstege,F.C. (2005). Genome-wide analyses reveal RNA polymerase II located upstream of genes poised for rapid response upon *S. cerevisiae* stationary phase exit. *Mol.Cell.* 2(18):171-183.
113. Raser,J.M. y O'Shea,E.K. (2004). Control of stochasticity in eukaryotic gene expression. *Science.* 5678(304):1811-1814.
114. Raser,J.M. y O'Shea,E.K. (2005). Noise in gene expression: origins, consequences, and control. *Science.* 5743(309):2010-2013.
115. Regenberg,B., Grotkjaer,T., Winther,O., Fausboll,A., Akesson,M., Bro,C., Hansen,L.K., Brunak,S. y Nielsen,J. (2006). Growth-rate regulated genes have profound impact on interpretation of transcriptome profiling in *Saccharomyces cerevisiae*. *Genome Biol.* 11(7):R107-
116. Ren,B., Robert,F., Wyrick,J.J., Aparicio,O., Jennings,E.G., Simon,I., Zeitlinger,J., Schreiber,J., Hannett,N., Kanin,E., Volkert,T.L., Wilson,C.J., Bell,S.P. y Young,R.A. (2000). Genome-wide location and function of DNA binding proteins. *Science* 5500(290):2306-2309.
117. Ringrose,L. (2007). Polycomb comes of age: genome-wide profiling of target sites. *Curr.Opin.Cell Biol.* 3(19):290-297.
118. Robinson,M.D., Grigull,J., Mohammad,N. y Hughes,T.R. (2002). FunSpec: a web-based cluster interpreter for yeast. *BMC.Bioinformatics.*(3:35. Epub;%2002 Nov 13.):35-
119. Rosaleny,L.E., Ruiz-Garcia,A.B., Garcia-Martinez,J., Perez-Ortin,J.E. y Tordera,V. (2007). The Sas3p and Gcn5p histone acetyltransferases are recruited to similar genes. *Genome Biol.* 6(8):R119-
120. Rougvie,A.E. y Lis,J.T. (1988). The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of *D. melanogaster* is transcriptionally engaged. *Cell.* 6(54):795-804.

121. Rudra,D., Mallick,J., Zhao,Y. y Warner,J.R. (2007). Potential interface between ribosomal protein production and pre-rRNA processing. *Mol.Cell Biol.* 13(27):4815-4824.
122. Rudra,D., Zhao,Y. y Warner,J.R. (2005). Central role of Ifh1p-Fhl1p interaction in the synthesis of yeast ribosomal proteins. *EMBO J.* 3(24):533-542.
123. Saguez,C., Olesen,J.R. y Jensen,T.H. (2005). Formation of export-competent mRNP: escaping nuclear destruction. *Curr.Opin.Cell Biol.* 3(17):287-293.
124. Sambrook,J. y Russell,D.W. (2001). *Molecular Cloning: A Laboratory Manual, Third Edition.* Cold Spring Harbor Laboratory Press
125. Sandoval,J., Rodriguez,J.L., Tur,G., Serviddio,G., Pereda,J., Boukaba,A., Sastre,J., Torres,L., Franco,L. y Lopez-Rodas,G. (2004). RNAPol-ChIP: a novel application of chromatin immunoprecipitation to the analysis of real-time gene transcription. *Nucleic Acids Res.* 11(32):e88-
126. Santangelo,G.M. (2006). Glucose signaling in *Saccharomyces cerevisiae*. *Microbiol.Mol.Biol.Rev.* 1(70):253-282.
127. Santiago,T.C., Purvis,I.J., Bettany,A.J. y Brown,A.J. (1986). The relationship between mRNA stability and length in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 21(14):8347-8360.
128. Scrutton,M.C., Wu,C.W. y Goldthwait,D.A. (1971). The presence and possible role of zinc in RNA polymerase obtained from *Escherichia coli*. *Proc.Natl.Acad.Sci.U.S.A.* 10(68):2497-2501.
129. Sekinger,E.A. y Gross,D.S. (2001). Silenced chromatin is permissive to activator binding and PIC recruitment. *Cell.* 3(105):403-414.
130. Sexton,T., Schober,H., Fraser,P. y Gasser,S.M. (2007). Gene regulation through nuclear organization. *Nat.Struct.Mol.Biol.* 11(14):1049-1055.

## Bibliografía

131. Sheth,U. y Parker,R. (2003). Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science* 5620(300):805-808.
132. Steinmetz,E.J., Warren,C.L., Kuehner,J.N., Panbehi,B., Ansari,A.Z. y Brow,D.A. (2006). Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol.Cell.* 5(24):735-746.
133. Struhl,K. (2007). Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nat.Struct.Mol.Biol.* 2(14):103-105.
134. Sun,X.X., Dai,M.S. y Lu,H. (2008). Mycophenolic acid activation of p53 requires ribosomal proteins L5 and L11. *J.Biol.Chem.* 18(283):12387-12392.
135. Sunnerhagen,P. (2007). Cytoplasmic post-transcriptional regulation and intracellular signalling. *Mol.Genet.Genomics* 4(277):341-355.
136. Svejstrup,J.Q. (2004). The RNA polymerase II transcription cycle: cycling through chromatin. *Biochim.Biophys.Acta* 1-3(1677):64-73.
137. Tamkun,J.W. (2007). Stalled polymerases and transcriptional regulation. *Nat.Genet.* 12(39):1421-1422.
138. Thompson,C.M. y Young,R.A. (1995). General requirement for RNA polymerase II holoenzymes in vivo. *Proc.Natl.Acad.Sci.U.S.A* 10(92):4587-4590.
139. Tipper,D.J. (1973). Inhibition of yeast ribonucleic acid polymerases by thiolutin. *J.Bacteriol.* 1(116):245-256.
140. Tolhuis,B., de,W.E., Muijrs,I., Teunissen,H., Talhout,W., van,S.B. y van,L.M. (2006). Genome-wide profiling of PRC1 and PRC2 Polycomb chromatin binding in *Drosophila melanogaster*. *Nat.Genet.* 6(38):694-699.
141. Vanti,M., Gallastegui,E., Respaldiza,I., Rodriguez-Gil,A., Gomez-Herreros,F., Jimeno-Gonzalez,S., Jordan,A. y Chavez,S.



- (2009). Yeast genetic analysis reveals the involvement of chromatin reassembly factors in repressing HIV-1 basal transcription. *PLoS.Genet.* 1(5):e1000339-
142. Vu,L., Siddiqi,I., Lee,B.S., Josaitis,C.A. y Nomura,M. (1999). RNA polymerase switch in transcription of yeast rDNA: role of transcription factor UAF (upstream activation factor) in silencing rDNA transcription by RNA polymerase II. *Proc.Natl.Acad.Sci.U.S.A.* 8(96):4390-4395.
143. Wade,J.T., Hall,D.B. y Struhl,K. (2004). The transcription factor Ifh1 is a key regulator of yeast ribosomal protein genes. *Nature.* 7020(432):1054-1058.
144. Wang,Y., Liu,C.L., Storey,J.D., Tibshirani,R.J., Herschlag,D. y Brown,P.O. (2002). Precision and functional specificity in mRNA decay. *Proc.Natl.Acad.Sci.U.S.A* 9(99):5860-5865.
145. Warner,J.R. (1999). The economics of ribosome biosynthesis in yeast. *Trends Biochem.Sci.* 11(24):437-440.
146. Xu,Z., Wei,W., Gagneur,J., Perocchi,F., Clauder-Munster,S., Camblong,J., Guffanti,E., Stutz,F., Huber,W. y Steinmetz,L.M. (2009). Bidirectional promoters generate pervasive transcription in yeast. *Nature.*
147. Zeitlinger,J., Stark,A., Kellis,M., Hong,J.W., Nechaev,S., Adelman,K., Levine,M. y Young,R.A. (2007). RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nat.Genet.* 12(39):1512-1516.
148. Zenklusen,D., Larson,D.R. y Singer,R.H. (2008). Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat.Struct.Mol.Biol.* 12(15):1263-1271.