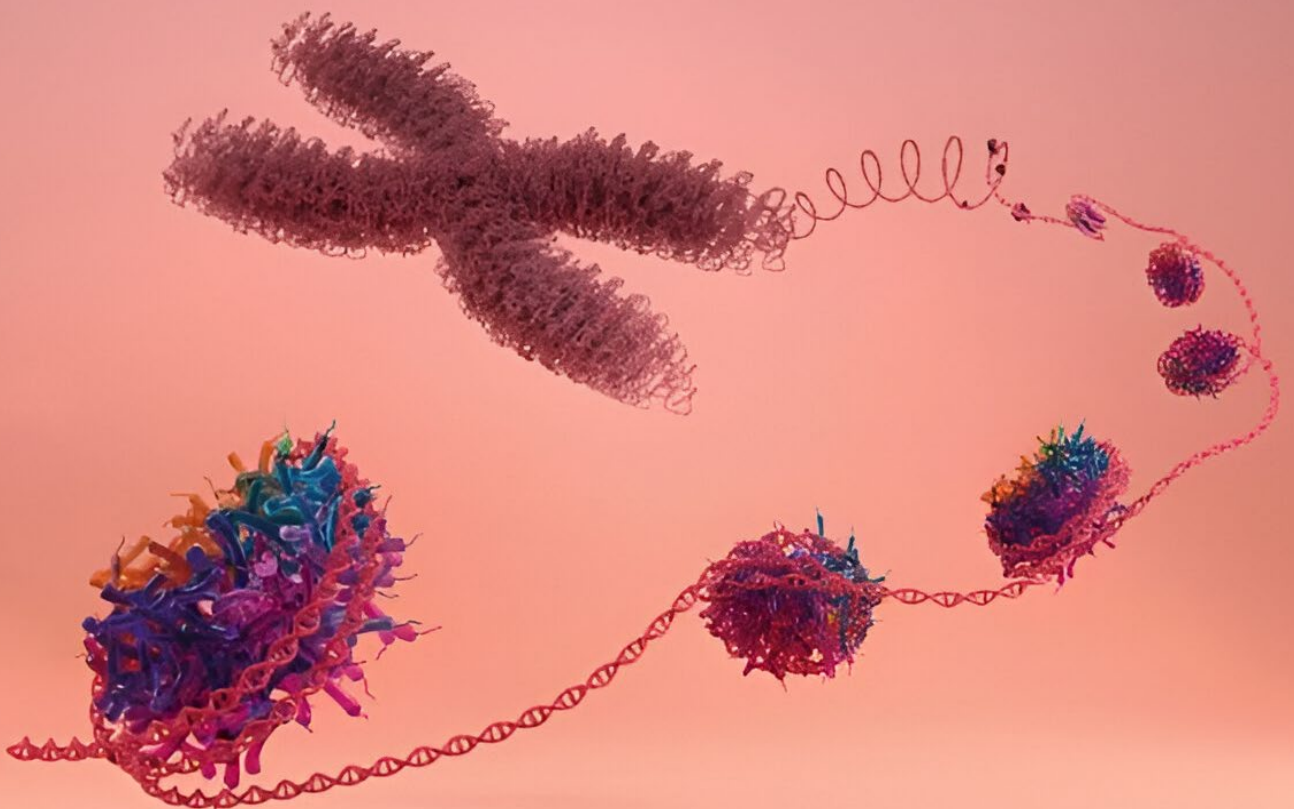


# Regulación del posicionamiento nucleosómico mediante la secuencia del DNA y los remodeladores de cromatina en *Schizosaccharomyces pombe*

Tesis Doctoral  
Isabel Vaquero Eusebio

2024





**Regulación del posicionamiento  
nucleosómico mediante la secuencia del  
DNA y los remodeladores de cromatina en  
*Schizosaccharomyces pombe***

**Tesis Doctoral**

**Isabel Vaquero Eusebio**

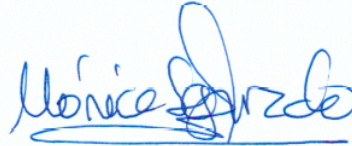
**Salamanca  
2024**



**La Dra. Mónica Segurado Carrascal, Profesora Contratada Doctora Permanente de la Universidad de Salamanca**

CERTIFICA:

Que la memoria titulada "**Regulación del posicionamiento nucleosómico mediante la secuencia del DNA y los remodeladores de cromatina en *Schizosaccharomyces pombe***" presentada por la graduada Dña. Isabel Vaquero Eusebio para optar al grado de Doctor en Biología por la Universidad de Salamanca, dentro del programa de Doctorado en Biología Funcional y Genómica, ha sido realizada bajo la dirección del Dr. Francisco Antequera Márquez y la Dra. Alicia García Martínez en el Centro Mixto Instituto de Biología Funcional y Genómica, CSIC/Universidad de Salamanca. Y para que así conste, expide el siguiente certificado en Salamanca, a 20 de Febrero de 2024.

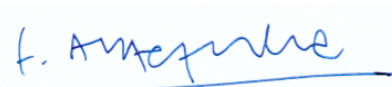


Fdo. Mónica Segurado Carrascal

**El Dr. Francisco Antequera Márquez, Profesor de Investigación y la Dra. Alicia García Martínez, Técnica Titulada Superior, ambos pertenecientes al Consejo Superior de Investigaciones Científicas**

CERTIFICAN:

Que la memoria titulada "**Regulación del posicionamiento nucleosómico mediante la secuencia del DNA y los remodeladores de cromatina en *Schizosaccharomyces pombe***" presentada por la graduada Dña. Isabel Vaquero Eusebio para optar al grado de Doctor en Biología por la Universidad de Salamanca, dentro del programa de Doctorado en Biología Funcional y Genómica, ha sido realizada bajo su dirección en el Centro Mixto Instituto de Biología Funcional y Genómica, CSIC/Universidad de Salamanca. Y para autorizar su presentación y evaluación por el tribunal correspondiente, firman el siguiente certificado en Salamanca, a 20 de Febrero de 2024.



Fdo. Francisco Antequera Márquez



Fdo. Alicia García Martínez



Esta Tesis Doctoral ha sido financiada por un contrato predoctoral del Programa de Ayudas para la Contratación Predoctoral de Personal Investigador (2019-2023) cofinanciado por la Junta de Castilla y León y por el Fondo Social Europeo.



*Ninguna ciencia, en cuanto a  
ciencia, engaña; el engaño está en  
quien no la sabe*

*Miguel de Cervantes Saavedra*



*A mis abuelos Juan y Asun, porque  
comenzasteis este camino y, aunque no  
hayáis podido llegar hasta el final, sé que lo  
celebráis conmigo.*



## AGRADECIMIENTOS

En primer lugar, quiero agradecer a Paco la oportunidad que me ha brindado durante estos años. Gracias por tus consejos, tus enseñanzas y tu paciencia, pero sobre todo por el aprendizaje transmitido, no me hubiese imaginado todo esto cuando llegué hace unos años al laboratorio.

Gracias también a Rodrigo por contribuir con sus ideas a lo largo de estos años en los seminarios de grupo.

Pero, por supuesto, dar las gracias a todas mis compañeras del laboratorio 2.5, sois FAMILIA. Gracias a Loreto por compartir este trayecto conmigo de principio a fin, los atardeceres en el labo no hubieran sido lo mismo sin ti. Millones de gracias a las Lauras por vuestra ayuda tan importante como necesaria en "cosas de ordenadores", sin vuestros *scripts* esta tesis no hubiese tenido fin. Gracias a Laura Rodríguez por nuestras charletas de la tarde que indudablemente, hacían que el día terminara mucho mejor. A Mar, gracias por tu sabiduría, por tu energía incansable y por tus consejos de mamá. A Alicia, gracias, gracias y gracias porque esta tesis es tan tuya como mía. Gracias por enseñarme lo que es el trabajo en el laboratorio y la importancia de hacerlo en equipo, sin tu ayuda las cinéticas hubieran sido interminables. Gracias por tus palabras de ánimo y tus abrazos cuando las cosas se torcían. Has sido y serás mi ejemplo a seguir. Gracias a Marta y a Laura Marín por los cafés y quedadas de compartir risas y lamentos, pero, sobre todo, gracias por Malta 2.0 por hacer de él, un viaje inolvidable. A todas vosotras chicas, ¡Gracias por formar parte de todo esto!

Gracias a Adri por ser más que compañeros de pasillo, gracias por tus ánimos y por las risas que, a tu lado, son infinitas. Por las cosas de Alcalá que solo sabes si eres de Alcalá.

Inés "*the best*", gracias por aparecer en mi vida de casualidad y quedarte en ella para siempre. Gracias por los momentos vividos, las eternas conversaciones por teléfono, las risas y planes alternativos que te caracterizan. Eres de lo mejor que me llevo de aquí.

A Gloria, porque nadie me entiende mejor que tú. Gracias por los fines por Salamanca solucionando nuestros problemas en la terraza de un bar. Gracias por tu apoyo y tus incansables palabras de ánimo. ¡Nos vemos en Dinamarca!

Gracias a todas esas personas, que, a pesar de la distancia, siguen estando ahí cuando las necesito. Gracias en especial a Jenni por esos audios infinitos que con el paso del tiempo se han convertido en *podcasts*, gracias por seguir aguantando mis chapas tantos años. A mis chicas de la "chupipandi" por estar siempre dispuestas a una quedada *express* los domingos para desayunar

y, en definitiva, por crecer a mi lado. Juntas resolvemos mejor los problemas. Gracias a Carlos y a María porque vuestros abrazos quitan los males a cualquiera. A Marisa por todos nuestros viajes y vivencias compartidas, Santander siempre será mi segunda casa. A David por estar siempre dispuesto a escucharme y a resolver todas mis dudas biológicas y no tan biológicas. En general, gracias a todos por hacer mi vida más bonita.

Por último, dar gracias a toda mi familia por estar siempre ahí cuando los necesito, aunque no entendáis ni la mitad de los que os cuento. A mi hermana Paula, gracias por ser tú siempre y por cuidar de mí aun siendo la pequeña. A mis padres, a vosotros os debo todo lo que soy. Gracias por tener una solución para todos los problemas. Gracias por apoyarme en todo, por las conversaciones infinitas y tus palabras de ánimo, mamá, y por tus besos y abrazos calmantes, papá.

En especial a David, eres de lo mejor que tengo en esta vida. Gracias por seguir siendo el sostén emocional de cada día durante todos estos años.

# ÍNDICE





<b>INTRODUCCIÓN</b>	7
<b>1. ORGANIZACIÓN DE LA CROMATINA Y ESTRUCTURA DE LOS NUCLEOSOMAS</b>	9
<b>2. DISTRIBUCIÓN DE LOS NUCLEOSOMAS EN EL GENOMA</b>	10
Posicionamiento y ocupación de los nucleosomas	10
Regiones libres de nucleosomas o NDRs	11
<b>3. DETERMINANTES DEL POSICIONAMIENTO NUCLEOSÓMICO</b>	12
Posicionamiento estadístico	12
Factores de transcripción	13
Secuencia del DNA	13
Secuencias que favorecen o dificultan el posicionamiento nucleosómico	14
<i>Signatures</i> nucleosómicas	15
Remodeladores de cromatina	16
Familia de remodeladores SWI/SNF	18
Complejo RSC	20
Complejo SWI/SNF	21
Familia de remodeladores INO80	21
Familia de remodeladores CHD	22
Chaperonas de histonas y otros remodeladores de cromatina	23
<b>OBJETIVOS</b>	25
<b>RESULTADOS</b>	29
<b>1. EL PAPEL DE LA SECUENCIA DEL DNA EN EL POSICIONAMIENTO DE LOS NUCLEOSOMAS</b>	31
Diseño de secuencias de DNA y análisis de su capacidad para posicionar nucleosomas	31
Matrices de puntuación con frecuencia de mono, di y trinucleótidos	33
Modificación del gen <i>ura4</i> de <i>S. japonicus</i> según las matrices de puntuación	34
Modificación de la ORF del gen <i>ura4</i> de <i>S. japonicus</i> según las matrices de puntuación	35
Análisis de la composición de bases del conjunto de versiones modificadas del gen <i>ura4</i>	38

Posicionamiento nucleosómico mediante desbalance de A y T en el gen <i>ura4</i> de <i>S. pombe</i> y de <i>S. japonicus</i>	39
Posicionamiento en una secuencia artificial mediante desbalance de A y T	41
Desplazamiento de la fase del patrón nucleosómico mediante modificación de la secuencia de DNA	43
Relación entre contenido en G+C y posicionamiento y ocupación nucleosómica	46
Análisis de la relación entre posicionamiento nucleosómico y la transcripción génica	48
Medida de vidas medias de los mRNAs de las distintas versiones del gen <i>ura4</i>	48
Análisis de los niveles de transcripción sentido y antisentido de las distintas versiones del gen <i>ura4</i>	51
<b>2. ANÁLISIS FUNCIONAL DE REMODELADORES DE CROMATINA EN <i>S. pombe</i></b>	53
Construcción de mutantes condicionales de diferentes remodeladores de cromatina en <i>S. pombe</i>	53
Caracterización funcional de la organización nucleosómica y de la expresión génica en presencia y ausencia de remodeladores de cromatina	55
Familia de remodeladores SWI/SNF	55
Dinámica de las NDRs en los mutantes condicionales <i>snf21-sAID</i> y <i>snf22-sAID</i>	56
Análisis de la expresión génica en los mutantes condicionales <i>snf21-sAID</i> y <i>snf22-sAID</i>	59
Características de las NDRs reguladas por RSC y SWI/SNF	64
Familia de remodeladores INO80	65
Dinámica de las NDRs en el mutante condicional <i>ino80-sAID</i>	66
Análisis de la expresión génica en el mutante condicional <i>ino80-sAID</i>	69
Características de las NDRs reguladas por INO80	72
Familia de remodeladores CHD	73
Dinámica de los nucleosomas en los mutantes condicionales <i>hrp1-sAID</i> y <i>hrp3-sAID</i>	74

Análisis de la expresión génica en los mutantes condicionales <i>hrp1-sAID</i> y <i>hrp3-sAID</i>	80
<b>DISCUSIÓN</b>	87
<b>1. LAS MODIFICACIONES EN LA SECUENCIA DEL DNA AFECTAN AL POSICIONAMIENTO DE NUCLEOSOMAS</b>	89
Contribución del posicionamiento nucleosómico a la transcripción	93
<b>2. LOS REMODELADORES DE CROMATINA EN LA ORGANIZACIÓN DE LOS NUCLEOSOMAS</b>	96
Familia SWI/SNF	97
Familia INO80	99
Familia CHD	100
<b>CONCLUSIONES</b>	103
<b>MATERIALES Y MÉTODOS</b>	107
<b>1. MICROORGANISMOS EMPLEADOS</b>	109
<b>2. MEDIOS DE CULTIVO</b>	109
<b>3. GENERACIÓN DE SECUENCIAS QUE POSICIONAN NUCLEOSOMAS</b>	109
Generación de las matrices de puntuación	109
Sesgo de uso de codones	110
Síntesis de las moléculas de DNA	110
<b>4. TÉCNICAS PARA LA CONSTRUCCION DE LAS CEPAS</b>	111
Sistema <i>degrón</i> inducible por auxinas	111
Transformación de microorganismos	111
<b>5. TÉCNICAS DE AISLAMIENTO Y MANIPULACIÓN DEL DNA</b>	112
Obtención de plásmido de DNA de <i>E. coli</i>	112
Obtención de DNA mononucleosómico de <i>S. pombe</i>	112
Obtención de mapas de nucleosomas mediante secuenciación masiva	113
PCR	114
PCR cuantitativa (qPCR o RT-PCR)	114
<b>6. TÉCNICAS DE AISLAMIENTO Y MANIPULACIÓN DE RNA</b>	114
Obtención de RNA total	114
Transcripción reversa o retrotranscripción (RT)	115

<b>7. TÉCNICAS DE OBTENCIÓN Y MANIPULACIÓN DE PROTEINAS</b>	115
Obtención de extractos proteicos	115
<i>Western Blot</i> (WB)	115
<b>8. ESTUDIO DE ESTABILIDADES Y VIDAS MEDIAS DE mRNAs</b>	116
<b>9. ANÁLISIS DE TRANSCRIPCIÓN MEDIANTE <i>RNA-seq</i></b>	117
<b>10. HERRAMIENTAS BIOINFORMÁTICAS</b>	117
<b>BIBLIOGRAFÍA</b>	119
<b>ABREVIATURAS</b>	141
<b>ANEXO I</b>	145
<b>ANEXO II</b>	149

# INTRODUCCIÓN





## 1. ORGANIZACIÓN DE LA CROMATINA Y ESTRUCTURA DE LOS NUCLEOSOMAS

El genoma de los organismos eucariotas está formado por millones de pares de bases que se empaquetan en el núcleo celular. Para ello, debe sufrir un alto grado de compactación formando una estructura llamada cromatina (**Figura 1**).

Las unidades fundamentales de la cromatina son los nucleosomas, los cuales facilitan el empaquetamiento del genoma en el núcleo y a su vez, modulan el acceso de elementos reguladores a la molécula de DNA.

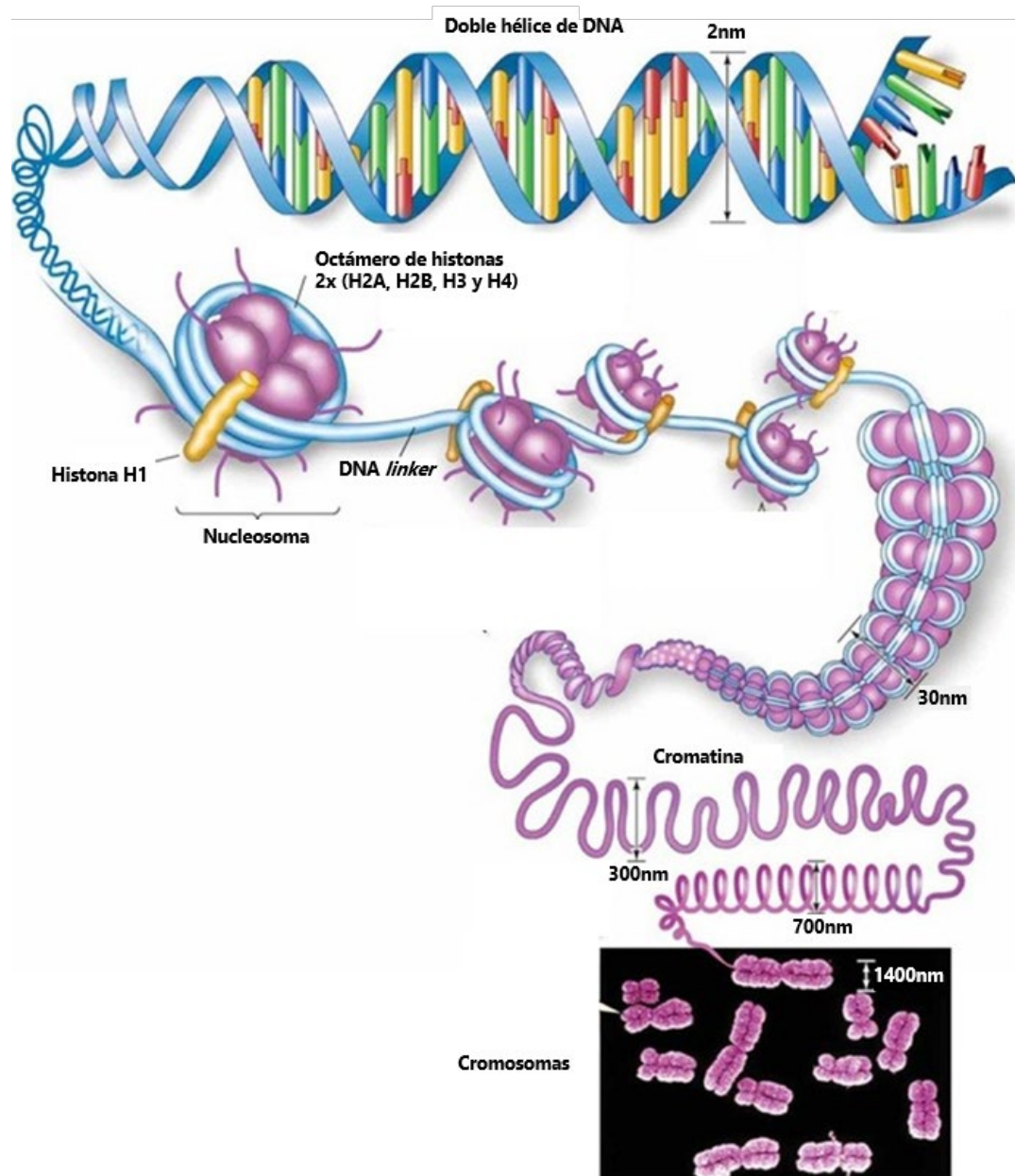


Figura 1: Niveles de organización de la cromatina.

Los nucleosomas están formados por un octámero de proteínas histonas alrededor del cual se enrollan 147 pb de DNA, que dan 1.7 vueltas en una superhélice levógira (Luger et al., 1997; Baldi et al., 2018). Los nucleosomas están conectados mediante un fragmento de DNA espaciador o *linker*, cuya longitud varía dependiendo de la especie (Chereji et al., 2018). Por ejemplo, en *S. pombe*, tiene un tamaño de unas 4 pb (Moyle-Heyrman et al., 2013), en *S. cerevisiae* de unas 10 pb (Wang et al., 2008) y, en humanos alcanza un tamaño de hasta 60 pb (Wang et al., 2021a).

Cada nucleosoma está formado por dos moléculas de cada una de las histonas H2A, H2B, H3 y H4. Dos heterodímeros H3/H4 se asocian entre sí formando un tetrámero, situado en el centro del nucleosoma y, a cada lado se ensambla un dímero de H2A/H2B. Sin embargo, solo hay una única copia de la histona H1, la cual se une al DNA *linker* (Veron et al., 2006). Los contactos entre las histonas y el DNA cerca de la diada son más estables que las interacciones con el resto del DNA en 14 puntos a lo largo del nucleosoma (revisado en Markert & Luger 2020).

La carga neta del octámero es positiva gracias a la elevada cantidad de aminoácidos básicos de las histonas, que se asocian con los grupos fosfato de la cadena de DNA con carga negativa. Esta interacción iónica es la principal fuerza de unión entre ambos elementos, aunque también existen puentes de hidrógeno y fuerzas hidrofóbicas entre ellos (revisado en McGinty & Tan 2015).

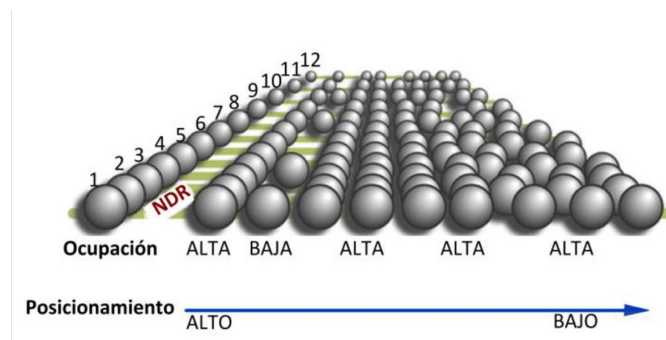
Los extremos amino-terminales de las histonas son flexibles y se proyectan hacia el exterior del *core* nucleosómico. Estas colas poseen muchos residuos susceptibles de sufrir modificaciones postraduccionales, que pueden alterar la unión entre el nucleosoma y el DNA y, además, son reconocidas por proteínas y complejos reguladores. Estas, juegan un papel importante en la regulación de procesos como la replicación, transcripción, recombinación y reparación (revisado en Kornberg & Lorch 2020). Por ejemplo, la di y trimetilación de la histona H3 en la lisina 9 (H3K9) es una marca característica de silenciamiento de la heterocromatina para la mayoría de los organismos eucariotas, entre ellos *S. pombe*. Sin embargo, este silenciamiento se compensa con la acetilación de las colas de H3K9, llevada a cabo por acetiltransferasas, lo que sugiere que este proceso de metilación-acetilación de H3K9 actúa como un interruptor de la actividad de la heterocromatina (Allshire & Ekwall 2015).

## 2. DISTRIBUCIÓN DE LOS NUCLEOSOMAS EN EL GENOMA

### Posicionamiento y ocupación de los nucleosomas

Los términos de posicionamiento y ocupación son importantes para entender la distribución de los nucleosomas en el genoma cuando se analizan en poblaciones de células. El posicionamiento hace referencia a la localización que ocupa cada nucleosoma con respecto a la

secuencia del DNA. Un nucleosoma está bien posicionado cuando se sitúa en la misma posición en todas o en la mayoría de las células de la población. Si, por el contrario, se localiza sobre secuencias distintas en células diferentes, el nucleosoma no está posicionado. La ocupación se refiere a la frecuencia con la que una región del genoma está asociada a un nucleosoma (**Figura 2**).



**Figura 2. Diferencias entre ocupación y posicionamiento nucleosómico.** Diagrama de distintos tipos de posicionamiento y ocupación nucleosómica. Las dos características son independientes dado que una región del genoma puede tener alta ocupación y bajo posicionamiento y viceversa. Se representan con esferas grises los nucleosomas que se encuentran sobre la misma secuencia del DNA en una población de 12 células. Modificado de Pugh (2010).

### Regiones libres de nucleosomas o NDRs

Las regiones desprovistas de nucleosomas o NDRs (*Nucleosome Depleted Regions*), se identificaron inicialmente como regiones hipersensibles a la digestión con nucleasas (Groudine et al., 1981; Felsenfeld et al., 1996). Se localizan, en general, en el extremo 5' de los genes, aguas arriba del sitio de inicio de la transcripción (TSS). Las NDRs están flanqueadas por dos nucleosomas que suelen estar bien posicionados, denominados -1 y +1 (Lantermann et al., 2010; Givens et al. 2012; Deniz et al., 2016). El nucleosoma +1 es el primer nucleosoma dentro de la región transcrita y suele incluir el TSS (Kim et al., 2021). En la levadura *Saccharomyces cerevisiae*, muchos genes también presentan una NDR en el extremo 3', de las cuales, algunas corresponden al extremo 5' del gen siguiente (Shivaswamy et al., 2008; Vaillant et al., 2010), hecho que es menos frecuente en *S. pombe* (Lantermann et al., 2010; Soriano et al., 2013) quizás debido a que la densidad de genes es menor.

Las NDRs colocalizan con los promotores de muchos genes y por eso están enriquecidas en secuencias consenso de unión de factores de transcripción. (Wunderlich & Mirny 2009; Soriano et al., 2013; Mivelaz et al., 2020).

### 3. DETERMINANTES DEL POSICIONAMIENTO NUCLEOSÓMICO

#### Posicionamiento estadístico

El concepto de posicionamiento estadístico se propuso hace mucho tiempo para explicar el posicionamiento de los nucleosomas a lo largo del genoma (Kornberg & Stryer 1988). Este modelo postula la existencia de barreras físicas a partir de las cuales se irían posicionando pasivamente los nucleosomas uno tras otro manteniendo una distancia regular entre ellos. Si el espacio que queda disponible entre dos barreras permite incorporar un número entero de nucleosomas, estos tendrán un espaciamiento regular. Si, por el contrario, la distancia no permite la localización de un número entero de nucleosomas, aquellos más cercanos a las barreras permanecerán bien posicionados respecto a la secuencia del DNA y a medida que nos alejamos de ellas, la fase del posicionamiento se irá perdiendo (Vaillant et al., 2010; Krietenstein et al., 2016; revisado en Singh & Mueller-Planitz 2021).

Estas barreras incluyen a las NDRs, bien porque su secuencia de DNA tiene propiedades mecánicas de flexibilidad que desfavorecen su ensamblaje en nucleosomas, o bien porque a ellas pueden unirse factores de transcripción que compiten con los nucleosomas por la ocupación de esa región. También, podrían considerarse barreras, secuencias de DNA que unen nucleosomas con una elevada afinidad como ocurre con el nucleosoma +1, a partir del cual, se posicionaría el resto de los nucleosomas (Deniz et al., 2016; Chereji et al., 2018). Este modelo de posicionamiento está bastante estudiado en *S. cerevisiae*, pero no está tan bien caracterizado en *S. pombe* (Lantermann et al., 2010). Este hecho se ha demostrado experimentalmente hace unos años en nuestro laboratorio cuando se sustituyó la ORF (*marco de lectura abierto*) del gen *SPBC16G5.03* de *S. pombe* que tiene una expresión constitutiva y no es esencial, por el gen procariótico *KanMX6*. El gen *SPBC16G5.03* presenta siete nucleosomas bien posicionados y está asociado a una NDR aguas arriba del TSS. Este reemplazo de la ORF no alteró la posición del nucleosoma +1 del gen reemplazado, sin embargo, se mostró un perfil desorganizado del resto de nucleosomas sobre el gen *KanMX6* (González et al., 2016). Por lo tanto, al menos en este caso, no se cumple el posicionamiento estadístico que postula que la pérdida del posicionamiento de los nucleosomas a lo largo de las ORFs es dependiente de la desestabilización del nucleosoma +1 (Mavrich et al., 2008a).

Aunque es posible que el posicionamiento estadístico contribuya en algunos casos a la distribución de los nucleosomas, en la actualidad, muchos estudios concluyen que esta depende de la acción combinada de tres factores principales: los factores de transcripción, la secuencia del

DNA y los remodeladores de cromatina (Segal & Widom 2009; revisado en Singh & Mueller-Planitz 2021). La contribución de cada uno de ellos se explica con más detalle a continuación.

### **Factores de transcripción**

Los factores transcripción se unen a secuencias consenso que, aunque se encuentran dispersas por todo el genoma, presentan una mayor densidad en las regiones promotoras (Hughes & de Boer 2013; Weingarten-Gabbay & Segal 2014; Avsec et al., 2021) que pueden ser regiones de baja ocupación nucleosómica como las NDRs, o simplemente las regiones 5' de los genes. Se ha demostrado que la delección de algunos factores de transcripción causa la desaparición de algunas NDRs, como ocurre en el mutante Sap1 en *S. pombe* (Tsankov et al., 2011) y los mutantes Reb1 y Abf1 en *S. cerevisiae* (Badis et al., 2008; Hartley & Madhani 2009). Sin embargo, no todas las NDRs con sitios de unión para esos factores se cierran en su ausencia, probablemente debido a la unión de otros factores al mismo promotor. Por ejemplo, solo el 19% de las NDRs a las que se unen los factores de transcripción Atf1 y Pcr1 identificados mediante inmunoprecipitación de cromatina se cierran en ausencia de los dos factores en *S. pombe* (Soriano et al., 2013).

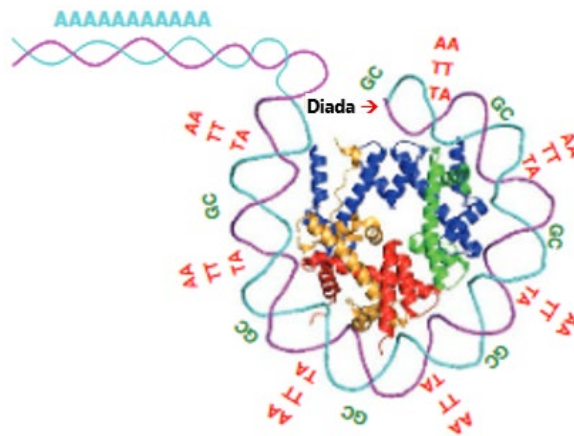
Además, los factores de transcripción también pueden contribuir al posicionamiento de los nucleosomas de manera indirecta, reclutando remodeladores de cromatina a regiones específicas del genoma (Yen et al., 2012). Por ejemplo, factores de transcripción como Rap1 y Reb1 se unen preferentemente a sus sitios de reconocimiento localizados en la superficie del nucleosoma -1, hecho que hace que se recluten los remodeladores de cromatina RSC e INO80 (Koerber et al., 2009; Singh & Mueller-Planitz 2021).

### **Secuencia del DNA**

La secuencia del DNA puede contribuir a determinar las posiciones de los nucleosomas, aunque existe cierta controversia sobre su grado de participación. Como se ha explicado anteriormente, algunos autores apoyan el modelo estadístico (Kornberg & Stryer 1988). Otros autores proponen que la posición de cada nucleosoma viene determinada por la secuencia de nucleótidos a los que se une (Segal et al., 2006; Kaplan et al., 2009). Sin embargo, las características de la secuencia del DNA que influyen en el posicionamiento de los nucleosomas varían de unas especies a otras (Quintales et al., 2015a) e incluso, dentro de una misma especie (González et al., 2016; García et al., 2024), lo que supone la existencia de un código degenerado capaz de adaptarse a muchas secuencias diferentes del genoma y a la vez, lo suficientemente preciso para ser específico de cada especie.

### Secuencias que favorecen o dificultan el posicionamiento nucleosómico

En promedio, las secuencias asociadas con nucleosomas bien posicionados generalmente poseen los dinucleótidos AA/TA/TT en el surco menor de la doble hélice de DNA orientado hacia el interior, mientras que los dinucleótidos GG/GC/CC son más frecuentes en el surco menor cuando se encuentra orientado hacia el exterior del octámero de histonas (**Figura 3**). Además, esta preferencia por los dinucleótidos AA/TA/TT se repite periódicamente cada 10 pb que oscilan en fase entre sí y fuera de fase con dinucleótidos GC periódicos cada 10 pb (Struhl & Segal 2013; revisado en Singh & Mueller-Planitz 2021). Esta periodicidad no se detecta en nucleosomas individuales, sino en el promedio de muchos miles de nucleosomas a lo largo del genoma.



**Figura 3: Estructura de un nucleosoma.** Las histonas H3 (azul), H4 (verde), H2A (amarillo) y H2B (rojo), forman el octámero sobre el que se enrollan 147 nucleótidos de DNA de doble cadena. La diada es el punto medio del DNA mononucleosómico. En promedio, los dinucleótidos AA/TA/TT se localizan en el surco menor cuando está orientado hacia el interior, mientras los dinucleótidos GG/GC/CC son más frecuentes en el surco menor cuando se encuentra orientado hacia el exterior del octámero de histonas. Modificado de Struhl & Segal (2013). La ilustración sólo representa la mitad del DNA nucleosómico.

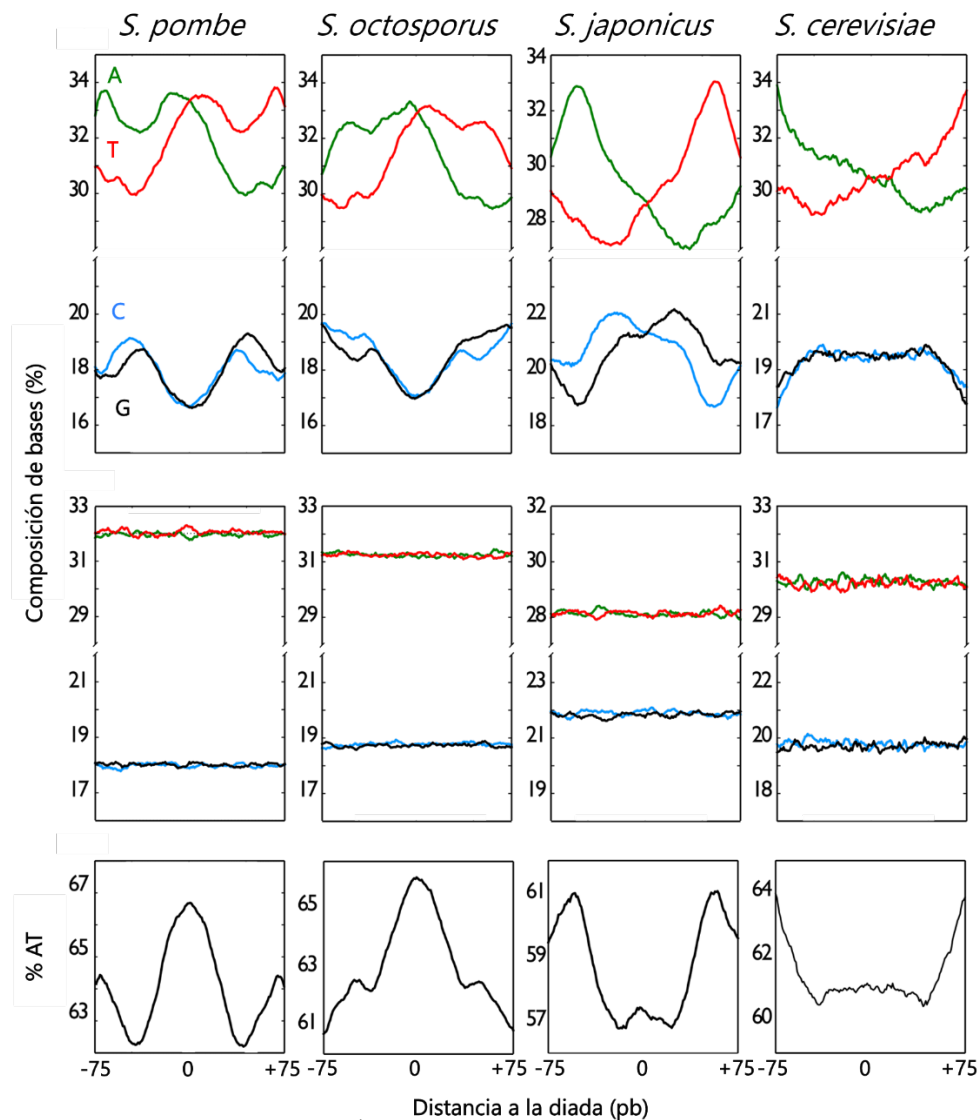
La composición de la secuencia está directamente relacionada con la rigidez y la flexibilidad de la molécula de DNA (Basu et al., 2021a). Según estudios recientes, el dinucleótido TA tiene el valor de flexibilidad más positivo, lo que indica que la presencia de TA hace que el DNA sea más flexible en promedio, sin embargo, el dinucleótido CG tiene el valor de flexibilidad más negativo, lo que indica una mayor rigidez en la cadena de DNA (Basu et al., 2021b). La epigenética también tiene un papel relevante en la estructura del nucleosoma porque la metilación de dinucleótidos CG hace más rígida la molécula de DNA alterando la estabilidad de los nucleosomas (Basu et al., 2022).

Todo ello indica que la flexibilidad de la cadena de DNA es esencial para la formación de un complejo con el octámero de histonas, mientras que las secuencias que ofrecen más resistencia a la torsión se relacionan con exclusión nucleosómica. Ejemplos de estas últimas son los poli (dA:dT) y los poli (dG:dC) (Tillo et al., 2009; revisado en Singh & Mueller-Planitz 2021). Así, en *S. cerevisiae*, las NDRs están enriquecidas en estas secuencias poli (dA:dT), sin embargo, en *S. pombe* no aparecen en ellas con mayor frecuencia que en el resto del genoma (García et al., 2017), que sí que presenta un mayor contenido promedio de nucleótidos de A y T en torno a la diada del nucleosoma (Quintales et al., 2015a). Por lo tanto, secuencias idénticas en su composición de bases, pueden tener distinta capacidad para posicionar nucleosomas en el contexto genómico, dependiendo de la especie en la que se encuentren.

### ***Signatures nucleosómicas***

Las *signatures* nucleosómicas son los patrones de distribución que presentan los cuatro nucleótidos a lo largo de las 147 pb del DNA nucleosómico (Quintales et al., 2015a). Las *signatures* de especies diferentes (**Figura 4**) difieren entre sí, aunque, tienen en común un mayor contenido de A en el extremo 5' y de T en el extremo 3'. Además, los perfiles de distribución de A y C son imagen especular de los de T y G, respectivamente, respecto de la posición de la diada. Las *signatures* de *S. pombe* y *S. octosporus*, tienen un mayor contenido de A y T en la zona de la diada que las de *S. japonicus* y *S. cerevisiae* (**Figura 4**).

La característica más importante de las *signatures* nucleosómicas es que contienen información para posicionar nucleosomas *in vivo*. Este hecho se demuestra porque cuando se diseñaron moléculas artificiales de DNA de secuencia arbitraria que incorporaban las *signatures* de *S. pombe* o *S. cerevisiae*, las moléculas resultantes eran capaces de posicionar nucleosomas en las regiones predichas en cada una de las dos especies, pero no cuando se integraban en la otra (González et al., 2016).



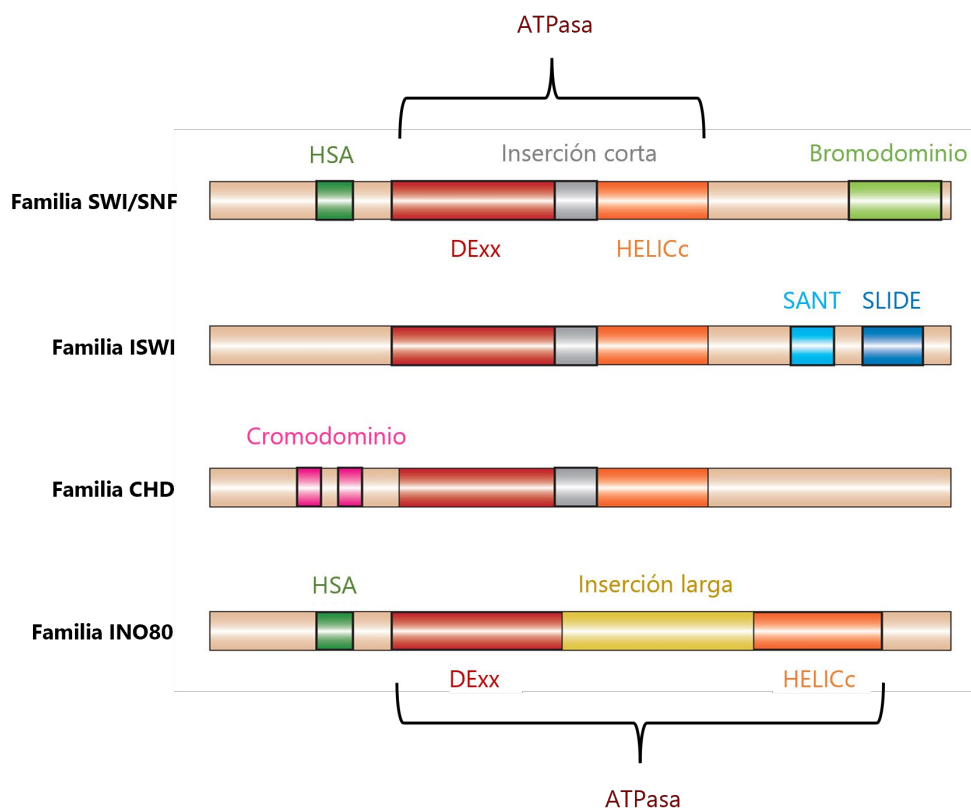
**Figura 4: Signatures nucleosómicas.** En los diagramas de la fila superior se muestra la distribución de los cuatro nucleótidos en secuencias de DNA nucleosómico de *S. pombe*, *S. octosporus*, *S. japonicus* y *S. cerevisiae*, alineadas a la diada del nucleosoma (posición 0). En los diagramas centrales, se muestra la distribución de los cuatro nucleótidos en el mismo número de secuencias de DNA del mismo tamaño que en el diagrama de la parte superior, pero seleccionadas al azar en el genoma. En la fila inferior, se muestra el contenido en A+T del DNA nucleosómico de los diagramas superiores. Modificado de Quintales et al (2015a).

### Remodeladores de cromatina

Los remodeladores de cromatina son complejos multiproteicos que emplean la hidrólisis del ATP para romper las interacciones entre el DNA y el octámero de histonas (Hughes & Rando 2015; revisado en Clapier et al., 2017; Markert & Luger 2020). De esta manera, pueden desalojar los nucleosomas del genoma o desplazarlos a una localización diferente sin necesidad de desensamblar su estructura (Nodelman & Bowman 2021).

Aún no está bien definido cuál es el mecanismo que dirige a los remodeladores de cromatina a cada *locus* específico. Es posible que esta función dependa de las modificaciones postraduccionales de las colas de las histonas, como la trimetilación de la lisina 4 en la histona H3 (H3K4), reconocida por los remodeladores CHD1 en humanos (revisado en Singh & Mueller-Planitz 2021), las interacciones físicas con factores de transcripción como Rap1 con RSC en *S. cerevisiae* (Yadon et al., 2013) e incluso, la propia secuencia del DNA nucleosómico, a través del módulo del *core* y la subunidad Arp8 del remodelador INO80 en *S. cerevisiae* (Oberbeckmann et al., 2021).

Los remodeladores de cromatina se pueden dividir en cuatro familias según su dominio ATPasa (**Figura 5**): SWI/SNF (la familia SWI/SNF incluye los complejos SWI/SNF y RSC), INO80, CHD e ISWI (Klein-Brill et al., 2019). Desde el punto de vista de la función que realizan, se dividen en tres grupos: los que desplazan el nucleosoma +1 alejándolo del TSS (*pushers*), los que lo acercan (*pullers*), y los que mantienen la periodicidad de los nucleosomas a lo largo de los genes (*spacers*) (Kubik et al., 2019; Singh et al., 2021).



**Figura 5. Familias de remodeladores de cromatina.** Todas las familias de remodeladores de cromatina tienen una subunidad ATPasa dividida en dos módulos similares a RecA (*RecA-like*): DExx (módulo 1, rojo) y HELICc (módulo 2, naranja) separados por una inserción corta (gris) en el caso de las familias SWI/SNF, ISWI y CHD o larga (amarillo) en el caso de la familia INO80. A ambos lados del dominio catalítico, poseen otras regiones necesarias para su función como son los bromodominios (verde claro) y cromodominios (rosa) encargados de reconocer modificaciones epigenéticas de las

histonas. El módulo SANT-SLIDE (azul), únicamente presente en la familia ISWI, interacciona con el DNA nucleosómico. El dominio HSA (verde oscuro), presente en las familias SWI/SNF e INO80, participa en la regulación de la actividad del dominio ATPasa. Modificada de Clapier & Cairns (2009).

*RecA-like* es un dominio estructural presente en las ATPasas tales como helicasas, que utiliza la energía de la unión de nucleótidos y la hidrólisis de ATP para realizar un trabajo mecánico. Las ATPasas pueden contener diferente número de dominios *RecA-like* pero todos ellos regulan la función del dominio ATPasa y modulan la interacción con otras proteínas (revisado en Clapier et al., 2017). DExx y HELICc son dominios localizados en los extremos N-terminal y C-terminal respectivamente dentro de la subunidad ATPasa (Caruthers & McKay 2002). Lo que caracteriza e individualiza a cada una de las cuatro familias de remodeladores de cromatina son los dominios únicos, adicionales a los módulos DExx y HELICc y que, además, son responsables de la función específica de cada una de ellas (revisado en Singh & Mueller-Planitz 2021). Los remodeladores de las familias SWI/SNF, ISWI y CHD presentan una inserción corta dentro del dominio ATPasa, mientras que los remodeladores de la familia INO80 presentan una inserción larga. Además, la familia SWI/SNF posee un bromodominio y un dominio HSA (helicasa SANT). El bromodominio reconoce los residuos de lisinas acetiladas de las histonas que es imprescindible para llevar a cabo la acción de este remodelador, y el dominio HSA, por su parte, se une a la actina o a proteínas relacionadas con la actina (ARP) para colaborar en la regulación de la actividad del dominio ATPasa (revisado en Clapier et al., 2017).

El elemento estructural identificador de la familia ISWI es el módulo SANT-SLIDE encargado de interaccionar con el DNA nucleosómico (Ryan et al., 2011). Por su parte, la familia CHD o *Chromodomain Helicase DNA-binding* presenta cromodominios, motivos estructurales cortos (aproximadamente 50 aminoácidos) en tándem en el extremo N-terminal que son los encargados de reconocer dimetilaciones o trimetilaciones de H3K4 (Morettini et al., 2011). Por último, la familia INO80 posee el dominio HSA al igual que la familia SWI/SNF (Clapier & Cairns 2009).

Dado que *S. pombe*, el organismo modelo utilizado en este trabajo, carece de los remodeladores de la familia ISWI (Shim et al., 2012), en los apartados siguientes se describirán las familias de remodeladores presentes en esta levadura. La estructura de los remodeladores en *S. pombe* ha sido muy poco estudiada y por eso, la mayor parte de la información hace referencia a *S. cerevisiae*.

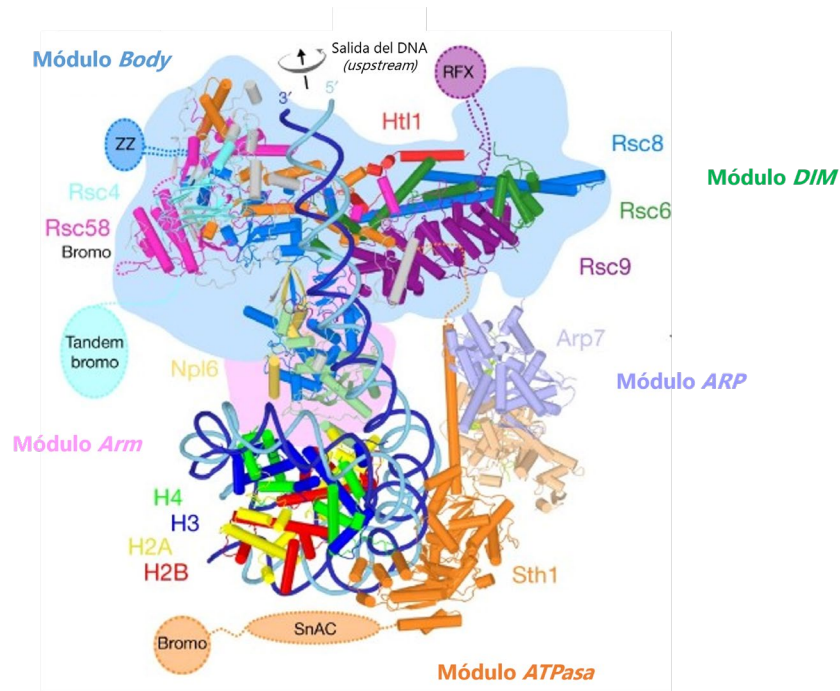
### **Familia de remodeladores SWI/SNF**

La familia de remodeladores SWI/SNF (*Switch defective/Sucrose Non-Fermenting*) pertenece al grupo de los denominados *pushers*, encargados de alejar al nucleosoma +1 del

promotor de manera que, en su ausencia, este nucleosoma se desplaza hacia la NDR y reduce su tamaño (Kubik et al., 2019).

Esta familia incluye los complejos SWI/SNF y RSC. En *S. cerevisiae*, estos complejos están formados por 17 y 11 subunidades, respectivamente. En *S. pombe*, están formados por 12 y 13 subunidades, de las cuales comparten seis (**Tabla 1**) (Monahan et al., 2008). En *S. pombe*, 8 de las 13 subunidades de RSC están codificadas por genes esenciales mientras que en SWI/SNF, solamente 4 son esenciales (revisado en Prajapati et al., 2020). Una diferencia importante entre los dos complejos de *S. pombe* es que la subunidad catalítica de RSC, Snf21 (*STH1* en *S. cerevisiae*) es esencial, mientras que la subunidad correspondiente al complejo SWI/SNF, Snf22 (*SNF2* en *S. cerevisiae*), no lo es.

En *S. cerevisiae*, las 17 subunidades del complejo RSC se organizan en cinco módulos (*DIM*, *Body*, *ARP*, *Arm* y *ATPasa*) (**Figura 6**), cada uno de ellos con funciones diferentes. El módulo principal, *Body*, está formado por siete proteínas, cinco de las cuales poseen bromodominios que contactan con grupos acetilo de las colas de las histonas mientras que el módulo que interactúa con el DNA (*DIM*) contacta con la secuencia del promotor para llevar a cabo la actividad remodeladora de nucleosomas (Wagner et al., 2020). El módulo *Arm*, es el encargado de interactuar con la superficie del nucleosoma y el módulo *ATPasa* lleva a cabo la hidrólisis de ATP necesaria para la acción remodeladora (Han et al., 2020). El complejo SWI/SNF en *S. cerevisiae* está formado por 11 subunidades que se agrupan en cuatro módulos (*Body*, *ARP*, *Arm* y *ATPasa*) según la función que desempeñan. En este complejo remodelador, el módulo principal (*Body*) se organiza a su vez en tres módulos (*Spine*, *Hinge* y *Core*) con funciones muy específicas pero todos ellos contribuyen al ensamblaje del complejo.



**Figura 6. Estructura del complejo RSC.** Estructura del complejo remodelador de cromatina RSC de *S. cerevisiae*. Lo forman 17 subunidades agrupadas en cinco módulos diferentes: *Body* (azul), *DIM* (verde), *Arm* (rosa), *ARP* (lila) y *ATPasa* (naranja) y nueve dominios flexibles que los conectan entre sí. El nucleosoma se encuentra situado entre el módulo *Arm* y el módulo *ATPasa*. Las cuatro histonas que forman el nucleosoma se representan en verde, azul, amarillo y rojo. La flecha indica la dirección de salida de la molécula de DNA. Modificada de Wagner et al. (2020).

## Complejo RSC

El complejo de RSC (*Remodeling the Structure of Chromatin*) tiene una función más general en la regulación de la transcripción, participando tanto en la activación como en la represión transcripcional (Monahan et al., 2008). De hecho, se considera el complejo remodelador más importante para la transcripción desde levaduras a humanos (revisado en Magaña-Acosta & Valadez-Graham 2020) porque afecta a un número de genes mayor que SWI/SNF (revisado en Prajapati et al., 2020). Este complejo remodelador de cromatina se asocia al correcto mantenimiento de las NDRs, estableciendo las posiciones de los nucleosomas +1 y -1 (Rawal et al., 2018), tanto en *S. pombe* como en *S. cerevisiae*, contribuyendo al mantenimiento de la posición preferente del TSS (Yague-Sanz et al., 2017; Kubik et al., 2018). Además, se ha propuesto que la función del complejo RSC es muy importante para la progresión del ciclo celular. Así, su subunidad catalítica, Snf21, se expresa principalmente en la fase G2-M del ciclo celular y su expresión baja en la fase S (Yamada et al., 2008).

Este remodelador de cromatina, se encuentra en NDRs asociado a factores de transcripción como Rap1 (revisado en Singh & Mueller-Planitz 2021), un factor de transcripción pionero que une al DNA nucleosómico y colabora con el remodelador de cromatina RSC para desplazar al

nucleosoma +1, permitiendo así el acceso a otros factores de transcripción (Mivelaz et al., 2020). Rap1, regula la expresión de más del 90% de los genes ribosomales mediante el reclutamiento de factores de transcripción adicionales como Hmo1, Fhl1 e Ifh1 en *S. cerevisiae* (Knight et al., 2014).

### Complejo SWI/SNF

Estudios recientes en *S. cerevisiae*, han puesto de manifiesto que la función remodeladora del complejo SWI/SNF se limita a una serie de promotores discretos, principalmente a genes que poseen caja TATA en sus promotores, no teniendo un efecto global sobre todas las NDRs del genoma. Así, defectos en la función de la subunidad catalítica, Snf22, tienen un impacto específico sobre las NDRs de genes con una elevada tasa de transcripción, que reducen su tamaño. Algunos estudios, sugieren que SWI/SNF podría unirse a la RNA polimerasa II (RNAPII) directamente o facilitar la unión de esta a los promotores y promover así el proceso de transcripción (Wilson et al., 1996) como ha sido descrito para el promotor del gen codificante para el factor de transcripción Gcn4 (Rawal et al., 2018).

### Familia de remodeladores INO80

La familia de remodeladores INO80 (*Inositol requiring 80*) pertenece al grupo de los denominados *pullers*, y tiene como diana de su actividad al nucleosoma +1, igual que los complejos RSC y SWI/SNF, aunque realizando sobre el mismo funciones opuestas a la que llevan a cabo los complejos citados anteriormente. Así, defectos en la función del complejo INO80, determinan el incremento de tamaño de las NDRs. Como los remodeladores anteriores, el complejo INO80, tampoco actúa sobre el nucleosoma -1 (Kubik et al., 2019), sin embargo, también se le puede considerar como un *spacer* porque participa en el espaciamiento de nucleosomas a lo largo de los genes (Kubik et al., 2019; Cheon et al., 2021; Singh et al., 2021).

El complejo INO80 está muy conservado en la evolución. En *S. pombe* está formado por 14 subunidades, siendo la subunidad catalítica Ino80, esencial para la viabilidad celular (revisado en Prajapati et al., 2020). Este remodelador de cromatina, se encuentra en las NDRs asociado a factores de transcripción como Reb1. En ellas, se posiciona entre el factor de transcripción y el nucleosoma +1, ajustando a este último en su posición óptima para facilitar la activación de la transcripción (Oberbeckmann et al., 2021).

INO80 también se encuentra en otras localizaciones genómicas como son los centrómeros donde colabora con el remodelador Hrp1 (familia CHD) para llevar a cabo el intercambio de la histona H3 por la variante centromérica CENP-A (Choi et al., 2017). Además, está muy representado en las regiones de los TSS, e interacciona físicamente con el complejo de elongación

de la RNAPII para modular el sitio de pausa de esta en genes asociados a INO80 en *S. cerevisiae* (Cheon et al., 2021).

### Familia de remodeladores CHD

La familia de remodeladores CHD (*Chromodomain Helicase DNA-binding*) pertenece al grupo de los *spacers*. Los remodeladores CHD actúan sobre los nucleosomas que ocupan posiciones a lo largo de los genes y mantienen el correcto espaciado entre ellos (Kubik et al., 2019). Se desconoce cuál es la composición en subunidades de este complejo remodelador en *S. pombe* y si está formado por subunidades adicionales a la subunidad catalítica. En *S. cerevisiae*, Chd1, actúa como un remodelador monomérico, sin embargo, en metazoos, los remodeladores CHD son mucho más complejos y llevan a cabo además de las funciones de espaciado de nucleosomas, la incorporación de la variante histónica H3.3 (revisado en Clapier et al., 2017).

En *S. pombe*, la familia CHD está formada por las proteínas Mit1, Hrp1 y Hrp3 (Flaus et al., 2006). La doble delección de Hrp1 y Hrp3 no compromete la viabilidad celular, aunque estos mutantes presentan un fenotipo severo en el posicionamiento y ocupación nucleosómica, además de un aumento de la transcripción antisentido con relación a los mutantes de delección simples. La ausencia de las tres proteínas es letal (Pointner et al., 2012).

Mit1 es una proteína que está conservada evolutivamente, aunque no aparece en *S. cerevisiae* (Lantermann et al., 2010). Forma parte del complejo SHREC, que en *S. pombe* está implicado en la represión de la transcripción de las regiones heterocromáticas como el *locus* del *mating type* (Creamer et al., 2014).

Hrp1 está asociado a las regiones centroméricas y es necesario para mantener el nivel de la variante histónica CENP-A, de tal manera que interacciona directamente con el centrómero durante la fase S del ciclo celular (Walfridsson et al., 2005; Yamada et al., 2008).

Hrp3 regula el posicionamiento y estabilidad de los nucleosomas, evitando así la transcripción críptica (Touat-Todeschini et al., 2012; Dong et al., 2021). La ausencia de un remodelador Hrp3 funcional produce un descenso de la ocupación nucleosómica alrededor del TSS, desde el nucleosoma +1 y, un aumento de esta en las regiones 3' de los genes, en torno al sitio de terminación de la transcripción (TTS), lo que sugiere que aquellas zonas que más se transcriben son las más susceptibles a sufrir la pérdida de posicionamiento de nucleosomas. Esto, pone en manifiesto el papel clave de Hrp3 en el posicionamiento de los nucleosomas de las regiones transcritas, sin verse afectadas las regiones intergénicas (Hennig et al., 2012; Shim et al., 2012). Con relación a su función asociada a la transcripción génica, este remodelador interacciona

con la RNAPII, facilitando el proceso de elongación para posicionar los nucleosomas tras el paso de la misma (Ocampo et al., 2019).

**Tabla 1: Subunidades de las familias de remodeladores de *S. pombe*.** El complejo de remodelación SWI/SNF está formado por 12 subunidades entre las que se encuentra su subunidad catalítica (Snf22), RSC está formado por 13 subunidades. Snf21 es su subunidad catalítica. SWI/SNF y RSC comparten seis subunidades (asterisco). El complejo INO80 está formado por 14 subunidades, recibiendo su subunidad catalítica el mismo nombre que el complejo. En el caso de la familia CHD, no se ha descrito si sus miembros están compuestos por una o varias subunidades.

SWI/SNF		CHD	INO80
SWI/SNF	RSC	CHD	INO80
Snf22	Snf21	?	Ino80
Sol1	Rsc1		les1
Snf5	Rsc9		les3
Ssr1*	Rsc4		les5
Arp9*	Ssr1*		Nhp10
Snf59	Arp9*		Arp4
Ssr2*	Ssr2*		Arp8
Arp42*	Arp42*		Taf14
Ssr3*	Ssr3*		les4
Ssr4*	Sfh1		les2
Snf30	Rsc58		Rvb1
Tfg3	Ssr4*		Rvb2
	Rsc7		Arp5
			les6

### Chaperonas de histonas y otros remodeladores de cromatina

Existe otro grupo de proteínas con actividad remodeladora, como las chaperonas de histonas. Estas son proteínas encargadas de transportar los tetrámeros de H3-H4 y los dímeros de H2A-H2B al DNA naciente, donde los remodeladores como ISWI y CHD ejercen sus funciones (revisado en Clapier et al., 2017). Las AAA-ATPasas son un grupo de chaperonas dependientes de ATP, dentro de las cuales se encuentran Abo1 y Abo2 en *S. pombe* que están codificadas por los genes ATAD2 y ATAD2 $\beta$  en humanos. En *S. cerevisiae* una misma chaperona realiza estas funciones, denominada Yta7 (Gal et al., 2016; Chacin et al., 2021).

Abo1 es una chaperona específica de las histonas H3 y H4. Durante la replicación del genoma, Abo1 deposita copias de H3 y H4 a medida que aumenta la longitud del DNA naciente

para lo que necesita al menos 80 pb de DNA (Kang et al., 2021). La ausencia de esta chaperona produce cambios en la estructura de la cromatina y fallos en la segregación cromosómica (Gal et al., 2016). Además, la delección de *abo1* altera los niveles de la dimetilación y trimetilación de H3K9, causando defectos en el silenciamiento en las regiones subtelo méricas y pericentroméricas de las regiones heterocromáticas (Dong et al., 2020).

Abo2 es una chaperona de histonas menos estudiada, cuyas funciones están aún por esclarecer (Cho et al., 2019). Los genes *abo1* y *abo2* no son esenciales en *S. pombe*, pero la ausencia de ambos causa la muerte celular (Gal et al., 2016).

Además, existen otros tipos de chaperonas como el complejo HIR (HIRA en humanos). En *S. pombe* está codificado por dos genes: *hip1* y *slm9* que son subunidades del complejo Hip1 e Hip2, respectivamente. Tanto Hip1 como Slm9 son chaperonas de histonas H3-H4 directamente relacionadas con el correcto ensamblaje de los nucleosomas. De hecho, se ha demostrado que la ausencia de un complejo HIR funcional determina la aparición de nucleosomas unidos a chaperonas en las NDRs (Wang et al., 2021b).

Otro complejo proteico con actividad remodeladora menos estudiado en *S. cerevisiae* es Fun30 homólogo de Swi2/Snf2 (Eapen et al., 2012). En *S. pombe* existen tres parálogos: Fft1, Fft2 y Fft3. Fun30 está asociado a la RNAPII y colabora con chaperonas de histonas para facilitar el desensamblaje de nucleosomas y, por lo tanto, la elongación de la transcripción (Lee et al., 2017). Además, Fun30 es el principal remodelador de nucleosomas que promueve la resección de 5' a 3' de los extremos del DNA en el entorno de las roturas de doble cadena o DSB (*Double-Strand Breaks*) de manera dependiente de las nucleasas Exo1 y Sgs1 (Chen et al., 2012). Fun30 también participa en el silenciamiento de los retrotransposones TF2 en *S. pombe*, manteniendo la ocupación nucleosómica en las regiones promotoras LTR (*Long Terminal Repeat*) (Persson et al., 2016).

# OBJETIVOS





Los objetivos de este trabajo son:

1. Establecer qué elementos de la secuencia del DNA contribuyen al posicionamiento de los nucleosomas en el genoma de *Schizosaccharomyces pombe*.
2. Identificar las diferencias funcionales de las diferentes familias de remodeladores de cromatina en *Schizosaccharomyces pombe*.



# RESULTADOS





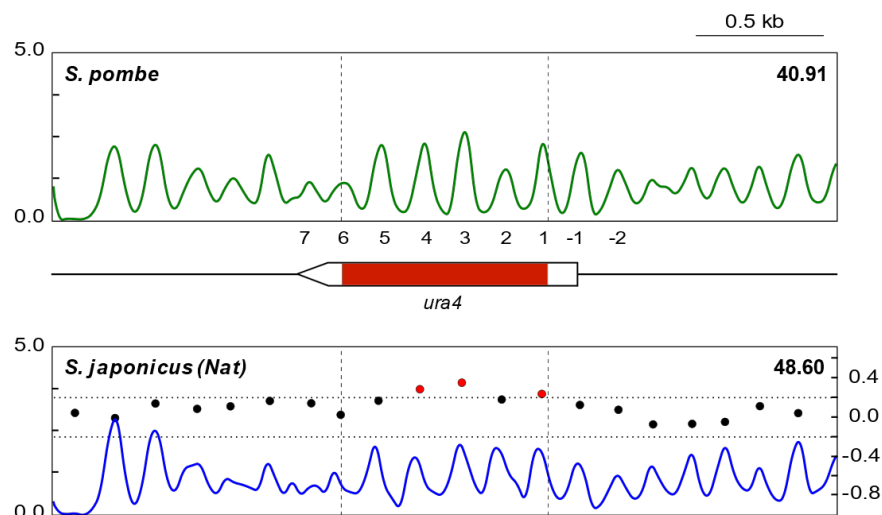
## 1. EL PAPEL DE LA SECUENCIA DEL DNA EN EL POSICIONAMIENTO DE LOS NUCLEOSOMAS

### Diseño de secuencias de DNA y análisis de su capacidad para posicionar nucleosomas

Como se ha comentado en la Introducción, el patrón específico de distribución de los cuatro nucleótidos a lo largo del DNA nucleosómico, constituye las *signatures* nucleosómicas. Una característica destacable de la *signature* de *S. pombe* es el mayor contenido en adenina en el extremo 5' y de timina en el 3', siendo la diada el punto de inflexión de la abundancia relativa de ambas bases (**Figura 4**). González et al. (2016), demostraron que el reemplazamiento de genes de *S. pombe* por sus ortólogos de otras especies causaba la pérdida de posicionamiento, pero la modificación de sus codones por sinónimos que determinasen una secuencia final más parecida a la *signature* nucleosómica de *S. pombe*, recuperaba el posicionamiento nucleosómico. Estos resultados pusieron en valor la relevancia de la secuencia del DNA en el posicionamiento de los nucleosomas, pero a su vez planteaban algunas preguntas que hemos intentado responder en este trabajo. Una de las dificultades más importantes en el estudio de las *signatures* es que no son detectables en las secuencias individuales de DNA asociadas con los aproximadamente 45.000 nucleosomas bien posicionados del genoma de *S. pombe*. Sin embargo, González et al. (2016) también demostraron que la secuencia de los nucleosomas individuales es importante porque cuando intercambiaron el orden de los codones sinónimos del gen *ura4* de *S. pombe* sin modificar su capacidad de codificación ni la composición promedio de bases, el patrón de posicionamiento nucleosómico se perdía. Otros experimentos demostraron que, además, la información era redundante y estaba dispersa a lo largo de las 150 pb del DNA mononucleosómico, ya que modificaciones en la secuencia del gen *ura4* en regiones discretas de los nucleosomas individuales (diada, DNA *linker* y posiciones intermedias entre diada y DNA *linker*), tenían muy poco impacto en su posicionamiento. A partir de esta información y con el objetivo de intentar esclarecer qué elementos específicos de la secuencia del DNA mononucleosómico pueden ser relevantes para el posicionamiento de los nucleosomas, diseñamos una serie de experimentos que se explican a continuación.

Utilizamos como secuencia diana modificable, la del gen *ura4* de *S. japonicus*. Este gen, cuando sustituye en el genoma de *S. pombe* al gen *ura4* endógeno, se organiza en seis nucleosomas peor posicionados que los del gen *ura4* de *S. pombe* (González et al., 2016; **Figura 7**). Para analizar el efecto de la secuencia en el posicionamiento, sintetizamos diferentes versiones de la ORF o del gen completo (incluyendo las 5' y 3' UTRs), que se describirán a continuación y

reemplazamos con ellas el gen *ura4* endógeno de *S. pombe*. El posicionamiento de los nucleosomas se estudió mediante digestión de cromatina con nucleasa micrococcal (*MNasa*). El DNA mononucleosómico purificado se secuenció mediante secuenciación masiva y posteriormente se llevó a cabo el tratamiento bioinformático de los resultados para representar los mapas genómicos de ocupación nucleosómica que se van a mostrar a lo largo de todos los Resultados. En estos mapas, los nucleosomas bien posicionados se definen como picos discretos flaqueados por valles profundos, y separados con una periodicidad de aproximadamente 150 pb. Para cuantificar las diferencias de posicionamiento entre distintos mapas, empleamos la herramienta bioinformática DANPOS (Chen et al., 2013). En concreto, empleamos el parámetro de *fuzziness* (*borrosidad*) que cuantifica el grado en que cada fragmento de DNA nucleosómico se desvía respecto al promedio de todos los fragmentos secuenciados de ese nucleosoma. Previamente en nuestro laboratorio, habíamos calibrado este parámetro entre duplicados biológicos para establecer el valor umbral a partir del cual las diferencias podían considerarse significativas. Así, cuando las diferencias de *borrosidad* del mismo nucleosoma entre dos mapas son mayores de 2 desviaciones estándar de la media de la diferencia entre duplicados biológicos, podemos considerar que no está bien posicionado. El promedio de la *borrosidad* en valor absoluto o *fuzziness score* (indicado en cada Figura) para cada nucleosoma individual del gen *ura4*, nos permitió comparar el posicionamiento de cada una de las versiones diseñadas. Los detalles experimentales, se describen en la sección de Materiales y Métodos.



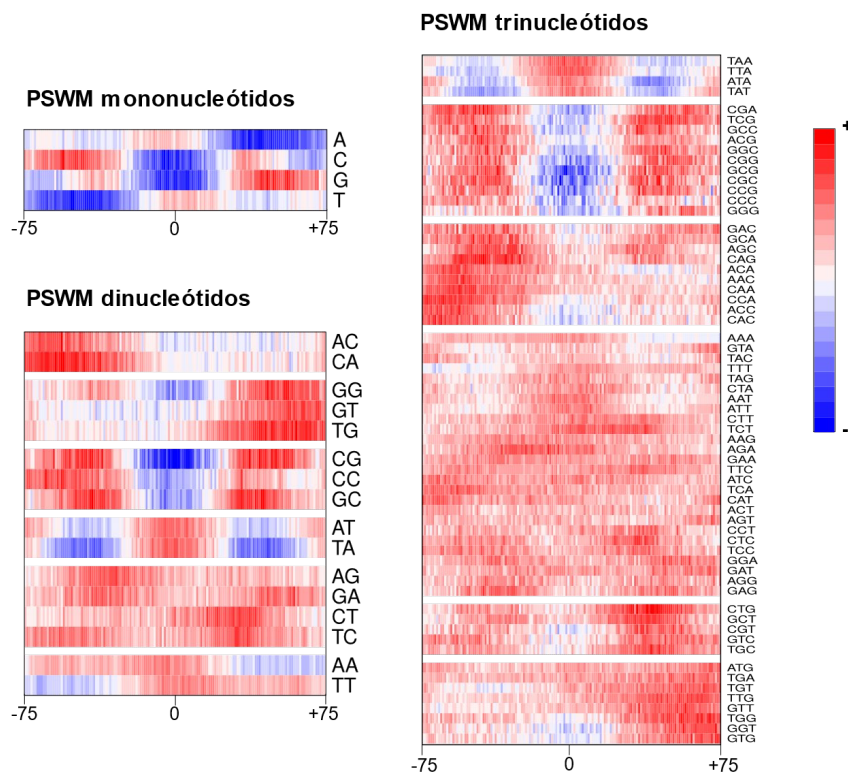
**Figura 7. Posicionamiento de nucleosomas del gen *ura4* de *S. japonicus* en *S. pombe*.** Análisis de ocupación de *MNase-seq* de la región de *ura4* de una cepa WT de *S. pombe* (panel superior, verde), y de una cepa que contiene la versión silvestre de la ORF del gen *ura4* de *S. japonicus* en su lugar (panel inferior, azul). Bajo el panel superior, se representa la región del gen *ura4* analizada. La región codificante es del mismo tamaño en las dos especies y se representa en rojo. El eje Y de la izquierda indica la ocupación nucleosómica normalizada al promedio genómico. Las líneas discontinuas verticales indican la posición de la ORF del gen *ura4* de las dos especies. El eje Y de la derecha, del panel inferior, indica el

$\log_2$  de la diferencia de *borrosidad* del mapa correspondiente a la cepa que contiene el gen *ura4* de *S. japonicus* respecto al de *S. pombe*. Los puntos rojos representan los valores superiores a  $2^0$  (línea horizontal discontinua) de la diferencia entre la versión modificada y el WT (Modificada de González et al., 2016). Los valores de la esquina superior derecha indican la borrosidad en valor absoluto de los nucleosomas +1 al +6.

### Matrices de puntuación con frecuencia de mono, di y trinucleótidos

Para extraer la información contenida en las *signatures* nucleosómicas, en el trabajo de González et al. (2016), se generaron "Matrices de puntuación de posición específica" (*PSWM: Position-Specific Weight Matrix*) que incorporaban la frecuencia de cada dinucleótido a lo largo de los perfiles agregados de 38.154 secuencias mononucleosómicas de *S. pombe*. A partir de esta información, las matrices permiten asignar una puntuación numérica a secuencias individuales, que será mayor cuanto más se aproxime su composición de bases a la *signature* promedio de la especie que se utiliza como referencia.

En este trabajo, hemos utilizado esta misma metodología para el diseño de las distintas versiones del gen *ura4* de *S. japonicus*, con las modificaciones que se detallan a continuación. Así, hemos generado matrices de puntuación con las secuencias de aproximadamente 50.000 nucleosomas bien posicionados de células silvestres de *S. pombe*, a partir de un mapa genómico de ocupación obtenido mediante digestión de la cromatina con *MNasa*, seguida de secuenciación masiva mediante el protocolo de *paired-end* de Illumina. A partir de estas secuencias, calculamos la frecuencia de mono-, di- y trinucleótidos a lo largo de los 150 nucleótidos del DNA nucleosómico (**Figura 8**). Como las matrices resultantes contienen niveles diferentes de información, utilizamos cada una de ellas para modificar la secuencia del gen *ura4* de *S. japonicus* y diseñamos las versiones *ura4SJ\_1mer*, *ura4SJ\_2mer* y *ura4SJ\_3mer* (**Figura 9**), en las que se modificó la secuencia completa del gen, incluyendo las UTRs en 5' y 3'. La finalidad de este experimento era comparar la capacidad de cada matriz para modificar el gen *ura4* de *S. japonicus* de forma que adoptase el posicionamiento del gen *ura4* endógeno de *S. pombe* al reemplazarlo por cada una de las tres versiones modificadas.



**Figura 8: Mapas de calor de las matrices de puntuación construidas con la frecuencia de mono, di y trinucleótidos de *S. pombe*.** El eje X indica la posición de cada nucleótido relativa a la diada del nucleosoma (0). El eje Y indica la puntuación en *log-odd* de los 4 mononucleótidos, 16 dinucleótidos y 64 trinucleótidos a lo largo del DNA mononucleosómico. Esta puntuación se calcula como el cociente de la frecuencia de cada mono, di o trinucleótido en cada posición a lo largo del DNA nucleosómico con respecto a su frecuencia genómica. La barra de la derecha representa la escala de color asociada con los valores numéricos de la matriz. Para simplificar su comprensión se han sustituido los valores numéricos por los signos "+" y "-", que indican el gradiente de puntuaciones desde las más altas (en tonos rojos) a las más bajas (tonos azules) en las tres matrices.

### Modificación del gen *ura4* de *S. japonicus* según las matrices de puntuación

La **Figura 9B** muestra que el gen modificado con la matriz de mononucleótidos (*ura4SJ\_1mer*), es incapaz de posicionar de manera que únicamente el nucleosoma +1 se mantiene comparable al del gen *ura4* de *S. pombe* silvestre (**Figura 9A**). Además, la ocupación nucleosómica, reflejada en la altura de los picos es muy reducida respecto al promedio genómico. En la versión *ura4SJ\_3mer* (**Figura 9D**), la pérdida de posicionamiento afecta a los nucleosomas +2, +4 y +5 aunque el nivel de ocupación es mayor que en el caso anterior. En la secuencia *ura4SJ\_2mer* (**Figura 9C**), los nucleosomas muestran un patrón más regular y una ocupación comparable a la del gen *ura4* de *S. pombe* silvestre. Además, los valores promedio de *borrosidad* calculados para cada una de las tres versiones, coinciden con el grado de posicionamiento (**Figura 9**) de manera que la construcción generada con la matriz de mononucleótidos (*ura4SJ\_1mer*) es

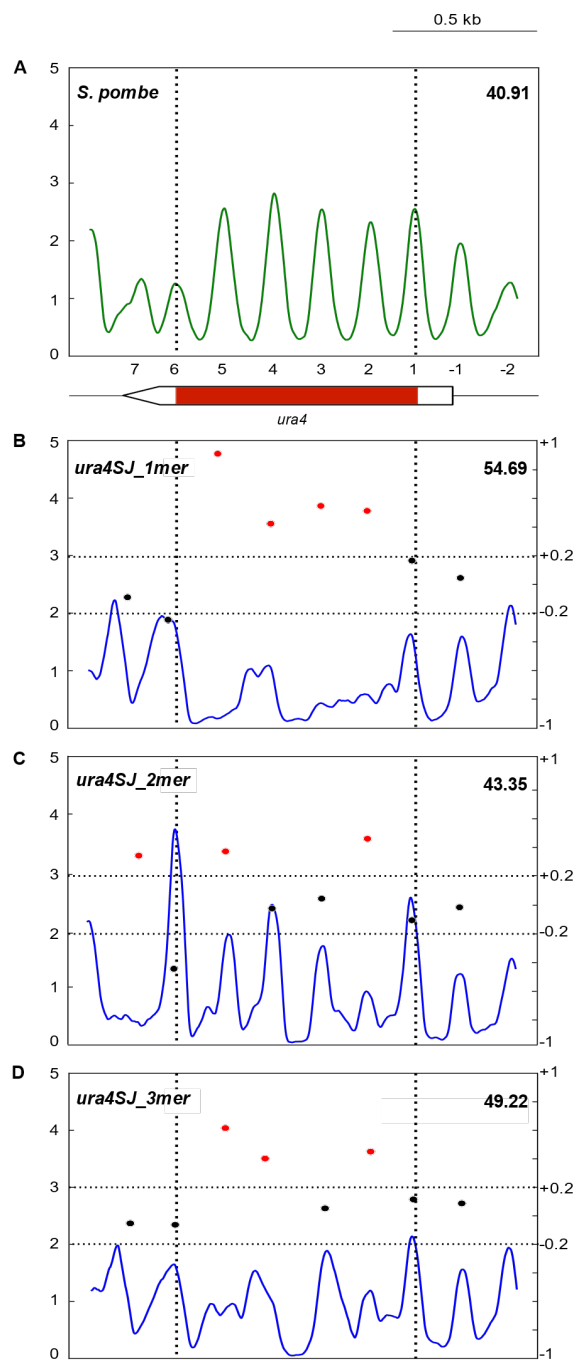
la peor posicionada, y la generada con la matriz de dinucleótidos (*ura4SJ\_2mer*) es la que más se asemeja a la ORF silvestre.

Como control de que las diferencias entre las distintas construcciones se deben a las distintas secuencias del gen *ura4* y no a diferencias en el grado de digestión de la *MNasa* ni a la secuenciación, comparamos mediante DANPOS una región del cromosoma 2 no relacionada con el gen *ura4* entre las tres cepas. La **Figura Suplementaria 1** (Anexo I) muestra como en estos tres casos y en todas las construcciones que se presentarán en los siguientes apartados, el posicionamiento de los nucleosomas en esta región control es comparable en todas las cepas.

Por tanto, la matriz de puntuación que incorpora la frecuencia de dinucleótidos es la herramienta óptima para modificar secuencias de DNA con el objetivo de que se asemejen lo más posible a la *signature* de la especie receptora.

### **Modificación de la ORF del gen *ura4* de *S. japonicus* según las matrices de puntuación**

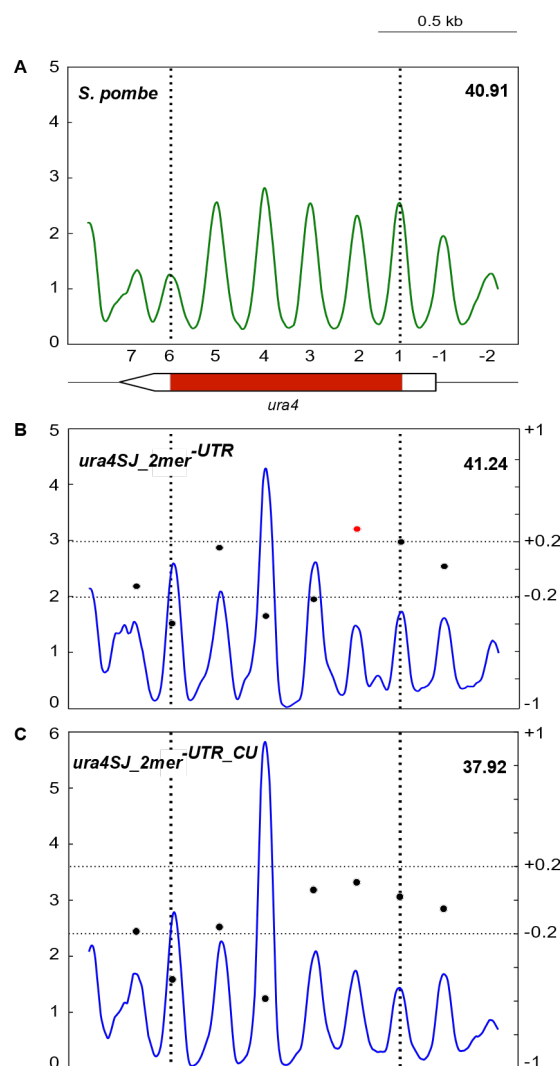
La secuencia modificada *ura4SJ\_2mer*, se empaqueta en nucleosomas relativamente bien posicionados (**Figura 9C**) pero no recapitula exactamente el posicionamiento del gen *ura4* endógeno de *S. pombe* (**Figura 9A**). Una posible causa es la modificación de la secuencia de los nucleosomas +1 y +6 que incluyen las regiones reguladoras UTR 5' y 3'. Para comprobarlo, diseñamos una secuencia en la que modificamos solo la ORF del gen *ura4* de *S. japonicus* utilizando la matriz de puntuación con frecuencia de dinucleótidos (**Figura 8**). En la construcción resultante, *ura4SJ\_2mer*<sup>-UTR</sup>, el DNA nucleosómico de +1 y +6 está modificado únicamente en la parte correspondiente a la ORF y mantiene la versión silvestre en la parte de las UTRs. El análisis de posicionamiento (**Figura 10B**), indica que esta versión mejoraba el posicionamiento de los nucleosomas si lo comparáramos con las versiones *ura4SJ\_1mer*, *ura4SJ\_2mer* y *ura4SJ\_3mer* (**Figura 9B, 9C y 9D**) ya que solo el nucleosoma +2 tenía peor posicionamiento que la versión silvestre del gen *ura4* de *S. pombe* (**Figura 10B**).



**Figura 9. Modificación de secuencias de *S. japonicus* a partir de las matrices de puntuación con frecuencia de mono, di, y trinucleótidos de la Figura 8.** Mapas de ocupación de nucleosomas generados por *MNase-seq* de la región del gen *ura4* de *S. pombe* (A), y comparación con las versiones *ura4SJ\_1mer* (B), *ura4SJ\_2mer* (C) y *ura4SJ\_3mer* (D). Los símbolos y las escalas utilizados son iguales que los de la Figura 7.

Otro elemento que podría estar relacionado con el posicionamiento de los nucleosomas es el sesgo de uso de codones. Este hace referencia a las diferencias entre especies en la frecuencia de uso de codones sinónimos de manera que un codón se utiliza con una frecuencia mayor o menor de la esperada en función del número de codones de cada aminoácido (Sharp et al., 1988; Kanaya et al., 1999; Duret 2000; revisado en Parvathy et al., 2022). El sesgo en el uso de codones

se ha relacionado con la optimización de la traducción y los niveles de expresión de los genes (Yang et al., 2021; Zhao et al., 2021), y de manera indirecta con el posicionamiento de los nucleosomas (González et al., 2016). Para estudiar la relación con este último, se utilizó la matriz de puntuación con frecuencia de dinucleótidos (**Figura 8**) incorporando una penalización para la elección de aquellos codones de uso menos frecuente para *S. pombe* (ver Materiales y Métodos; frecuencia de uso de codones de *S. pombe* extraído de Nakamura et al., 2000). En esta penalización, se multiplica la frecuencia de cada dinucleótido de la matriz empleada por la frecuencia de aparición de ese codón para ponderar por las frecuencias de uso de codones. La versión resultante *ura4SJ\_2mer<sup>-UTR\_CU</sup>* presenta todos sus nucleosomas bien posicionados (**Figura 10C**).



**Figura 10. Modificación de secuencias codificantes del gen *ura4* de *S. japonicus* utilizando matrices de puntuación con frecuencia de dinucleótidos.** Mapas de ocupación de nucleosomas generados por *MNase-seq* de la región del gen *ura4* de *S. pombe* (**A**), y comparación con la versión con las UTRs 5' y 3' no modificadas (**B**) (*ura4SJ\_2mer<sup>-UTR</sup>*) y con la que incorpora el sesgo de codones (**C**) (*ura4SJ\_2mer<sup>-UTR\_CU</sup>*). Los símbolos y las escalas utilizados son iguales que los de la Figura 7.

Los resultados de este apartado sugieren que la incorporación de la *signature* nucleosómica en las UTRs puede tener un efecto negativo sobre el posicionamiento de los nucleosomas, quizás debido a las alteraciones en la arquitectura del promotor. Asimismo, parece que el sesgo de codones puede contribuir al posicionamiento de nucleosomas. Es posible que, durante la evolución, la selección positiva de los codones de uso más frecuente no solamente haya contribuido a la optimización de la traducción, sino también al empaquetamiento de los genes en nucleosomas posicionados.

### **Análisis de la composición de bases del conjunto de versiones modificadas del gen *ura4***

Las secuencias del gen *ura4* analizadas hasta ahora, se organizan en nucleosomas posicionados de forma desigual (**Figuras 9 y 10**). Con el fin de encontrar alguna correlación entre las diferentes versiones y su capacidad de posicionamiento, analizamos la composición conjunta de bases de las mitades proximal y distal respecto al sitio de inicio de la transcripción de los 4 nucleosomas (2 al 5) incluidos enteramente en la ORF en cada una de las versiones del gen. Para ello, utilizamos como referencia las diadas extraídas de los mapas de ocupación nucleosómica de una cepa silvestre de *S. pombe* (Soriano et al., 2013).

**Tabla 2. Análisis de la composición de bases del conjunto de versiones modificadas del gen *ura4*.** Se muestra el número de bases de A, C, G y T en las mitades proximal y distal del DNA mononucleosómico correspondiente a la suma del contenido de los cuatro nucleosomas del +2 al +5, de las versiones silvestres y modificadas del gen *ura4*. Cada flecha hacia arriba o hacia abajo indica un aumento o una disminución de 10 nucleótidos en relación a la secuencia del gen *ura4* WT de *S. pombe*. Las versiones que posicionan bien los nucleosomas están marcadas en color amarillo y las que no posicionan en gris.

<b>Secuencia (mitad proximal)</b>	<b>A</b>	<b>C</b>	<b>G</b>	<b>T</b>
<i>ura4_pombe_WT</i>	83	61	69	87
<i>ura4_japonicus_WT</i>	64	68	85	83
<i>ura4SJ_1mer</i>	82	96 ↑↑↑	73	49 ↓↓↓
<i>ura4SJ_2mer</i>	103 ↑↑	73 ↑	66	58 ↓↓↓
<i>ura4SJ_3mer</i>	92 ↑	72 ↑↑	75	61 ↓↓
<i>ura4SJ_2mer<sup>UTR</sup></i>	113 ↑↑↑	66	65	56 ↓↓↓
<i>ura4SJ_2mer<sup>UTR_CU</sup></i>	124 ↑↑↑↑	55	60	61 ↓↓

Secuencia (mitad distal)	A	C	G	T
<i>ura4_pombe_WT</i>	81	54	75	90
<i>ura4_japonicus_WT</i>	72	70	78	80
<i>ura4SJ_1mer</i>	75	56	89 ↑	80 ↓
<i>ura4SJ_2mer</i>	72	48	90 ↑	90
<i>ura4SJ_3mer</i>	77	60	77	86
<i>ura4SJ_2mer<sup>UTR</sup></i>	72	53	73	102 ↑
<i>ura4SJ_2mer<sup>UTR_CU</sup></i>	83	36 ↓	74	107 ↑

Como se muestra en la **Tabla 2**, las versiones que mejor posicionan nucleosomas (marcadas en amarillo) presentan un incremento de A y un descenso de T en la mitad proximal del DNA mononucleosómico. Además, en la mitad distal se producen variaciones mínimas en relación a las que se producen en la mitad proximal. Así, las versiones modificadas de *ura4* de *S. japonicus* que mejor posicionan (*ura4SJ\_2mer*, *ura4SJ\_2mer<sup>-UTR</sup>* y *ura4SJ\_2mer<sup>-UTR\_CU</sup>*), muestran un incremento de 30-40 adeninas en la mitad proximal y una disminución comparable del contenido en timina. En general, los valores de G y C en las dos mitades son similares a los de *ura4* silvestre de *S. pombe* excepto en las versiones modificadas que no posicionan nucleosomas (**Tabla 2**, marcadas en gris) lo que podría tener un efecto negativo sobre el posicionamiento.

La conclusión de este análisis es que en las versiones que posicionan nucleosomas, se produce un desbalance entre los niveles de A y T entre la mitad proximal y distal del DNA mononucleosómico. No obstante, este desbalance es más acusado en la mitad proximal que en la distal, y parece no ir acompañado de variaciones en el contenido de C y G, que, aunque ligeras, sí se observan en las *signatures*. Estos resultados también sugieren que el parámetro relevante para el posicionamiento es la asimetría en el contenido en A y T entre las dos mitades del DNA nucleosómico y que quizás no fuese necesario incorporar el perfil de la *signature* de acuerdo con la matriz de puntuación tal y como habíamos hecho en los análisis de las **Figuras 9 y 10**.

### Posicionamiento nucleosómico mediante desbalance de A y T en el gen *ura4* de *S. pombe* y de *S. japonicus*

Para comprobar esta posibilidad, modificamos la secuencia de las ORFs de los genes *ura4* de *S. pombe* y *S. japonicus* reemplazando los codones silvestres de cada una por sus sinónimos de manera que la composición final de los cuatro nucleótidos en la mitad proximal y distal de cada nucleosoma fuese lo más similar posible a la de las secuencias que posicionaban

nucleosomas (marcadas en amarillo) en la **Tabla 2**. Esta modificación se hizo directamente sin tener en cuenta el perfil de la *signature* ni las matrices de puntuación.

Las versiones de *ura4* generadas se denominaron *ura4SP\_AT* y *ura4SJ\_AT*. El cambio de codones en la modificación *ura4SP\_AT* supuso un incremento de 42 nucleótidos de A en las mitades proximales respecto del *ura4* de *S. pombe*, acompañado de un descenso de unos 10 nucleótidos de T, G y C. En las mitades distales, la incorporación de codones sinónimos supuso un incremento de 41 nucleótidos de T, y un descenso de 11 nucleótidos de A y de 23 de C. La versión *ura4SJ\_AT* presentaba en las mitades proximales, un incremento de 58 nucleótidos de A acompañados de un descenso de 29 nucleótidos de T y, 14 de G y 15 de C. En las mitades distales se incrementó en 38 nucleótidos las T, y descendieron los niveles de A en 18 y C en 15 nucleótidos (**Tabla 3**).

**Tabla 3. Análisis de la composición de bases del conjunto de versiones modificadas empleando el desbalance de A y T sobre el gen *ura4*.** Se muestra el número de bases de A, C, G y T en las mitades proximal y distal del DNA mononucleosómico correspondiente a la suma del contenido de los cuatro nucleosomas del +2 al +5, de las versiones silvestres y modificadas con el desbalance de A y T del gen *ura4* de *S. pombe* y *S. japonicus*. Cada flecha hacia arriba o hacia abajo indica un aumento o una disminución de 10 nucleótidos en relación a la secuencia del gen *ura4* WT de *S. pombe*. Las versiones que posicionan bien los nucleosomas están marcadas en color amarillo.

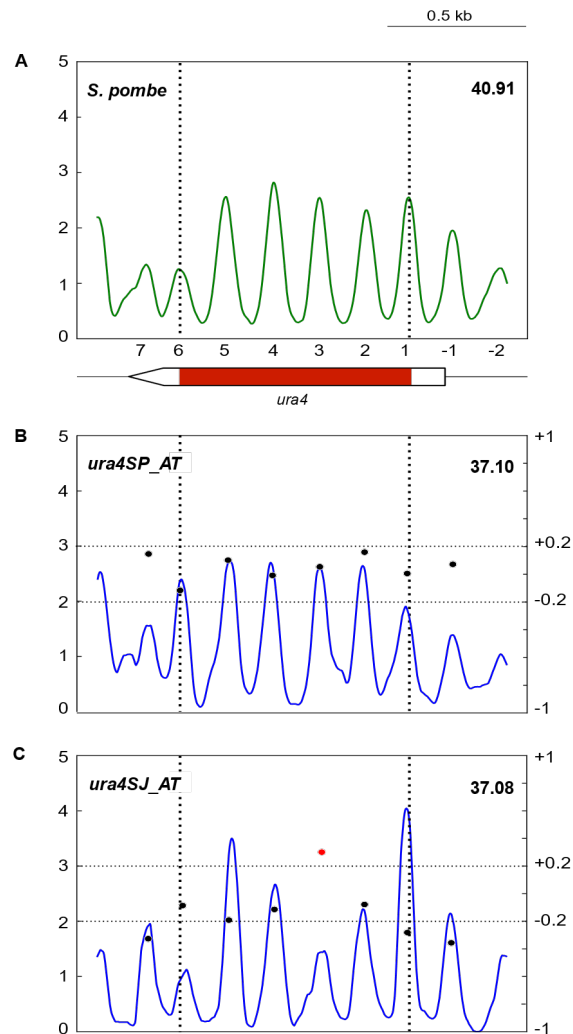
Secuencia (mitad proximal)	A	C	G	T
<i>ura4_pombe_WT</i>	83	61	69	87
<i>ura4_japonicus_WT</i>	64	68	85	83
<i>ura4SP_AT</i>	125 ↑↑↑↑	50 ↓	54 ↓	71 ↓
<i>ura4SJ_AT</i>	141 ↑↑↑↑↑	46 ↓	55 ↓	58 ↓↓

Secuencia (mitad distal)	A	C	G	T
<i>ura4_pombe_WT</i>	81	54	75	90
<i>ura4_japonicus_WT</i>	72	70	78	80
<i>ura4SP_AT</i>	70 ↓	31 ↓↓	68	131 ↑↑↑↑
<i>ura4SJ_AT</i>	63 ↓	39 ↓	70	128 ↑↑↑

Una vez sintetizadas las distintas versiones del gen *ura4*, reemplazamos con ellas el gen *ura4* endógeno de *S. pombe* y generamos el mapa de ocupación nucleosómica mediante *MNase-seq*. La **Figura 11B** muestra que la versión *ura4SP\_AT* se organiza en nucleosomas tan bien posicionados como los del gen *ura4* silvestre de *S. pombe* (**Figura 11A**). Por su parte, la versión *ura4SJ\_AT* tiene un patrón de posicionamiento mejor que el de *ura4* silvestre de *S. japonicus*

(Figura 7) y es similar a la del de *S. pombe*, donde solo el nucleosoma +3 pierde el posicionamiento (Figura 11C).



**Figura 11. Modificación de la secuencia de los genes *ura4* de *S. pombe* y de *S. japonicus* incorporando el desbalance de A y T entre las mitades proximal y distal del DNA mononucleosómico.** Mapas de ocupación de nucleosomas generados por *MNase-seq* de la región del gen *ura4* de *S. pombe* (A) y comparación con las versiones *ura4SP\_AT* (B) y *ura4SJ\_AT* (C). Los símbolos y las escalas utilizados son iguales a los de la Figura 7.

### Posicionamiento en una secuencia artificial mediante desbalance de A y T

Para verificar de una manera más radical la capacidad de posicionar nucleosomas mediante el desbalance de A y T, generamos una secuencia aleatoria de 918 pb sin capacidad codificante empleando el programa *Random DNA Generator* (<http://www.faculty.ucr.edu/~mmaduro/random.htm>). Esta secuencia (*sec.artificial\_control*) contenía un 47% de GC, que es similar al promedio genómico de *S. pombe* (43%). Por su tamaño, esta secuencia podría incorporar seis nucleosomas, considerando 150 pb la longitud del DNA mononucleosómico y 3 pb el DNA *linker*. Como la secuencia no era codificante, la dividimos en tripletes y los consideramos como codones para utilizar el mismo método de modificación de la

composición de bases que utilizamos en con las ORFs de los genes *ura4* de *S. japonicus* y de *S. pombe*. La secuencia resultante (*sec.artificial\_AT*) tenía un contenido promedio de GC del 38%. En el análisis de la composición de bases de las mitades proximal y distal de los seis nucleosomas de la *sec.artificial\_AT* observamos que en la mitad proximal, se producía un incremento de 76 nucleótidos de A, frente al descenso de 34 nucleótidos de T, 17 de C y 25 de G. En la mitad distal, por su parte, se produce un incremento de 76 nucleótidos de T, y una disminución de 39, 27 y 10 nucleótidos de A, C y G, respectivamente, empleando la *sec.artificial\_control* como referencia (**Tabla 4**).

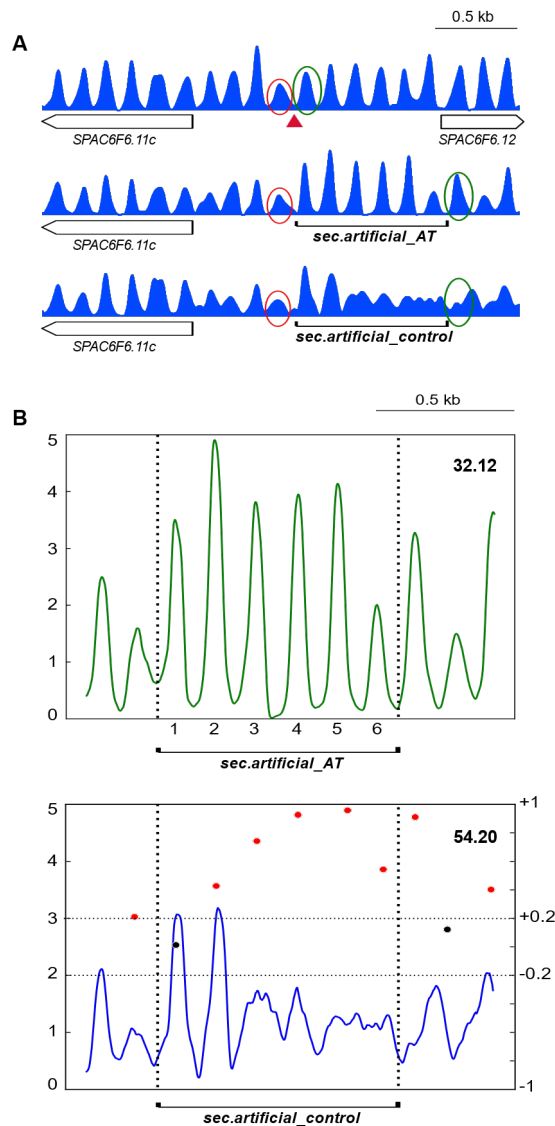
**Tabla 4. Composición de bases de las versiones artificiales modificadas mediante desbalance de A y T.** Se muestra el número de bases de A, C, G y T en las mitades proximal y distal del DNA mononucleosómico correspondiente a la suma del contenido de los nucleosomas +1 al +6, de las versiones de la secuencia artificial control y modificada integradas en el genoma de *S. pombe*. Cada flecha hacia arriba o hacia abajo indica un aumento o una disminución de 10 nucleótidos en relación a la secuencia artificial control.

<b>Secuencia (mitad proximal)</b>	<b>A</b>	<b>C</b>	<b>G</b>	<b>T</b>
<i>Sec.artificial_control</i>	106	116	101	127
<i>Sec.artificial_AT</i>	182 ↑↑↑↑↑↑	99 ↓	76 ↓↓	93 ↓↓

<b>Secuencia (mitad distal)</b>	<b>A</b>	<b>C</b>	<b>G</b>	<b>T</b>
<i>Sec.artificial_control</i>	117	105	101	127
<i>Sec.artificial_AT</i>	78 ↓↓↓	78 ↓↓	91 ↓	203 ↑↑↑↑↑↑

Ambas secuencias se integraron entre dos nucleosomas bien posicionados, en una región intergénica del cromosoma 1, entre los genes *SPAC6F6.11cy* y *SPAC6F6.12*. Los mapas de ocupación nucleosómica (**Figura 12A**) muestran que la versión *sec.artificial\_control* es incapaz de posicionar nucleosomas mientras que la *sec.artificial\_AT* está organizada en seis nucleosomas que ocupan las posiciones predichas por la composición de bases. El análisis cuantitativo realizado mediante DANPOS, corrobora estos resultados (**Figura 12B**).



**Figura 12. Posicionamiento nucleosómico sobre una secuencia artificial mediante desbalance de A y T. A)** Perfil de ocupación nucleosómica en la región intergénica entre los genes *SPAC6F6.11c-SPAC6F6.12* en una cepa WT de *S. pombe* (panel superior), en la cepa con la inserción de la secuencia artificial modificada, *sec.artificial\_AT* (panel medio), y en la cepa con la inserción de *sec. artificial\_control* (panel inferior). El triángulo rojo indica el lugar de inserción de las secuencias artificiales en la cepa silvestre. Los óvalos rojo y verde indican los nucleosomas de la cepa WT entre los que se insertaron las secuencias artificiales. **B)** Mapas de ocupación nucleosómica generados por *MNase-seq* de la *sec.artificial\_AT* (panel superior, verde) y comparación con la *sec.artificial\_control*. Los símbolos y las escalas utilizados son iguales que los de la Figura 7.

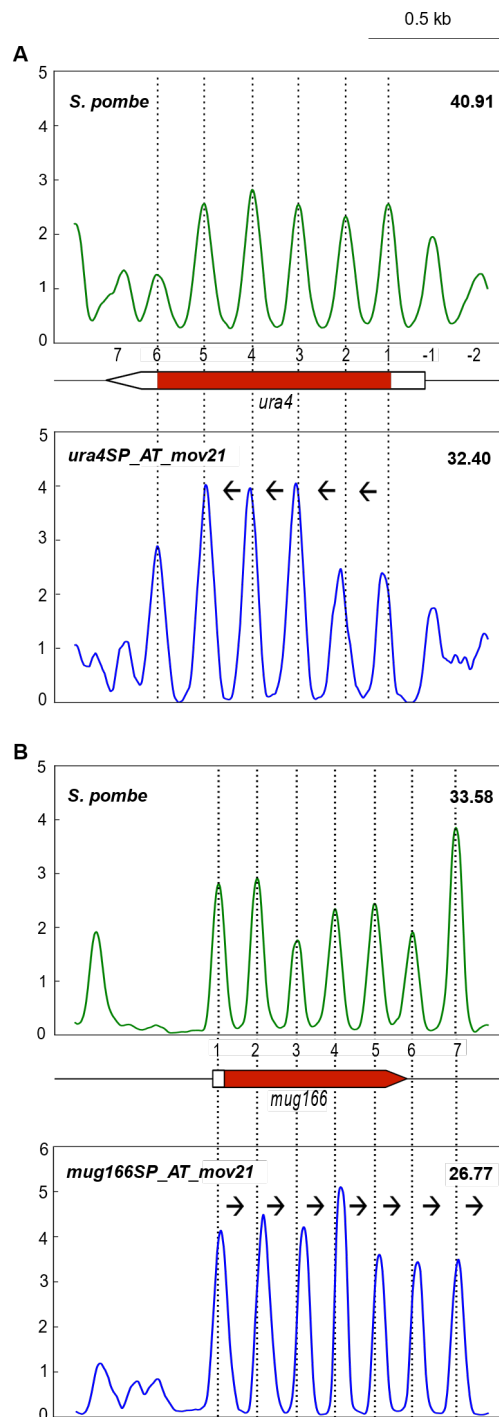
### Desplazamiento de la fase del patrón nucleosómico mediante modificación de la secuencia de DNA

Para poner a prueba mediante una aproximación diferente el potencial del contenido en A y T, intentamos desplazar la posición de los nucleosomas a lo largo de un gen, pero manteniendo la distancia regular entre ellos. Para ello, aplicamos el mismo criterio que en el apartado anterior,

pero aplicando un desfase de 21 pb respecto al patrón endógeno de los seis nucleosomas a lo largo de los genes *ura4* y *mug166* de *S. pombe* y comprobamos si se mantenía la periodicidad del patrón nucleosómico, pero con el desplazamiento incorporado. En ambos casos, no se modificaron las primeras 21 pb del primer nucleosoma solapante con la ORF en el sentido de la transcripción, pero a partir de ellas, se incorporó el desbalance del contenido de A y T por mitades utilizando codones sinónimos en regiones de 150 pb separados por 3 pb de DNA *linker* que no se modificaron.

El análisis de posicionamiento de las versiones *ura4SP\_AT\_mov21* (**Figura 13A**) y de *mug166SP\_AT\_mov21* (**Figura 13B**), muestra como en los dos casos se mantiene el perfil de seis nucleosomas bien posicionados, pero con las posiciones de las diadas desplazadas en la dirección prevista según el desfase que habíamos introducido. La resolución de los mapas de ocupación no es suficiente para estimar el desplazamiento con resolución de nucleótido, pero el resultado indica claramente que las diadas de los nucleosomas +1 al +4 de la construcción *ura4SP\_AT\_mov21* se han desplazado unos 15-20 nucleótidos aproximadamente respecto a las de la versión silvestre. La falta de desplazamiento de los nucleosomas +5 y +6 podría deberse a que el DNA *linker* que habíamos establecido era solo de 3 nucleótidos y el desplazamiento de los nucleosomas a lo largo del gen obligaría a los que estaban posicionados en 3' (cuya secuencia no se había modificado) a desplazarse también. Por ese motivo, la NDR en 3' presente en el gen *mug166* podría absorber este desfase, siendo en este caso el desplazamiento promedio es de unas 20 pb, que se corresponde con el desfase esperado en la secuencia modificada.

Los resultados de este apartado sugieren que el desbalance en el contenido de A y T entre la mitad proximal y distal del DNA mononucleosómico es suficiente para posicionar nucleosomas en *S. pombe in vivo* en el contexto genómico.

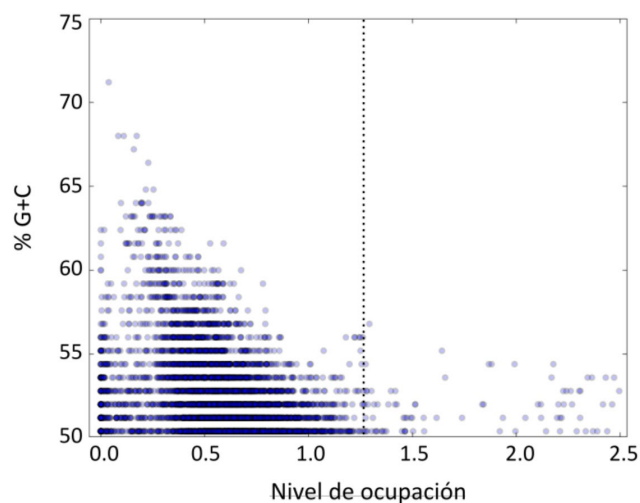


**Figura 13. Desplazamiento del patrón nucleosómico mediante modificación de la fase del contenido en A y T. A)**

Mapa de ocupación de nucleosomas generados por *MNase-seq* del gen *ura4* de *S. pombe*, WT (panel verde) y comparación con la versión *ura4SP\_AT\_mov21*. **B)** Mismo tipo de análisis para el gen *mug166* de *S. pombe*, WT (panel verde) y su comparación con la versión *mug166SP\_AT\_mov21*. Las líneas discontinuas marcan las diadas de los nucleosomas en la cepa WT. Las flechas indican el sentido del desplazamiento de las diadas de los nucleosomas en las versiones modificadas.

### Relación entre contenido en G+C y posicionamiento y ocupación nucleosómica

La especificidad de especie de las *signatures* nucleosómicas sugiere que secuencias similares pueden contribuir de forma diferencial al posicionamiento de nucleosomas en distintos organismos. En este sentido, se ha descrito que secuencias ricas en poli (dA:dT) excluyen nucleosomas en el genoma de *S. cerevisiae*, donde colocalizan con las regiones promotoras en las NDRs (Iyer & Struhl 1995; Segal & Widom 2009; Zhang et al., 2011; Raveh-sadka et al., 2012; García et al., 2017). Sin embargo, en *S. pombe* solamente el 18% de las NDRs tiene una frecuencia en poli (dA:dT) mayores de 7 pb mayor que el promedio genómico (Lantermann et al., 2010; García et al., 2017). Asimismo, la integración de secuencias ricas en A+T en *S. cerevisiae* puede generar NDRs (Zhang & Reese 2007; Raveh-sadka et al., 2012; Small et al., 2014), mientras que en *S. pombe*, trabajos previos en nuestro laboratorio han demostrado que hay una relación inversa entre contenido en G+C y la ocupación nucleosómica (**Figura 14**). De acuerdo con estos datos, en las versiones modificadas del gen *ura4* presentadas en apartados anteriores parece existir una relación directa entre el contenido en G+C de la secuencia modificada y una pérdida de posicionamiento y ocupación de sus nucleosomas (el contenido en G+C de *ura4SJ\_1mer* es de un 52%, frente al 38% de *ura4SJ\_2mer<sup>UTR-CU</sup>* y el 43% de *ura4* silvestre de *S. pombe*) (**Tabla 2**).



**Figura 14. Relación entre el contenido G+C y ocupación nucleosómica en *S. pombe*.** El gráfico muestra la distribución en todo el genoma de *S. pombe* de 6051 regiones de 125 pb que tienen un contenido promedio en nucleótidos de G y C superior o igual al 50% y el nivel promedio de ocupación nucleosómica que presenta cada una. La ocupación promedio del DNA mononucleosómico en *S. pombe* es de 1.25 (línea de puntos), calculada a partir de los datos de secuenciación masiva publicados en el trabajo de Soriano et al. (2013). (Figura de la Tesis Doctoral de Rebeca Serrano, 2014).

Para analizar en más detalle la relación entre contenido en G+C y posicionamiento nucleosómico, diseñamos una nueva versión del gen *ura4* de *S. pombe* en la que sustituimos los codones silvestres de la ORF por los sinónimos de mayor contenido en G o C. La secuencia

resultante, *ura4SP\_GC*, presenta un contenido en G+C total del 65% frente al 43% del gen *ura4* de *S. pombe* silvestre (Tabla 5).

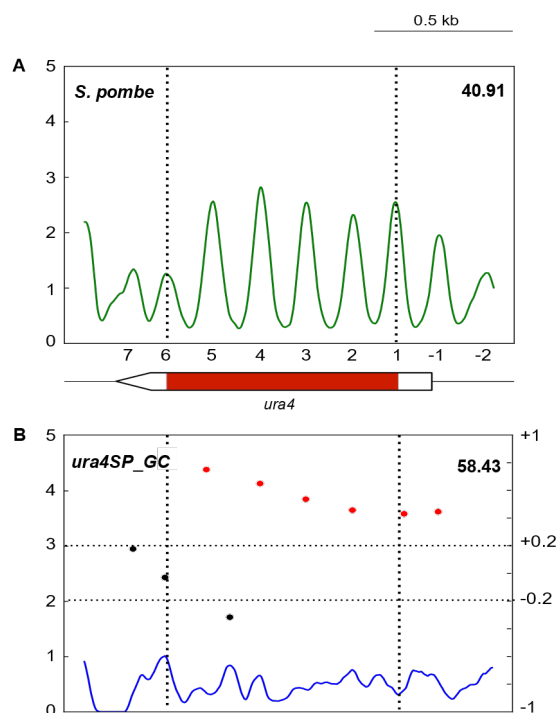
**Tabla 5. Composición de bases promedio de la versión *ura4SP\_GC*.** Se muestra el número de bases de A, C, G y T en las mitades proximal y distal del DNA mononucleosómico correspondiente a la suma del contenido de los cuatro nucleosomas del +2 al +5, de las versiones silvestre y *ura4SP\_GC*. Cada flecha hacia arriba o hacia abajo indica un aumento o una disminución de 10 nucleótidos en relación a la secuencia del gen *ura4* WT de *S. pombe*.

Secuencia (mitad proximal)	A	C	G	T
<i>ura4_pombe_WT</i>	83	61	69	87
<i>ura4SP_GC</i>	61 ↓↓	95 ↑↑↑	102 ↑↑↑	42 ↓↓↓↓

Secuencia (mitad distal)	A	C	G	T
<i>ura4_pombe_WT</i>	81	54	75	90
<i>ura4SP_GC</i>	62 ↓	92 ↑↑↑	99 ↑↑	47 ↓↓↓↓

Como puede observarse, al incrementar el contenido en G+C, no solo se pierde el posicionamiento de los nucleosomas del gen *ura4*, sino que, además, hay una gran reducción de la ocupación respecto al promedio genómico (Figura 15B). Estos resultados confirman que en *S. pombe*, secuencias con un alto contenido en G+C excluyen nucleosomas *in vivo*.



**Figura 16. Posicionamiento de nucleosomas del gen *ura4* de *S. pombe* con un elevado contenido en G+C.** Mapas de ocupación de nucleosomas generados por *MNase-seq* de la región de *ura4* de *S. pombe* (A) y comparación con la versión *ura4SP\_GC* (B). Los símbolos y las escalas utilizados son iguales a los de la Figura 7.

### **Análisis de la relación entre posicionamiento nucleosómico y la transcripción génica**

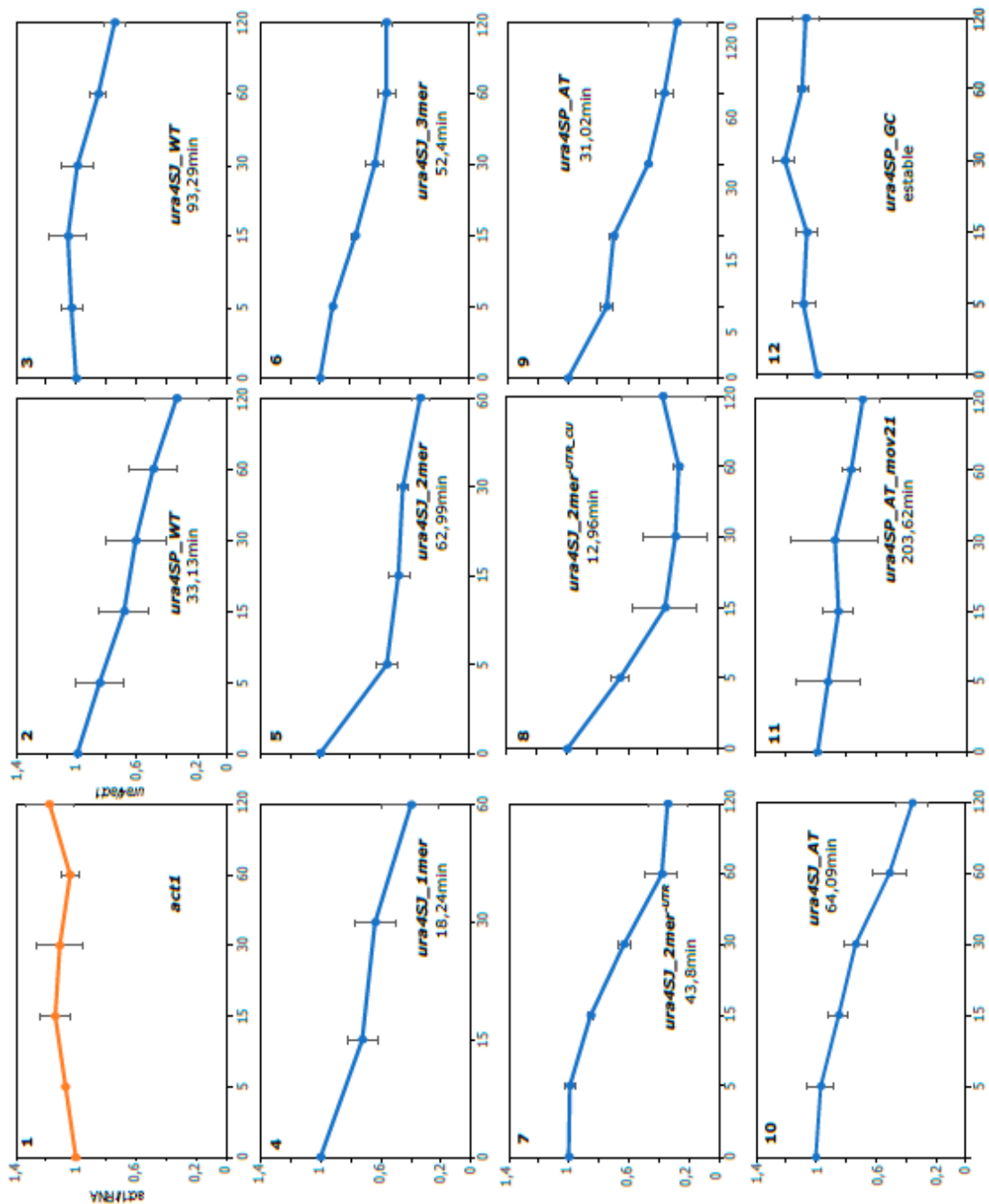
Una de las funciones básicas de los nucleosomas más allá de su capacidad de compactación del genoma, es la regulación de la transcripción. Así, las pérdidas de ocupación y posicionamiento nucleosómico favorecen el aumento de la transcripción antisentido, como ocurre en algunos mutantes de remodeladores de cromatina (Pointner et al., 2012; Shim et al., 2012). Además, la presencia de nucleosomas en regiones promotoras se relaciona con la represión transcripcional (Lorch et al., 1987; Soriano et al., 2013; Barozzi et al., 2014), y la pérdida de ocupación de los nucleosomas -1 y +1, favorece la transcripción críptica en estas regiones al quedar expuestos sitios de unión de factores de transcripción lo que puede reducir la expresión de los genes asociados a ellas (Gill et al., 2020). De esta manera, la relación entre transcripción y posicionamiento de los nucleosomas en torno a las regiones promotoras parece clara, pero se desconoce el efecto de la pérdida de posicionamiento de los nucleosomas a lo largo de los genes. Para analizarlo, hemos utilizado las diferentes versiones del gen *ura4* de *S. pombe* y de *S. japonicus* (Figuras 9, 10, 11, 13 y 15) con diferentes grados de posicionamiento nucleosómico para analizar los niveles de RNA en cada una de ellas.

### **Medida de vidas medias de los mRNAs de las distintas versiones del gen *ura4***

Los niveles basales de mRNA en la célula están determinados por el equilibrio entre su tasa de síntesis y de degradación (Pérez-Ortín 2007). Las tasas de degradación son específicas de cada transcrito (Herrick et al., 1990; Sachs 1993) y varían en un rango de hasta 100 veces (revisado en Pérez-Ortín et al., 2013a). Además, están controladas por elementos presentes en la propia secuencia del mRNA, localizados en las UTR en 3' (Shaw & Kamen 1986; Heaton et al., 1992). Estos elementos son diana para la unión de proteínas que modulan la interacción del mRNA con la maquinaria celular de degradación (revisado en Pérez-Ortín et al., 2013a y en Yamada & Akimitsu 2018).

Para poder comparar los niveles de RNA de las diez versiones diferentes del gen *ura4* que hemos generado, necesitamos calcular la vida media de cada uno de ellos. Normalmente, se emplean drogas que inhiben a la RNAPII como la tiolutina (Jiménez et al., 1973; Herrick et al., 1990) de manera que puede estimarse la reducción del nivel de cada mRNA a partir de la adición de la droga al medio de cultivo. Sin embargo, el gen *ura4*, al igual que otros genes, se induce en

presencia de tiolutina (Lauinger et al., 2017). Como estrategia alternativa, hemos aprovechado que el gen *ura4* de *S. pombe* deja de expresarse cuando se elimina la glucosa del medio de cultivo (Kim et al., 2002). Así, se partió de cultivos en fase exponencial en medio mínimo (MM) con glucosa como fuente de carbono, y se recogió una muestra control en estas condiciones, y otras a los 15, 30, 60, 90 y 120 minutos tras retirar la glucosa del medio. Posteriormente, se extrajo el RNA, se sintetizaron los cDNAs correspondientes, y se cuantificaron mediante qPCR, utilizando el mRNA del gen *act1* como normalizador. Los niveles de RNA correspondientes a las muestras sin glucosa se normalizaron al tiempo control con glucosa (0 h) (**Figura 16**). A partir de estas gráficas, se calcularon las vidas medias de cada una de las versiones del gen *ura4* (Materiales y Métodos). La estabilidad del normalizador *act1* en ausencia de glucosa se analizó previamente (**Figura 16, panel 1**), utilizando el 28S rRNA como control de carga, dado que los rRNAs son mucho más estables que los mRNAs (Abelson et al., 1974; Gillery et al., 1995; Sunyer-Figueres et al., 2018). Para todas las versiones de *ura4* modificadas en la ORF (*ura4SJ\_2mer<sup>UTR</sup>*, *ura4SJ\_2mer<sup>UTR-CU</sup>*, *ura4SJ\_AT*, *ura4SP\_AT*, *ura4SP\_GC*), se utilizó la misma pareja de oligonucleótidos que anillaba en la UTR en 3'. Para aquellas con modificaciones en la UTR (*ura4SJ\_1mer*, *ura4SJ\_2mer*, *ura4SJ\_3mer*, *ura4SP\_ATmov21*), se diseñaron parejas de oligonucleótidos específicas. Como era de esperar, en función de las diferencias en la secuencia, existe gran variabilidad entre las vidas medias de las versiones modificadas de *ura4* (**Figura 16, paneles 2-12**), oscilando entre aquellas muy estables como *ura4SP\_AT\_mov21* (**panel 11**) con 203 min, o aquellas con vidas medias muy cortas, como *ura4SJ\_2mer<sup>-UTR-CU</sup>* (**panel 8**) con aproximadamente 13 min.

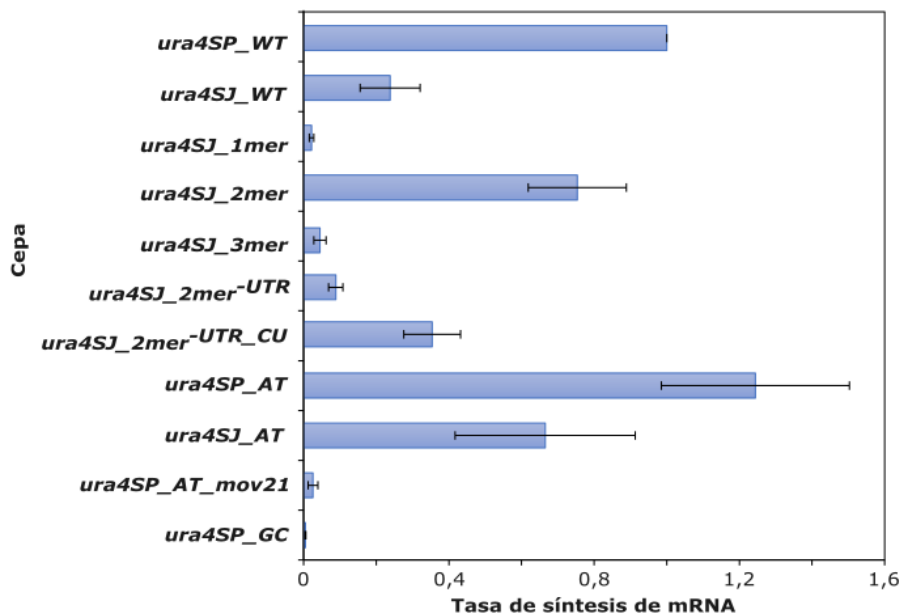


**Figura 16. Estabilidad y vidas medias de las distintas versiones de *ura4*.** Las gráficas muestran los niveles de mRNA de las distintas versiones de *ura4* medidos por RT-qPCR en ausencia de glucosa, y relativizados al tiempo 0 (con glucosa). El eje Y muestra los niveles de mRNA del gen *act1* frente a los niveles de 28S rRNA para el panel 1 (naranja), y los niveles de la versión *ura4* correspondiente frente a los niveles de *act1* para los paneles 2-13 (azul). El eje X indica los minutos tras la retirada de la glucosa en el medio. Cada valor es resultado de al menos tres réplicas biológicas. Las barras de error representan la desviación estándar.

### **Análisis de los niveles de transcripción sentido y antisentido de las distintas versiones del gen *ura4***

Para analizar los niveles de transcripción de las distintas versiones de *ura4*, aislamos RNA total, sintetizamos cDNA con polaridad de cadena con oligonucleótidos específicos y utilizamos el gen *act1* como normalizador. Para comparar los niveles de RNA de las distintas versiones, calculamos la tasa de síntesis de cada mensajero (*synthesis rate* o SR) tal y como describieron Pérez-Ortín et al. (2013b) (ver Materiales y Métodos) que tiene en cuenta los valores de vida media de los transcritos (**Figura 16**) y la concentración de mRNA calculada por qPCR.

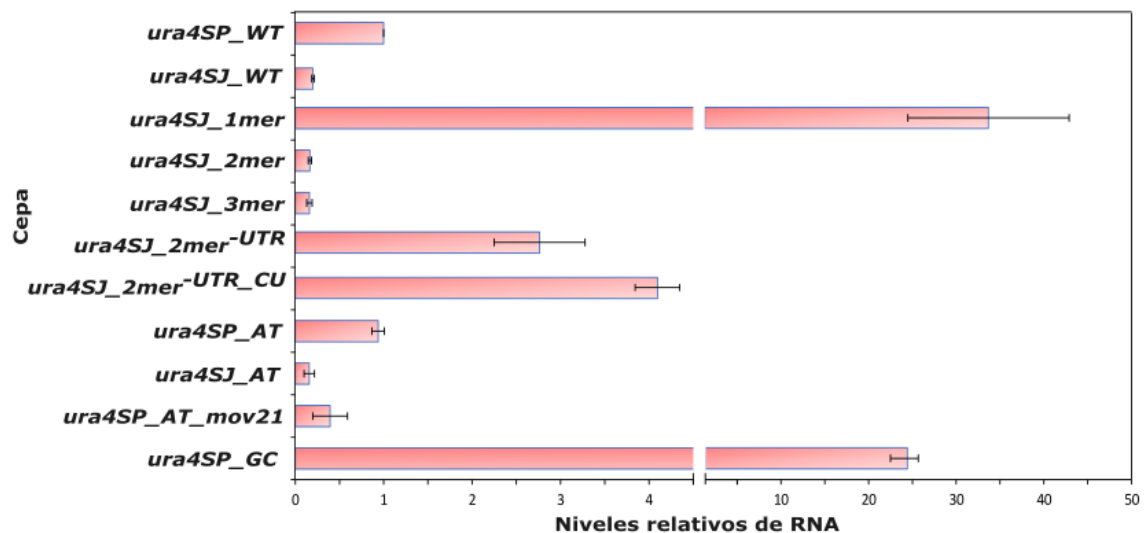
En la **Figura 17** se muestra que las versiones *ura4SP\_AT*, *ura4SJ\_AT* y *ura4SJ\_2mer*, con un posicionamiento nucleosómico comparable al del gen *ura4* silvestre de *S. pombe*, presentan a su vez niveles comparables de tasa de síntesis de sus mensajeros. Sin embargo, las versiones *ura4SJ\_2mer<sup>-UTR</sup>*, *ura4SJ\_2mer<sup>-UTR\_CU</sup>* y *ura4SP\_AT\_mov21*, con un posicionamiento nucleosómico similar a las anteriores, tienen reducidas drásticamente sus tasas de síntesis. Esto mismo, ocurre con las tasas correspondientes a las versiones *ura4SJ\_1mer*, *ura4SJ\_3mer* y *ura4SP\_GC*, que tienen un mal posicionamiento nucleosómico. Así pues, no parece existir una correlación clara entre los niveles de transcripción y el posicionamiento de los nucleosomas en las versiones de *ura4* analizadas. Una posible explicación a la disminución de la tasa de síntesis de los mRNAs de las versiones *ura4SJ\_1mer*, *ura4SJ\_3mer* y *ura4SP\_GC*, es la pérdida de ocupación nucleosómica en sus nucleosomas, observándose en las tres modificaciones regiones de hipersensibilidad a la *MNasa* o nuevas NDRs que podrían funcionar como promotores crípticos desde los que se inicie la transcripción aberrante (ver Discusión). En el caso de las versiones *ura4SJ\_2mer<sup>-UTR</sup>* y *ura4SJ\_2mer<sup>-UTR\_CU</sup>*, algunos de sus nucleosomas tienen una mejor ocupación que los del propio *ura4* silvestre de *S. pombe*, lo que podría interpretarse como que presentan una mayor afinidad por el DNA, que podría dificultar el tránsito de la RNAPII. El resultado obtenido para *ura4SP\_AT\_mov21* podría estar relacionado con la alteración de la posición del nucleosoma +1 (ver Discusión).



**Figura 17. Tasa de síntesis de los mRNA de las distintas versiones del gen *ura4* analizadas en este trabajo.** Los histogramas representan la tasa de síntesis de los mensajeros de las distintas versiones de *ura4*, normalizados al gen *act1*, e incorporando las vidas medias de los mismos presentadas en la Figura 17. Los valores están relativizados a los del gen *ura4*WT de *S. pombe*. Los valores representan la media de entre 3 y 4 réplicas biológicas. Las barras de error representan la desviación estándar de la media (SEM).

Experimentos de *NET-seq*, han revelado que el 68% de los genes codificantes de proteínas de *S. pombe* presentan transcripción antisentido y se ha observado que el aumento de estos está asociado con un descenso de la transcripción sentido (Wery et al., 2018). El gen *ura4* silvestre presenta un antisentido que, probablemente, resulta de la extensión de la transcripción del gen convergente *tam14*, aunque los niveles de la transcripción sentido del gen *ura4* son aproximadamente 35 veces superiores a los de antisentido. Para analizar si las alteraciones en la cromatina de las distintas versiones de *ura4* afectan a los niveles de transcripción antisentido, analizamos sus niveles mediante RT-qPCR, utilizando oligonucleótidos específicos para sus UTR en 5'. En la **Figura 18** se presentan los niveles de antisentido de todas las versiones de *ura4* relativizados a los niveles del silvestre. Como puede observarse, en las versiones *ura4SJ\_2mer<sup>-UTR</sup>*, *ura4SJ\_2mer<sup>-UTR\_CU</sup>*, hay un incremento de los niveles de RNA antisentido de hasta 4 veces respecto a la cepa silvestre, aunque los niveles más elevados los presentan las versiones *ura4SJ\_1mer* y *ura4SP\_GC*, donde se produce un incremento de hasta 35 veces respecto a los niveles del gen *ura4* silvestre de *S. pombe*. Sin embargo, el resto de las versiones modificadas presentan niveles comparables a los del silvestre. Este incremento del antisentido para estas dos versiones correlaciona con la pérdida completa de posicionamiento y ocupación nucleosómica que presentan (**Figuras 9 y 15**).

Además, en estas cepas los niveles de antisentido son superiores a los del sentido (unas 200 veces para *ura4SJ\_1mer* y unas 85 veces para *ura4SP\_GC*). Es improbable que la pérdida de posicionamiento de nucleosomas de estas modificaciones se deba a una tasa de transcripción muy elevada, como se ha descrito para el 10% de los genes del genoma de *S. pombe* (Koeber et al., 2009; Soriano et al., 2013), ya que los niveles de antisentido son comparables a los niveles de sentido de la cepa silvestre. Estos resultados apoyan la idea de que la bajada de ocupación nucleosómica, puede permitir el acceso a factores de transcripción a sitios putativos de unión presentes más allá del promotor de *ura4*. Otra posibilidad es que la configuración de los nucleosomas presentes en 3' del gen *ura4* contribuya a la terminación del gen convergente *tam14*, y que la pérdida de posicionamiento en ambas cepas comprometa dicha terminación (ver Discusión).



**Figura 18. Niveles de transcripción antisentido de las distintas versiones de *ura4* analizadas en este trabajo.** El diagrama de barras representa los niveles de antisentido de las distintas versiones de *ura4*, normalizados al gen *act1*. Los valores están relativizados a los niveles de antisentido del gen *ura4*WT de *S. pombe*. Los valores representan la media entre 3 y 4 réplicas biológicas. Las barras de error representan la desviación estándar de la media (SEM).

## 2. ANÁLISIS FUNCIONAL DE REMODELADORES DE CROMATINA EN *S. pombe*

### Construcción de mutantes condicionales de remodeladores de cromatina en *S. pombe*

Como se ha detallado en la Introducción, los remodeladores de cromatina tienen un papel relevante en el posicionamiento de los nucleosomas y en la regulación de la transcripción (Ocampo et al., 2019; revisado en Lin et al., 2020). La mayor parte de la información sobre el funcionamiento de estos complejos proviene de trabajos realizados con mutantes de delección (Monahan et al., 2008; Yamada et al., 2008; Pointner et al., 2012), lo cual presenta algunas

limitaciones: 1) las subunidades catalíticas de algunos de estos complejos son esenciales (RSC, INO80) y, para su estudio hay que recurrir a mutantes en subunidades accesorias cuyo impacto en la función del complejo puede no ser comparable a la delección de aquellas; 2) es difícil distinguir qué fenotipos de los observados son consecuencia directa o indirecta de la falta de funcionamiento del complejo remodelador; y 3) el uso de mutantes de delección impide analizar el papel de los remodeladores en diversos procesos celulares y la posible reversión de los fenotipos si se volviese a introducir un remodelador funcional en la célula. Para solventar estos inconvenientes, algunos estudios han recurrido al uso de mutantes condicionales, casi siempre derivados del sistema de *degrón*, en organismos como *S. cerevisiae* o *Drosophila* (Klein-Brill et al., 2019; Jo et al., 2021; Hendy et al., 2022), sin embargo, hasta la fecha no se ha realizado ningún estudio similar en *S. pombe*.

Recientemente, se ha optimizado un sistema de generación de mutantes condicionales específico para *S. pombe* basado en un sistema de *degrón* inducible por auxina (AID) (Zhang et al., 2022). En este método la proteína diana se etiqueta con un epítipo llamado *short AID* (sAID), que consiste en tres copias en tándem de una secuencia de 36 aminoácidos de la proteína IAA17 de *Arabidopsis*. Este epítipo es más pequeño (108 pb) que el que se usa habitualmente en el sistema de *degrón* (AID), que tiene una longitud de 229 pb. La segunda mejora es el uso como inductor de un análogo de la auxina, el 5-adamantil-IAA, que es funcional a concentraciones miles de veces menores que las utilizadas para la auxina. Este compuesto es reconocido por la proteína F-box OsTIR1 de *Oryza sativa*, que se expresa constitutivamente en *S. pombe* bajo el promotor del gen *ura4* y alberga una mutación en el bolsillo de unión que le permite reconocer el inductor a bajas concentraciones.

Con el objetivo de estudiar la dinámica de la cromatina y los efectos sobre la transcripción en ausencia de la función catalítica de los complejos remodeladores RSC, SWI/SNF, INO80 y CHD, hemos empleado este sistema de *degrón* para construir mutantes condicionales de las subunidades catalíticas de cada uno de ellos: *snf21-sAID* para RSC, *snf22-sAID* para SWI/SNF, *ino80-sAID* para INO80, y *hrp1-sAID* y *hrp3-sAID* para CHD. Una vez construidos, tal y como se indica en Materiales y Métodos, partimos de un cultivo de cada uno de ellos en fase exponencial de crecimiento en medio rico, y recogimos una muestra inicial correspondiente al tiempo 0 (0 h). A continuación, añadimos el inductor 5-adamantil-IAA al medio, y recogimos muestras a diferentes tiempos para calibrar el tiempo necesario para la desaparición de cada remodelador. Después de ese periodo, los cultivos se transfirieron al mismo medio a la misma temperatura, pero en ausencia del inductor y se recogieron muestras para estimar el tiempo necesario para restaurar los niveles de cada proteína similares a los del tiempo 0. En cada una de las muestras

analizamos la ausencia y reaparición de cada remodelador mediante *Western-Blot*, el posicionamiento nucleosómico mediante *MNase-seq* y, los niveles de RNA mediante *RNA-seq* (ver Materiales y Métodos).

## Caracterización funcional de la organización nucleosómica y de la expresión génica en presencia y ausencia de remodeladores de cromatina

### Familia de remodeladores SWI/SNF

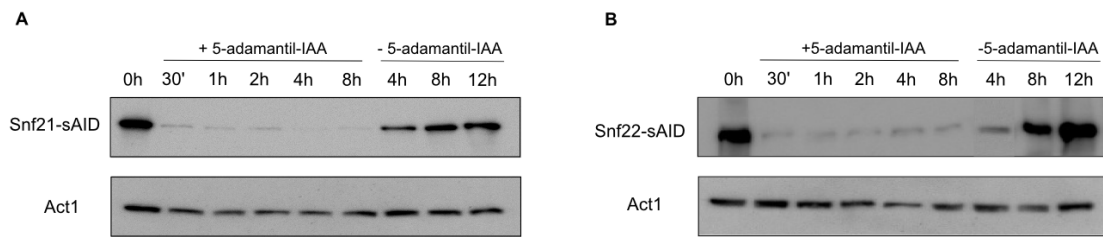
La familia SWI/SNF pertenece al grupo de los remodeladores denominados *pushers*, cuya función es mantener el posicionamiento correcto del nucleosoma +1. En *S. cerevisiae*, la ausencia de este remodelador causa el desplazamiento de este nucleosoma hacia la NDR, con lo cual se reduce su tamaño (Rawal et al., 2018; Kubik et al., 2019), y afecta a la expresión de un gran número de genes (Wilson et al., 1996; Monahan et al., 2008; Ocampo et al., 2019).

Para estudiar la dinámica de la cromatina y la transcripción regulada por los remodeladores RSC y SWI/SNF, pertenecientes a esta familia, utilizamos los mutantes condicionales *snf21-sAID* y *snf22-sAID* mencionados anteriormente. Recogimos una muestra a las 0 horas (control), y a los 30 minutos, 1, 2, 4 y 8 horas en presencia del inductor, y a las 4, 8 y 12 horas tras su retirada del medio de cultivo (**Figura 29**).



**Figura 29: Diagrama de las cinéticas *snf21-sAID* y *snf22-sAID*.** Se muestran los tiempos correspondientes a las cinéticas de *snf21-sAID* y *snf22-sAID* a los cuales se han recogido muestras para *RNA-seq* en presencia y en ausencia de 5-adamantil-IAA. En cada tiempo también se recogieron muestras para *MNase-seq*.

La **Figura 20** muestra la presencia de las dos proteínas a los diferentes tiempos. En ninguno de los dos casos se detecta proteína a los 30 minutos en presencia de 5-adamantil-IAA, mientras que la recuperación se produce gradualmente a partir de las 4 horas de la retirada del inductor.

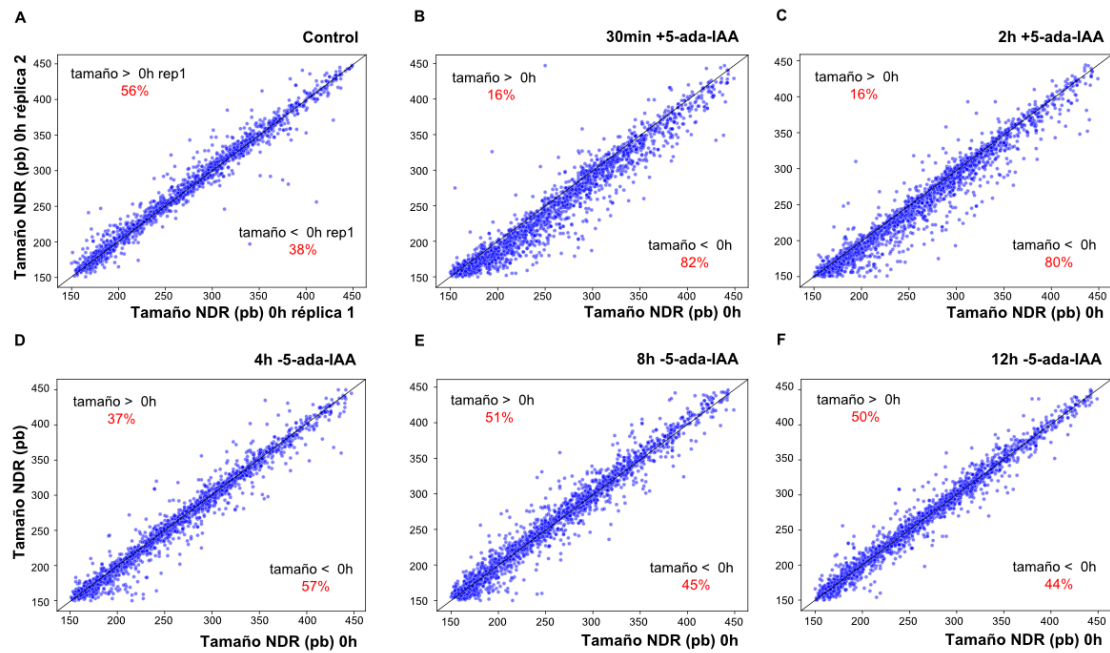


**Figura 20. Niveles de la proteína Snf21 y Snf22 en presencia y ausencia de 5-adamantil-IAA.** *Western Blot* de extractos totales de proteínas de las cepas *snf21-sAID* (A) y *snf22-sAID* (B). Se indican los tiempos en presencia y ausencia del inductor. Las membranas se hibridaron con el anticuerpo *anti-sAID* (blot superior). Como control de carga, se analizaron los niveles de actina (blot inferior) con un anticuerpo *anti-act1*.

### Dinámica de las NDRs en los mutantes condicionales *snf21-sAID* y *snf22-sAID*

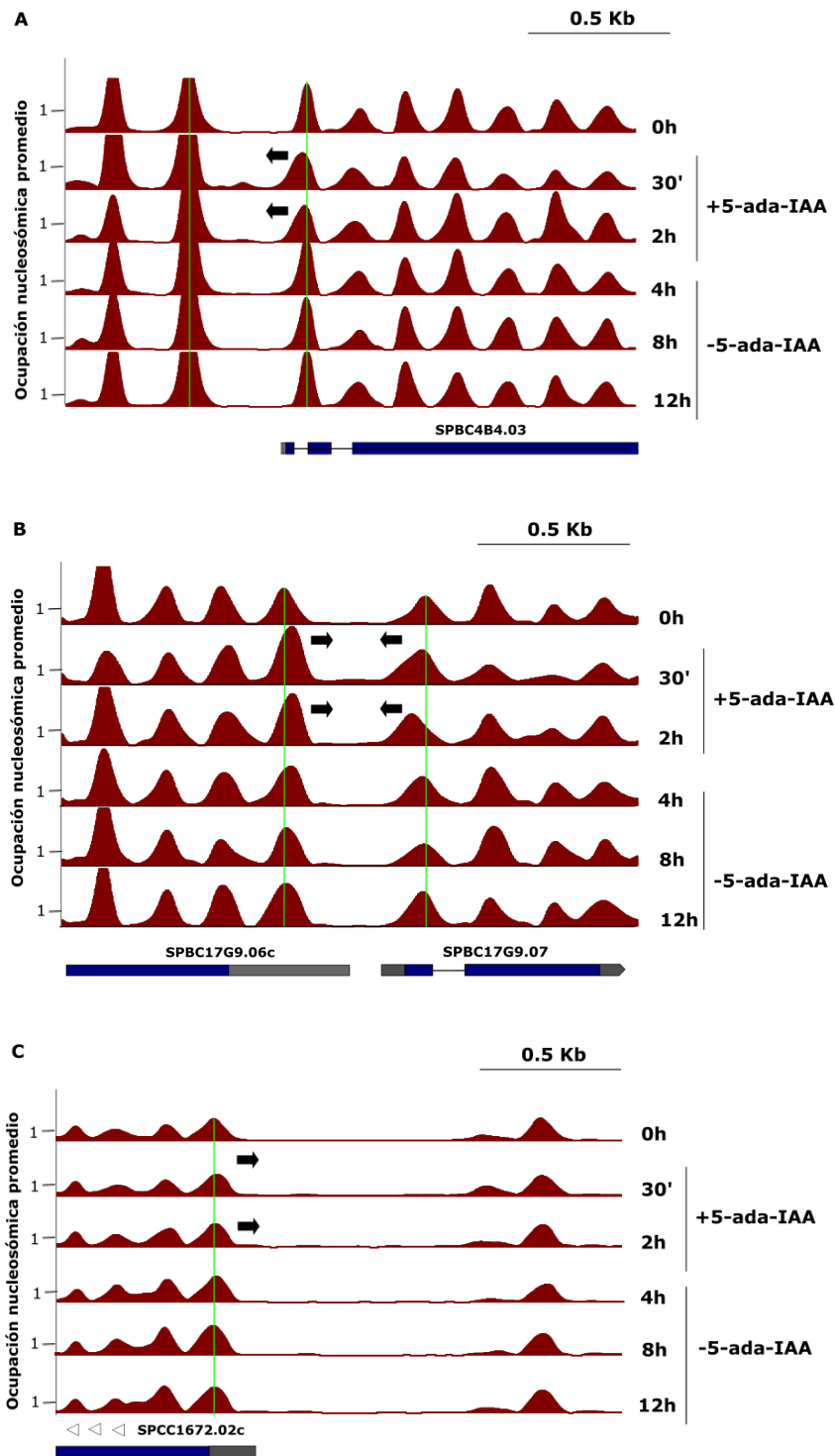
Para estudiar el efecto sobre el tamaño de las NDRs en ausencia de los remodeladores RSC y SWI/SNF, generamos mapas de ocupación nucleosómica mediante *MNase-seq* en las muestras correspondientes a los tiempos mostrados en la **Figura 20**.

Analizamos el tamaño de 1974 NDRs presentes en todas las muestras de la cinética calculando la distancia entre las coordenadas de las diadas de los nucleosomas +1 y -1. En la **Figura 21**, se representan los gráficos de dispersión que comparan el tamaño de las NDRs de cada tiempo respecto a la muestra control del mutante *snf21-sAID*. El panel A, muestra la variación experimental de tamaño de NDRs entre dos réplicas biológicas de la muestra control de 0 horas. A los 30 minutos de añadir 5-adamantil-IAA, se produce una reducción del tamaño del 82% de las NDRs (panel B) que se mantiene en la muestra de 2 horas (panel C). Cuatro horas después de la retirada del inductor, los tamaños de las NDRs se han recuperado en gran parte (panel D) y, a las 8 y 12 horas, vuelven a ser comparables a los de la muestra control (paneles E y F). Estos resultados indican que para mantener la posición del nucleosoma +1 es necesaria la presencia continua de RSC ya que el efecto de la ausencia de su subunidad catalítica Snf21 se produce muy rápidamente. Los resultados también indican que el efecto es reversible y coincide con la aparición de la proteína Snf21 4 horas después de inactivar el *degrón* (**Figura 20A**).



**Figura 21. Dinámica de la reducción del tamaño de las NDRs en la cinética de *snf21-sAID*.** Los gráficos de dispersión representan la variación de tamaño de 1974 NDRs analizadas. Se muestran las comparaciones entre dos réplicas biológicas de la muestra control (A), y entre el control y las muestras correspondientes a los 30 minutos (B) y 2 horas tras la adición de 5-adamantil-IAA (C), y, a las 4, 8 y 12 horas tras su retirada (D-F). Los porcentajes en rojo encima o debajo de la diagonal en cada panel, indican la proporción de NDRs que aumentan o disminuyen de tamaño respecto al control de 0 horas.

La ausencia de la proteína Snf22 del complejo remodelador SWI/SNF tiene un impacto menor sobre el tamaño de las NDRs que la de Snf21 causando solamente la reducción del 34% de ellas (Datos no mostrados). Este fenotipo, igual que en el mutante *snf21-sAID*, se genera rápidamente después de la inactivación del remodelador y también es reversible. Del 34% de NDRs que reducen su tamaño, el 75% también se reducen en *snf21-sAID*. No obstante, cuando se analizan las NDRs afectadas en los dos casos, se observa que la reducción de tamaño causada por la inactivación de RSC es mayor que en el caso de SWI/SNF siendo en promedio de unos 17 frente a 8 nucleótidos, respectivamente. Para ilustrar estos resultados con algunos ejemplos específicos, la **Figura 22** muestra los mapas de ocupación nucleosómica de las regiones genómicas correspondientes al gen *SPBC4B4.03* (**Figura 22A**) y a los genes divergentes *SPBC17G9.06c* y *SPBC17G9.07* (**Figura 22B**) para la cinética de *snf21-sAID*, la región del gen *SPCC1672.02c* (**Figura 22C**) para la de *snf22-sAID*. De acuerdo con los resultados de la **Figura 21**, el nucleosoma +1 se desplaza hacia la NDR en ausencia de estos remodeladores sin que esto afecte a la posición del resto de nucleosomas en los que se organizan los genes.



**Figura 22. Dinámica de la reducción y recuperación del tamaño de las NDRs.** Mapas de nucleosomas de las mismas muestras que en las Figuras 21 y 22. Desplazamiento del nucleosoma +1 en el gen *SPBC4B4.03* en *snf21-sAID* (RSC) (**A**), en los genes divergentes *SPBC17G9.06c* y *SPBC17G9.07* en *snf21-sAID* (RSC) (**B**), y en el gen *SPCC1672.02c* en *snf22-sAID* (SWI/SNF) (**C**). Las flechas indican el desplazamiento del nucleosoma +1 hacia la NDR en ausencia de cada remodelador. Las líneas verdes verticales marcan la diada de los nucleosomas +1 y -1 en la muestra de 0 h. El eje de la izquierda indica la ocupación nucleosómica promedio de todo el genoma.

Estos resultados demuestran que los complejos RSC y SWI/SNF actúan específicamente sobre el nucleosoma +1, contribuyendo al mantenimiento de su posición y, como consecuencia, al tamaño de las NDRs. Por otro lado, no hemos observado diferencias significativas a las 0 horas en el tamaño promedio de las NDR afectadas por uno u otro complejo remodelador.

### **Análisis de la expresión génica en los mutantes condicionales *snf21-sAID* y *snf22-sAID***

Para estudiar cómo afecta a la expresión génica la reducción del tamaño de las NDRs en ausencia de los remodeladores RSC y SWI/SNF, secuenciamos RNA total de dos réplicas biológicas de la muestra control antes de añadir el inductor (0 horas), y de muestras de 2 horas después de añadirlo y 12 horas tras retirarlo. Para el análisis de los datos, empleamos la herramienta bioinformática *RANA-seq* (Prieto & Barrio 2020), que alinea las secuencias sobre el genoma de referencia con HISAT2, y analiza la expresión diferencial con DESeq2. Se consideraron genes sobreexpresados aquellos cuyo nivel de expresión era superior a 1,5 veces el nivel de expresión de la muestra control y, subexpresados aquellos con un nivel inferior a 0,75 veces, que presentaban un p-valor de significación estadística inferior a 0,05 (ver Materiales y Métodos).

En el caso del mutante *snf21-sAID*, a las 2 horas de añadir el inductor (*snf21-sAID* +5-adamantil-IAA 2h) se produce la subexpresión de 1203 genes, de los que 34 mantienen la subexpresión en la muestra de 12 horas después de retirarlo. Además, se sobreexpresan 290 genes, de los cuales 39 se mantienen sobreexpresados a las 12 horas (*snf21-sAID* -5-adamantil-IAA 12h). En esta muestra, se observan 93 genes subexpresados y 234 sobreexpresados cuyo patrón de expresión no se encontraba alterado en la muestra de 2 horas (**Figura 23A**).

Los datos de expresión correspondientes a la cinética del mutante *snf22-sAID* (**Figura 23B**), indican que 2 horas en presencia del inductor (*snf22-sAID* +5-adamantil-IAA 2h) causan la subexpresión de 398 genes y la sobreexpresión de 71 de los cuales 48 y 18 mantienen su patrón de expresión desregulado en la muestra de 12 horas sin 5-adamantil-IAA (*snf22-sAID* -5-adamantil-IAA 2h). Además, se subexpresan 48 genes y se sobreexpresan 33 genes respecto a la muestra de 2 horas.

El mantenimiento de algunos genes desregulados después de retirar el inductor podría ser debido a que 12 horas no son suficientes para restablecer un patrón de expresión silvestre una vez que se ha comprometido la función de los complejos RSC o SWI/SNF. Sería interesante establecer si el nivel de expresión de esos genes se recupera después de más tiempo o si, por el contrario, han sufrido modificaciones en su estructura. Por otra parte, los genes con expresión

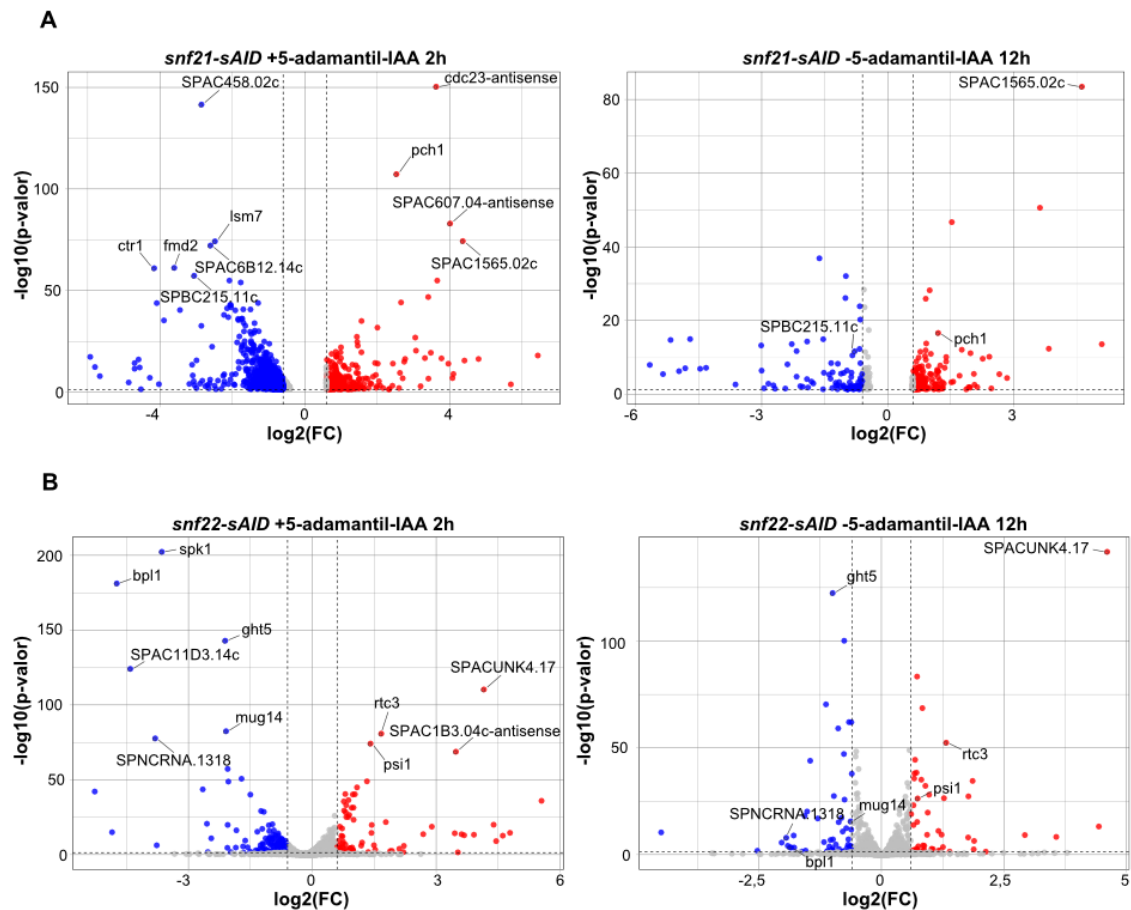
diferencial en la muestra de 12 horas respecto a la de 2 horas, probablemente representan genes que se han desregulado entre las 2 y las 8 horas que se mantuvo el cultivo en presencia de 5-adamantil-IAA antes de retirarlo (**Figura 19**).



**Figura 23. Análisis de expresión diferencial en la cinética de apagado y recuperación de los mutantes condicionales *snf21-sAID* y *snf22-sAID*.** Diagramas de Venn de los genes expresados diferencialmente en las muestras de 2 horas en presencia y 12 horas en ausencia de 5-adamantil-IAA en la cinética de *snf21-sAID* (**A**) y de *snf22-sAID* (**B**). Se consideran genes diferencialmente expresados aquellos cuyo nivel de expresión es superior a 1,5 veces el nivel de expresión en la muestra control (sobreexpresados) o con un nivel inferior a 0,75 veces el nivel del control (subexpresados).

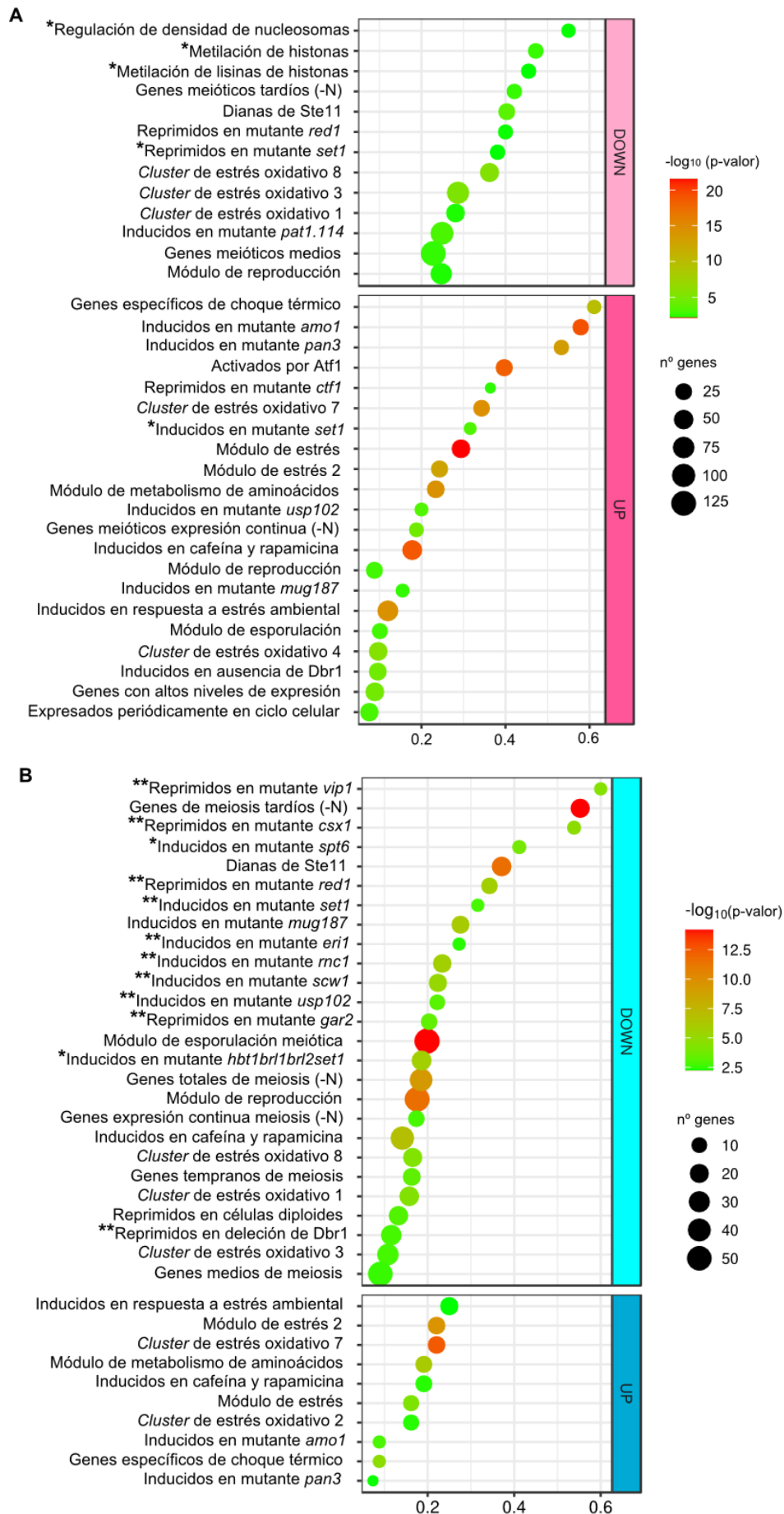
Los resultados del análisis de expresión diferencial muestran que la ausencia de Snf21 afecta a un número de genes mayor que la de Snf22 (**Figura 23**) y, correlaciona con el mayor efecto de la ausencia de Snf21 que de Snf22 en el mantenimiento del tamaño y número de las NDRs afectadas. En ambos casos, más de un 80% de los genes desregulados se subexpresan. De estos genes, el 71% en Snf21 y el 42% en Snf22, están asociados con NDRs en los que el desplazamiento del nucleosoma +1 reduce el tamaño de la NDR (**Figuras 21** y **22**). Por otro lado, los genes subexpresados con modificación de tamaño de su NDR representan un porcentaje muy pequeño del total de NDRs afectadas por la ausencia de estos complejos remodeladores, lo que sugiere que muchas NDRs reducen su tamaño sin que esto determine un cambio detectable en la transcripción de sus genes asociados. En la **Figura 24A** (*snf21-sAID*) se representan con mayor detalle los datos expresión génica de la **Figura 23A** en diagramas de volcán. Se indican algunos ejemplos de genes específicos sobreexpresados en presencia del inductor (panel izquierdo) que recuperan sus niveles de expresión tras 12 horas en su ausencia a excepción de los genes *pch1* o *SPAC1565.02c* (panel derecho). Por otro lado, en la **Figura 24B**, se representan los diagramas de

volcán correspondientes a la **Figura 24B** (*snf22-sAID*). Entre los genes destacados, la mayoría recuperan sus niveles de expresión a las 12 horas en ausencia del inductor o se encuentran en el límite de hacerlo como ocurre con el gen *psi1* (panel derecho). Sin embargo, el gen *SPACUNK4.17* aumenta sus niveles de expresión a las 12 horas en ausencia del inductor, lo que podría corresponder con los mayores tiempos en presencia del compuesto inductor (**Figura 20**).



**Figura 24. Diagramas de volcán de expresión diferencial de los genes de la Figura 24.** Se representan los genes desregulados en la cinética del mutante condicional *snf21-sAID* (**A**) y del mutante *snf22-sAID* (**B**). En rojo se indican los genes sobreexpresados, en azul los subexpresados y en gris aquellos cuyos niveles de expresión no se modifican. El eje X representa en  $\log_2$  el cambio de expresión, y el eje Y representa en  $-\log_{10}$  el p-valor. Las líneas discontinuas verticales marcan los umbrales de expresión y la línea discontinua horizontal marca el umbral de significación estadística.

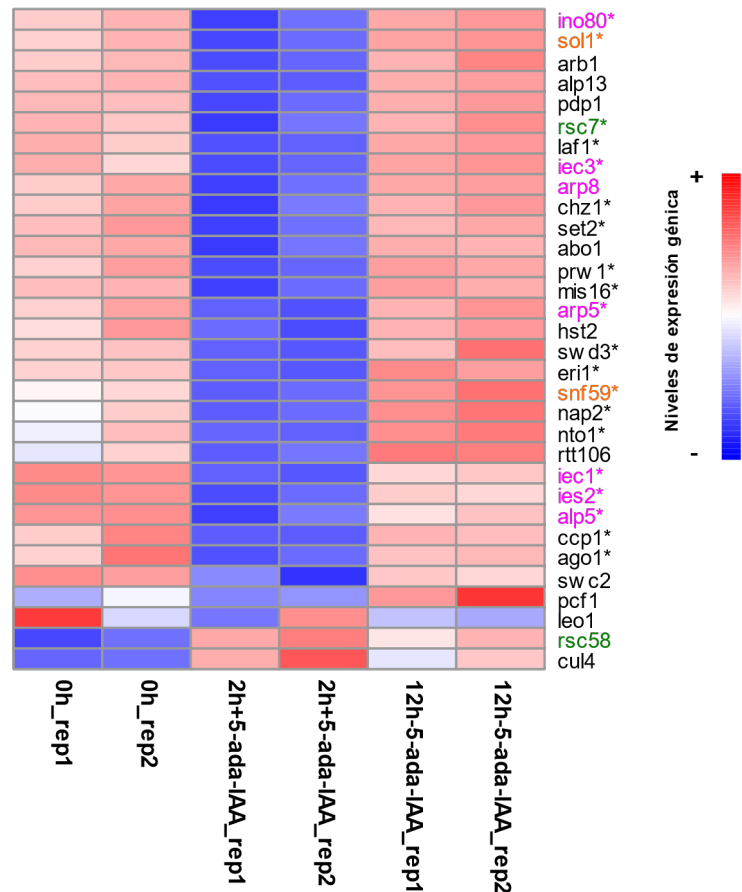
A continuación, se realizó una caracterización funcional de los genes diferencialmente expresados (**Figuras 23 y 24**) mediante un análisis de enriquecimiento funcional con la herramienta Angeli (Bitton et al., 2015). En la **Figura 25** se muestra el resumen de dicho análisis para los genes desregulados a las 2 horas en ausencia de Snf21 (**panel A**) y de Snf22 (**panel B**).



**Figura 25. Análisis de enriquecimiento de los genes desregulados en ausencia de Snf21 y Snf22.** En el eje Y se representan las categorías de *Gene Ontology*(GO) enriquecidas en la lista de genes sobreexpresados (UP) y subexpresados (DOWN) a las 2 horas en presencia de 5-adamantil-IAA respecto a las 0 horas, para los mutantes *snf21-sAID* (A) y *snf22-sAID* (B). El tamaño de los círculos indica el número de genes presentes en esa categoría y el color la significación estadística ordenada de mayor (rojo) a menor (verde). En el eje X se representa la proporción de genes de nuestra lista en relación al número total de genes en cada categoría funcional. Se marcan las categorías relacionadas con la cromatina (\*) y con el metabolismo de mRNAs (\*\*).

Para ambos remodeladores, los genes sobreexpresados están constituidos en su mayor parte por genes de respuesta a estrés. Así, en el mutante condicional *snf21-sAID* (RSC), entre las categorías estadísticamente más significativas se encuentra la correspondiente al módulo de estrés y la relativa a los genes inducidos en cafeína y rapamicina, y en *snf22-sAID* (SWI/SNF) el módulo de estrés 2 y el *cluster* de estrés oxidativo 7. Es importante destacar que en los genes sobreexpresados en ausencia de complejos RSC y SWI/SNF funcionales, no hemos observado modificación del tamaño de las NDRs ni desplazamiento del nucleosoma +1 y que, por lo tanto, la sobreexpresión de estos genes no parece estar causada directamente por la ausencia de acción del remodelador sobre los mismos. En cuanto a los genes subexpresados, ambos mutantes presentan un enriquecimiento en categorías funcionales similares, relacionadas con la respuesta a estrés oxidativo y la regulación meiótica. No obstante, en cada mutante, se subexpresan genes que pertenecen a categorías específicas. Así, en *snf21-sAID* existe un enriquecimiento en aquellas relacionadas con ensamblaje de cromatina y epigenética (**Figura 25A**), mientras que en *snf22-sAID* aparecen sobrerrepresentados genes relacionados con el metabolismo de mRNAs (**Figura 25B**), y con cromatina (**Figura 25B**).

Es importante destacar que 32 de los 127 genes subexpresados en ausencia de RSC codifican subunidades de distintos complejos remodeladores, o enzimas con actividad chaperona de histonas, como se indica en el mapa de calor presentado en la **Figura 26**. Es muy llamativo que los genes subexpresados incluyan hasta siete genes codificantes de distintas subunidades del complejo INO80 (*ino80*, *iec3*, *arp8*, *arp5*, *iec1*, *ies2* y *alp5*), dos de SWI/SNF (*sol1* y *snf59*) y una del propio RSC (*rsc7*). Además, en estos genes se produce una disminución de tamaño de la NDR como consecuencia de la ausencia de Snf21, que revierte en la muestra de 12 horas paralelamente al patrón de expresión. Eso sugiere que RSC podría actuar como un regulador directo de otros complejos remodeladores de cromatina, modulando la síntesis de los mensajeros codificantes de algunas de sus subunidades, y que, por lo tanto, cuando no es funcional podría verse comprometida la función de otros complejos (ver Discusión).



**Figura 26. Mapa de calor de genes desregulados en el mutante *snf21-sAID* relacionados con remodelación de cromatina.** Se representan los TPM (transcritos por millón de Kb) correspondientes a genes relacionados con remodelación de cromatina de dos réplicas biológicas (rep1 y rep2) de las muestras de 0h (control), 2h en presencia del inductor, y 12h tras su retirada. Los niveles de expresión se representan en una escala de color de rojo (niveles altos de expresión) a azul (niveles bajos de expresión). Se indican los genes correspondientes a subunidades del complejo INO80 (rosa), de SWI/SNF (naranja) y de RSC (verde). Los genes en los que además se produce un desplazamiento del nucleosoma +1, están marcados con un asterisco (\*).

### Características de las NDRs reguladas por RSC y SWI/SNF

En *S. cerevisiae*, el remodelador RSC interacciona físicamente con los factores de transcripción Rap1 y Reb1 (Gavin et al., 2002; Wu et al., 2018). Rap1 es un factor de transcripción pionero que se une al DNA nucleosómico, y junto con RSC, contribuye a mantener la posición estable del nucleosoma +1 para facilitar el acceso a otros factores de transcripción como Hmo1, Fhl1 e Ifh1 (Mivelaz et al., 2020). Reb1 se localiza en las NDRs en las inmediaciones del nucleosoma -1 y regula genes transcritos por la RNA Pol I y RNA Pol II, actuando como un factor de terminación (Koerber et al., 2009; Rhee & Pugh, 2011; Kasinathan et al., 2014). En *S. pombe*, el factor Rap1 no está conservado, y en su lugar, actúan los factores de transcripción Crf1 (Ifh1 en *S. cerevisiae*) y Fhl1. Estos factores, interaccionan entre ellos y Fhl1 se une al motivo conservado HomOE

(Maldonado et al., 2023). Reb1 sí está presente en *S. pombe* y su motivo de unión también ha sido caracterizado (Zhao et al., 1997).

Para estudiar la posible asociación entre estos factores y RSC, decidimos analizar los motivos de unión de factores de transcripción en las NDRs de genes subexpresados en ausencia de Snf21. Para ello, utilizamos dos algoritmos de búsqueda de secuencias estadísticamente sobrerrepresentadas: DREME (de la plataforma MEME, Bailey et al., 2015) y XXmotif (Luehr et al., 2012). Las secuencias obtenidas se compararon con las listas de motivos de *S. pombe* descritos en la literatura con el programa TOMTOM (de la misma plataforma MEME), e identificamos los motivos Homole y Reb1 entre aquellos con una significación estadística más alta. La presencia de estos motivos en las NDRs de genes diana del remodelador RSC sugiere que podría haber una asociación entre este remodelador y estos factores de transcripción en *S. pombe*.

El complejo remodelador SWI/SNF actúa principalmente sobre genes cuyos promotores poseen caja TATA en *S. cerevisiae* (Kubik et al., 2019), mientras que en *S. pombe*, esta correlación todavía no ha sido establecida. Utilizando los algoritmos mencionados en el párrafo anterior, buscamos la presencia de caja TATA en las NDRs cuyo tamaño se reduce en ausencia de Snf22 y la encontramos en 807 NDRs de un total de 876, lo que representa el 92% de NDRs con caja TATA en el genoma de *S. pombe*.

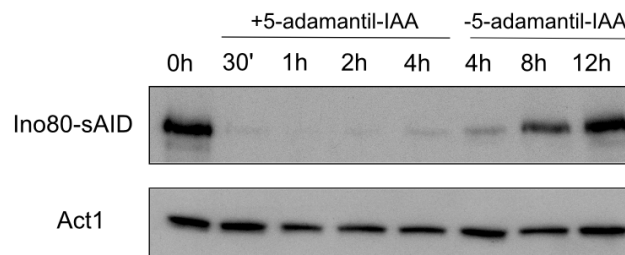
En resumen, los complejos RSC y SWI/SNF actúan específicamente sobre el nucleosoma +1, contribuyendo al mantenimiento de las NDRs. Sin embargo, la ausencia de Snf21 afecta a un mayor número de genes (1203 subexpresados y 290 sobreexpresados) que Snf22 (398 subexpresados y 71 sobreexpresados). De los cuales, 191 y 32, lo están también en ausencia de Snf21, lo que sugiere que estos complejos remodeladores son parcialmente redundantes aunque también realizan funciones específicas. En ambos casos, más de un 80% de los genes desregulados se subexpresan. Es llamativo que en ausencia de RSC, se subexpresen siete subunidades del complejo INO80, entre ellas su subunidad catalítica Ino80, lo que sugiere que RSC podría actuar como un regulador de la actividad de este complejo.

### **Familia de remodeladores INO80**

La familia de remodeladores INO80 pertenece al grupo de los denominados *pullers*, con funciones opuestas a los complejos RSC y SWI/SNF. Esto quiere decir que cuando no es funcional, se produce un incremento del tamaño de las NDRs (Kubik et al., 2019). En *S. cerevisiae*, INO80 contribuye al mantenimiento de la posición estable del nucleosoma +1 y, además, tiene función

como *spacer* porque modula el posicionamiento de nucleosomas a lo largo de los genes (Kubik et al., 2019; Cheon et al., 2021; Singh et al., 2021).

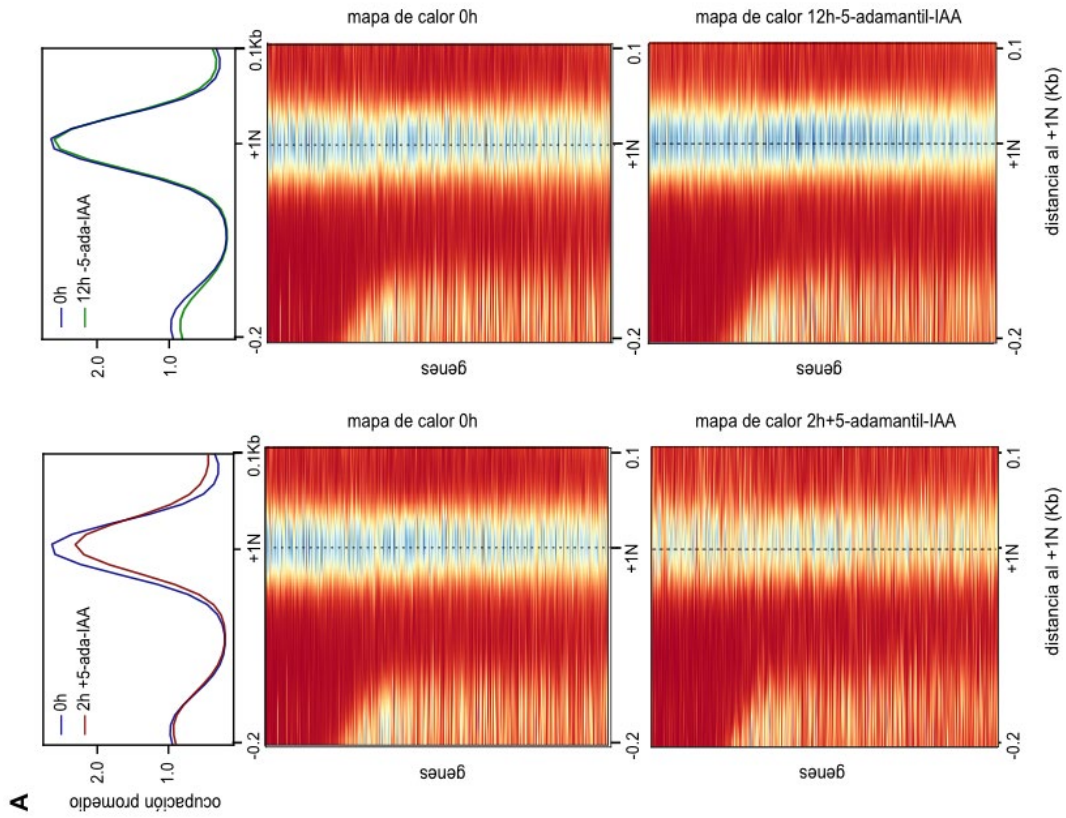
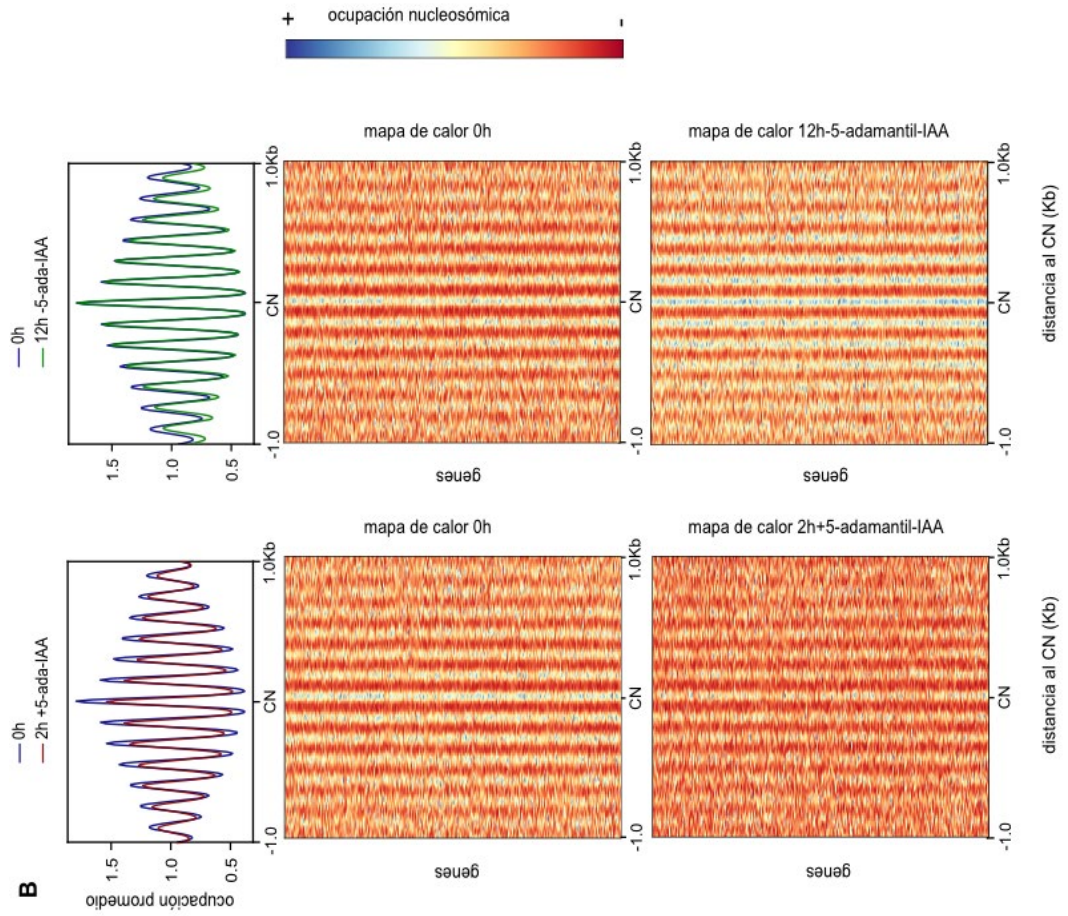
Para estudiar la dinámica de la cromatina y la transcripción regulada por INO80 en *S. pombe*, construimos un mutante condicional *ino80-sAID* y llevamos a cabo experimentos con un diseño similar al de los complejos RSC y SWI/SNF (**Figura 19**). Como se muestra en la **Figura 27**, a los 30 minutos en presencia de 5-adamantil-IAA no se detectan niveles de Ino80, mientras que la recuperación se produce de forma gradual desde las 4 hasta las 12 horas tras la retirada del inductor.



**Figura 27. Niveles de la proteína Ino80 en presencia y ausencia de 5-adamantil-IAA.** *Western Blot* de extractos totales de proteína de la cepa *ino80-sAID*. Los tiempos de cada muestra y la información relativa a los anticuerpos son iguales a los de la Figura 21.

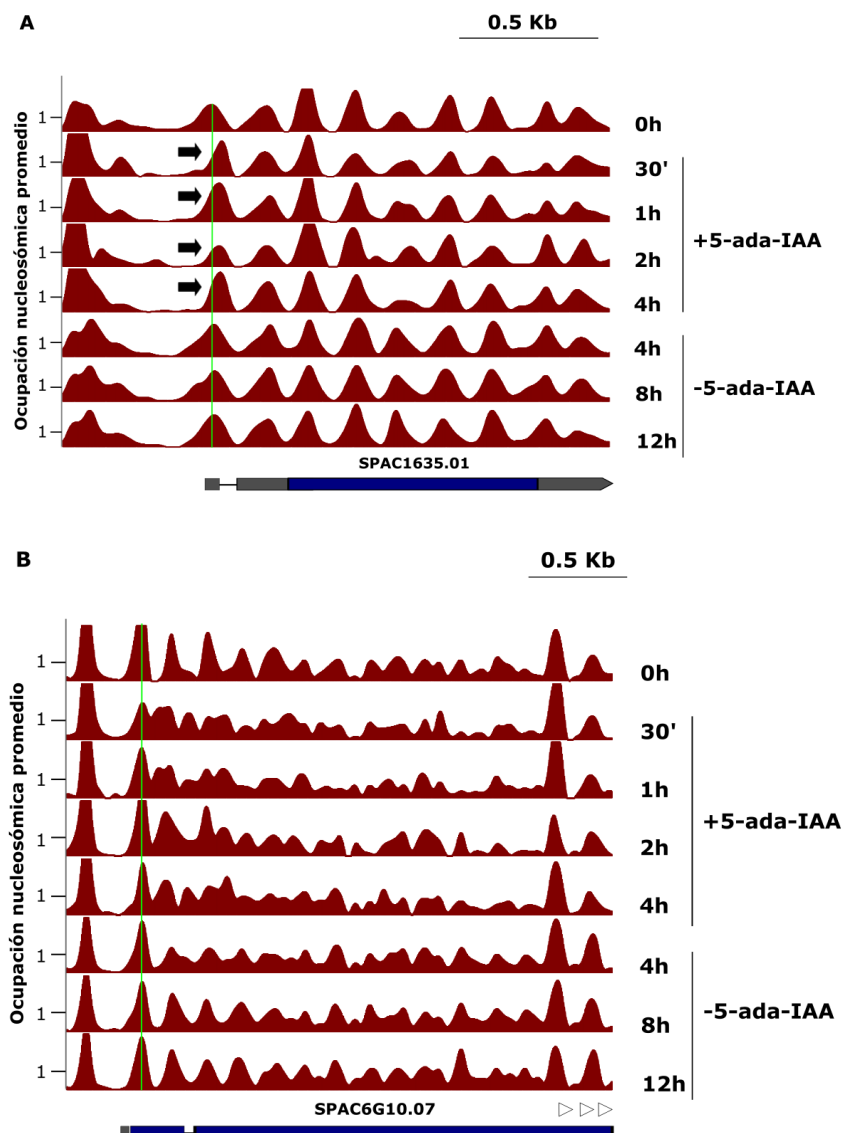
### Dinámica de las NDRs en el mutante condicional *ino80-sAID*

Para estudiar el efecto sobre el tamaño de las NDRs del complejo remodelador INO80, generamos mapas de ocupación nucleosómica mediante *MNase-seq*, en los tiempos mostrados en la **Figura 28**. Para comparar la distribución promedio de los nucleosomas a lo largo de todos los genes, representamos los perfiles utilizando como coordenadas de referencia las diadas de los nucleosomas +1 (+1N) y centrales (CN). Como puede observarse en la **Figura 28A**, a las dos horas en presencia de la molécula inductora, se produce un desplazamiento del nucleosoma +1, de unos 10 nucleótidos en promedio, que tiene como consecuencia un incremento del tamaño de la NDR, mientras que este fenotipo revierte a las 12 horas sin inductor, lo cual indica su actividad como *puller*. Además, la inactivación de INO80 altera el posicionamiento de los nucleosomas de la región central de los genes y se recupera a las 12 horas de la retirada de 5-adamantil-IAA (**Figura 28B**) lo cual indica su actividad como *spacer*.



**Figura 28. Dinámica del cambio de tamaño de las NDRs y posicionamiento de los nucleosomas a lo largo de los genes dependiente de INO80. A)** En la parte superior de la figura, se muestran los perfiles agregados de ocupación nucleosómica de todos los genes del genoma para las muestras control de 0 horas (azul), 2 horas en presencia de 5-adamantil-IAA (rojo) y 12 horas tras su retirada (verde). Se muestra una región de 300 pb de los genes de *S. pombe* alineada a la diada del nucleosoma +1 (+1N). En la parte inferior de la figura, se muestran los mapas de calor de ocupación nucleosómica de los mismos genes ordenados de mayor a menor tamaño de sus NDRs. **B)** Representación de una región de 2 Kb de los mismos genes que en A, alineados a la diada de los nucleosomas centrales (CN).

Como ejemplo de estas alteraciones en una región específica del genoma, se muestran los mapas de ocupación nucleosómica de la región del gen *SPAC1635.01* (**Figura 29A**) donde se observa el aumento de tamaño de la NDR y el desplazamiento del nucleosoma +1 cuando el inductor está presente y, la recuperación del posicionamiento cuando se retira. En la **Figura 29B**, se muestra la región correspondiente al gen *SPAC6G10.07* donde se observa el fenotipo de pérdida de posicionamiento de los nucleosomas a lo largo del gen, sobre todo en la zona central en ausencia de Ino80. En este gen, sin embargo, el nucleosoma +1 no modifica su posición.

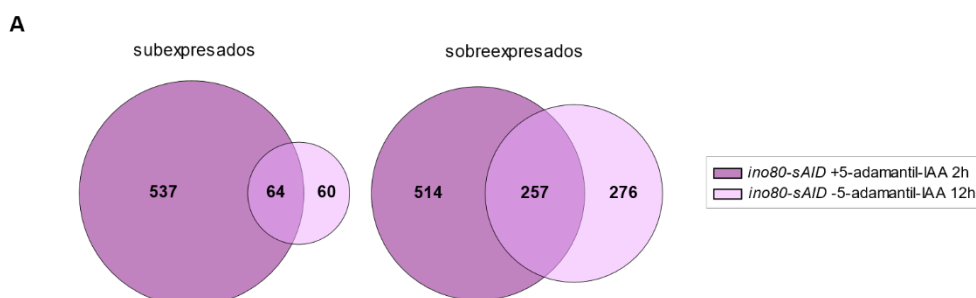


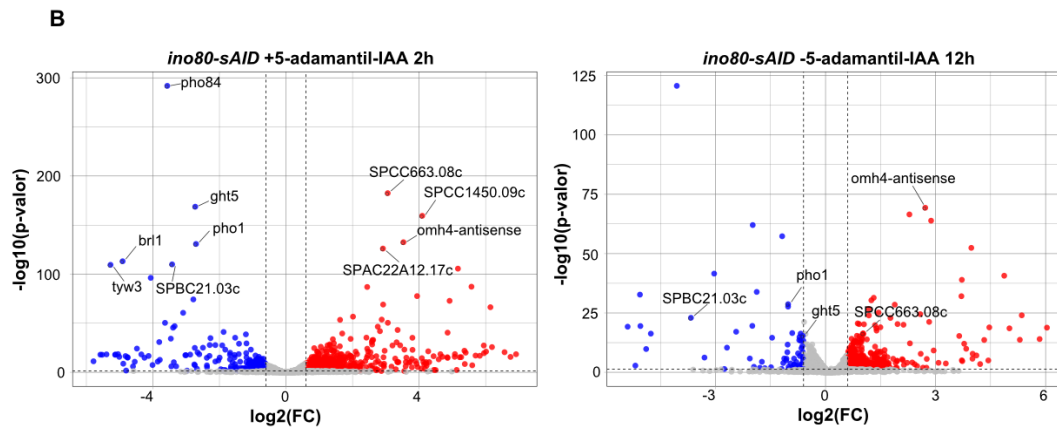
**Figura 29. Dinámica de las NDRs y posicionamiento de nucleosomas en la cinética de *ino80-sAID*.** Mapas de nucleosomas de dos regiones genómicas del cromosoma 1 de *S. pombe* en las mismas muestras que en la Figura 28. **A)** Región del gen *SPAC1635.01* donde se muestra el aumento del tamaño de su NDR (flechas) y su posterior recuperación. **B)** Región del gen *SPAC6G10.07* donde se muestra la pérdida de posicionamiento de los nucleosomas a lo largo del gen en el mutante *ino80-sAID*. Las líneas verdes verticales marcan la diada de los nucleosomas +1. El eje de la izquierda indica la ocupación nucleosómica promedio de todo el genoma.

### Análisis de la expresión génica en el mutante condicional *ino80-sAID*

Para estudiar la transcripción génica en ausencia de Ino80, secuenciamos mediante *RNA-seq*, RNA total correspondiente a los mismos tiempos que para los remodeladores RSC y SWI/SNF (0 horas, 2 horas en presencia de 5-adamantil-IAA y 12 horas tras su retirada).

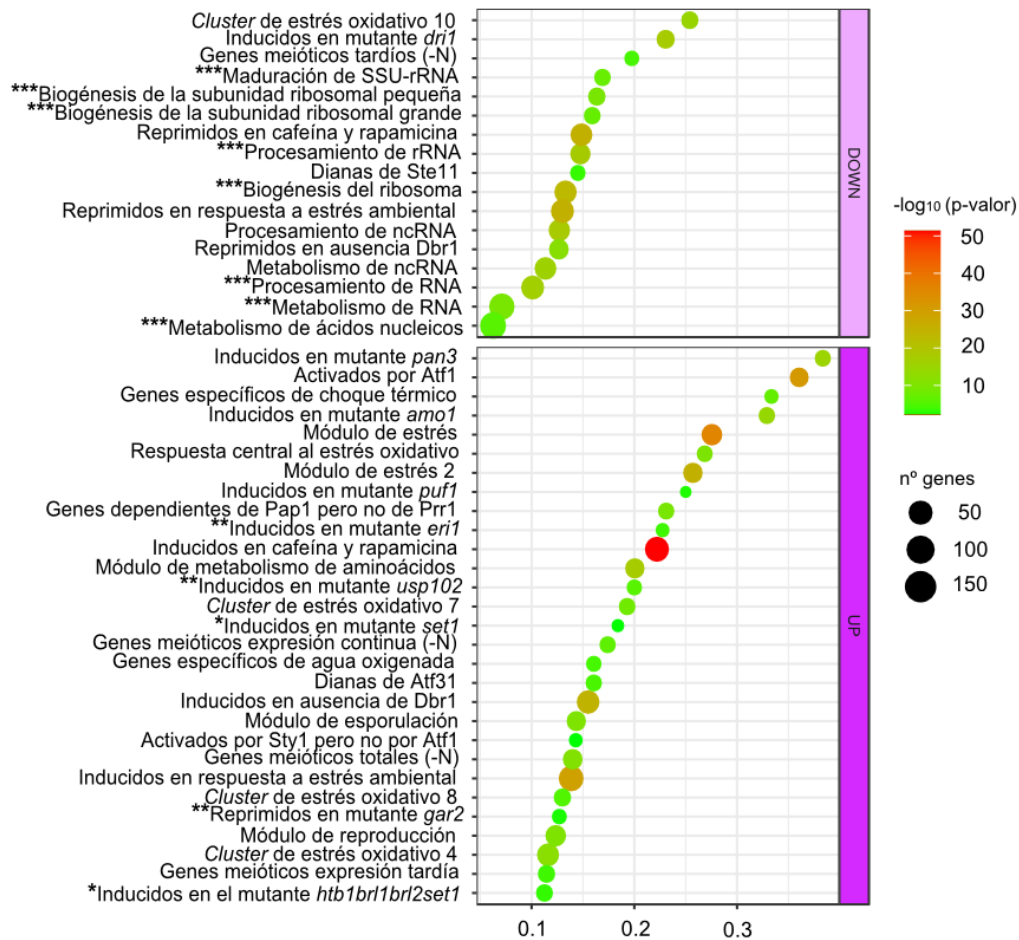
Los diagramas de Venn de la **Figura 30A** muestran que después de 2 horas en presencia del inductor se produce la subexpresión de 601 genes de los cuales 64 se mantienen desregulados 12 horas después de retirar el inductor. Además, se produce la sobreexpresión de 771 genes, de los cuales, 257 se mantienen sobreexpresados a las 12 horas tras su retirada. De estos 771, casi la mitad (362) son ncRNAs, hecho que no ocurría en ausencia de RSC ni de SWI/SNF. Los genes que no recuperan su patrón de expresión 12 horas después de retirar el 5-adamantil-IAA indican que ese periodo podría no ser suficiente para recuperar un fenotipo silvestre o, como se comentó en el caso de RSC y SWI/SNF, que la estructura de la cromatina de algunos genes haya sufrido daños irreparables. Asimismo, en la muestra de 12 horas aparecen 60 genes subexpresados y 276 genes sobreexpresados que no se encontraban en la muestra de 2 horas, y que podrían representar aquellos que se habrían desregulado entre las 2 y 4 horas antes de retirar el inductor (**Figura 27**). La **Figura 30B** representa los diagramas de volcán de estos datos de expresión, mostrándose algunos ejemplos de genes específicos cuyo patrón de expresión se modifica con una significación estadística alta ( $p$ -valor < 0.05). Entre los genes sobreexpresados en presencia del inductor (panel izquierdo), aparecen destacados algunos de los que recuperan sus niveles silvestres de expresión a las 12 horas en su ausencia, a excepción del antisentido del gen *ohm4* o el gen *SPCC663.08c* (panel derecho). En cuanto a los genes subexpresados, se muestran ejemplos de genes que recuperan sus niveles de expresión o se mantienen al límite de hacerlo a las 12 horas en ausencia del inductor como ocurre con *pho1* o *ght5*.





**Figura 30. Análisis de expresión diferencial en la cinética de apagado y recuperación de *ino80-sAID*.** **A)** Diagrama de Venn de los genes diferencialmente expresados. **B)** Diagrama de volcán de los genes de A.

Los datos de *RNA-seq* demuestran que la ausencia del remodelador INO80 determina la subexpresión y sobreexpresión de un número de genes muy similar (601 y 771). Cuando analizamos el desplazamiento del nucleosoma +1 en la muestra de 2 horas en presencia del inductor, observamos que del total de genes subexpresados y sobreexpresados, aproximadamente en la mitad aumenta el tamaño de su NDR (42% y 47%, respectivamente). Estos resultados indican que, en esos genes, la ausencia del remodelador desestabiliza el nucleosoma +1 y causa tanto la activación como la represión de la expresión, mientras que la otra mitad de los genes no serían diana directa del remodelador. En cuanto a los genes no codificantes (ncRNA) sobreexpresados, el 45% corresponde a transcritos antisentido en cuyos genes se produce desplazamiento del nucleosoma +1 en ausencia de Ino80, así como una pérdida de posicionamiento en los nucleosomas adyacentes. Sin embargo, en el 55% restante solamente se pierde el posicionamiento de los nucleosomas más allá del +1. Esto indica que, para algunos genes, INO80 funciona únicamente como espaciador para mantener el posicionamiento de nucleosomas (**Figura 29B**), mientras que, en otros, solo funciona como estabilizador del nucleosoma +1 (**Figura 29A**) y, para otros, puede combinar ambas funciones.



**Figura 31. Análisis de enriquecimiento de los genes desregulados en ausencia de Ino80.** Misma explicación que para la Figura 26. Se marcan las categorías relacionadas con la cromatina (\*), con el metabolismo de mRNAs (\*\*) y con biogénesis de ribosomas y metabolismo de rRNAs (\*\*\*).

En el análisis de enriquecimiento funcional de los genes sobreexpresados en ausencia de Ino80 tras 2 horas en presencia de 5-adamantil-IAA (**Figura 31**), se aprecia un enriquecimiento de aquellos relacionados con la respuesta a estrés: genes inducidos en cafeína y rapamicina, módulo de estrés, genes inducidos por Atf1, genes específicos de choque térmico, módulo de estrés 2, módulos de estrés oxidativo 4, 7 y 8, y genes inducidos en respuesta a estrés ambiental. El enriquecimiento de estas categorías funcionales es un fenotipo que comparte con los mutantes de los remodeladores RSC y SWI/SNF descritos en el apartado anterior, y que probablemente refleja cómo la ausencia de estos complejos activa las rutas de respuesta a estrés. Además de estos genes, en ausencia de Ino80 se sobreexpresan también aquellos que pertenecen a categorías relacionadas con cromatina y epigenética y con el metabolismo de mRNA (**Figura 31**). En cuanto a los genes subexpresados resulta muy llamativa la presencia de varias categorías funcionales relacionadas con la biogénesis de los ribosomas y el metabolismo del rRNA. Estos resultados sugieren que en ausencia de un complejo INO80 funcional la biogénesis de ribosomas

en la célula puede estar comprometida, como consecuencia de defectos en el procesamiento del rRNA. En este sentido, en *S. cerevisiae* se ha descrito que INO80 junto con el remodelador ISW2 (no conservado en *S. pombe*), se asocia con el locus del rDNA, regulando la densidad de nucleosomas en el mismo y su replicación, de manera que, la pérdida de función de estos dos complejos determina un descenso del número de repeticiones del rDNA que no afecta a su transcripción (Cutler et al., 2018). Es posible que INO80 participe en un mecanismo de regulación similar en *S. pombe*, que estamos estudiando actualmente.

### Características de las NDRs reguladas por INO80

En *S. cerevisiae*, se han presentado dos modelos contrapuestos para explicar la interacción de INO80 con las NDRs. Por un lado, Oberbeckmann et al. (2021), propusieron que la asociación del remodelador estaría mediada por el factor de transcripción Reb1, mientras que Chen et al. (2022), determinaron que era independiente de este factor. Para estudiar si en *S. pombe* existe una asociación entre factores de transcripción e INO80, realizamos una búsqueda de secuencias sobrerrepresentadas en las NDRs de los genes desregulados en ausencia del remodelador. Concretamente, caracterizamos los posibles motivos de unión de factores de transcripción en la muestra de 2 horas en presencia de 5-adamantil-IAA, siguiendo la misma estrategia que para los complejos RSC y SWI/SNF descrita anteriormente.

En las NDRs de genes subexpresados encontramos sobrerrepresentados los motivos HomolD, FLEX, y el correspondiente al factor de transcripción Sap1. HomolD es un motivo al que se unen factores de transcripción que regulan la expresión de genes ribosomales (Witt et al., 1995) y su presencia en estas NDRs está directamente relacionada con la desregulación de genes relacionados con la biogénesis del ribosoma en ausencia de Ino80 (**Figura 31**). FLEX se ha caracterizado como motivo de unión de factores de transcripción que regulan la expresión de genes meióticos (Alves-Rodrigues et al., 2016), categoría representada también en el análisis de enriquecimiento funcional mostrado arriba (**Figura 31**). Por su parte, Sap1 es un factor ubicuo que interacciona con el 6,8% del genoma de *S. pombe*, principalmente en las NDRs (Hickey et al., 2015) que, sin embargo, no encontramos en los genes subexpresados en ausencia de RSC y SWI/SNF. Por otro lado, en las NDRs de genes sobreexpresados, encontramos el motivo CRE, al que pueden unirse los factores de transcripción Atf1 y Pcr1 (Eshaghi et al., 2010), encargados de la regulación de la expresión de genes de respuesta a estrés.

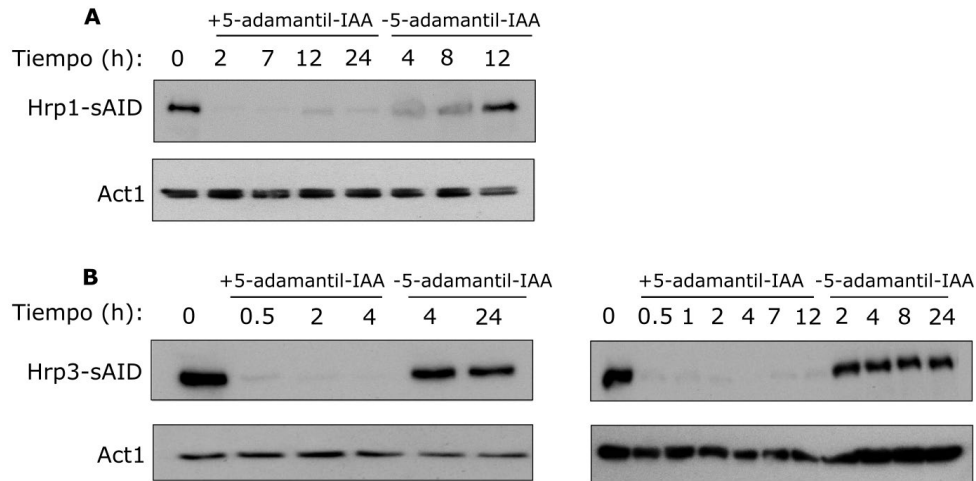
Los motivos de unión de factores de transcripción encontrados en las NDRs de los genes desregulados en ausencia de Ino80, están relacionados con la expresión de grupos de genes mostrados en la **Figura 31**. Sin embargo, no hemos encontrado enriquecimiento para el motivo

de unión del factor Reb1, ni para factores pioneros, a diferencia de lo que ocurre para RSC. Esto sugiere que en *S. pombe* la interacción de INO80 con las NDRs podría producirse de forma directa, sin necesidad de estar mediada por factores pioneros como Reb1 que se uniesen previamente a las mismas o podría llevarse a cabo por la acción de otros factores de transcripción.

### Familia de remodeladores CHD

La familia de remodeladores CHD pertenece al grupo de los *spacers* y como tal, mantiene el posicionamiento de los nucleosomas a lo largo de los genes (Kubik et al., 2019). En *S. pombe*, los remodeladores Hrp1 y Hrp3 pertenecen a la familia CHD. Hrp1, contribuye al mantenimiento del espaciado internucleosómico en los genes (Pointner et al., 2012). Además, es necesario para la condensación de los cromosomas previa a su segregación (Yoo et al., 2000), para mantener elevados los niveles de CENP-A (Walfridsson et al., 2005; Yamada et al., 2008) y, participa en la terminación de la transcripción (Alén et al., 2002). Por su parte, la ausencia de Hrp3 produce una pérdida del posicionamiento de los nucleosomas, un descenso de la ocupación en las regiones transcritas y favorece la transcripción críptica y antisentido (Hennig et al., 2012; Shim et al., 2012; Touat-Todeschini et al., 2012; Dong et al., 2021). Además, se ha descrito su relación con el complejo Mediador, importante regulador de la RNAPII, a través de la interacción física descrita con su subunidad Med15 (Khorosjutina et al., 2010).

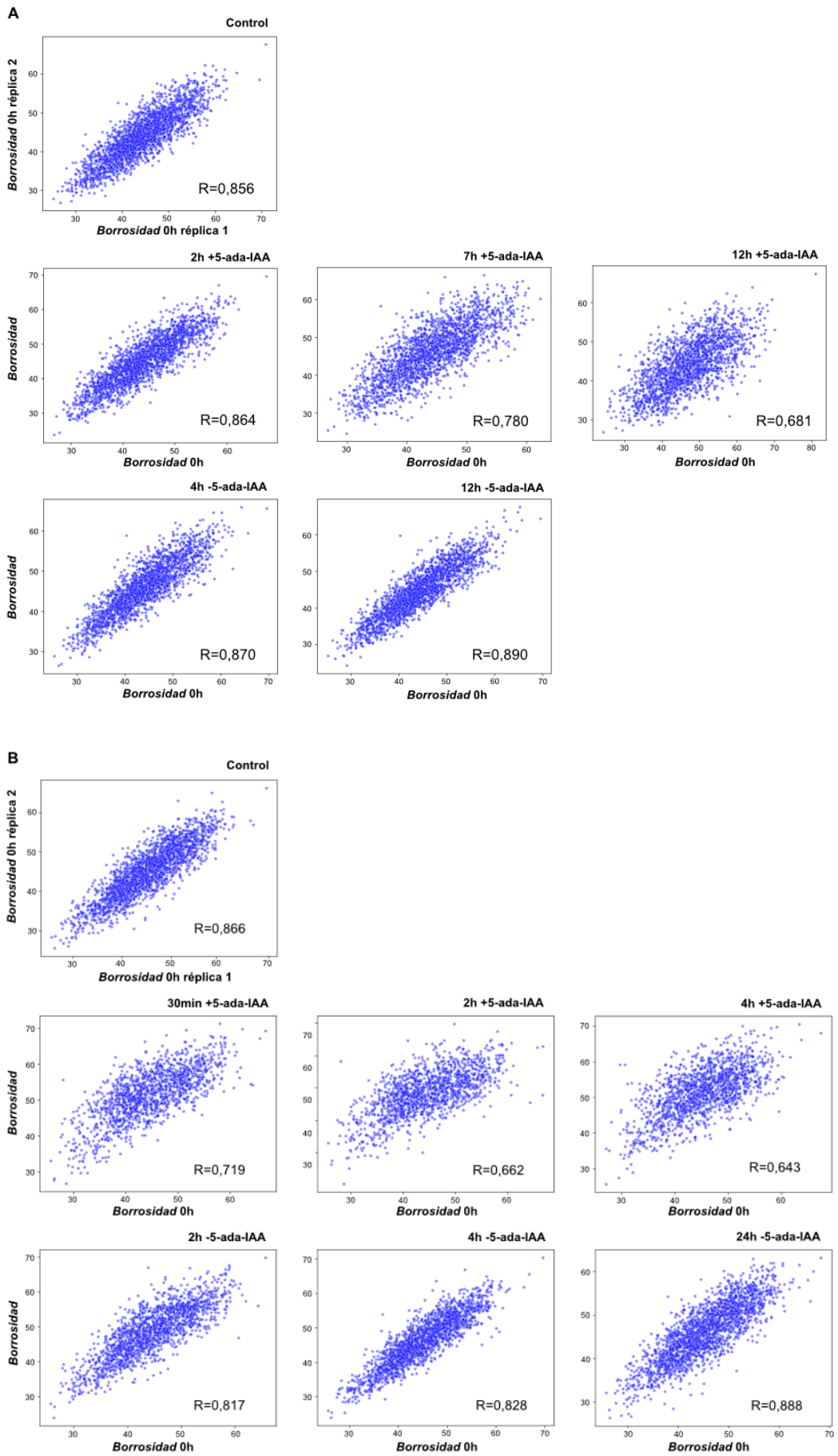
Para estudiar la dinámica de la cromatina y la transcripción regulada por los dos remodeladores CHD, construimos los mutantes condicionales *hrp1-sAID* y *hrp3-sAID*. Inicialmente, analizamos por *Western blot* los niveles de cada proteína a lo largo del tiempo en las cinéticas de degradación y recuperación. Para Hrp1, se recogieron muestras a las 0 h (control), a las 2, 7, 12 y 24 horas en presencia del inductor, y a las 4, 8 y 12 horas tras su retirada. Para Hrp3, se recogieron muestras en dos experimentos independientes a las 0 h (control), 30 minutos, 1, 2, 4 y 7 horas en presencia de 5-adamantil-IAA, y a las 2, 4, 8 y 24 horas tras su retirada. Como puede observarse en la **Figura 32A**, tras 2 horas en presencia del inductor no se detectan niveles de Hrp1, mientras que la recuperación se produce de forma gradual desde las 4 hasta las 12 horas tras su retirada. En el caso de Hrp3, 30 minutos en presencia de 5-adamantil-IAA son suficientes para que la proteína desaparezca por completo, y 2 horas tras la retirada del inductor, para que los niveles de proteína sean similares a los de la muestra control (**Figura 32B**).



**Figura 32. Niveles de las proteínas Hrp1 y Hrp3 en presencia y ausencia de 5-adamantil-IAA.** **A)** *Western Blot* de extractos de proteínas totales para la cepa *hrp1-sAID* (**A**) y *hrp3-sAID* (**B**). Se indican las horas en presencia y ausencia del inductor. La información relativa a los anticuerpos es igual a los de la Figura 21.

### Dinámica de los nucleosomas en los mutantes condicionales *hrp1-sAID* y *hrp3-sAID*

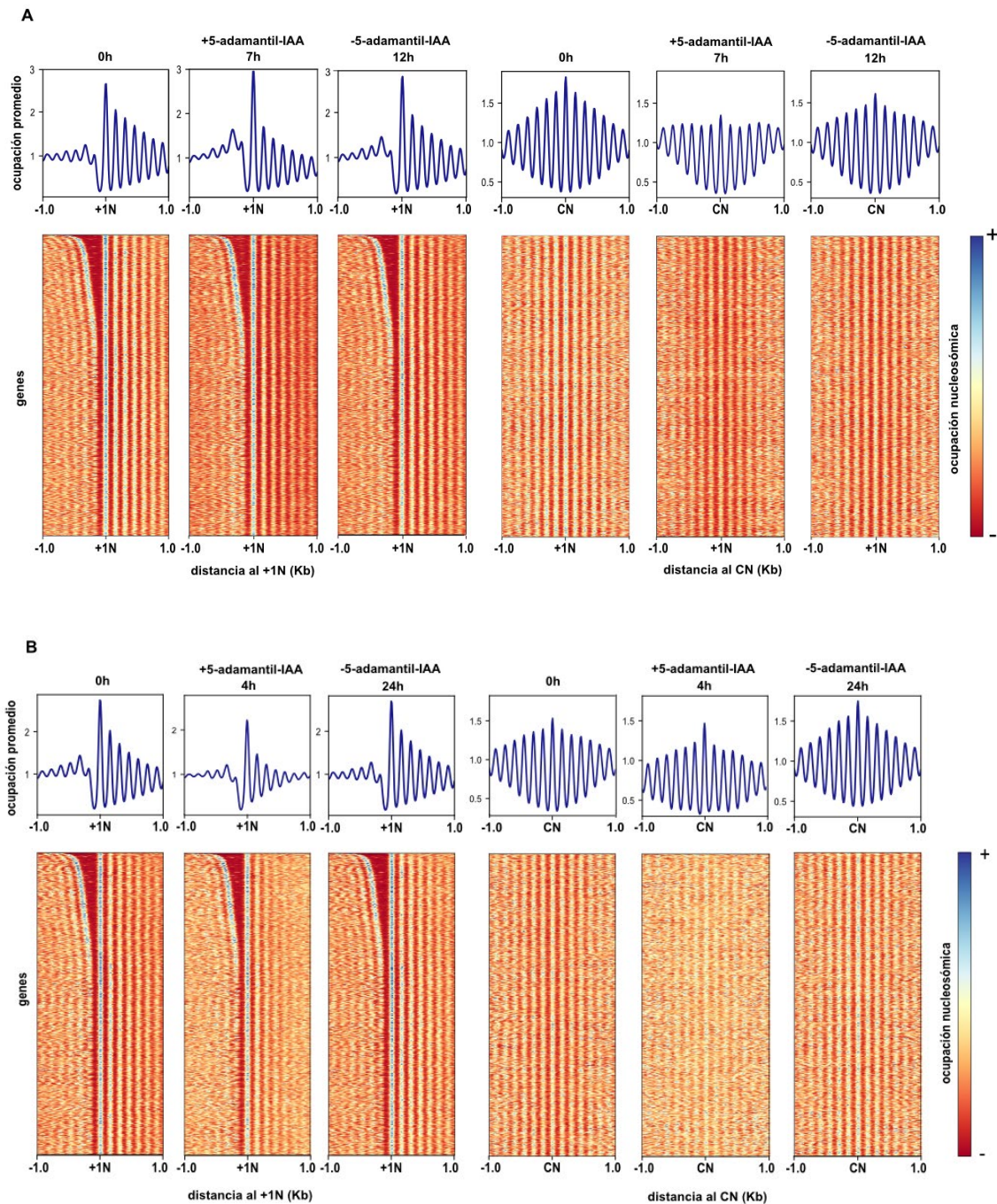
Para estudiar los cambios en ocupación y posicionamiento nucleosómico mediante *MNase-seq*, empleamos los mismos tiempos que en la **Figura 32**, estimando la *borrosidad* con el programa DANPOS como hicimos en el apartado 1 de esta sección de Resultados (Pág. 32). En la **Figura 33**, se representa la dispersión correspondiente a los valores de *borrosidad* de un nucleosoma localizado en la región central de cada gen a cada tiempo de las cinéticas de *hrp1-sAID* y de *hrp3-sAID* respecto a la muestra de 0 h (control). La correlación entre muestras se midió utilizando el coeficiente de correlación de Pearson ( $R$ ). Como puede observarse en los paneles superiores de la **Figura 33 (A y B)**, hay una elevada correlación entre dos réplicas biológicas de la muestra control ( $R=0,856$  y  $R=0,866$ , respectivamente), corroborando la fiabilidad de nuestras mediciones. En las muestras de 2, 7 y 12 horas en ausencia de Hrp1, se produce un incremento progresivo de la *borrosidad* nucleosómica ( $R=0,864$  a las 2h y  $R=0,681$  a las 12h) (**Figura 33A**). Aunque 2 horas después de añadir el inductor no se detectan niveles de proteína por *Western blot* (**Figura 32A**), el posicionamiento nucleosómico es similar al del control, lo que indica que la ausencia de Hrp1 no tiene un efecto inmediato sobre la dinámica de los nucleosomas. Sin embargo, 4 horas después de retirar el inductor parecen suficientes para que se revierta el fenotipo de pérdida de posicionamiento (**Figura 33A**) aunque los niveles de proteína detectables aún son considerablemente inferiores a los niveles del control (**Figura 32A**).



**Figura 33. Dinámica de pérdida y recuperación de posicionamiento de los nucleosomas en los mutantes *hrp1-sAID* y *hrp3-sAID*.** **A)** Gráficos de dispersión de la *borrosidad* para *hrp1-sAID* de 1963 nucleosomas que ocupan posiciones centrales entre la muestra de 0 horas y las muestras correspondientes a las 2, 7 y 12 horas tras la adición de 5-adamantil-IAA (paneles centrales) y, 4 y 12 horas tras su retirada (paneles inferiores). **B)** Gráficos de dispersión para *hrp3-sAID* de la *borrosidad* de 2021 nucleosomas centrales entre la muestra de 0 horas y las muestras correspondientes a 30 minutos, 2 y 4 horas tras la adición de 5-adamantil-IAA (paneles centrales) y 2, 4 y 24 horas tras su retirada (paneles inferiores). En cada mutante, se han comparado dos réplicas biológicas del control a 0 h para determinar la variabilidad experimental (paneles superiores). El valor R representa el coeficiente de correlación de Pearson, como medida de la dispersión entre muestras.

En el caso de *hrp3-sAID*, la **Figura 33B** muestra cómo tras 30 minutos en presencia de 5-adamantil-IAA la *borrosidad* de los nucleosomas aumenta ( $R=0,719$ ), alcanzándose la máxima dispersión en la muestra de 4 horas ( $R=0,643$ ). Cuatro horas después de retirar el inductor, los niveles de *borrosidad* de los nucleosomas son similares a los de la muestra control ( $R=0,828$ ) y sólo disminuyen ligeramente después de 24 horas ( $R=0,888$ ). Estos resultados indican que el impacto sobre el posicionamiento de los nucleosomas de la pérdida de Hrp3 es más rápido que para Hrp1, que requería hasta 12 horas en presencia de 5-adamantil-IAA para obtener valores de *borrosidad* similares a los observados para Hrp3 en 2 horas.

Estos resultados indicaban que en los mutantes condicionales *hrp1-sAID* y *hrp3-sAID* se produce una pérdida reversible del posicionamiento de 1963 y 2021 nucleosomas centrales correspondientes a tales genes, respectivamente. Para visualizar el efecto sobre la distribución promedio de todos los nucleosomas a lo largo de todos los genes de *S. pombe*, alineamos el mapa de ocupación de cada uno a la diada de los nucleosomas +1 e, independientemente, a la de los nucleosomas centrales. En el primer caso (**Figura 34A**), se observa que la ausencia de Hrp1 (muestra 7h +5-adamantil-IAA) no tiene impacto en el mantenimiento de las NDRs ni en la ocupación y posicionamiento de los nucleosomas +1. Sin embargo, sí se aprecia una disminución en la ocupación de los nucleosomas adyacentes al +1 en el perfil agregado y una reducción del posicionamiento en el mapa de calor, que se visualiza en la mayor dispersión de las columnas que indican la posición de los nucleosomas adyacentes al nucleosoma central. El fenotipo observado revierte tras 12 horas en ausencia de 5-adamantil-IAA (**Figura 34A**).

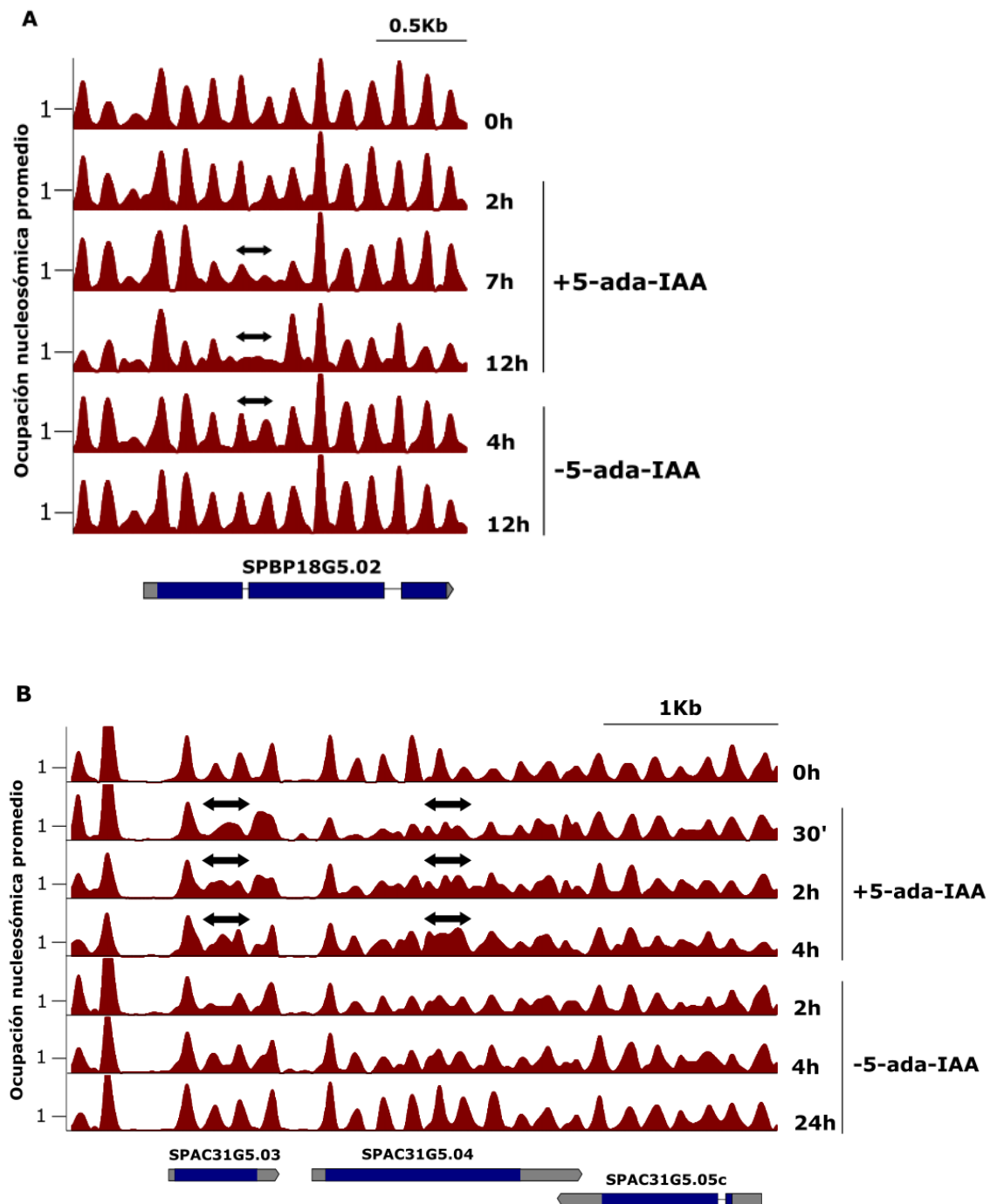


**Figura 34. Dinámica de la pérdida y recuperación del posicionamiento y ocupación de los nucleosomas en los mutantes *hrp1-sAID* (A) y *hrp3-sAID* (B).** En la parte superior de las figuras, se muestran los perfiles agregados de ocupación nucleosómica (azul) para las muestras control (0 h), en presencia y ausencia de 5-adamantil-IAA. En la parte inferior de A y B, se muestran los mapas de calor que representan la ocupación nucleosómica de cada uno de los genes de los perfiles agregados superiores de una región de 2 Kb ordenados de mayor a menor tamaño de su NDR y alineados a la diada del nucleosoma +1 (+1N) o a la de los nucleosomas centrales (CN).

En el caso de Hrp3 (**Figura 34B**), observamos una pérdida de ocupación y posicionamiento nucleosómico mucho mayor que en Hrp1 tal y como se observaba en los análisis de *borrosidad* de la **Figura 33**. Esta alteración es mayor a partir del nucleosoma +1 aunque este también está

afectado, a diferencia del mutante *hrp1-sAID*. Este fenotipo es aún más evidente en la alineación de los nucleosomas a lo largo de los genes utilizando la diada del nucleosoma central como referencia. Tanto el perfil agregado de ocupación nucleosómica como el mapa de calor en presencia del inductor (**Figura 34A**), muestran una alteración muy significativa de la ocupación y del posicionamiento nucleosómico. Además, en estos perfiles, la reversión del fenotipo tras la retirada del inductor no parece completa, observándose todavía una menor ocupación y un peor posicionamiento que en la muestra control en el mutante *hrp1-sAID*, que en el *hrp3-sAID*. Esto puede deberse a que el tiempo de recuperación para Hrp3 fue de 24 horas mientras que para Hrp1 fue de 12 horas.

Como ejemplo del efecto de la pérdida y recuperación del posicionamiento nucleosómico en genes específicos, la **Figura 35A**, muestra los mapas de ocupación en la región del gen *SPBP18G5.02* en el mutante *hrp1-sAID* donde se observa el mismo fenotipo que el descrito para el promedio genómico. Estos resultados sugieren que la diana preferente del remodelador Hrp1, serían los nucleosomas que ocupan posiciones centrales de los genes. La **Figura 35B** muestra los mapas de ocupación de la región que comprende los genes *SPAC31G5.03*, *SPAC31G5.04* y *SPAC31G5.05c* del mutante *hrp3-sAID*. A las 4 horas en presencia de 5-adamantil-IAA, se produce una pérdida de posicionamiento en los nucleosomas que ocupan posiciones centrales de los genes *SPAC31G5.03* y *SPAC31G5.04* y, una reducción de la ocupación del nucleosoma +1. En la muestra de 24 horas en ausencia del inductor, el patrón nucleosómico es comparable al de la muestra control de 0 horas. Por lo tanto, ambos remodeladores, caracterizados como *spacers* tienen como nucleosomas diana aquellos que ocupan posiciones intermedias en los genes y, por tanto, en su ausencia se produce la pérdida del posicionamiento de ellos, siendo este fenotipo más acusado en el mutante de Hrp3 que, además, puede afectar al nucleosoma +1.

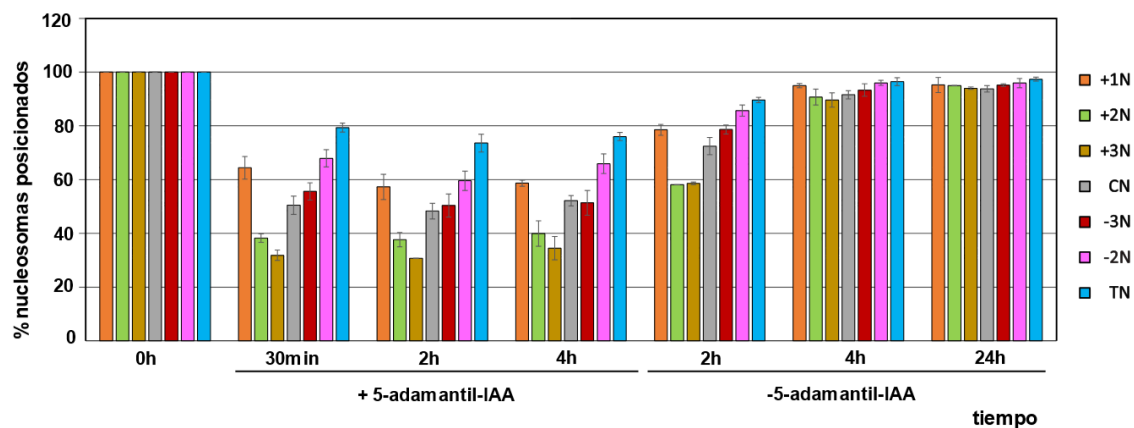


**Figura 35. Dinámica de la pérdida de posicionamiento y ocupación de los nucleosomas en los mutantes *hrp1-sAID* y *hrp3-sAID*.** **A)** Mapas de nucleosomas de las mismas muestras que en la Figura 33A de la región genómica del gen *SPBP18G5.02* para el mutante condicional *hrp1-sAID*. **B)** Mapas de nucleosomas de las mismas muestras que en la Figura 33B de la región genómica donde se encuentran los genes *SPAC31G5.03*, *SPAC31G5.04* y *SPAC31G5.05c* para el mutante condicional *hrp3-sAID*. Las flechas bidireccionales indican los nucleosomas que pierden el posicionamiento en presencia de 5-adamantil-IAA. El eje X indica la ocupación nucleosómica promedio de todo el genoma.

Para estudiar de forma más precisa los nucleosomas diana sobre los que podía actuar Hrp3, seleccionamos siete grupos no solapantes de nucleosomas a lo largo de los genes de *S. pombe*. Estos grupos de unos 2500 nucleosomas diferentes, incluían los nucleosomas +1 (+1N), los dos

siguientes (+2N y +3N), los que ocupan la posición central de los genes (CN), los que solapan con el TTS (TN) y los dos anteriores a estos (-2N y -3N). Cuantificamos con DANPOS los niveles de *borrosidad* promedio de cada grupo a lo largo de la cinética del mutante condicional *hrp3-sAID*. En la **Figura 36**, puede observarse como a los 30 minutos de añadir el inductor al cultivo, aproximadamente un 40% de los nucleosomas +1 de *S. pombe* pierden el posicionamiento y hasta un 60-70% de los nucleosomas +2 y +3.

Sin embargo, la ausencia de Hrp3 solamente afecta a un 20% de los nucleosomas que solapan con el TTS (TN). Este fenotipo se revierte por completo a las 4 horas de retirar el inductor, tal y como se muestra en las **Figuras 33** y **34**. Estos resultados sugieren que Hrp3 puede contribuir al posicionamiento de los nucleosomas en los genes desde el TSS, aunque se vuelve prescindible o menos relevante en torno al TTS.

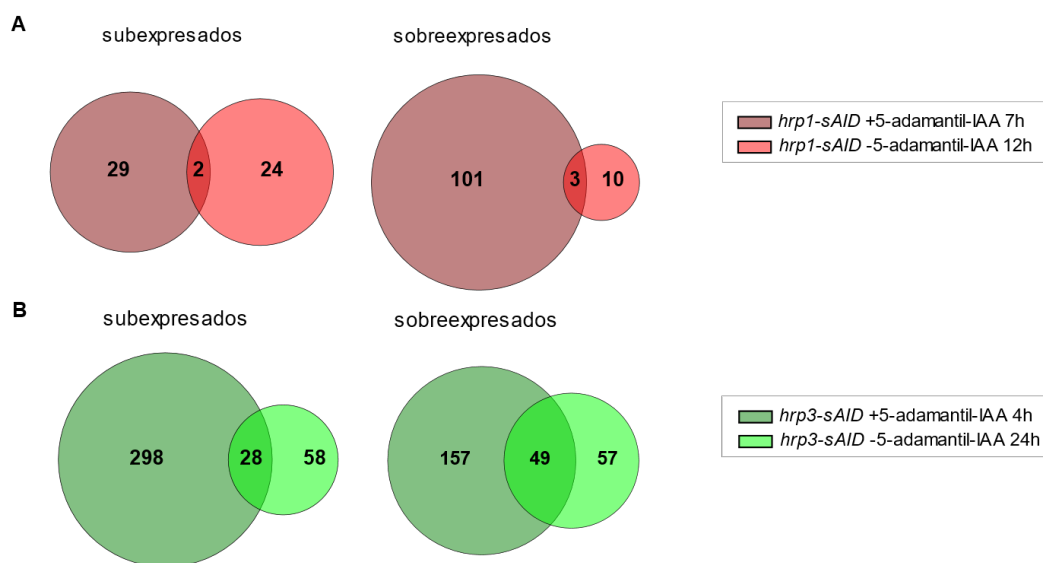


**Figura 36. Medida de *borrosidad* en distintos grupos de nucleosomas a lo largo de los genes en el mutante condicional *hrp3-sAID*.** Los histogramas indican el porcentaje de nucleosomas que se mantienen posicionados en los siete grupos de nucleosomas descritos en el texto e indicados a la derecha, a las 0 horas, 30 minutos, 2 y 4 horas en presencia del inductor y 2, 4 y 24 horas en su ausencia. Las barras de error indican la desviación estándar de las cuantificaciones realizadas en dos réplicas biológicas de cada muestra.

### **Análisis de la expresión génica en los mutantes condicionales *hrp1-sAID* y *hrp3-sAID***

Con el objetivo de estudiar si la alteración en el empaquetamiento de la cromatina que causa la ausencia de Hrp1 está asociada a cambios en la transcripción de los genes, secuenciamos mediante *RNA-seq*, el RNA total de las muestras de 0 horas (control), de 7 horas en presencia del inductor, y de 12 horas tras su retirada. En el caso de Hrp3, secuenciamos la muestra control (0 h), la de 4 horas en presencia del inductor y la de 24 horas después de su retirada. Se secuenciaron dos réplicas biológicas de cada muestra y analizamos los datos con las mismas herramientas bioinformáticas que para los remodeladores anteriores.

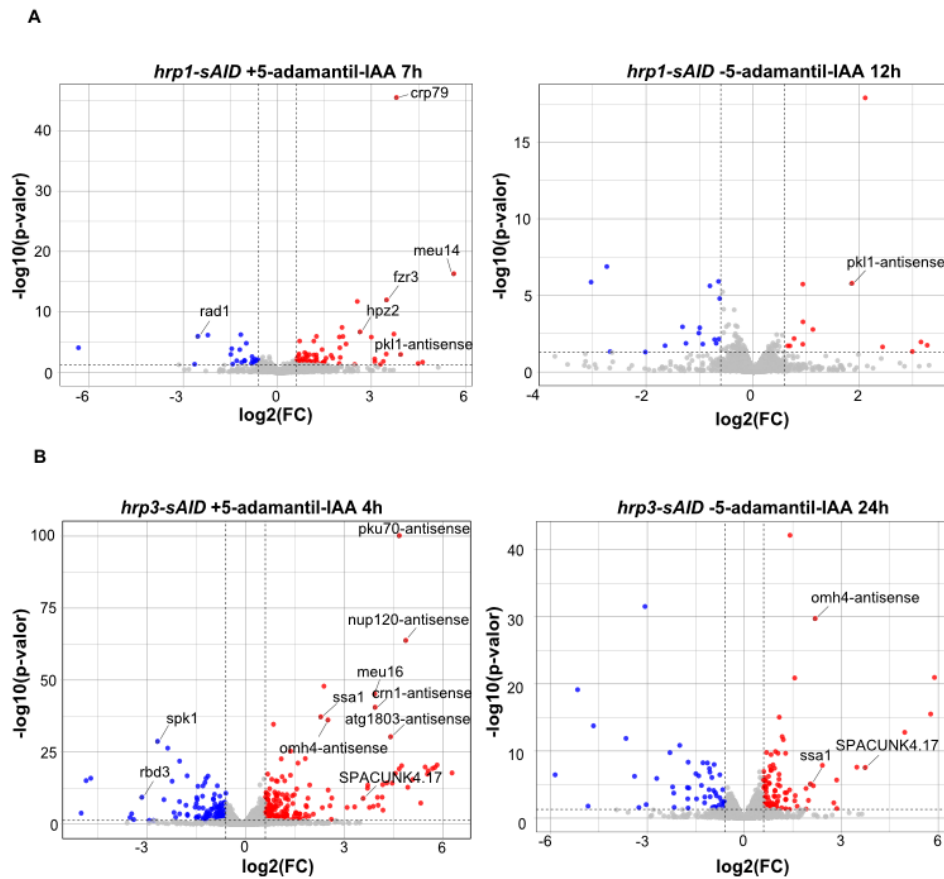
Como puede observarse en los diagramas de Venn en la **Figura 37A**, en el mutante *hrp1-sAID* a las 7 horas en presencia del inductor se produce la subexpresión de 31 genes de los que solo 2 se mantienen subexpresados a las 12 horas después de su retirada. Además, en la misma muestra se sobreexpresan 104 genes, de los cuales solamente 3 mantienen la sobreexpresión 12 horas después de retirar el inductor. En la muestra de recuperación también observamos 24 genes subexpresados y 10 sobreexpresados, cuyo patrón de expresión no se encontraba alterado en la muestra de 7 horas en ausencia del remodelador. Sin embargo, en ausencia de Hrp3, se produce la subexpresión de 326 genes y la sobreexpresión de 206 (**Figura 37B**). En este caso, cuando la molécula inductora se retira del medio, encontramos 86 genes subexpresados y 106 sobreexpresados, lo que sugiere que 24 horas en ausencia de 5-adamantil-IAA podrían no ser suficientes para que se recuperasen sus niveles de expresión basal. De estos dos grupos, 28 y 49, respectivamente, mostraban su expresión alterada en la muestra de 4 horas. El resto, se corresponderían con genes con niveles de expresión desregulados a partir de las 4 horas de incubación con 5-adamantil-IAA.



**Figura 37. Análisis de expresión diferencial en la cinética de apagado y recuperación de *hrp1-sAID* y *hrp3-sAID*.** A) Diagrama de Venn para representar los genes diferencialmente expresados en las muestras de 7 horas en presencia del inductor y 12 horas tras su retirada en *hrp1-sAID* (A) y en las muestras de 4 horas en presencia del inductor y 24 horas tras su retirada en *hrp3-sAID* (B).

En la **Figura 38A**, se representan los datos de la **Figura 37A** (*hrp1-sAID*) utilizando diagramas de volcán para visualizar las diferencias entre algunos genes específicos. Entre los más sobreexpresados en presencia del inductor (panel izquierdo), aparecen destacados genes como *crp79*, *meu14*, *fzr3*, *hpz2* y el antisentido de *pk11*, que recuperan sus niveles silvestres de expresión tras 12 horas en su ausencia, a excepción del antisentido de *pk11* (panel derecho). El gen *rad1* es

un ejemplo de los genes más subexpresados a las 7 horas en presencia del inductor, pero recupera su nivel de expresión en la muestra de 12 horas en su ausencia. Por otro lado, en la **Figura 38B** se representan los diagramas de volcán correspondientes a los datos de la **Figura 37B** (*hrp3-sAID*).

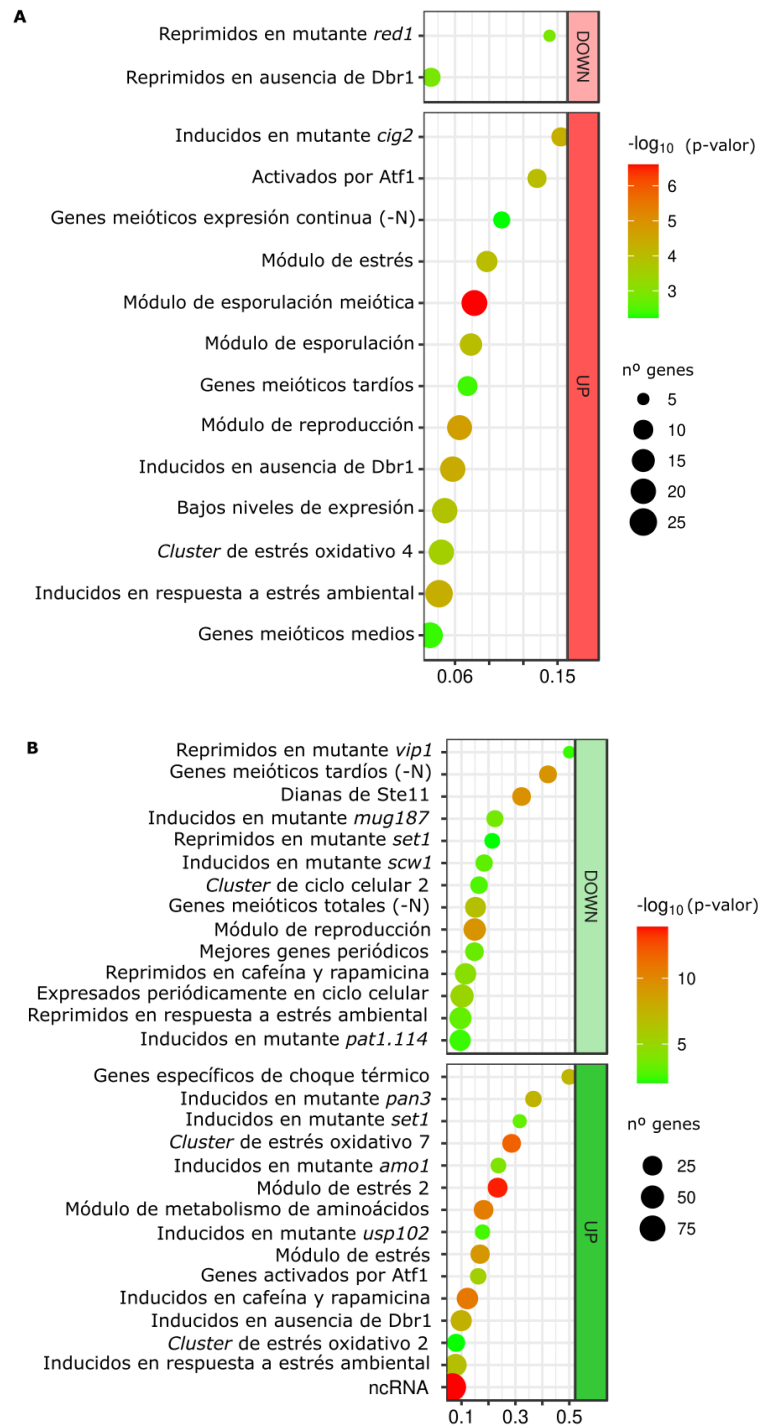


**Figura 38. Diagramas de volcán de los datos de expresión diferencial en la cinética de apagado y recuperación de *hrp1-sAID* y *hrp3-sAID*.** Diagramas de volcán de los genes sobre y subexpresados de la Figura 38. En rojo se indican los genes sobreexpresados, en azul los subexpresados y en gris aquellos cuyos niveles de expresión no se modifican. El eje X representa en  $\log_2$  el cambio de expresión, y el eje Y representa en  $-\log_{10}$  el p-valor. Las líneas discontinuas verticales marcan los umbrales de expresión y la línea discontinua horizontal marca el umbral de significación estadística.

Como muestra de genes sobreexpresados (panel izquierdo), se destacan los antisentidos de los genes *pku70*, *nup120*, *crn1*, *atg1803* y *omh4*. En relación a esto, aproximadamente el 50% de los genes que se sobreexpresan en ausencia de Hrp3 corresponden a antisentidos y transcritos no codificantes, de acuerdo con las observaciones previas en el mutante *hrp3Δ* (Shim et al. 2012; Pointner et al., 2012). Además, aparecen sobreexpresados otros genes como *meu14*, *ssa1* y *SPACUNK4.17*. En la muestra de 24 horas en ausencia del inductor (panel derecho), estos genes, muestran niveles de expresión similares a los de la muestra control, a excepción del antisentido de *omh4* o de los genes *SPACUNK4.17* o *ssa1*. En cuanto a los genes subexpresados, se muestran

como ejemplo *spk1* y *rbd3*, cuyos niveles de expresión sí se recuperan tras 24 horas sin 5-adamantil-IAA. De forma conjunta, estos resultados sugieren que la ausencia de Hrp1, afecta a la expresión de un número más pequeño de genes que la ausencia de Hrp3.

A continuación, con el fin de realizar una caracterización funcional de los genes expresados diferencialmente mostrados en las **Figuras 37 y 38**, se realizó un análisis de enriquecimiento funcional (**Figura 39**) utilizando la misma herramienta que para los remodeladores anteriores.



**Figura 39. Análisis de enriquecimiento de los genes desregulados en ausencia de Hrp1 y Hrp3.** En el eje Y se representan las diferentes categorías de *Gene Ontology* (GO) de expresión génica enriquecidas en la lista de genes sobreexpresados (UP) y subexpresados (DOWN) a las 7 horas para Hrp1 (**A**) y a las 4 horas para Hrp3 (**B**) en presencia de 5-adamantil-IAA en comparación con el tiempo de 0 horas. El tamaño de los círculos se refiere al número de genes presentes en esa categoría y el color a la significación estadística ordenada de mayor (rojo) a menor (verde). En el eje X se representa la proporción de genes de nuestra lista en relación al número total de genes que forman cada categoría funcional.

En los genes sobreexpresados en *hrp1-sAID* a las 7 horas en presencia de 5-adamantil-IAA (**Figura 39A**), se aprecia un enriquecimiento en varias categorías de *Gene Ontology* (GO) de las cuales, la más significativa es la formada por el módulo de esporulación meiótica, seguida del módulo de reproducción formado por genes meióticos o genes de envuelta nuclear entre otros. Se muestran, además, otras categorías relacionadas con la respuesta a estrés como el *cluster* de estrés oxidativo o el módulo de estrés (**Figura 39A**). Dentro de los módulos de esporulación meiótica y del módulo de reproducción, se encuentran los genes diferencialmente sobreexpresados con una significación estadística alta que se muestran en la **Figura 38A** (*crp79, meu14, fzf3, hpz2*). Por el contrario, en los genes subexpresados, encontramos enriquecidas las categorías de GO de genes reprimidos en el mutante *red1* y en ausencia de la proteína Dbr1. Ambas categorías están relacionadas con el metabolismo de mRNAs, aunque ninguna de ellas tiene una significación estadística tan elevada como la de las categorías pertenecientes a los genes sobreexpresados.

La caracterización funcional de los genes con patrón de expresión alterado en ausencia de Hrp3 se muestra en la **Figura 39B**. En los genes sobreexpresados se aprecia un enriquecimiento en varias categorías de GO de expresión génica de las cuales, la más significativa es la formada por los RNAs no codificantes (como los antisentidos mencionados arriba), seguida del módulo de estrés 2. Se muestran además otras categorías relacionadas con estreses como el módulo de estrés, módulo de genes de choque térmico, *clusters* relacionados con el estrés oxidativo, genes activados por Atf1 (factor de transcripción activador de genes de respuesta a estrés) y genes inducidos en respuesta a estrés ambiental. Por el contrario, en los genes subexpresados, encontramos un enriquecimiento en grupos relacionados con los procesos de esporulación y meiosis, tales como los módulos de genes meióticos tardíos, dianas de Ste11, inducidos en el mutante *mug187*, genes meióticos totales, *clusters* de ciclo celular, módulo de reproducción, genes expresados periódicamente en el ciclo celular y los inducidos en el mutante *pat1.114*. Hay que tener en cuenta, que la sobreexpresión de antisentidos en ausencia de Hrp3 conlleva la subexpresión de los transcritos sentidos correspondientes a dichos genes. Esto supone que entre las categorías funcionales que aparecen enriquecidas para los genes sobreexpresados y

subexpresados encontremos algunas opuestas entre estos dos grupos, como ocurre para el caso de los genes de respuesta a cafeína y rapamicina o los genes de respuesta a estrés ambiental.

Estos resultados nos indican que la ausencia de Hrp1 produce una sobreexpresión de genes meióticos y de estrés, mientras que la ausencia del remodelador Hrp3 lo hace de muchos transcritos antisentido y de genes relacionados con respuestas a estrés. Sin embargo, los genes que se subexpresan se asocian a la regulación del ciclo celular y de la meiosis.



# DISCUSIÓN





## 1. LAS MODIFICACIONES EN LA SECUENCIA DEL DNA AFECTAN AL POSICIONAMIENTO DE NUCLEOSOMAS

El desarrollo de técnicas de secuenciación masiva ha permitido la obtención de mapas de ocupación nucleosómica de genoma completo en muchos organismos, incluyendo levaduras (Tsankov et al., 2010; Givens et al., 2012), *Drosophila* (Chereji et al., 2019), *C. elegans* (Valouev et al., 2008), el pez medaka (Sasaki et al., 2009) y células humanas (Gaffney et al., 2012; Valouev et al., 2011). Todos ellos, presentan sus nucleosomas posicionados en la mayor parte del genoma, indicando que estos ocupan posiciones estables respecto a la secuencia del DNA.

Desde el punto de vista de la secuencia del DNA, hay dos mecanismos posibles para llevar a cabo el posicionamiento de los nucleosomas. El primero de ellos, propone que la secuencia del DNA posee información para formar unas barreras físicas a partir de las cuales se irían posicionando pasivamente los nucleosomas, guardando una distancia regular entre ellos independientemente de la secuencia del DNA asociada a ellos (Kornberg & Stryer, 1988; Vaillant et al., 2010; Krietenstein et al., 2016; revisado en Singh & Mueller-Planitz 2021). Estas barreras físicas podrían ser factores de transcripción unidos a los promotores o el nucleosoma +1 que se encuentra unido de forma muy estable a la secuencia del DNA y a partir del cual se posicionarían los nucleosomas a lo largo del gen (Deniz et al., 2016; Chereji et al., 2018). Si este modelo de posicionamiento funcionase en *S. pombe*, indicaría que siempre que las barreras físicas se mantuviesen, los cambios en la secuencia de DNA asociada a los nucleosomas no deberían alterar su posición. Sin embargo, cuando se reemplazó la ORF del gen *SPBC16G5.03* por la del marcador KanMX6, sin modificar la posición del nucleosoma +1 del gen reemplazado ni la NDR asociada, el resultado mostró un patrón desorganizado sobre el gen *KanMX6* (González et al., 2016).

El segundo modelo propone que el posicionamiento podría estar determinado por la secuencia (Segal et al., 2006; Kaplan et al., 2009). Cuando González et al. (2016) modificaron la secuencia del DNA del gen *ura4* reemplazando los codones endógenos por los sinónimos con un contenido en G+C muy diferente, el patrón nucleosómico a lo largo del gen se perdía por completo, lo que indica que la secuencia del DNA, al menos en parte, tiene una función relevante sobre la organización del patrón nucleosómico. Sin embargo, la correcta localización de los nucleosomas no puede ser una función exclusiva de la secuencia del DNA ya que en ausencia de remodeladores de cromatina como Hrp3 en *S. pombe* o CHD1 en *S. cerevisiae*, el posicionamiento se pierde por completo en (Pointner et al., 2012; Kubik et al., 2019).

No obstante, las *signatures* nucleosómicas descritas en Quintales et al. (2015a), demostraron que la secuencia contiene información para posicionar nucleosomas incluso en

moléculas artificiales de DNA y que, además, son específicas de especie. Dado que tanto la secuencia del DNA como los remodeladores de cromatina tienen papeles relevantes en la organización de los nucleosomas, el proyecto de esta Tesis Doctoral ha estado dedicado a analizar la contribución relativa de cada uno de ellos en la levadura *Schizosacharomyces pombe*.

En la primera parte del trabajo hemos analizado en detalle los requerimientos de la secuencia del DNA para especificar la localización de los nucleosomas en el genoma. Para ello, hemos incorporado las *signatures* nucleosómicas a la secuencia del gen *ura4* de *S. pombe* y de *S. japonicus*. En primer lugar, los resultados indican que mientras las versiones obtenidas a partir de matrices de mononucleótidos (*ura4SJ\_1mer*) y trinucleótidos (*ura4SJ\_3mer*) presentaban una pérdida completa del posicionamiento de sus nucleosomas (**Figura 9B y 9D**), la versión del gen *ura4* obtenida a partir de la matriz de dinucleótidos (*ura4SJ\_2mer*) es la que mejor posicionamiento de nucleosomas presenta (**Figura 9C**).

Sin embargo, el hecho de que la secuencia modificada *ura4SJ\_2mer* no recapitulara exactamente el posicionamiento de los nucleosomas del gen *ura4* de *S. pombe* silvestre (**Figura 9**) ni el del gen *ura4* de *S. japonicus* modificado del trabajo de González et al. (2016), podría deberse al hecho de haber introducido modificaciones en las secuencias reguladoras. Así, llevamos a cabo la remasterización de la ORF del gen *ura4* de *S. japonicus*, excluyendo de esta, las UTRs en 5' y 3', presentes en la versión *ura4SJ\_2mer*. El análisis de posicionamiento mostrado en la **Figura 10B**, indica que la versión *ura4SJ\_2mer*<sup>-UTR</sup> conseguía mejorar el posicionamiento de los nucleosomas respecto a las versiones *ura4SJ\_1mer*, *ura4SJ\_2mer* y *ura4SJ\_3mer* que presentaban sus UTRs modificadas (**Figura 9**). Esto sugeriría que la información relevante para el posicionamiento de los nucleosomas estaba presente en la región codificante del gen.

Otro elemento interesante que podría estar relacionado con el posicionamiento de nucleosomas es el sesgo de uso de codones. Fue descrito por Grantham et al (1980) y hace referencia a una mayor frecuencia de uso de algunos codones respecto a sus sinónimos en el genoma. Este sesgo está presente en los genomas de levaduras hasta humanos (revisado en Liu, 2020; Yang et al., 2021) y está relacionado con la eficiencia de la transcripción y la traducción de los mRNAs (Zhao et al., 2021; Zhou et al., 2016). Además, existe una asociación entre el uso de codones y la estabilidad del mRNA que afecta directamente a la tasa de elongación de la traducción, ya que los codones más frecuentes están decodificados por los tRNAs más abundantes (Harigaya & Parker, 2016). Así, el sesgo de codones también puede afectar al plegamiento de la proteína y su correcto funcionamiento (Zhou et al., 2016; Yang et al., 2021). Además de estas funciones bien caracterizadas del sesgo de uso de codones, una menos

estudiada es si tiene alguna influencia en el posicionamiento de nucleosomas. Si así fuese, al modificar los codones veríamos una alteración del patrón nucleosómico. En la **Figura 10C** se muestra cómo el posicionamiento de los nucleosomas mejora cuando se incorpora el sesgo de uso de codones a la secuencia de DNA en relación a la secuencia que no lo tiene incorporado (**Figura 10B**). Por lo tanto, parece que el sesgo de uso de codones, además de su papel en la regulación de las tasas de transcripción y traducción, también contribuye al posicionamiento de los nucleosomas y es posible que, si esa función tiene valor adaptativo, haya sido seleccionada a lo largo de la evolución.

Llegados a este punto, nos preguntamos que tenían en común todas las secuencias de DNA que posicionaban nucleosomas y que, en principio, estaría ausente en las que no eran capaces de hacerlo. Dado que no hay una correlación entre el contenido A+T global de la secuencia del DNA y el posicionamiento, era razonable suponer que la información relevante debía estar en la distribución de los nucleótidos a lo largo de las 150 pb del DNA nucleosómico.

Para comprobarlo, analizamos las posibles diferencias en la composición de bases en cada mitad de los 150 nucleótidos del DNA asociado con los nucleosomas. Esta posibilidad se basaba en el trabajo reciente de García et al. (2024) en el que describieron que la distribución de los nucleótidos es diferente en la mitad proximal y distal de los nucleosomas respecto al TSS a lo largo de los genes en *S. pombe*. Los resultados revelaron una composición asimétrica en el contenido de A y T en las dos mitades en las cuatro versiones del gen *ura4* capaces de posicionar nucleosomas (*ura4SJ\_2mer*, *ura4SJ\_2mer<sup>-UTR</sup>* y *ura4SJ\_2mer<sup>-UTR,CU</sup>*) que no estaba presente en las versiones que no posicionaban nucleosomas (*ura4SJ\_1mer* y *ura4SJ\_3mer*) (**Tabla 2**). Por lo tanto, al incorporar la *signature* nucleosómica en la secuencia de DNA, estábamos introduciendo una asimetría en la distribución de A y T a la vez, que se reducía el contenido en G y C en algunas posiciones. Esto, sugiere que la diferencia en la composición de nucleótidos en las dos mitades del DNA nucleosómico podría ser suficiente para posicionar nucleosomas y que, quizás, no fuese necesario incorporar el perfil de la *signature* definido por la matriz de puntuación.

Para comprobar esta posibilidad, modificamos la secuencia de las ORFs de los genes *ura4* de *S. pombe* y *S. japonicus* sustituyendo los codones silvestres de cada gen por sus sinónimos, de manera que la composición final de los cuatro nucleótidos en la mitad proximal y distal de cada nucleosoma se asemejara lo más posible a la composición de las secuencias que posicionan nucleosomas (**Tabla 2**). La versión *ura4SP\_AT* presenta un patrón similar al del gen *ura4* silvestre de *S. pombe* (**Figura 11B**), lo que indica que, a pesar de la sensibilidad del posicionamiento a los cambios en la secuencia del DNA, se pueden incorporar muchas modificaciones sin que afecte a

la localización de los nucleosomas. Sin embargo, la versión *ura4SJ\_AT* (**Figura 11C**) mejora el patrón de posicionamiento del gen *ura4* de *S. japonicus* comparado con su versión silvestre (**Figura 7**). Una característica a tener en cuenta es que, al introducir los cambios en los codones endógenos para modificar el contenido en A y T, la mitad proximal permite más cambios que la mitad distal en el gen *ura4* de ambas especies. Es probable que, aunque en las secuencias nativas individuales no se detecta esta asimetría en el contenido de nucleótidos, algunos codones estén optimizados en determinadas posiciones del DNA nucleosómico con el fin de favorecer el posicionamiento. En relación a ello, García et al. (2024) generaron dos versiones modificadas del gen *ura4* de *S. pombe* en la secuencia de cada una de las mitades del DNA nucleosómico y observaron que los cambios en la mitad distal conducían a la pérdida completa del patrón nucleosómico mientras que los cambios en la mitad proximal no lo alteraban. Este hecho, junto a la diferente permisividad a los cambios en cada una de las mitades de las versiones del gen *ura4* de este trabajo, indica que no contribuyen de la misma manera al posicionamiento de los nucleosomas.

La demostración más radical de que la asimetría en el contenido en A y T es suficiente para posicionar nucleosomas en *S. pombe* independientemente de la *signature* nucleosómica, es que aplicando este criterio conseguimos posicionar nucleosomas en secuencias arbitrarias (**Figura 12**) e incluso desplazar su distribución manteniendo su periodicidad (**Figura 13**). La secuencia aleatoria de partida (*sec.artificial\_control*), poseía un contenido en AT similar al promedio genómico de *S. pombe* y la modificamos por tripletes siguiendo la misma estrategia que para las secuencias codificantes. La versión resultante (*sec.artificial\_AT*) está organizada en nucleosomas bien posicionados mientras que en la secuencia *sec.artificial\_control* hay una pérdida completa del posicionamiento (**Figura 12**). Por lo tanto, esta asimetría entre el contenido de A y T por mitades era suficiente para llevar a cabo el posicionamiento de nucleosomas tanto de secuencias naturales como artificiales en *S. pombe*.

En segundo lugar, modificamos la ORF del gen *ura4* de *S. pombe* aplicando la asimetría en el contenido de A y T por mitades con un desfase de 21 pb en el sentido de la transcripción. Los resultados de posicionamiento de la versión *ura4SP\_AT\_mov21* (**Figura 13A**), muestra los seis nucleosomas bien posicionados, pero con las diadas desplazadas en el sentido de la transcripción, tal y como habíamos introducido en desfase en la secuencia. Todo ello nos indica la gran importancia de la secuencia del DNA en el posicionamiento de los nucleosomas incluso en una posición diferente a su localización endógena. Por lo tanto, la modificación del DNA nucleosómico mediante el desbalance de A y T por mitades está de acuerdo con el hecho de que existen muchas

maneras diferentes de cómo la secuencia de DNA puede contribuir a posicionar nucleosomas, siempre y cuando se cumpla la asimetría presente entre ambas mitades en *S. pombe*.

Si una determinada composición de secuencia favorece el posicionamiento nucleosómico, esperaríamos que hubiese otras que tuviesen el efecto contrario. Está ampliamente documentado que las secuencias de poli (dA:dT) excluyen nucleosomas en *S. cerevisiae* (Lantermann et al., 2010; Segal & Widom 2009). Sin embargo, este no es el caso de *S. pombe*, donde esas secuencias se encuentran con la misma frecuencia en las NDRs que en el resto del genoma (García et al., 2017). Las secuencias que no posicionaban nucleosomas en la **Tabla 2**, tenían en común que además de no cumplir el desbalance de A y T en cada mitad, presentaban un incremento en el contenido de G y C. Con esta información como premisa, diseñamos la versión *ura4SP\_GC* con un contenido en G+C total del 65% frente al 43% de la secuencia original (**Tabla 5**). Como se muestra en la **Figura 15**, el incremento del contenido en G+C, provoca una pérdida del posicionamiento de los nucleosomas del gen *ura4*, además una disminución de la ocupación.

Estos resultados confirman que en *S. pombe*, secuencias con un alto contenido en G+C se incorporan con mayor dificultad en nucleosomas mientras que aquellas que presentan una asimetría en el contenido de A y T entre las dos mitades, favorecen el posicionamiento. Las secuencias que posicionan nucleosomas suelen poseer los dinucleótidos AA/TA/TT en el surco menor de la doble hélice si se encuentra orientado hacia el interior del octámero de histonas y los dinucleótidos GG/GC/CC cuando está orientado hacia el exterior (Struhl & Segal 2013; revisado en Singh & Mueller-Planitz 2021). Esta distribución de los dinucleótidos a lo largo del DNA nucleosómico está directamente relacionada con sus propiedades de rigidez y flexibilidad (Basu et al., 2020). Sin embargo, las secuencias enriquecidas en poli (dA:dT) y poli (dG:dC), confieren mayor resistencia a la torsión e impiden la formación del nucleosoma (Tillo et al., 2009; Basu et al., 2021; revisado en Singh & Mueller-Planitz 2021). Por lo tanto, al modificar la secuencia, estamos aumentando o disminuyendo su flexibilidad y con ello, alterando la capacidad de posicionar nucleosomas.

### **Contribución del posicionamiento nucleosómico a la transcripción**

Como se ha detallado en el apartado anterior, la asimetría en la composición del DNA nucleosómico podría conferir diferentes propiedades a la afinidad de la interacción entre el DNA y el octámero de histonas que, a su vez, podría afectar a la direccionalidad de la transcripción (García et al., 2024).

Durante la transcripción, la RNAPII se detiene en localizaciones específicas de los nucleosomas que coinciden con los sitios de mayor interacción entre el DNA y las histonas (Kujirai

& Kurumizaka, 2020). Esta pausa de la RNAPII permite generar un *loop* de DNA intranucleosómico que es importante para mantener la posición de los nucleosomas en la secuencia del DNA. La formación de este *loop* depende de la secuencia de DNA de la mitad distal, y su función es retener al nucleosoma para evitar su expulsión mientras la RNAPII transcribe la mitad proximal (Kujirai et al., 2020; Flipovski et al., 2022; Kulaeva et al., 2013; Kulaeva et al., 2009). En relación a estas observaciones, García et al. (2024) han descubierto que las dos mitades del DNA nucleosómico contribuyen de diferente manera a la estabilidad de los nucleosomas *in vivo* ya que en experimentos donde se ha modificado la mitad distal de los nucleosomas del gen *ura4* e *his7* de *S. pombe*, se pierde el posicionamiento de los nucleosomas y la expresión de los genes se reduce. Por ello, la asimetría en la secuencia de DNA nucleosómico tiene un papel importante sobre el posicionamiento de los nucleosomas y sobre el proceso de transcripción.

En línea con estas observaciones, hemos encontrado que las versiones modificadas *ura4SJ\_2mer* (**Figura 9C**), *ura4SP\_AT* y *ura4SJ\_AT* (**Figura 11**) presentan un posicionamiento nucleosómico similar al del gen *ura4* silvestre de *S. pombe* y niveles de expresión comparables (**Figura 17**). Sin embargo, aunque las versiones *ura4SJ\_2mer*<sup>-UTR</sup> y *ura4SJ\_2mer*<sup>-UTR.CU</sup> (**Figura 10**), tienen un posicionamiento nucleosómico similar a las anteriores, su nivel de expresión está muy reducido (**Figura 17**). Esto podría deberse a que las modificaciones implementadas en sus secuencias aumentan la afinidad entre el octámero de histonas y el DNA, lo que podría dificultar el paso de la RNAPII. Además, en las versiones *ura4SJ\_2mer*<sup>-UTR</sup> y *ura4SJ\_2mer*<sup>-UTR.CU</sup> se produce un aumento de los niveles de RNA antisentido de hasta cuatro veces respecto a los del *ura4* silvestre de *S. pombe*. En relación a ello, existen estudios que determinan que la transcripción en regiones promotoras es intrínsecamente bidireccional y que, dependiendo de los factores de transcripción, remodeladores de cromatina y otros complejos proteicos, ha evolucionado para dirigir la correcta dirección de la transcripción (Jin et al., 2017; Rondón et al., 2009). Cuando esta regulación no tiene lugar, se generan transcritos antisentido que alteran la expresión génica y las funciones celulares (Camblong et al., 2007; Gullerova & Proudfoot, 2008). Por ello, los nucleosomas de las versiones *ura4SJ\_2mer*<sup>-UTR</sup> y *ura4SJ\_2mer*<sup>-UTR.CU</sup> (**Figura 10**), que muestran unas posiciones más estables que el silvestre, podrían dificultar el paso de la RNAPII en la dirección sentido y ser más permisivas en la dirección antisentido, tal y como se observa en las **Figuras 17** y **18**.

En el análisis de posicionamiento de la versión *ura4SP\_AT\_mov21* mostrado en la **Figura 13A**, los nucleosomas están bien posicionados pero desplazados en el sentido de la transcripción. Además, muestra reducida su tasa de síntesis de mRNA (**Figura 17**), lo que podría estar relacionado, al igual que en las versiones anteriores, con una mayor unión del DNA por el

octámero de histonas en relación con la del *ura4* silvestre de *S. pombe*. Este descenso de los niveles de mRNA podría estar causado por el desplazamiento de la posición del nucleosoma +1 al introducir el desfase en la secuencia. Existen varias evidencias que apoyan la importancia de una determinada distancia entre la posición del nucleosoma +1 y el TSS para el correcto ensamblaje del PIC (*Preinitiation Complex*) y el inicio de la transcripción. Algunos autores han establecido la importancia de la posición de la diada del nucleosoma +1 a unas 60 pb en promedio, en posición 3' del TSS en levaduras (Albert et al., 2007; Nocetti & Whitehouse 2016). Además, esta posición es crítica para la unión de los factores encargados de desencadenar el inicio de la transcripción (Nocetti & Whitehouse 2016; Kubik et al., 2018). Estudios más recientes *in vivo*, han determinado que si la distancia entre el TSS y el borde del nucleosoma +1 es muy corta (10 pb) o muy larga (40 pb), el PIC no se activa y la transcripción no se produce. Por lo tanto, la distancia entre el TSS y el borde del nucleosoma +1 tiene que ser de unas 18 pb para facilitar el acceso al PIC (Abril-Garrido et al., 2023). En línea con estas evidencias, podría ser que en la versión *ura4SP\_AT\_mov21*, al presentar desplazados sus nucleosomas, entre ellos el nucleosoma +1, la distancia entre este y el TSS no sea la correcta para desencadenar un proceso de transcripción óptimo, lo que explicaría el descenso en sus niveles de mRNA (**Figura 17**).

Por otro lado, las observaciones descritas en el apartado anterior nos confirman que las secuencias de DNA que presentan un alto contenido en G+C, excluyen nucleosomas en *S. pombe*. Esta pérdida completa del paisaje nucleosómico, acompañada de un descenso de la ocupación de los nucleosomas en las versiones *ura4SJ\_1mer* (**Figura 9B**) y *ura4SP\_GC* (**Figura 15**), implica una reducción drástica de sus tasas de síntesis de mRNA (**Figura 17**). Este descenso de ocupación de los nucleosomas en estas versiones modificadas del gen *ura4*, podría dejar expuestos posibles promotores crípticos, ya que los sitios de unión de factores de transcripción se encuentran distribuidos por todo el genoma y una de las funciones de los nucleosomas es limitar su acceso (Lorch & Kornberg 2020). A partir de estas regiones se iniciaría la transcripción aberrante tal y como ocurre en ausencia de los remodeladores de cromatina Hrp1 y Hrp3 de *S. pombe* (Hennig et al., 2012; Pointner et al., 2012; Shim et al., 2012). Esto explicaría el disparo de los niveles de RNA antisentido en estas versiones (**Figura 18**). La regulación de los niveles de RNA antisentido es muy importante en un genoma tan compacto como el de *S. pombe* donde la densidad promedio es de un gen cada 2-2.5 Kb y donde muchos promotores son bidireccionales (Wery et al., 2018; Atkinson et al., 2018).

En *S. pombe*, el proceso de terminación de la transcripción no es tan eficiente ni está tan bien definido como en *S. cerevisiae* y, por tanto, la RNAPII tiende a avanzar más allá de la señal de poliadenilación antes de su liberación (Wery et al., 2018; Liu et al., 2017; Booth et al., 2016;

Mata, 2013; Gullerova & Proudfoot, 2008). En relación a ello, el 70% de los genes codificantes de proteínas en *S. pombe* presentan transcripción antisentido que correlaciona con un descenso en la transcripción sentido (Wery et al., 2018; Atkinson et al., 2018). Otra posibilidad es que la transcripción antisentido a lo largo del gen *ura4* en las versiones *ura4SJ\_1mer* (**Figura 9B**) y *ura4SP\_GC* (**Figura 15**) se derive de la terminación defectuosa del gen convergente *tam14*, causada por el descenso de la ocupación nucleosómica en la región 3' del gen *ura4* y que el antisentido que hemos detectado sea en realidad la continuación del mRNA del gen *tam14*.

Por lo tanto, podemos concluir que la función del posicionamiento de los nucleosomas, además de facilitar la compactación del genoma es, la regulación de la transcripción y como tal, favorecer el paso de la RNAPII en un sentido e impedirlo en el contrario.

## 2. LOS REMODELADORES DE CROMATINA EN LA ORGANIZACIÓN DE LOS NUCLEOSOMAS

Los organismos eucariotas poseen complejos proteicos conocidos como remodeladores de cromatina, los cuales, mediante la hidrólisis del ATP y en colaboración con otros factores, desempeñan un papel activo en la regulación y el empaquetamiento del DNA en el genoma (Misteli 2020). La organización fundamental del genoma en cromatina se lleva a cabo por un recambio continuo de nucleosomas mediado por los remodeladores de cromatina, polimerasas, topoisomerasas, histonas y modificaciones en el DNA (revisado en Eustermann et al., 2023). El correcto posicionamiento de los nucleosomas desempeña una función esencial en la regulación de la accesibilidad al DNA, influyendo así en actividades genómicas como la transcripción, la replicación y la reparación del DNA (revisado en Lin et al., 2020; Ocampo et al., 2019). Las tres familias de remodeladores (SWI/SNF, INO80 y CHD) de *S. pombe*, se agrupan en tres categorías según su función: *pushers*, *pullers* y *spacers* (Singh et al., 2021; Kubik et al., 2019). Es importante resaltar que *pushers* y *pullers* no mantienen las NDRs libres de nucleosomas ya que en su ausencia solo se produce un pequeño desplazamiento del nucleosoma +1 hacia la NDR.

Empleando la técnica de *degrón* mejorada para *S. pombe*, descrita por Zhang et al. (2022), hemos estudiado la dinámica de la cromatina y su efecto en la transcripción en ausencia de la subunidad catalítica de cada uno de los complejos remodeladores. Esta técnica nos ha permitido estudiar el efecto de la inhibición de la actividad catalítica de cada remodelador sin tener que recurrir a la delección de las subunidades. Este sistema permite, además, volver a expresar un remodelador funcional para estudiar la dinámica de reversión del fenotipo causado por su ausencia.

## Familia SWI/SNF

En primer lugar, estudiamos el papel de la familia de remodeladores SWI/SNF, formada por los complejos SWI/SNF y RSC, cuya función principal es mantener el posicionamiento del nucleosoma +1. En ausencia de estos remodeladores, se produce un desplazamiento este nucleosoma hacia la NDR, provocando una reducción de su tamaño (Kubik et al., 2019; Rawal et al., 2018), lo que afecta a la expresión de un gran número de genes en *S. cerevisiae* (Ocampo et al., 2019; Monahan et al., 2008; Wilson et al., 1996). Sin embargo, hasta la fecha, se desconocen sus implicaciones en *S. pombe* y, si estos remodeladores desempeñan funciones redundantes. Es importante señalar, que en *S. cerevisiae*, RSC es al menos 10 veces más abundante que el complejo SWI/SNF y, además, es esencial a diferencia de SWI/SNF que no lo es (Cairns et al., 1996), lo que podría explicar que RSC regule la expresión de muchos más genes que SWI/SNF (Ocampo et al., 2019; Rawal et al., 2018; Sudarsanam et al., 2000).

Como se muestra en la **Figura 19**, a los 30 minutos en presencia del inductor, no se detectan niveles de las proteínas Snf21 ni Snf22. Sin embargo, se reduce el tamaño del 82% de las de las NDRs en *snf21-sAID* (RSC) y del 34% en *snf22-sAID* (**Figura 20**), el 75% de las cuales también se reduce en ausencia de Snf21 (RSC). Estos resultados demuestran la importancia de los complejos RSC y SWI/SNF en el mantenimiento de la posición estable del nucleosoma +1 y, por lo tanto, del tamaño de las NDRs en *S. pombe*. En ausencia de Snf21 (RSC) y Snf22 (SWI/SNF), se produce un desplazamiento del nucleosoma +1 hacia la NDR siendo la reducción promedio causada por la inactivación de RSC de 17 nucleótidos y la de SWI/SNF de 8. Teniendo en cuenta que en la inmensa mayoría de los genes de *S. pombe* el TSS está localizado a unos 15-20 nucleótidos en el interior del DNA del nucleosoma +1, ese desplazamiento podría ocultar sitios de unión para algunos factores de transcripción y causar la subexpresión de los genes asociados a estos remodeladores, como se comentará más adelante.

Por otro lado, aunque no se observaron diferencias significativas en el tamaño de las NDRs afectadas por uno u otro complejo remodelador en los tiempos analizados para *S. pombe*, estudios recientes en *S. cerevisiae* indican que las NDRs más pequeñas ( $\approx 100$  pb) muestran menos impacto frente a la ausencia de RSC y SWI/SNF en comparación con las NDRs más grandes (120-200 pb) (Chen et al., 2022). Esto podría estar relacionado con el gran tamaño de estos complejos que requieren al menos 50 pb libres de nucleosomas para su unión al DNA (Chen et al., 2022; Wagner et al., 2020; Brahma & Henikoff 2019).

Otro factor a tener en cuenta es que el reclutamiento de estos complejos se produce de diferente manera. En el caso de RSC, depende de la secuencia del DNA (Krietenstein et al., 2016).

Se ha descrito que RSC desplaza los nucleosomas de las secuencias ricas en poli (dA:dT) que son muy abundantes en las regiones promotoras de *S. cerevisiae* (Rossi et al., 2021; Klein-Brill et al., 2019; Krietenstein et al., 2016; Yuan et al., 2005). Sin embargo, estos elementos no están sobrerrepresentados en las NDRs de *S. pombe* (García et al., 2017), por lo que es probable que el reclutamiento de RSC ocurra de manera diferente en esta levadura. En cambio, en *S. cerevisiae*, el complejo SWI/SNF actúa principalmente sobre genes con caja TATA (Woike et al., 2023; Kubik et al., 2019; Butryn et al., 2018). Nuestros resultados indican que en *S. pombe*, el 80% de las 876 NDRs que poseen caja TATA disminuyen su tamaño en ausencia de Snf22, sugiriendo que este remodelador tiene un impacto directo sobre los genes que tienen esa secuencia en sus promotores.

Además de la secuencia del DNA, existe una interacción física entre los remodeladores de cromatina y determinados factores de transcripción. En *S. cerevisiae*, la asociación entre RSC y los factores de transcripción Rap1 y Reb1 está muy bien establecida (Wu et al., 2018; Gavin et al., 2002). Aunque Rap1 no está conservado en *S. pombe*, los factores de transcripción Crf1 (Ilfh1 en *S. cerevisiae*) y Fhl1 actúan directamente en su lugar, interactuando entre sí y, además, Fhl1 se une al motivo conservado Homole (Maldonado et al., 2023). La identificación de este motivo y de sitios de unión para Reb1 en las NDRs que reducen su tamaño en ausencia de RSC sugiere que estos factores de transcripción podrían contribuir al reclutamiento de este remodelador en *S. pombe*.

Aunque los complejos RSC y SWI/SNF actúan específicamente sobre el nucleosoma +1, la ausencia de Snf21 afecta a un mayor número de genes que Snf22. No obstante, 191 genes subexpresados y 32 genes sobreexpresados en ausencia de Snf22, también muestran alterados sus niveles de expresión en ausencia de Snf21. Esto sugiere, una posible redundancia funcional entre estos complejos remodeladores, aunque cada uno desempeña funciones específicas ya que la ausencia de RSC no puede ser compensada por SWI/SNF (Kubik et al., 2019).

Es importante destacar que en ausencia de RSC, siete subunidades del complejo remodelador INO80, incluida su subunidad catalítica Ino80 (esencial para el funcionamiento de este complejo remodelador) (**Figura 25**), están entre los 1203 genes subexpresados. Este resultado sugiere que RSC podría actuar como un regulador de la actividad de INO80 y, por lo tanto, algunos genes podrían sobreexpresarse en ausencia de RSC como consecuencia de la pérdida de la función de INO80. Así, de los 290 genes que se sobreexpresan en ausencia de Snf21, 149 también lo hacen en ausencia de INO80. Es importante destacar que, en estos 149 genes, no se encuentran motivos de unión para Rap1 ni Reb1 pese a que se ha descrito su presencia en los genes regulados por RSC e INO80 en *S. cerevisiae*.

## Familia INO80

En segundo lugar, estudiamos la familia de remodeladores INO80 que desempeña funciones opuestas a los complejos RSC y SWI/SNF en la dinámica de las NDRs. En su ausencia, se produce un desplazamiento del nucleosoma +1 hacia el exterior de la NDR como se muestra en el ejemplo de la **Figura 28A** que podría hacer accesibles sitios adicionales de unión para factores de transcripción y desregular el nivel de expresión de los genes asociados.

En este punto, nos preguntamos qué características debían tener las NDRs dianas para INO80 en *S. pombe*. En *S. cerevisiae*, existen dos modelos diferentes que explican la interacción de INO80 con las NDRs. El primero de ellos, propone que la asociación del remodelador INO80 con las NDRs esta mediada por el factor de transcripción Reb1 (Oberbeckmann et al., 2021). En segundo lugar, Chen et al. (2022) determinaron que INO80 estaba asociado a Reb1 pero que el reclutamiento del remodelador era independiente de este factor. En cualquier caso, INO80 necesita un mínimo de 30 pb libres en la NDR para llevar a cabo el movimiento de los nucleosomas, siendo 70 pb el tamaño óptimo (Oberbeckmann et al., 2021; Udugama et al., 2011).

En nuestros resultados, a los tiempos analizados, no hemos encontrado enriquecimiento para el motivo de unión del factor Reb1 ni para factores pioneros, a diferencia de lo observado para el remodelador RSC. La ausencia de sitios de unión del factor Reb1 o de otras secuencias sobrerrepresentadas en las NDRs que aumentan su tamaño en ausencia de Ino80, sugiere que en *S. pombe*, a diferencia de *S. cerevisiae*, la interacción del remodelador con las NDRs podría ocurrir de forma directa, sin estar mediada por factores que lo reclutasen.

INO80, además de contribuir al mantenimiento de la posición del nucleosoma +1, se ha descrito como *spacer* en *S. cerevisiae* ya que modula el posicionamiento de nucleosomas a lo largo de algunos genes (Cheon et al., 2021; Singh et al., 2021; Kubik et al., 2019; Udugama et al., 2011). En *S. pombe*, INO80 también tiene esa función tal y como se muestra en la **Figura 27B** y en el ejemplo de la **Figura 28B**. En nuestros resultados mostramos como el 42% de los genes subexpresados y el 47% de los genes sobreexpresados muestran, además, un desplazamiento en el nucleosoma +1. Esto sugiere que el desplazamiento del nucleosoma +1 puede tanto activar como reprimir la expresión génica. El resto de los genes afectados no muestran desplazamiento en el nucleosoma +1, por lo que podrían no ser diana directa del remodelador, al menos en los tiempos analizados.

Entre los genes sobreexpresados en ausencia de Ino80 (**Figura 29**), destaca la categoría de RNAs no codificantes (ncRNAs). El 45% de estos, corresponden a transcritos antisentido de

genes en los que se produce un desplazamiento del nucleosoma +1 y una pérdida del posicionamiento de los nucleosomas a lo largo del gen, mientras que el 55% restante son transcritos antisentido en los que únicamente se produce una pérdida del posicionamiento de los nucleosomas a lo largo de los genes. Por tanto, estas dos funciones del remodelador INO80 conservadas en *S. pombe*, son independientes.

### Familia CHD

El remodelador INO80, como *spacer*, podría realizar funciones solapantes con los remodeladores de la familia CHD. En *S. cerevisiae* únicamente existe un único miembro de esta familia, llamado Chd1, que actúa como monómero (revisado en Eustermann et al., 2023). Sin embargo, en *S. pombe*, existen tres remodeladores de esta familia, Mit1, Hrp1 y Hrp3. Hrp1 no solo regula el espaciamiento de los nucleosomas (Pointner et al., 2012), sino que también es necesario para la condensación cromosómica (Yoo et al., 2000), la carga de CENP-A en el centrómero (Yamada et al., 2008; Walfridsson et al., 2005), y la terminación de la transcripción (Alén et al., 2002). Por otro lado, estudios con el mutante de delección *hrp3* han demostrado que su ausencia provoca una pérdida del posicionamiento y de ocupación nucleosómica que está asociada a un aumento de la transcripción críptica y antisentido (Dong et al., 2021; Touat-Todeschini et al., 2012; Shim et al., 2012; Hennig et al., 2012).

Empleando los mutantes condicionales *hrp1-sAID* y *hrp3-sAID*, hemos observado que la pérdida del posicionamiento de los nucleosomas ocurre inmediatamente después de la inactivación de los remodeladores y que esta se revierte cuando vuelven a estar presentes (**Figura 32**). El análisis cuantitativo de *borrosidad* indica que, en ausencia de Hrp1, se pierde el posicionamiento de aproximadamente el 20% de nucleosomas que ocupan posiciones centrales en los genes, mientras que la ausencia de Hrp3 afecta hasta al 50%, lo que sugiere que estos remodeladores actúan de manera diferente. Hrp1 parece ser esencial para mantener el espaciamiento adecuado de los nucleosomas a lo largo de los genes más allá del +1 (**Figura 33A**), mientras que Hrp3, aunque en menor medida, también afecta a la ocupación y posicionamiento de este nucleosoma (**Figura 33B y 35**).

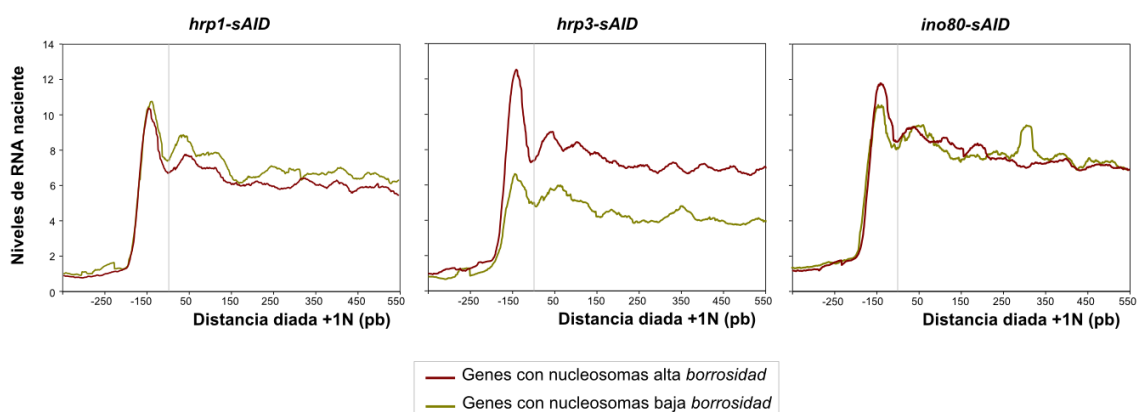
Cuando se analizaron los niveles de RNA en las cinéticas realizadas con ambas cepas, se observó que la ausencia de Hrp1 produce la desregulación de un número mucho menor de genes en comparación con la ausencia de Hrp3 (135 frente a 532). Además, los grupos de genes desregulados no son solapantes, como se deduce de los análisis de enriquecimiento funcional de la **Figuras 38**.

Por lo tanto, aunque Hrp1 y Hrp3 comparten funciones básicas en el mantenimiento del posicionamiento de los nucleosomas, también tienen roles específicos y diferenciados en la regulación génica y en la estructura de la cromatina en *S. pombe*.

Dado que no se observa solapamiento entre los genes con expresión desregulada en ausencia de Hrp1 y Hrp3, se analizó si existía alguna relación entre los genes afectados por ausencia de Hrp3 y los afectados por la ausencia de Ino80 en relación a su función espaciadora y que, además, presentasen cambios en sus niveles de expresión. Sin embargo, tampoco encontramos solapamiento entre estos grupos de genes, lo que sugiere que cada uno de los tres remodeladores con actividad espaciadora (Ino80, Hrp1 y Hrp3) tiene como diana grupos de genes diferentes.

Además, cuando analizamos la tasa de transcripción asociada a genes cuyo posicionamiento nucleosómico está comprometido en ausencia de cada uno de estos remodeladores (**Figura 40**), observamos que los nucleosomas más sensibles a la ausencia de Hrp3 pertenecen a genes con una tasa de transcripción elevada, mientras que esto no sucede para Hrp1 ni para Ino80.

En este mismo sentido, en el mutante *chd1Δ* de *S. cerevisiae* se produce una desorganización más severa de la cromatina en aproximadamente el 10% de los genes con mayor tasa de transcripción (Yen et al., 2012), hecho relacionado con que Chd1 está asociado con el complejo de elongación de la RNAPII (Kujirai & Kurumizaka 2020; revisado en Prajapati et al., 2020). La función de Chd1 en *S. cerevisiae*, podría ser equivalente a la función realizada por Hrp3 en *S. pombe*. Es posible que el papel de Hrp1 se limite a unos pocos genes, y que su función principal sea la carga de CENP-A en los centrómeros (Walfridsson et al., 2005; Yamada et al., 2008). Por otro lado, el papel de Ino80 como espaciador podría estar determinado por otras características en la secuencia de DNA o por la interacción de otros factores, que sería interesante estudiar en un futuro.



**Figura 40. Tasa de transcripción en genes con nucleosomas mal posicionados en los mutantes condicionales *hrp1-sAID*, *hrp3-sAID* e *ino80-sAID*.** Los gráficos representan los niveles promedio de RNA nascente medido por *NET-seq* (Wery et al., 2018) en 800 genes con nucleosomas de alta (rojo) y baja (verde) *borrosidad* de los mutantes condicionales *hrp1-sAID*, *hrp3-sAID* e *ino80-sAID* a las 7, 4 y 2 horas respectivamente, en presencia del inductor. Los datos se alinearon a la diada de los nucleosomas +1.

Por lo tanto, en *S. cerevisiae*, se ha establecido que RSC e INO80 actúan como barreras para mantener las NDRs y localizar al nucleosoma +1 en una posición estable. Este nucleosoma actúa como referencia para que ISW1 y CHD1 espacien el resto de los nucleosomas (Chereji & Clark 2018; Ocampo et al., 2016). Sin embargo, *S. pombe* carece de remodeladores ISW1 por lo que esa función espaciadora podría ser asumida por INO80, que además es esencial en *S. pombe*, a diferencia de *S. cerevisiae* donde no lo es. Este hecho, destaca la adaptación y diversificación de los mecanismos de remodelación de la cromatina entre especies, donde la ausencia de ciertos remodeladores en una especie puede ser compensada en otra por la presencia y la adaptación funcional de otros remodeladores.

# CONCLUSIONES





Las conclusiones de este trabajo son:

1. El contenido promedio en adenina y timina en las mitades proximal y distal del DNA nucleosómico es diferente. Esta composición asimétrica contiene la información necesaria y suficiente para posicionar nucleosomas *in vivo* en *Schizosaccharomyces pombe*.
2. Una de las funciones del posicionamiento nucleosómico es regular la direccionalidad de la transcripción porque la alteración del posicionamiento y el descenso de la ocupación nucleosómica a lo largo de los genes favorece la transcripción antisentido.
3. Los dos complejos remodeladores de cromatina *pushers* en *S. pombe*, RSC y SWI/SNF, son solo parcialmente redundantes, ya que la ausencia de función de RSC tiene un mayor impacto sobre el tamaño de las NDRs y causa la desregulación transcripcional de un mayor número de genes que la de SWI/SNF.
4. El remodelador RSC actúa como un regulador transcripcional de INO80 modulando la expresión de los genes que codifican 7 de sus 14 subunidades.
5. El complejo remodelador de cromatina INO80 puede funcionar independientemente como *puller* y *spacer* en *S. pombe*.
6. Los remodeladores de la familia CHD, Hrp1 y Hrp3 regulan el espaciamiento nucleosómico en grupos de genes diferentes con poco solapamiento entre ellos. Además, Hrp3 regula preferentemente genes con tasas de transcripción más elevadas y contribuye al posicionamiento del nucleosoma +1 independientemente de los nucleosomas a lo largo de los genes.



# **MATERIALES Y MÉTODOS**





## 1. MICROORGANISMOS EMPLEADOS

Para desarrollar este trabajo, se han construido diferentes cepas de *S. pombe* a partir de la cepa *h-972*. Las cepas que contienen diferentes modificaciones en la ORF del gen *ura4*, se han construido empleando el mutante *h- leu1-32 ura4 DS/E* que presenta unas 300 pb delecionadas dentro de la ORF del gen *ura4*. Las cepas de *S. pombe* en las que se ha integrado una secuencia artificial no codificante, se han construido a partir del mutante *h- ura4-d18*, que posee una deleción completa del gen *ura4*. Por otro lado, se ha empleado la cepa DH5 $\alpha$  de *Escherichia coli* (*F- endA1 gyr96 hsdR17  $\Delta$ lacU169 f80lacZ $\Delta$ M15 recA1 relA1 supE44-thi-1*) para la clonación y amplificación de plásmidos.

## 2. MEDIOS DE CULTIVO

Las cepas de *S. pombe* empleadas se cultivaron en medio rico (YES) o en medio mínimo (MM) suplementado con 225 mg/L de los aminoácidos necesarios a la temperatura de 32°C. *E. coli* se cultivó en medio rico LB (0,5% extracto de levadura, 1% bacto triptona y 1% NaCl) (Bertani 1951) a 37°C. La composición de los medios de cultivo de *S. pombe* que se han empleado en este trabajo ha sido descrita por Moreno et al. (1991). El crecimiento en medio líquido se determinó midiendo la turbidez del cultivo por su absorbancia o densidad óptica a 595 nm (DO<sub>595</sub>).

## 3. GENERACIÓN DE SECUENCIAS QUE POSICIONAN NUCLEOSOMAS

### Generación de las matrices de puntuación

Para extraer la información contenida en las *signatures* nucleosómicas, en el trabajo de González et al. 2016, se generaron "Matrices de puntuación de posición específica" (PSWM: *Position-Specific Weight Matrix*) que incorporaban la frecuencia de cada dinucleótido a lo largo de 38.154 secuencias de DNA mononucleosómico de *S. pombe* obtenidas por secuenciación mediante el protocolo de *single read*. En este protocolo, cada fragmento de DNA mononucleosómico se secuencia únicamente por uno de sus extremos, lo que supone que la definición de la posición de las diadas nucleosómicas es menos precisa que cuando se secuencian estos fragmentos mediante el protocolo de *paired end*, en el que ambos extremos se leen. En este trabajo, hemos empleado el protocolo de *paired end*, y obtenido 49.998 nucleosomas bien posicionados con los que hemos generado tres nuevas matrices de puntuación, en las que se calculó la frecuencia promedio de mono, di, y trinucleótidos a lo largo del DNA mononucleosómico. Por tanto, las dimensiones de estas tres PSWM fueron de 4 (mononucleótidos), 16 (dinucleótidos) y 64 (trinucleótidos) por 150 pb del DNA

mononucleosómico. Los valores para cada posición (i) y para cada mono, di o trinucleótido (N) se calcularon de acuerdo con la siguiente expresión:

$$\text{Puntuación}_i^N = \log_2 \frac{\text{Freq}_i^N}{\text{Freq}_{\text{genómica}}^N}$$

donde  $\text{Freq}_i^N$  es la frecuencia de mono, di o trinucleótido N en la posición i de todas las secuencias alineadas y  $\text{Freq}_{\text{genómica}}^N$  es la frecuencia promedio del mismo mono, di o trinucleótido a lo largo de todo el genoma.

Con esta información, las matrices permitían asignar a secuencias individuales una puntuación numérica, que sería mayor cuanto más se aproximase su composición de bases al promedio en cuestión.

### Sesgo de uso de codones

Para el diseño de la versión *ura4SJ\_2mer*<sup>UTR.CU</sup> de este trabajo, se ha empleado el sesgo de uso de codones descrito para *S. pombe*. Para ello, se incorporó una penalización para la elección de codones de uso menos frecuentes. El uso de codones de *S. pombe* fue extraído de Nakamura et al. (2000) y se encuentra en el repositorio de <http://www.kazusa.or.jp/codon/>. Para incorporar la penalización, se ha multiplicado la frecuencia de cada dinucleótido de la matriz empleada por la frecuencia de aparición de ese codón según la siguiente expresión:

$$S_{ABC/X}^i = (S_{AB}^i + S_{BC}^{i+1})U_{ABC} + S_{CX}^{i+2} \frac{U_{ABC} \cdot U_{XYZ}}{2}$$

donde  $S_{ABC/X}^i$  es la puntuación para el codón ABC en la posición i del nucleosoma, XYZ es el codón que aparece a continuación en la secuencia,  $S_{AB}^i$  se refiere a la puntuación del dinucleótido AB en la posición i del nucleosoma,  $U_{ABC}$  se refiere a la frecuencia genómica del codón ABC y,  $U_{XYZ}$  se refiere a la frecuencia genómica del codón XYZ.

### Síntesis de las moléculas de DNA

Las moléculas de DNA sintético para las versiones *ura4SJ\_1mer*, *ura4SJ\_2mer*, *ura4SJ\_3mer*, *ura4SJ\_2mer*<sup>UTR</sup> y *ura4SJ\_2mer*<sup>UTR.CU</sup> fueron encargadas a GENEWIZ® o GeneArt™ (Life Technologies). Estas empresas envían las secuencias clonadas en plásmidos que se amplificaron en *E. coli*. El fragmento de interés se liberó del plásmido mediante digestión con enzimas de restricción y se empleó para transformar la levadura. Por otro lado, las moléculas de DNA sintético para las versiones *ura4SP\_2mer*, *ura4SP\_AT*, *ura4SJ\_AT*, *ura4SP\_GC*, *ura4SP\_AT\_mov21*, *mug166SP\_AT\_mov21*, *sec.artificial\_control* y *sec.artificial\_AT* fueron encargadas a IDT (Integrated

DNA Technologies). Cuando fue necesario ligar estos fragmentos a un marcador de resistencia a antibiótico, se clonaron en un plásmido donde posteriormente se introdujo el marcador. Los fragmentos se amplificaron por PCR con oligonucleótidos específicos y se emplearon para transformar la levadura. En el diseño de todas las secuencias se incluyeron los sitios de recombinación y de restricción necesarios. Todas las secuencias de DNA diseñadas en este trabajo, junto con las coordenadas genómicas en las que han sido integradas, se encuentran recogidas en el ANEXO II.

#### 4. TÉCNICAS PARA LA CONSTRUCCION DE LAS CEPAS

##### **Sistema *degrón* inducible por auxinas**

Las cepas para generar mutantes condicionales mediante el sistema de *degrón* inducible por auxinas para los genes de las subunidades catalíticas de los distintos complejos remodeladores de cromatina (*snf21*, *snf22*, *ino80*, *hrp1* y *hrp3*) se han construido siguiendo el método de Zhang et al. (2022). Estas cepas contienen el gen de interés etiquetado en su extremo C-terminal con el epítipo 3xsAID ligado al marcador de resistencia a geneticina y, además, un módulo formado por el gen *ura4* unido al receptor de auxinas modificado (OsTIR1-F74A) integrado en el locus endógeno del gen *ura4*. La molécula inductora empleada en este sistema es un análogo de la auxina o ácido indolacético (IAA) denominado 5-adamantyl-IAA (Sigma). En presencia de este compuesto, las proteínas etiquetadas se marcan para su degradación por el proteasoma. La comprobación final de los transformantes se llevó a cabo mediante PCR, secuenciación automática y *Western Blot*.

##### **Transformación de microorganismos**

Para la transformación de *S. pombe* se utilizó la técnica de electroporación descrita en Forsburg & Rhind 2006. En los casos en los que se integró el gen *ura4* modificado, los transformantes se seleccionaron en placas de MM. En los casos en los que se utilizó el gen marcador *KanMX4*, las células se recuperaron en placas de YES durante 24 h y posteriormente, para seleccionar los transformantes se hicieron réplicas en placas de YES con 100 mg/ml de sulfato de geneticina (G418; Gibco).

La transformación de *E. coli* se realizó según el protocolo descrito por Kushner (1978).

## 5. TÉCNICAS DE AISLAMIENTO Y MANIPULACIÓN DEL DNA

### Obtención de plásmido de DNA de *E. coli*

Para aislar el plásmido de DNA de *E. coli* a pequeña escala se utilizó el método de lisis alcalina (Birnboim & Doly 1979), mientras que para la obtención a gran escala se empleó el kit *NucleoSpin Plasmid* de *Macherey-Nagel*, siguiendo el protocolo de la casa comercial.

### Obtención de DNA mononucleosómico de *S. pombe*

Se utilizaron 200 ml de cultivo de *S. pombe* en YES a  $DO_{595nm}$  de 0.8. Las células se trataron con formaldehído al 0.5% y se incubaron a 25°C durante 20 min con agitación suave (60-80 rpm). Posteriormente, se inactivó el formaldehído con glicina hasta una concentración de 125 mM e incubando a 25°C durante 10 min con agitación continua (100-110 rpm).

Una vez transcurrido este tiempo, el cultivo se transfirió a tubos estériles de 50 ml que se centrifugaron a 3.500 rpm durante 5 min a 20°C. Las células sedimentadas se lavaron con agua destilada estéril. A continuación, se resuspendieron en 10 ml de solución de preincubación (ácido cítrico monohidrato 20 mM,  $Na_2HPO_4$  20 mM y EDTA pH 8 40 mM) con 21  $\mu$ l de  $\beta$ -mercaptoetanol 14.3 M (Sigma Aldrich) recién añadido y se incubaron durante 10 min a 30°C en un baño con agua en agitación (200 rpm). A continuación, se centrifugaron las células durante 5 min a 4°C y se resuspendieron en 5 ml de tampón Sorbitol/Tris (sorbitol 0.5M y Tris/HCl 50 mM pH 7.5) con  $\beta$ -mercaptoetanol 10 mM recién añadido. Seguidamente, se añadieron 8 mg de Zimoliasa 20 T (Amsbio) y se incubaron durante 25 min a 30°C en un baño de agua en agitación a 200 rpm (la cantidad de Zimoliasa así como el tiempo de incubación puede variar en función de la cepa o las condiciones fisiológicas). A continuación, se recogieron los protoplastos mediante centrifugación a 3.500 rpm durante 5 min y a 4°C y se lavaron con 20 ml de tampón Sorbitol/Tris sin  $\beta$ -mercaptoetanol.

Tras volver a centrifugar en las mismas condiciones, se resuspendieron en 3 ml de tampón NP (sorbitol 1M, NaCl 50 mM, Tris/HCl 10 mM pH 7.5,  $MgCl_2$  5 mM,  $CaCl_2$  1mM,  $\beta$ -mercaptoetanol 1 mM, espermidina 0.5 mM y NP-40 0.07%). A continuación, se dividió la muestra en cuatro alícuotas de 750  $\mu$ l, se añadió nucleasa micrococcal (*MNase*, Thermo) a cada una de ellas a concentraciones finales de 0, 60, 120 y 200 U/ml y se incubaron durante 30 min a 37°C en un baño de agua sin agitación. Pasado este tiempo, se paró la reacción con 100  $\mu$ l de la solución de parada (SDS 5% y EDTA 50 mM) y se incubaron todas las alícuotas con 0.5 mg/ml de proteinasa K (Sigma) a 65°C durante 12-14 h.

El DNA de todas la alícuotas se purificó con fenol:cloroformo:isoamílico (25:24:1) y se precipitó añadiendo 1 µg de glicógeno, 1/25 volúmenes de NaCl 5 M y 0.7 volúmenes de isopropanol 100% e incubando durante 1 h a -20°C y centrifugando durante 45 min a 4°C. Finalmente, se eliminó el sobrenadante, se lavó el precipitado con etanol 70%, se resuspendió en 50 µl de TE 1X y se incubó con 20 µg de RNAsa A 20 mg/ml durante 2-3 h a 37°C.

Tras la extracción del DNA, se realizó una electroforesis en gel de agarosa al 1.5% en TAE y se corrió hasta que las bandas de DNA mononucleosómico quedaron bien separadas del resto de fragmentos. Después, se cortaron las bandas correspondientes al DNA mononucleosómico de la digestión con *MNase* siempre que se hubiera alcanzado una proporción de 80-90% de mononucleosoma con respecto a dinucleosoma. Por último, se extrajo el DNA utilizando las columnas *Freeze 'N Squeeze DNA Gel Extraction Spin Columns* (BioRad).

Para generar las librerías necesarias para la secuenciación se empleó el kit *NEXTflex™ Rapid DNA-seq Kit 2.0* (BioScientific).

### **Obtención de mapas de nucleosomas mediante secuenciación masiva**

Las muestras de DNA mononucleosómico se secuenciaron masivamente empleando el equipo *NextSeq 500* (Illumina) utilizando el protocolo de *paired-end*. Las lecturas resultantes se alinearon empleando la herramienta Bowtie (Langmead et al., 2009) frente al genoma de *S. pombe* (ASM294v2.20). Posteriormente, para representar los datos obtenidos, se empleó la herramienta NUCwave descrita en Quintales et al. (2015b). Además, para obtener una métrica que nos permitiese valorar el posicionamiento de los nucleosomas, empleamos la herramienta bioinformática DANPOS (Chen et al., 2013), en concreto la función *dpos*, que permite comparar el posicionamiento de nucleosomas individuales entre diferentes mapas. Se empleó el valor de *fuzziness* o *borrosidad*, que hace referencia al grado en que cada lectura asociada a un nucleosoma concreto se desvía de su posición preferente. Previamente, en nuestro laboratorio, se habían calibrado las diferencias de *borrosidad* entre duplicados biológicos para establecer el valor umbral a partir del cual las diferencias se podrían considerar significativas. Así, cuando la diferencia para cada nucleosoma concreto se encuentra por encima de 2 desviaciones estándar de la media de la diferencia entre dos mapas de genoma completo diferentes, consideramos que dicho nucleosoma no se encuentra bien posicionado. Este parámetro de *borrosidad*, medido en  $\log_2FC$  (*Fold Change*), nos permite comparar el posicionamiento de cada una de nuestras construcciones con su versión silvestre.

Además, se ha empleado otro valor proporcionado por DANPOS, llamado *fuzziness score* o *puntuación de la borrosidad* que es el valor de *borrosidad* absoluto, expresado en unidades arbitrarias. Se ha obtenido este valor para cada uno de los nucleosomas en los que se empaqueta cada gen en estudio y se ha realizado un promedio, este valor numérico promedio se relaciona directamente con la calidad del posicionamiento para cada una de las versiones de este trabajo. Así, cuanto estos valores absolutos de borrosidad son superiores al valor de la versión empleada como control en cada caso, indica que sus nucleosomas se encuentran peor posicionados que los de la referencia.

## PCR

Para llevar a cabo las PCRs que se han realizado de manera rutinaria en el laboratorio, se han empleado dos polimerasas. Para la amplificación con alta fidelidad se ha empleado *Phusion High-Fidelity DNA Polymerase* (Thermo Scientific™), mientras que, para la amplificación desde DNA genómico, cDNA o de colonias de levadura se ha empleado la *Phire Green Hot Start II DNA Polymerase* (Thermo Scientific™). Ambas enzimas se han utilizado siguiendo las recomendaciones del fabricante.

Las reacciones de amplificación se han realizado en un termociclador T Personal (Biometra). El número de ciclos de extensión, así como la temperatura de anillamiento se fue variando, dependiendo del tamaño del fragmento a amplificar.

## PCR cuantitativa (qPCR o RT-PCR)

Las qPCRs se realizaron con el equipo *CFX96 Real-time System* (Bio-Rad) y se utilizaron los reactivos del paquete comercial SYBR® Premix Ex Taq™ (Takara), siguiendo las instrucciones del fabricante y los requerimientos específicos del termociclador. Para el análisis de datos se utilizó el programa Bio-Rad CFX Manager 2.0 (Bio-Rad).

## 6. TÉCNICAS DE AISLAMIENTO Y MANIPULACIÓN DE RNA

### Obtención de RNA total

Se recogieron 5 ml de cultivo de *S. pombe* en fase exponencial, lavándose las células sedimentadas con agua tratada con DEPC al 0.2%. A continuación, se extrajo el RNA con el kit comercial *NZY Total RNA Isolation Kit* (NZYTech) que incluye el tratamiento con *DNase I* (NZYTech) de las muestras, siguiendo las instrucciones del fabricante.

## Transcripción reversa o retrotranscripción (RT)

Para la síntesis de cDNA, se utilizaron 200 ng de RNA para la reacción de transcripción reversa con el kit comercial *Biotoools High Retrotranscriptase-Starter Kit* (Biotoools), siguiendo las recomendaciones del fabricante. Los niveles de cDNA se analizaron a continuación por qPCR.

## 7. TÉCNICAS DE OBTENCIÓN Y MANIPULACIÓN DE PROTEINAS

### Obtención de extractos proteicos

Para llevar a cabo la extracción de proteínas, las muestras fueron procesadas utilizando ácido tricloroacético (TCA). Se recogieron 10 ml de cultivo a  $DO_{595nm}$  de 0.8 - 1, a los que se añadió TCA frío a una concentración final del 10%. Las células se recogieron mediante centrifugación a 13.200 rpm durante 1 min a 4°C. Una vez retirado el sobrenadante, las células sedimentadas se resuspendieron en 1 ml de TCA frío al 20% y se volvieron a centrifugar. En este punto, la muestra puede congelarse a -80°C para su posterior procesamiento. Para la rotura celular, las células se resuspendieron en 100  $\mu$ l de TCA frío al 12.5%, se añadió un volumen equivalente de *glass beads* de 450  $\mu$ m de diámetro y se procesaron en una FastPrep aplicando 3 ciclos de 30 seg a velocidad 5.5. Seguidamente, se llevó a cabo una centrifugación de 20 min a 13.200 rpm y a 4°C con el fin de separar los extractos proteicos de los restos celulares. El precipitado obtenido se resuspendió en 1 ml de acetona fría (PanReac) y se disgregó empleando palillos redondos largos. Una vez que se obtuvo una mezcla homogénea, la muestra se centrifugó durante 5 min a 13.200 rpm a 4°C y el precipitado obtenido se dejó secar a 55°C durante 10 min. Finalmente, el precipitado seco se resuspendió en 75  $\mu$ l de un tampón alquilante (Tris/HCl 0.1 M pH 8, ETDA 1 mM y SDS 1%) y 25  $\mu$ l de tampón de carga para SDS-PAGE 4X.

### Western Blot (WB)

Los extractos proteicos obtenidos, previamente desnaturalizados, se cargaron en geles de poliacrilamida en presencia de SDS (SDS-PAGE) (Laemmli 1970) al 6-8% dependiendo del tamaño de la proteína de interés. La electroforesis se realizó a voltaje constante (100-120 V) en el tampón de carrera (Tris-base 25 mM, Glicina 250 mM y SDS 0.1%). Después de la separación, las proteínas se transfirieron a membranas de PVDF (Hybond-P, GE Healthcare) con el sistema *mini-Trans-Blot*® (Bio-Rad), aplicando un amperaje constante de 400 mA durante 2 h en solución de transferencia (50 mM Tris y 250 mM glicina).

Para el bloqueo de la membrana tras la transferencia, se utilizó una solución que contenía leche en polvo al 5% (p/v) y Tween®20 al 0.05% (v/v) en TBS, incubándose durante 1 h en

agitación constante (80 rpm). El anticuerpo primario se diluyó en una solución con leche en polvo al 1% (p/v) y Tween® 20 al 0.05% (v/v) en TBS. Las membranas se incubaron a 4°C con el anticuerpo primario durante toda la noche: anti-sAID (MBL, AB\_2890014) a una dilución 1:2000 o anti-actina (MP Biomedicals, AB\_2223041) a una dilución 1:5000. Al día siguiente, se realizaron tres lavados sucesivos de las membranas con Tween® 20 al 0.05% (v/v) en TBS durante 10 min. Después, se incubaron durante 1 h a temperatura ambiente con el anticuerpo secundario anti-mouse (GE Healthcare, AB\_772210) diluido 1:10000 en una solución similar a la empleada para diluir el anticuerpo primario. Seguidamente, se realizaron otros tres lavados de las membranas de la misma manera que los anteriores. Pasado este tiempo, se reveló la membrana usando *Pierce™ ECL Plus Western Blotting Substrate* (Thermo Scientific™) siguiendo las instrucciones del fabricante. Para el revelado se emplearon películas *Medical X-Ray Film Blue* (Agfa).

## 8. ESTUDIO DE ESTABILIDADES Y VIDAS MEDIAS DE mRNAs

Para calcular la vida media de las cepas con las diferentes versiones del gen *ura4*, se procedió a eliminar la glucosa del medio de cultivo tal y como describieron Kim et al. (2002). Para ello, se partió de un cultivo de 30 ml a  $DO_{595nm}$  de 0.8 crecido en MM a 32°C. Inicialmente, se recogió una muestra que correspondía al tiempo 0 h. A continuación, el cultivo se centrifugó, se resuspendió en MM sin glucosa y se tomaron muestras a diferentes tiempos (15, 30, 60, 90 y 120 min).

Posteriormente, se extrajo el RNA y se sintetizó el cDNA correspondiente según se explica en el apartado 6. Los niveles de cDNA se analizaron por qPCR utilizando los niveles de mRNA del gen *act1* como normalizador. Los niveles de RNA correspondientes a cada muestra sin glucosa se normalizaron al tiempo control con glucosa (0 h).

La vida media de los mensajeros de interés se calculó representando el  $\log_2$  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:

$$\frac{t1}{2} = \frac{1}{-pte}$$

## 9. ANÁLISIS DE TRANSCRIPCIÓN MEDIANTE *RNA-seq*

Para generar las librerías necesarias para la secuenciación de las muestras de *RNA-seq*, se empleó el kit *TrueSeq Stranded mRNA Librarie* (Illumina). Las muestras se secuenciaron masivamente en un equipo *NextSeq 500* (Illumina) utilizando el protocolo de *paired-end*.

Para el análisis de los datos, empleamos la herramienta bioinformática *RANa-seq* (Prieto & Barrio 2020), que alinea las lecturas con HISAT2, y analiza la expresión diferencial con DESeq2. Se seleccionaron como umbrales de significación un cambio mínimo absoluto representado por el logaritmo en base 2 ( $\log_2FC$ ) de 0.6 y un p-valor de 0.05. Esto significa que, para considerar un cambio como significativo, el logaritmo en base 2 del cambio debe ser al menos 0.6 en valor absoluto, y el p-valor asociado a la prueba estadística debe ser menor o igual a 0.05. Estos umbrales ayudan a establecer criterios rigurosos para determinar la significancia de los resultados en el contexto del análisis.

## 10. HERRAMIENTAS BIOINFORMÁTICAS

Los perfiles agregados y los mapas de calor que se muestran en el trabajo han sido pintados empleando las herramientas *computeMatrix* y *plotHeatmap* de Deeptools (Ramírez et al., 2016). Los diagramas de volcán se han obtenido con el programa VolcaNoseR (Goedhart et al., 2020). El resto de los gráficos mostrados han sido generados con la herramienta <https://www.bioinformatics.com.cn/srplot>.

Para la remasterización de secuencias de DNA a partir de matrices de puntuación, se ha empleado el programa escrito en Python *remaster.py* (<https://github.com/rodrigoSantamaria/remaster>).

La herramienta BEDTools (Quinlan & Hall 2010) ha sido empleada en este trabajo para manipular los archivos de coordenadas genómicas tipo *bed* con un gran volumen de datos.



# **BIBLIOGRAFÍA**





- Abelson, H. T., Johnson, L. F., Penman, S., & Green, H. (1974). Changes in RNA in Relation to Growth of the Fibroblast: II. The Lifetime of mRNA, rRNA, and tRNA in Resting and Growing Cells. *Cell* 1(4), 161-165.
- Abril-Garrido, J., Dienemann, C., Grabbe, F., Velychko, T., Lidschreiber, M., Wang, H., & Cramer, P. (2023). Structural basis of transcription reduction by a promoter-proximal +1 nucleosome. *Molecular Cell* 83(11), 1798-1809.
- Albert, I., Mavrich, T. N., Tomsho, L. P., Qi, J., Zanton, S. J., Schuster, S. C., & Pugh, B. F. (2007). Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature* 446(7135), 572–576.
- Alén, C., Kent, N. A., Jones, H. S., O'sullivan, J., Aranda, A., & Proudfoot, N. J. (2002). A Role for Chromatin Remodeling in Transcriptional Termination by RNA Polymerase II. *Molecular Cell* 10, 1441–1452.
- Allshire, R. C., & Ekwall, K. (2015). Epigenetic regulation of chromatin states in *Schizosaccharomyces pombe*. *Cold Spring Harbor Perspectives in Biology* 7(7), 1–25.
- Alves-Rodrigues, I., Ferreira, P. G., Moldón, A., Vivancos, A. P., Hidalgo, E., Guigó, R., & Ayté, J. (2016). Spatiotemporal Control of Forkhead Binding to DNA Regulates the Meiotic Gene Expression Program. *Cell Reports* 14(4), 885–895.
- Atkinson, S. R., Marguerat, S., Bitton, D. A., Rodríguez-López, M., Rallis, C., Lemay, J.-F., Cotobal, C., Malecki, M., Smialowski, P., Mata, J., Korber, P., Bachand, F., & Bähler, J. (2018). Long noncoding RNA repertoire and targeting by nuclear exosome, cytoplasmic exonuclease, and RNAi in fission yeast. *RNA* 24(9), 1195–1213.
- Avsec, Ž., Weilert, M., Shrikumar, A., Krueger, S., Alexandari, A., Dalal, K., Fropf, R., McAnany, C., Gagneur, J., Kundaje, A., & Zeitlinger, J. (2021). Base-resolution models of transcription-factor binding reveal soft motif syntax. *Nature Genetics* 53(3), 354–366.
- Badis, G., Chan, E. T., van Bakel, H., Pena-Castillo, L., Tillo, D., Tsui, K., Carlson, C. D., Gossett, A. J., Hasinoff, M. J., Warren, C. L., Gebbia, M., Talukder, S., Yang, A., Mnaimneh, S., Terterov, D., Coburn, D., Li Yeo, A., Yeo, Z. X., Clarke, N. D., ... Hughes, T. R. (2008). A Library of Yeast Transcription Factor Motifs Reveals a Widespread Function for Rsc3 in Targeting Nucleosome Exclusion at Promoters. *Molecular Cell* 32(6), 878–887.

- Bailey, T. L., Johnson, J., Grant, C. E., Noble, W. S. (2015). The MEME Suite. *Nucleic Acids Research*, 43, 39–49.
- Baldi, S., Krebs, S., Blum, H., & Becker, P. B. (2018). Genome-wide measurement of local nucleosome array regularity and spacing by nanopore sequencing. *Nature Structural and Molecular Biology* 25(9), 894–901.
- Barozzi, I., Simonatto, M., Bonifacio, S., Yang, L., Rohs, R., Ghisletti, S., & Natoli, G. (2014). Coregulation of Transcription Factor Binding and Nucleosome Occupancy through DNA Features of Mammalian Enhancers. *Molecular Cell* 54(5), 844–857.
- Basu, A., Bobrovnikov, D. G., Cieza, B., Arcon, J. P., Qureshi, Z., Orozco, M., & Ha, T. (2022). Deciphering the mechanical code of the genome and epigenome. *Nature Structural and Molecular Biology* 29(12), 1178–1187.
- Basu, A., Bobrovnikov, D. G., Qureshi, Z., Kayikcioglu, T., Ngo, T. T. M., Ranjan, A., Eustermann, S., Cieza, B., Morgan, M. T., Hejna, M., Rube, H. T., Hopfner, K. P., Wolberger, C., Song, J. S., & Ha, T. (2021a). Measuring DNA mechanics on the genome scale. *Nature* 589(7842), 462–467.
- Bertani, G. (1951). Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J Bacteriol* 62(3), 293–300.
- Basu, A., Bobrovnikov, D. G., & Ha, T. (2021b). DNA mechanics and its biological impact. *Journal of Molecular Biology* 433(6): 166861.
- Birnboim, H. C., & Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* 7, 1513–1523.
- Bitton, D. A., Schubert, F., Dey, S., Okoniewski, M., Smith, G. C., Khadayate, S., Pancaldi, V., Wood, V., & Bähler, J. (2015). AnGeli: A tool for the analysis of gene lists from fission yeast. *Frontiers in Genetics* 6: 330.
- Booth, G. T., Wang, I. X., Cheung, V. G., & Lis, J. T. (2016). Divergence of a conserved elongation factor and transcription regulation in budding and fission yeast. *Genome Research* 26(6), 799–811.
- Brahma, S., & Henikoff, S. (2019). RSC-Associated Subnucleosomes Define MNase-Sensitive Promoters in Yeast. *Molecular Cell* 73(2), 238–249.

- Brogaard, K., Xi, L., Wang, J. P., & Widom, J. (2012). A map of nucleosome positions in yeast at base-pair resolution. *Nature* 486(7404), 496–501.
- Butryn, A., Woike, S., Shetty, S.J., Auble, D.T., & Hopfner, K.P. (2018). Crystal structure of the full Swi2/Snf2 remodeler Mot1 in the resting state. *Elife* 7: e37774.
- Cairns, B.R., Lorch, Y., Li, Y., Zhang, M., Lacomis, L., Erdjument-Bromage, H., Tempst, P., Du, J., Laurent, B., & Kornberg, R.D. (1996). RSC, an essential, abundant chromatin-remodeling complex. *Cell* 87(7), 1249-1260.
- Camblong, J., Iglesias, N., Fickentscher, C., Dieppois, G., & Stutz, F. (2007). Antisense RNA Stabilization Induces Transcriptional Gene Silencing via Histone Deacetylation in *S. cerevisiae*. *Cell* 131(4), 706–717.
- Caruthers, J. M., & McKay, D. B. (2002). Helicase structure and mechanism. *Current Opinion Structural Biology* 12(1), 123-133.
- Chacin, E., Bansal, P., Reuswig, K. U., Diaz-Santin, L. M., Ortega, P., Vizjak, P., Gómez-González, B., Müller-Planitz, F., Aguilera, A., Pfander, B., Cheung, A. C. M., & Kurat, C. F. (2021). A CDK-regulated chromatin segregase promoting chromosome replication. *Nature Communications* 12(1): 5224.
- Chen, H., Kharerin, H., Dhasarathy, A., Kladde, M., & Bai, L. (2022). Partitioned usage of chromatin remodelers by nucleosome-displacing factors. *Cell Reports* 40(8): 111250.
- Chen, K., Xi, Y., Pan, X., Li, Z., Kaestner, K., Tyler, J., Dent, S., He, X., & Li, W. (2013). DANPOS: Dynamic analysis of nucleosome position and occupancy by sequencing. *Genome Research* 23(2), 341–351.
- Chen, X., Cui, D., Papusha, A., Zhang, X., Chu, C. D., Tang, J., Chen, K., Pan, X., & Ira, G. (2012). The Fun30 nucleosome remodeler promotes resection of DNA double-strand break ends. *Nature* 489(7417), 576–580.
- Chen, Y., Tokuda, J. M., Topping, T., Meisburger, S. P., Pabit, S. A., Gloss, L. M., & Pollack, L. (2017). Asymmetric unwrapping of nucleosomal DNA propagates asymmetric opening and dissociation of the histone core. *Proceedings of the National Academy of Sciences of the United States of America*, 114(2), 334–339.
- Cheon, Y., Han, S., Kim, T., Hwang, D., & Lee, D. (2021). The chromatin remodeler Ino80 mediates RNAPII pausing site determination. *Genome Biology* 22(1): 294.

- Chereji, R. V., Bryson, T. D., & Henikoff, S. (2019). Quantitative MNase-seq accurately maps nucleosome occupancy levels. *Genome Biology* 20(1), 198.
- Chereji, R.V., & Clark, D.J. (2018). Major Determinants of Nucleosome Positioning. *Biophys J* 114(10), 2279-2289.
- Chereji, R. V., Ramachandran, S., Bryson, T. D., & Henikoff, S. (2018). Precise genome-wide mapping of single nucleosomes and linkers in vivo. *Genome Biology* 19(1). 19.
- Cho, C., Jang, J., Kang, Y., Watanabe, H., Uchihashi, T., Kim, S. J., Kato, K., Lee, J. Y., & Song, J. J. (2019). Structural basis of nucleosome assembly by the Abo1 AAA+ ATPase histone chaperone. *Nature Communications* 10(1) : 5764.
- Choi, E. S., Cheon, Y., Kang, K., & Lee, D. (2017). The Ino80 complex mediates epigenetic centromere propagation via active removal of histone H3. *Nature Communications* 8(1): 529.
- Clapier, C. R., & Cairns, B. R. (2009). The biology of chromatin remodeling complexes. In *Annual Review of Biochemistry* 78, 273–304.
- Clapier, C. R., Iwasa, J., Cairns, B. R., & Peterson, C. L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. In *Nature Reviews Molecular Cell Biology* 18(7), 407–422.
- Creamer, K. M., Job, G., Shanker, S., Neale, G. A., Lin, Y., Bartholomew, B., & Partridge, J. F. (2014). The Mi-2 Homolog Mit1 Actively Positions Nucleosomes within Heterochromatin to Suppress Transcription. *Molecular and Cellular Biology* 34(11), 2046–2061.
- Cutler, S., Lee, L. J., & Tsukiyama, T. (2018). Chromatin remodeling factors *isw2* and *ino80* regulate chromatin, replication, and copy number of the *Saccharomyces cerevisiae* ribosomal DNA locus. *Genetics* 210(4), 1543–1556.
- De Bruin, L., Tompitak, M., Eslami-Mossallam, B., & Schiessel, H. (2016). Why do nucleosomes unwrap asymmetrically? *Journal of Physical Chemistry B*, 120(26), 5855–5863.
- Deniz, Ö., Flores, O., Aldea, M., Soler-López, M., & Orozco, M. (2016). Nucleosome architecture throughout the cell cycle. *Scientific Reports*, 6:19729.
- Dong, W., Oya, E., Zahedi, Y., Prasad, P., Svensson, J. P., Lennartsson, A., Ekwall, K., & Durand-Dubief, M. (2020). Abo1 is required for the H3K9me2 to H3K9me3 transition in heterochromatin. *Scientific Reports*, 10(1):6055.

- Dong, W., Prasad, P., Lennartsson, A., & Ekwall, K. (2021). The role of non-catalytic domains of Hrp3 in nucleosome remodeling. *International Journal of Molecular Sciences* 22(4), 1–12.
- Duret, L. (2000). tRNA gene number and codon usage in the *C. elegans* genome are co-adapted for optimal translation of highly expressed genes. *Trends Genetic* 16(7), 287-289.
- Eapen, V. V., Sugawara, N., Tsabar, M., Wu, W.-H., & Haber, J. E. (2012). The *Saccharomyces cerevisiae* Chromatin Remodeler Fun30 Regulates DNA End Resection and Checkpoint Deactivation. *Molecular and Cellular Biology* 32(22), 4727–4740.
- Eshaghi, M., Lee, J. H., Zhu, L., Poon, S. Y., Li, J., Cho, K. H., Chu, Z., Karuturi, R. K. M., & Liu, J. (2010). Genomic binding profiling of the fission yeast stress-activated MAPK sty1 and the bZIP transcriptional activator Atf1 in response to H<sub>2</sub>O<sub>2</sub>. *PLoS ONE* 5(7): e11620.
- Eustermann, S., Patel, A.B., Hopfner, KP, He, Y., & Korber, P. (2023). Energy-driven genome regulation by ATP-dependent chromatin remodellers. *Nature Reviews Molecular Cell Biology*.
- Felsenfeld, G., Boyes, J., Chung, J., Clark, D., & Studitsky, V. (1996). Chromatin structure and gene expression. *PNAS* 93 (18), 9384-9388.
- Filipovski, M., Soffers, J. H. M., Vos, S. M., & Farnung, L. (2022). Structural basis of nucleosome retention during transcription elongation. *Science* 376(6599), 1313–1316.
- Flaus, A., Martin, D. M. A., Barton, G. J., & Owen-Hughes, T. (2006). Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Research* 34(10), 2887–2905.
- Forsburg, S. L., & Rhind, N. (2006). Basic methods for fission yeast. *Yeast* 23(3), 173–183.
- Gaffney, D. J., McVicker, G., Pai, A. A., Fondufe-Mittendorf, Y. N., Lewellen, N., Michelini, K., Widom, J., Gilad, Y., & Pritchard, J. K. (2012). Controls of Nucleosome Positioning in the Human Genome. *PLoS Genetics* 8(11): e1003036.
- Gal, C., Murton, H. E., Subramanian, L., Whale, A. J., Moore, K. M., Paszkiewicz, K., Codlin, S., Bähler, J., Creamer, K. M., Partridge, J. F., Allshire, R. C., Kent, N. A., & Whitehall, S. K. (2016). Abo1, a conserved bromodomain AAA - ATPase, maintains global nucleosome occupancy and organisation. *EMBO Reports* 17(1), 79–93.

- García, A., Durán, L., Sánchez, M., González, S., Santamaría, R., & Antequera, F. (2024). Asymmetrical nucleosomal DNA signatures regulate transcriptional directionality. *Cell Reports* 43(1):113605.
- García, A., González, S., & Antequera, F. (2017). Nucleosomal organization and DNA base composition patterns. *Nucleus* 8(5), 469–474.
- Gavin, A.-C., Bo, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Rg Schultz, J. È., Rick, J. M., Michon, A.-M., Cruciat, C.-M., Remor, M., Ho, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D., & Superti-Furga, G. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415(6868), 141-147.
- Gill, J. K., Maffioletti, A., García-Molinero, V., Stutz, F., & Soudet, J. (2020). Fine Chromatin-Driven Mechanism of Transcription Interference by Antisense Noncoding Transcription. *Cell Report* 31(5):107612.
- Gillery, P., Georges, N., Wegrowski, J., Randoux, A., & Borel, J.-P. (1995). Protein synthesis in collagen lattice-cultured fibroblasts is controlled at the ribosomal level. *FEBS Letters* 357(3), 287-289.
- Givens, R. M., Lai, W. K. M., Rizzo, J. M., Bard, J. E., Mieczkowski, P. A., Leatherwood, J., Huberman, J. A., & Buck, M. J. (2012). Chromatin architectures at fission yeast transcriptional promoters and replication origins. *Nucleic Acids Research* 40(15), 7176–7189.
- Goedhart, J., & Luijsterburg, M. S. (2020). VolcanoNoseR is a web app for creating, exploring, labeling and sharing volcano plots. *Scientific Reports* 10(1):20560.
- González, S., García, A., Vázquez, E., Serrano, R., Sánchez, M., Quintales, L., & Antequera, F. (2016). Nucleosomal signatures impose nucleosome positioning in coding and noncoding sequences in the genome. *Genome Research* 26(11), 1532–1543.
- Groudine, M., Eisenman, R., & Weintraub, H. (1981). Chromatin structure of endogenous retroviral genes and activation by an inhibitor of DNA methylation. *Nature* 292, 311–317.
- Gullerova, M., & Proudfoot, N. J. (2008). Cohesin Complex Promotes Transcriptional Termination between Convergent Genes in *S. pombe*. *Cell* 132(6), 983–995.
- Han, Y., Reyes, A. A., Malik, S., & He, Y. (2020). Cryo-EM structure of SWI/SNF complex bound to a nucleosome. *Nature* 579(7799), 452–455.

- Harigaya, Y., & Parker, R. (2016). Analysis of the association between codon optimality and mRNA stability in *Schizosaccharomyces pombe*. *BMC Genomics* 17(1):895.
- Hartley, P. D., & Madhani, H. D. (2009). Mechanisms that Specify Promoter Nucleosome Location and Identity. *Cell* 137(3), 445–458.
- Heaton, B., Decker, C., Muhlrud, D., Donahue, J., Jacobson, A., & Parker, R. (1992). Analysis of chimeric mRNAs derived from the STE3 mRNA identifies multiple regions within yeast mRNAs that modulate mRNA decay. *Nucleic Acids Research* 20(20), 5365-5373.
- Hendy, O., Serebreni, L., Bergauer, K., Muerdter, F., Huber, L., Nemčko, F., & Stark, A. (2022). Developmental and housekeeping transcriptional programs in *Drosophila* require distinct chromatin remodelers. *Molecular Cell* 82(19), 3598-3612.
- Hennig, B. P., Bendrin, K., Zhou, Y., & Fischer, T. (2012). Chd1 chromatin remodelers maintain nucleosome organization and repress cryptic transcription. *EMBO Reports* 13(11), 997–1003.
- Herrick, D., Parker, R., & Jacobson, A. (1990). Identification and Comparison of Stable and Unstable mRNAs in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* 10(5), 2269-2284.
- Hickey, A., Esnault, C., Majumdar, A., Chatterjee, A. G., Iben, J. R., McQueen, P. G., Yang, A. X., Mizuguchi, T., Grewal, S. I. S., & Levin, H. L. (2015). Single-nucleotide-specific targeting of the Tf1 retrotransposon promoted by the DNA-Binding protein sap1 of *Schizosaccharomyces pombe*. *Genetics* 201(3), 905–924.
- Hughes, A. L., & Rando, O. J. (2015). Comparative genomics reveals Chd1 as a determinant of nucleosome spacing in vivo. *G3* 5(9), 1889–1897.
- Hughes, T. R., & de Boer, C. G. (2013). Mapping yeast transcriptional networks. *Genetics* 195(1), 9–36.
- Iyer, V., & Struhl, K. (1995). Poly (dA:dT), a ubiquitous promoter element that stimulates transcription via its intrinsic DNA structure. *EMBO Journal* 14(11):2570-2579.
- Jimenez, A., Tipper, D. J., & Davies, J. (1973). Mode of action of thiolutin, an inhibitor of macromolecular synthesis in *Saccharomyces cerevisiae*. *Antimicrobial Agents and Chemotherapy* 3(6), 729–738.

- Jin, Y., Eser, U., Struhl, K., & Churchman, L. S. (2017). The Ground State and Evolution of Promoter Region Directionality. *Cell* 170(5), 889-898.
- Jo, H., Kim, T., Chun, Y., Jung, I., & Lee, D. (2021). A compendium of chromatin contact maps reflecting regulation by chromatin remodelers in budding yeast. *Nature Communications*, 12(1):6380.
- Kanaya, S., Yamada, Y., Kudo, Y., & Ikemura, T. (1999). Studies of codon usage and tRNA genes of 18 unicellular organisms and quantification of *Bacillus subtilis* tRNAs: gene expression level and species-specific diversity of codon usage based on multivariate analysis. *Gene* 238(1):143-55.
- Kang, Y., Cho, C., Lee, K. S., Song, J. J., & Lee, J. Y. (2021). Single-Molecule Imaging Reveals the Mechanism Underlying Histone Loading of *Schizosaccharomyces pombe* AAA+ ATPase Abo1. *Molecules and Cells* 44(2), 79-87.
- Kaplan, N., Moore, I. K., Fondufe-Mittendorf, Y., Gossett, A. J., Tillo, D., Field, Y., LeProust, E. M., Hughes, T. R., Lieb, J. D., Widom, J., & Segal, E. (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* 458(7236), 362-366.
- Kasinathan, S., Orsi, G. A., Zentner, G. E., Ahmad, K., & Henikoff, S. (2014). High-resolution mapping of transcription factor binding sites on native chromatin. *Nature Methods* 11(2), 203-209.
- Khorosjutina, O., Wanrooij, P. H., Walfridsson, J., Szilagyi, Z., Zhu, X., Baraznenok, V., Ekwall, K., & Gustafsson, C. M. (2010). A chromatin-remodeling protein is a component of fission yeast mediator. *Journal of Biological Chemistry* 285(39), 29729-29737.
- Kim, J. M., Visanpattanasin, P., Jou, V., Liu, S., Tang, X., Zheng, Q., Li, K. Y., Snedeker, J., Lavis, L. D., Lionnet, T., & Wu, C. (2021). Single-molecule imaging of chromatin remodelers reveals role of ATPase in promoting fast kinetics of target search and dissociation from chromatin. *ELife*, 10: e69387.
- Kim, M. J., Kim, M., & Park, S. D. (2002). Post-transcriptional Regulation of *ura4+* Gene Expression by Glucose in *Schizosaccharomyces pombe*. *Molecules and Cells* 14(3), 437-443.
- Klein-Brill, A., Joseph-Strauss, D., Appleboim, A., & Friedman, N. (2019). Dynamics of Chromatin and Transcription during Transient Depletion of the RSC Chromatin Remodeling Complex. *Cell Reports* 26(1), 279-292.

- Knight, B., Kubik, S., Ghosh, B., Bruzzone, M. J., Geertz, M., Martin, V., Déneraud, N., Jacquet, P., Ozkan, B., Rougemont, J., Maerkl, S. J., Naef, F., & Shore, D. (2014). Two distinct promoter architectures centered on dynamic nucleosomes control ribosomal protein gene transcription. *Genes and Development* 28(15), 1695–1709.
- Koerber, R. T., Rhee, H. S., Jiang, C., & Pugh, B. F. (2009). Interaction of Transcriptional Regulators with Specific Nucleosomes across the *Saccharomyces* Genome. *Molecular Cell* 35(6), 889–902.
- Kornberg, R. D., & Lorch, Y. (2020). Primary Role of the Nucleosome. *Molecular Cell* 79(3), 371–375).
- Kornberg, R. D., & Stryer, L. (1988). Statistical distributions of nucleosomes: nonrandom locations by a stochastic mechanism. *Nucleic Acids Research* 16(14), 6677–6690.
- Krietenstein, N., Wal, M., Watanabe, S., Park, B., Peterson, C. L., Pugh, B. F., & Korber, P. (2016). Genomic Nucleosome Organization Reconstituted with Pure Proteins. *Cell* 167(3), 709–721.
- Kubik, S., Bruzzone, M. J., Challal, D., Dreos, R., Mattarocci, S., Bucher, P., Libri, D., & Shore, D. (2019). Opposing chromatin remodelers control transcription initiation frequency and start site selection. *Nature Structural and Molecular Biology* 26(8), 744–754.
- Kubik, S., O’Duibhir, E., de Jonge, W. J., Mattarocci, S., Albert, B., Falcone, J. L., Bruzzone, M. J., Holstege, F. C. P., & Shore, D. (2018). Sequence-Directed Action of RSC Remodeler and General Regulatory Factors Modulates +1 Nucleosome Position to Facilitate Transcription. *Molecular Cell* 71(1), 89–102.
- Kujirai, T., Ehara, H., Fujino, Y., Shirouzu, M., Sekine, S., & Kurumizaka, H. (2018). Structural basis of the nucleosome transition during RNA polymerase II passage. *Science* 362, 595–598.
- Kujirai, T., & Kurumizaka, H. (2020). Transcription through the nucleosome. *Current Opinion in Structural Biology* 61, 42–49.
- Kulaeva, O. I., Gaykalova, D. A., Pestov, N. A., Golovastov, V. V., Vassilyev, D. G., Artsimovitch, I., & Studitsky, V. M. (2009). Mechanism of chromatin remodeling and recovery during passage of RNA polymerase II. *Nature Structural and Molecular Biology* 16(12), 1272–1278.

- Kulaeva, O. I., Hsieh, F. K., Chang, H. W., Luse, D. S., & Studitsky, V. M. (2013). Mechanism of transcription through a nucleosome by RNA polymerase II. *Biochimic et Biophysic Acta* 1829(1), 76–83.
- Kushner, S. (1978). An improve method for transformation *Escherichia coli* with ColE1 derived plasmids. *Genetic engineering*, H.W.B.a.S. Nicosia, ed. (Amsterdam: Elsevier/North-Holland Biomedical Press), pp. 17-23
- Laemmli, U. K. (1970). Cleavage of Structural Proteins during de Assembly of the Head of Bacteriophage T4. *Nature* 227(5259):680-685.
- Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* 10(3).
- Lantermann, A. B., Straub, T., Strålfors, A., Yuan, G. C., Ekwall, K., & Korber, P. (2010). *Schizosaccharomyces pombe* genome-wide nucleosome mapping reveals positioning mechanisms distinct from those of *Saccharomyces cerevisiae*. *Nature Structural and Molecular Biology* 17(2), 251–257.
- Lauinger, L., Li, J., Shostak, A., Cemel, I. A., Ha, N., Zhang, Y., Merkl, P. E., Obermeyer, S., Stankovic-Valentin, N., Schafmeier, T., Wever, W. J., Bowers, A. A., Carter, K. P., Palmer, A. E., Tschochner, H., Melchior, F., Deshaies, R. J., Brunner, M., & Diernfellner, A. (2017). Thiolutin is a zinc chelator that inhibits the Rpn11 and other JAMM metalloproteases. *Nature Chemical Biology* 13(7), 709–714.
- Lee, J., Shik Choi, E., David Seo, H., Kang, K., Gilmore, J. M., Florens, L., Washburn, M. P., Choe, J., Workman, J. L., & Lee, D. (2017). Chromatin remodeller Fun30Fft3 induces nucleosome disassembly to facilitate RNA polymerase II elongation. *Nature Communications* 8:14527.
- Lin, A., Du, Y., & Xiao, W. (2020). Yeast chromatin remodeling complexes and their roles in transcription. *Current Genetics* 66(4), 657–670.
- Liu, X., Hoque, M., Larochelle, M., Lemay, J. F., Yurko, N., Manley, J. L., Bachand, F., & Tian, B. (2017). Comparative analysis of alternative polyadenylation in *S. cerevisiae* and *S. pombe*. *Genome Research* 27(10), 1685–1695.
- Liu, Y. (2020). A code within the genetic code: Codon usage regulates co-translational protein folding. *Cell Communication and Signaling* 18(1):145.

- Lorch, Y., & Kornberg, R.D. (2017). Chromatin-remodeling for transcription. *Quarterly Reviews of Biophysics* 50: e5.
- Lorch, Y., Lapointe, J. W., & Kornberg, R. D. (1987). Nucleosomes Inhibit the Initiation of Transcription but Allow Chain Elongation with the Displacement of Histones. *Cell* 49, 203-210.
- Luehr, S., Hartmann, H., & Söding, J. (2012). The XXmotif web server for eXhaustive, weight matriX-based motif discovery in nucleotide sequences. *Nucleic Acids Research* 40(1), 104-109.
- Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., & Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251-260.
- Magaña-Acosta, M., & Valadez-Graham, V. (2020). Chromatin Remodelers in the 3D Nuclear Compartment. *Frontiers in Genetics* 11:600615.
- Maldonado, E., Morales-Pison, S., Urbina, F., Arias, C., Castillo, C., Jara, L., & Solari, A. (2023). The Product of the Fission Yeast *fhl1* Gene Binds to the HomolE Box and Activates In Vitro Transcription of Ribosomal Protein Genes. *International Journal of Molecular Sciences* 24(11).
- Markert, J., & Luger, K. (2021). Nucleosomes Meet Their Remodeler Match. *Trends in Biochemical Sciences* 46(1), 41–50.
- Mata, J. (2013). Genome-wide mapping of polyadenylation sites in fission yeast reveals widespread alternative polyadenylation. *RNA Biology* 10(8), 1407–1414.
- 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., & Pugh, B. F. (2008a). Nucleosome organization in the *Drosophila* genome. *Nature* 453(7193), 358–362.
- Mavrich, T. N., Ioshikhes, I. P., Venters, B. J., Jiang, C., Tomsho, L. P., Qi, J., Schuster, S. C., Albert, I., & Pugh, B. F. (2008b). A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Research* 18(7), 1073–1083.
- McGinty, R. K., & Tan, S. (2015). Nucleosome structure and function. *Chemical Reviews* 115(6), 2255–2273.
- Misteli, T. (2020). The Self-Organizing Genome: Principles of Genome Architecture and Function. *Cell* 183(1), 28-45.

- Mivelaz, M., Cao, A. M., Kubik, S., Zencir, S., Hovius, R., Boichenko, I., Stachowicz, A. M., Kurat, C. F., Shore, D., & Fierz, B. (2020). Chromatin Fiber Invasion and Nucleosome Displacement by the Rap1 Transcription Factor. *Molecular Cell* 77(3), 488-500.
- Monahan, B. J., Villén, J., Marguerat, S., Bähler, J., Gygi, S. P., & Winston, F. (2008). Fission yeast SWI/SNF and RSC complexes show compositional and functional differences from budding yeast. *Nature Structural and Molecular Biology* 15(8), 873–880.
- Moreno, S., Klar, A., Nurse, P. (1991). Molecular genetic analysis of fission yeast *Schizosaccharomyces pombe*. *Methods Enzymol* 194, 795–823.
- Morettini, S., Tribus, M., Zeilner, A., Sebald, J., Campo-Fernandez, B., Scheran, G., Wörle, H., Podhraski, V., Fyodorov, D. V., & Lusser, A. (2011). The chromodomains of CHD1 are critical for enzymatic activity but less important for chromatin localization. *Nucleic Acids Research* 39(8), 3103–3115.
- Moyle-Heyrman, G., Zaichuk, T., Xi, L., Zhang, Q., Uhlenbeck, O. C., Holmgren, R., Widom, J., & Wang, J. P. (2013). Chemical map of *Schizosaccharomyces pombe* reveals species-specific features in nucleosome positioning. *Proceedings of the National Academy of Sciences of the United States of America* 110(50), 20158–20163.
- Nakamura, Y., Gojobori, T., & Ikemura, T. (2000). Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Research* 28(1).
- Ngo, T. T. M., Zhang, Q., Zhou, R., Yodh, J. G., & Ha, T. (2015). Asymmetric unwrapping of nucleosomes under tension directed by DNA local flexibility. *Cell* 160(6), 1135–1144.
- Nocetti, N., & Whitehouse, I. (2016). Nucleosome repositioning underlies dynamic gene expression. *Genes and Development* 30(6), 660–672.
- Nodelman, I. M., & Bowman G. D. (2021). Biophysics of Chromatin Remodeling. *Annu Rev Biophys.* 50, 73-93.
- Oberbeckmann, E., Krietenstein, N., Niebauer, V., Wang, Y., Schall, K., Moldt, M., Straub, T., Rohs, R., Hopfner, K. P., Korber, P., & Eustermann, S. (2021). Genome information processing by the INO80 chromatin remodeler positions nucleosomes. *Nature Communications* 12(1).
- Ocampo, J., Chereji, R.V., Eriksson, P.R., & Clark D.J. (2016). The ISW1 and CHD1 ATP-dependent chromatin remodelers compete to set nucleosome spacing in vivo. *Nucleic Acids Res* 44(10), 4625-4635.

- Ocampo, J., Chereji, R. V., Eriksson, P. R., & Clark, D. J. (2019). Contrasting roles of the RSC and ISW1/CHD1 chromatin remodelers in RNA polymerase II elongation and termination. *Genome Research* 29(3), 407–417.
- Parvathy, S. T., Udayasuriyan, V., & Bhadana, V. (2022). Codon usage bias. *Molecular Biology Reports* 49(1), 539–565.
- Pérez-Ortín, J. E. (2007). Genomics of mRNA turnover. *Briefings in Functional Genomics and Proteomics* 6(4), 282–291.
- Pérez-Ortín, J. E., Alepuz, P. M., & Moreno, J. (2007). Genomics and gene transcription kinetics in yeast. *Trends in Genetics* 23(5), 250–257.
- Pérez-Ortín, J. E., Medina, D. A., Chávez, S., & Moreno, J. (2013). What do you mean by transcription rate? The conceptual difference between nascent transcription rate and mRNA synthesis rate is essential for the proper understanding of transcriptomic analyses. *BioEssays* 35(12), 1056–1062.
- Persson, J., Steglich, B., Smialowska, A., Boyd, M., Bornholdt, J., Andersson, R., Schurra, C., Arcangioli, B., Sandelin, A., Nielsen, O., & Ekwall, K. (2016). Regulating retrotransposon activity through the use of alternative transcription start sites. *EMBO Reports* 17(5), 753–768.
- Pointner, J., Persson, J., Prasad, P., Norman-Axelsson, U., Strålfors, A., Khorosjutina, O., Krietenstein, N., Peter Svensson, J., Ekwall, K., & Korber, P. (2012). CHD1 remodelers regulate nucleosome spacing in vitro and align nucleosomal arrays over gene coding regions in *S. pombe*. *EMBO Journal* 31(23), 4388–4403.
- Prajapati, H. K., Ocampo, J., & Clark, D. J. (2020). Interplay among ATP-dependent chromatin remodelers determines chromatin organization in yeast. *Biology* 9(8), 1–23.
- Prieto, C., & Barrios, D. (2020). RaNA-Seq: Interactive RNA-Seq analysis from FASTQ files to functional analysis. *Bioinformatics* 36(6), 1955–1956.
- Pugh, B. F. (2010). A preoccupied position on nucleosomes. *Nature Structural and Molecular Biology* 17(8):923.
- Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26(6), 841–842.

- Quintales, L., Soriano, I., Vázquez, E., Segurado, M., & Antequera, F. (2015a). A species-specific nucleosomal signature defines a periodic distribution of amino acids in proteins. *Open Biology* 5(4):140218.
- Quintales, L., Vázquez, E., & Antequera, F. (2015b). Comparative analysis of methods for genome-wide nucleosome cartography. *Briefings in Bioinformatics* 16(4), 576–587.
- Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dündar, F., & Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research* 44(1), 160–165.
- Raveh-Sadka, T., Levo, M., Shabi, U., Shany, B., Keren, L., Lotan-Pompan, M., Zeevi, D., Sharon, E., Weinberger, A., & Segal, E. (2012). Manipulating nucleosome disfavoring sequences allows fine-tune regulation of gene expression in yeast. *Nature Genetics* 44(7), 743–750.
- Rawal, Y., Chereji, R. V., Valabhoju, V., Qiu, H., Ocampo, J., Clark, D. J., & Hinnebusch, A. G. (2018). Gcn4 Binding in Coding Regions Can Activate Internal and Canonical 5' Promoters in Yeast. *Molecular Cell* 70(2), 297–311.
- Rhee, H. S., & Pugh, B. F. (2011). Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell* 147(6), 1408–1419.
- Rondón, A. G., Mischo, H. E., Kawauchi, J., & Proudfoot, N. J. (2009). Fail-Safe Transcriptional Termination for Protein-Coding Genes in *S. cerevisiae*. *Molecular Cell* 36(1), 88–98.
- Rossi, M.J., Kuntala, P.K., Lai, W.K.M., Yamada, N., Badjatia, N., Mittal, C., Kuzu, G., Bocklund, K., Farrell, N.P., Blanda, T.R., Mairose, J.D., Basting, A.V., Mistretta, K.S., Rocco, D.J., Perkinson, E.S., Kellogg, G.D., Mahony, S., & Pugh, F. (2021). A high-resolution protein architecture of the budding yeast genome. *Nature* 592, 309–314.
- Ryan, D. P., Sundaramoorthy, R., Martin, D., Singh, V., & Owen-Hughes, T. (2011). The DNA-binding domain of the Chd1 chromatin-remodelling enzyme contains SANT and SLIDE domains. *EMBO Journal* 30(13), 2596–2609.
- Sachs, A. B. (1993). Messenger RNA Degradation in Eukaryotes Review. *Cell* 74, 413–421.
- Sasaki, S., Mello, C. C., Shimada, A., Nakatani, Y., Hashimoto, S. I., Ogawa, M., Matsushima, K., Sam, G. G., Kasahara, M., Ahsan, B., Sasaki, A., Saito, T., Suzuki, Y., Sugano, S., Kohara, Y., Takeda, H., Fire, A., & Morishita, S. (2009). Chromatin-associated periodicity in genetic variation downstream of transcriptional start sites. *Science* 323(5912), 401–404.

- Satchwell, S. C., Drew, H. R., & Travers, A. A. (1986). Sequence periodicities in chicken nucleosome core DNA. *Journal of Molecular Biology* 191, 659-675.
- Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thåström, A., Field, Y., Moore, I. K., Wang, J. P. Z., & Widom, J. (2006). A genomic code for nucleosome positioning. *Nature* 442(7104), 772–778.
- Segal, E., & Widom, J. (2009). What controls nucleosome positions? *Trends in Genetics* 25(8), 335–343.
- Sekinger, E. A., Moqtaderi, Z., & Struhl, K. (2005). Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. *Molecular Cell* 18(6), 735–748.
- Sharp, P. M., Cowe, E., Higgins, D. G., Shields, D. C., Wolfe, K. H., & Wright, F. (1988). Codon usage patterns in *Escherichia coli*, *Bacillus subtilis*, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Drosophila melanogaster* and *Homo sapiens*; a review of the considerable within-species diversity. *Nucleic Acids Research* 16(17), 8207-8211.
- Shaw, G., & Kamen, R. (1986). A Conserved AU Sequence from the 3' Untranslated Region of GM-CSF mRNA Mediates Selective mRNA Degradation. *Cell* 46(5), 659-667.
- Shim, Y. S., Choi, Y., Kang, K., Cho, K., Oh, S., Lee, J., Grewal, S. I. S., & Lee, D. (2012). Hrp3 controls nucleosome positioning to suppress non-coding transcription in eu-and heterochromatin. *EMBO Journal* 31(23), 4375–4387.
- Shivaswamy, S., Bhinge, A., Zhao, Y., Jones, S., Hirst, M., & Iyer, V. R. (2008). Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. *PLoS Biology* 6(3), 618–630.
- Singh, A. K., & Mueller-Planitz, F. (2021). Nucleosome Positioning and Spacing: From Mechanism to Function. *Journal of Molecular Biology* 433(6):166847.
- Singh, A. K., Schauer, T., Pfaller, L., Straub, T., & Mueller-Planitz, F. (2021). The biogenesis and function of nucleosome arrays. *Nature Communications* 12(1):7011.
- Small, E. C., Xi, L., Wang, J. P., Widom, J., & Licht, J. D. (2014). Single-cell nucleosome mapping reveals the molecular basis of gene expression heterogeneity. *Proceedings of the National Academy of Sciences of the United States of America* 111(24).

- Smolle, M., Venkatesh, S., Gogol, M. M., Li, H., Zhang, Y., Florens, L., Washburn, M. P., & Workman, J. L. (2012). Chromatin remodelers Isw1 and Chd1 maintain chromatin structure during transcription by preventing histone exchange. *Nature Structural and Molecular Biology* 19(9), 884–892.
- Soriano, I., Quintales, L., & Antequera, F. (2013). Clustered regulatory elements at nucleosome-depleted regions punctuate a constant nucleosomal landscape in *Schizosaccharomyces pombe*. *BMC Genomics* 14, 813.
- Struhl, K., & Segal, E. (2013). Determinants of nucleosome positioning. *Nature Structural and Molecular Biology* 20(3), 267–273.
- Sudarsanam, P., Iyer, V.R., Brown, P.O., & Winston, F. (2000). Whole-genome expression analysis of snf/swi mutants of *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America* 97(7), 3364–3369.
- Sunyer-Figueres, M., Wang, C., & Mas, A. (2018). Analysis of ribosomal RNA stability in dead cells of wine yeast by quantitative PCR. *International Journal of Food Microbiology* 270, 1–4.
- Tillo, D., & Hughes, T. R. (2009). G+C content dominates intrinsic nucleosome occupancy. *BMC Bioinformatics*, 10:442.
- Touat-Todeschini, L., Hiriart, E., & Verdel, A. (2012). Nucleosome positioning and transcription: Fission yeast CHD remodelers make their move. *EMBO Journal* 31(23), 4371–4372.
- Tsankov, A. M., Thompson, D. A., Socha, A., Regev, A., & Rando, O. J. (2010). The role of nucleosome positioning in the evolution of gene regulation. *PLoS Biology* 8(7).
- Tsankov, A., Yanagisawa, Y., Rhind, N., Regev, A., & Rando, O. J. (2011). Evolutionary divergence of intrinsic and trans-regulated nucleosome positioning sequences reveals plastic rules for chromatin organization. *Genome Research* 21(11), 1851–1862.
- Udugama, M., Sabri, A., & Bartholomew, B. (2011). The INO80 ATP-dependent chromatin remodeling complex is a nucleosome spacing factor. *Molecular and Cellular Biology* 31(4), 662–673.
- Vaillant, C., Palmeira, L., Chevereau, G., Audit, B., D'Aubenton-Carafa, Y., Thermes, C., & Arneodo, A. (2010). A novel strategy of transcription regulation by intragenic nucleosome ordering. *Genome Research* 20(1), 59–67.

- Valouev, A., Ichikawa, J., Tonthat, T., Stuart, J., Ranade, S., Peckham, H., Zeng, K., Malek, J. A., Costa, G., McKernan, K., Sidow, A., Fire, A., & Johnson, S. M. (2008). A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Research* 18(7), 1051–1063.
- Valouev, A., Johnson, S. M., Boyd, S. D., Smith, C. L., Fire, A. Z., & Sidow, A. (2011). Determinants of nucleosome organization in primary human cells. *Nature* 474(7352), 516–522.
- Veron, M., Zou, Y., Yu, Q., Bi, X., Selmi, A., Gilson, E., & Defossez, P. A. (2006). Histone H1 of *Saccharomyces cerevisiae* inhibits transcriptional silencing. *Genetics* 173(2), 579–587.
- Wagner, F. R., Dienemann, C., Wang, H., Stützer, A., Tegunov, D., Urlaub, H., & Cramer, P. (2020). Structure of SWI/SNF chromatin remodeler RSC bound to a nucleosome. *Nature* 579(7799), 448–451.
- Walfridsson, J., Bjerling, P., Thalen, M., Yoo, E. J., Park, S. D., & Ekwall, K. (2005). The CHD remodeling factor Hrp1 stimulates CENP-A loading to centromeres. *Nucleic Acids Research* 33(9), 2868–2879.
- Wang, J. P., Fondufe-Mittendorf, Y., Xi, L., Tsai, G.F., Segal, E., & Widom, J. (2008). Preferentially quantized linker DNA lengths in *Saccharomyces cerevisiae*. *PLoS Computer Biology* 4(9).
- Wang, S., Vogirala, V.K., Soman, A., Berezhnoy, N. V., Liu, Z. B., Wong, A. S. W., Korolev, N., Su, C.J., Sandin, S., & Nordenskiöld, L. (2021a). Linker histone defines structure and self-association behaviour of the 177 bp human chromatosome. *Science Reports* 11(380).
- Wang, T., Perazza, D., Boussouar, F., Cattaneo, M., Bougdour, A., Chuffart, F., Barral, S., Vargas, A., Liakopoulou, A., Puthier, D., Bargier, L., Morozumi, Y., Jamshidikia, M., Garcia-Saez, I., Petosa, C., Rousseaux, S., Verdel, A., & Khochbin, S. (2021b). ATAD2 controls chromatin-bound HIRA turnover. *Life Science Alliance*, 4(12).
- Warnecke, T., Batada, N. N., & Hurst, L. D. (2008). The impact of the nucleosome code on protein-coding sequence evolution in yeast. *PLoS Genetics* 4(11).
- Weingarten-Gabbay, S., & Segal, E. (2014). The grammar of transcriptional regulation. *Human Genetics* 133(6), 701–711.
- Wery, M., Gautier, C., Describes, M., Yoda, M., Vennin-Rendos, H., Migeot, V., Gautheret, D., Hermand, D., & Morillon, A. (2018). Native elongating transcript sequencing reveals global

- anti-correlation between sense and antisense nascent transcription in fission yeast. *RNA* 24(2), 196-208.
- Wilson, C. J., Chao, D. M., Kingston, R. E., & Young, R. A. (1996). RNA Polymerase II Holoenzyme Contains SWI/SNF Regulators Involved in Chromatin Remodeling. *Cell* 84(2):235-244.
- Wittl, I., Kwart1, M., Grob2, T., & Kaufer2, N. F. (1995). The tandem repeat AGGGTAGGGT is, in the fission yeast, a proximal activation sequence and activates basal transcription mediated by the sequence TGTGACTG. *Nucleic Acids Research* 23(21).
- Woike, S., Eustermann, S., Jung, J., Wenzl, S. J., Hagemann, G., Bartho, J., Lammens, K., Butryn, A., Herzog, F., & Hopfner, KP. (2023). Structural basis for TBP displacement from TATA box DNA by the Swi2/Snf2 ATPase Mot1. *Nature Structural and Molecular Biology* 30, 640–649.
- Wu, A. C. K., Patel, H., Chia, M., Moretto, F., Frith, D., Snijders, A. P., & van Werven, F. J. (2018). Repression of Divergent Noncoding Transcription by a Sequence-Specific Transcription Factor. *Molecular Cell* 72(6), 942-954.
- Wunderlich, Z., & Mirny, L. A. (2009). Different gene regulation strategies revealed by analysis of binding motifs. *Trends in Genetics* 25(10), 434–440.
- Yadon, A. N., Singh, B. N., Hampsey, M., & Tsukiyama, T. (2013). DNA Looping Facilitates Targeting of a Chromatin Remodeling Enzyme. *Molecular Cell* 50(1), 93–103.
- Yague-Sanz, C., Vázquez, E., Sánchez, M., Antequera, F., & Hermand, D. (2017). A conserved role of the RSC chromatin remodeler in the establishment of nucleosome-depleted regions. *Current Genetics* 63(2), 187–193.
- Yamada, K., Hirota, K., Mizuno, K.-I., Shibata, T., & Ohta, K. (2008). Essential roles of Snf21, a Swi2/Snf2 family chromatin remodeler, in fission yeast mitosis. *Genes Genet. Syst* 83(5), 361-372.
- Yamada, T., & Akimitsu, N. (2018). Contributions of regulated transcription and mRNA decay to the dynamics of gene expression. *Wiley Interdisciplinary Reviews: RNA* 10(1).
- Yang, Q., Lyu, X., Zhao, F., & Liu, Y. (2021). Effects of codon usage on gene expression are promoter context dependent. *Nucleic Acids Research* 49(2), 818–831.

- Yen, K., Vinayachandran, V., Batta, K., Koerber, R. T., & Pugh, B. F. (2012). Genome-wide nucleosome specificity and directionality of chromatin remodelers. *Cell* 149(7), 1461–1473.
- Yoo, J.E., Jin, Y.H., Jang, Y.K., Bjerling, P., Tabish, M., Hong, S.H., Ekwall, K., & Park, S.D. (2000). Fission yeast Hrp1, a chromodomain ATPase, is required for proper chromosome segregation and its overexpression interferes with chromatin condensation. *Nucleic acids research* 28(9), 2004–2011.
- Yuan, G.C., Liu, Y.J., Dion, M.F., Slack, M.D., Wu, L.F., Altschuler, S.J., Rando, O. J. (2005). Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* 309(5734), 626–630.
- Zhang, H., & Reese, J. C. (2007). Exposing the core promoter is sufficient to activate transcription and alter coactivator requirement at RNR3. *Proceedings of the National Academy of Sciences of the United States of America* 104, 8833–8838.
- Zhang, X. R., Zhao, L., Suo, F., Gao, Y., Wu, Q., Qi, X., & Du, L. L. (2022). An improved auxin-inducible degron system for fission yeast. *G3* 12(1).
- Zhang, Z., Wippo, C. J., Wal, M., Ward, E., Korber, P., & Pugh, B. F. (2011). A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. *Science* 332(6032), 977–980.
- Zhao, A., Guo, A., Liu, Z., & Pape, L. (1997). Molecular cloning and analysis of *Schizosaccharomyces pombe* Reb1p: sequence-specific recognition of two sites in the far upstream rDNA intergenic spacer. *Nucleic Acids Research* 25(4), 904–910.
- Zhao, F., Zhou, Z., Dang, Y., Na, H., Adam, C., Lipzen, A., Ng, V., Grigoriev, I. V., & Liu, Y. (2021). Genome-wide role of codon usage on transcription and identification of potential regulators. *Proceedings of the National Academy of Sciences of the United States of America* 118(6).
- Zhou, K., Gaullier, G., & Luger, K. (2019). Nucleosome structure and dynamics are coming of age. *Nature Structural and Molecular Biology* 26(1), 3–13.
- Zhou, Z., Danga, Y., Zhou, M., Li, L., Yu, C. H., Fu, J., Chen, S., & Liu, Y. (2016). Codon usage is an important determinant of gene expression levels largely through its effects on transcription. *Proceedings of the National Academy of Sciences of the United States of America* 113(41), 6117–6125.



# ABREVIATURAS





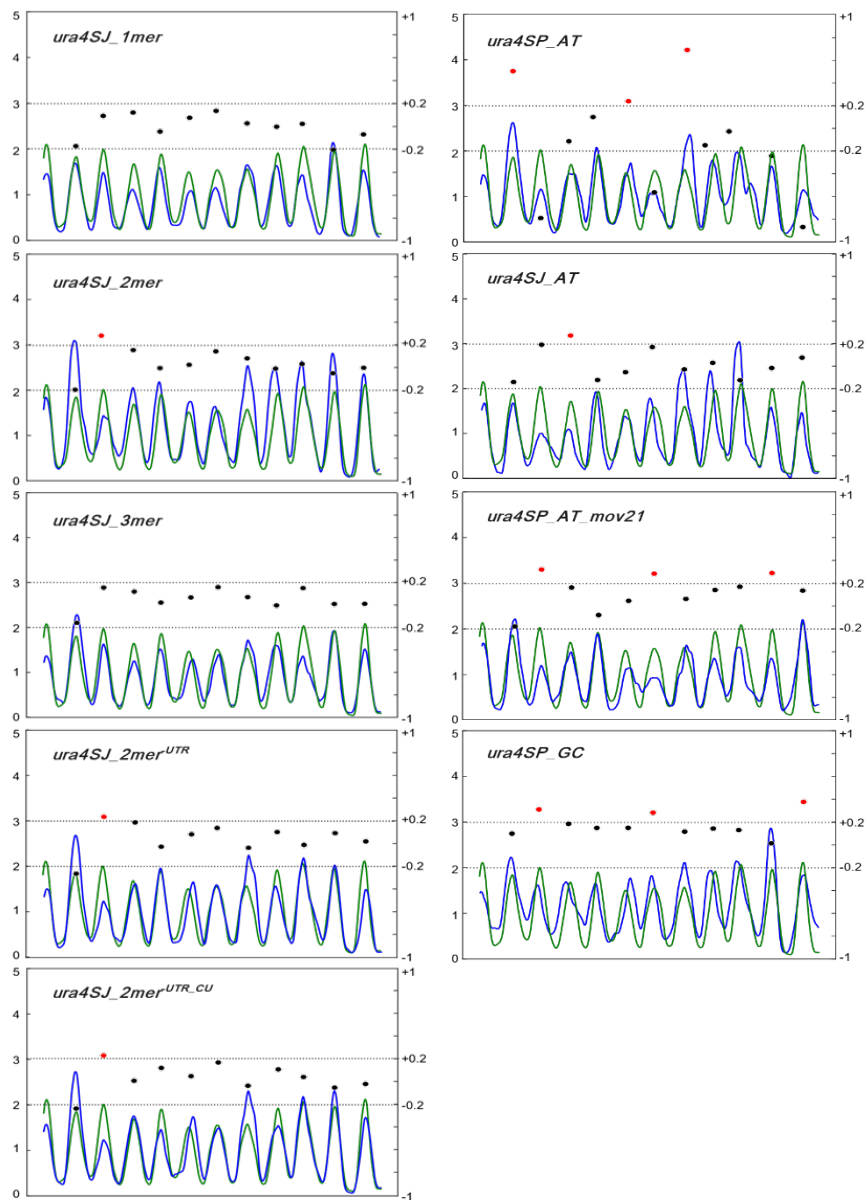
<b>A</b>	Adenina
<b>AID</b>	Sistema de <i>degrón</i> inducible por auxinas
<b>ARP</b>	Proteínas relacionadas con la actina
<b>ATP</b>	Adenosín trifosfato
<b>C</b>	Citosina
<b>cDNA</b>	ADN complementario
<b>DEPC</b>	Dietil Pirocarbonato
<b>DNA</b>	Ácido Desoxirribonucleico
<b>dNTP</b>	Desoxinucleótido trifosfato
<b>DO</b>	Densidad óptica
<b>DSB</b>	Roturas de doble cadena
<b>EDTA</b>	Ácido etilen diamino tetra acético
<b>G</b>	Guanina
<b>GO</b>	Grupo de ontología
<b>G418</b>	Geneticina
<b>h</b>	Horas
<b>H3K4</b>	Lisina 4 de la histona H3
<b>H3K9</b>	Lisina 9 de la histona H3
<b>IAA</b>	Ácido indolacético
<b>IGR</b>	Región intergénica
<b>kb</b>	Kilobase
<b>L</b>	Litro
<b>log</b>	Logaritmo
<b>LTR</b>	Repetición terminal larga
<b>M</b>	Molar
<b>mA</b>	miliAmperios
<b>mg</b>	Miligramo
<b>min</b>	Minutos
<b>ml</b>	Mililitros
<b>MM</b>	Medio mínimo
<b>mM</b>	Milimolar
<b>MNasa</b>	Nucleasa micrococcal
<b>mRNA</b>	ARN mensajero
<b>NDR</b>	Región libre de nucleosomas
<b>ng</b>	Nanogramos
<b>nm</b>	Nanómetros
<b>°C</b>	Grado Celsius
<b>O/N</b>	Durante la noche
<b>ORF</b>	Marco abierto de lectura
<b>pb</b>	Pares de bases
<b>PCR</b>	Reacción en cadena de la polimerasa
<b>PIC</b>	Complejo de preiniciación
<b>Pol</b>	Polimerasa
<b>PSWM</b>	Matriz de puntuación de posición específica
<b>qPCR</b>	PCR cuantitativa

<b>kDa</b>	kiloDalton
<b>RNA</b>	Ácido Ribonucleico
<b>RNAPII</b>	RNA polimerasa II
<b>rpm</b>	Revoluciones por minuto
<b>rRNA</b>	ARN ribosomal
<b>RT</b>	Temperatura ambiente
<b>sAID</b>	AID corto
<b>SDS</b>	Dodecil sulfato sódico
<b>SR</b>	Ratio de síntesis
<b>seg</b>	Segundos
<b>SEM</b>	Desviación estándar de la media
<b>Seq</b>	Secuenciación
<b>T</b>	Timina
<b>TAE</b>	Tris/Acetato/EDTA
<b>TCA</b>	Ácido tricloroacético
<b>TE</b>	Tris/EDTA
<b>TF</b>	Factor de transcripción
<b>tRNA</b>	RNA de transferencia
<b>TSS</b>	Sitio de inicio de la transcripción
<b>TTS</b>	Sitio de terminación de la transcripción
<b>U</b>	Unidades
<b>UTR</b>	Región no traducida
<b>V</b>	Voltios
<b>WB</b>	<i>Western Blot</i>
<b>YES</b>	Medio de extracto de levadura
<b>5-adamantyl-IAA</b>	Análogo de la auxina
<b>µg</b>	Microgramos
<b>µl</b>	Microlitros

# **ANEXO I**



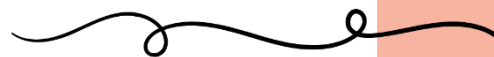




**Figura Suplementaria 1: Comparación de la distribución nucleosómica en una región control no modificada entre las diferentes versiones modificadas del gen *ura4* y su versión silvestre en *S. pombe*.** Comparación de la borrosidad de DANPOS de una región genómica del cromosoma II de 1.6 Kb comprendida entre las coordenadas 2171850 y 2173530, de los mapas empleados en las Figuras 9, 10, 11, 13 y 15. El perfil de nucleosomas en verde corresponde al de la cepa del gen *ura4* de *S. pombe* silvestre. Los distintos perfiles de nucleosomas en azul corresponden a cada una de las cepas modificadas del gen *ura4*. Los símbolos y las escalas utilizados son iguales que los de la Figura 7.



# **ANEXO II**





### Construcciones de la Figura 9

#### *ura4SJ\_1mer*

Integrado en *S. pombe* (Chr3:115780..116575)

TACCAAGAACCCCTCTTCTGCCTCGACCCGCAACTAGCGCTTCAAGAGCAAAGTA**ATG**AGTGATATAGCATTAA  
 AAACATATACTGAAAGAGCTAATGTTTCATCCGAACGCCGTCGCGAAGAAGCTGCTGCGGCTGATGGATGAGA  
 AGAAGTCGAACCTCTCCGTCGCCGTCGACCTACCAAGAAGAACCAGGTGCTCGAGCTCGTAGACAAAATAG  
 GACCAAGTATATGTTTATAAAAACTCATATTGATATTGTGGAGGACTTCGACGCGGACATGGTGCAGCAGTTG  
 GTGGCGTTGGCGGAGAAAACACAAATCCTCATCTTCGAGGACCGCAAGTTCGCCGACATCGGCAACACAGTA  
 AAATTACAATATAGTGCAGGAGTTTATAAAATTGCTTCTTGGGCTGACATACCAACTGCCACACGGTGCCGG  
 GGGAGGGGATTATTCGGGGCTGAAAGAAGTCGGCCTCCCCCTCGGCCGCGGCCTCCTCCTCTGGCCGAG  
 ATGAGCTCAAAGGAACATTAGCTACAGGATCTTACTCAAGCTACTTTGGAGTTGGCCGAGAAGCACAAAC  
 GACTTCTGCATGGGGTTTATTGCGCGGCGGGGTTCCCCGGCCTCAAATCCGACTTCATCCACATGACCCCCG  
 GGGTCGGCCTCGACGTAAAGGAGATGGATTAGGACAACAATATAGAACTCCTGAGGAGGTTATCTGCGAGT  
 CCCAGTCGGACATCATCATTGTGGGGCGGGGGTGTATGGGTCCGGCCGCGACGCCGCCAAGAGGCCGAG  
 CGCTACCGCAAGGCCGGCTGGGAAGCATATCAAAGAAGAATTTCAAACAATA**TAA**AAAAAACTAATGTTAA  
 GTTCTTCTGGCTCGTGATCGAGAAGGTGGATGCGTTGTTGCGTTTGTGTTT

#### *ura4SJ\_2mer*

Integrado en *S. pombe* (Chr3:115780..116575)

TACCAAGAACCACTCTTCTGCCTCGACCCGCAACTAGCGCTTCAAAAGTAAAGTA**ATG**TCAGATATAGCTTTAA  
 AAACCTATACTGAAAGAGCTAATGTGCATCCGAACGCCGTTGGCGAAGAAGTTGTTGCGGTTGATGGATGAGA  
 AGAAGTCCAACCTCAGCGTCGCCGTCGACCTACCAAGAAGAACCAAGTACTAGAATTAGTAGATAAAAATAG  
 GTCCTAGTATATGTTTATAAAAACTCATATTGACATCGTGGAGGACTTCGACGCGGACATGGTGCAGCAGTTG  
 GTGGCGTTGGCGGAAAAACACAAATCCTCATCTTGAAGACCGCAAGTTCGCAGACATAGGAAATACAGTA  
 AAATTACAATATAGTGTGGTGTATATAAAATTGCTTCTTGGGCGGACATCACGAACTGTCACACGGTGCCGGG  
 TGAGGGTATTATTAGTGGGTAAAAGAAGTCGGCCTCCCCCTCGGCCGCGGCCTCCTCCTACTAGCAGAAATG  
 AGTAGTAAAGGACTTTAGCTACTGGTCTTATACTCAGGCTACTCTGGAGCTGGCGGAGAAGCACAAATGATT  
 TTTGTATGGGGTTTATTGCACGTAGGCGATTCCCCGGCCTCAAAGCGACTTCATCCACATGACACCAGGAGT  
 AGGATTAGATGTAAAAGGAGATGGATTAGGTCAACAATATAGAACTCCTGAGGAGGTGATCTGCGAGTCGCA  
 GTCGGATATTATTATTGTGGGTAGGGTGTGTACGGCAGCGGCCGCGACGCCGCGCAAGAGGCAGAGCGCT  
 ACAGAAAAGCAGGATGGGAAGCTTATCAAAGAAGAATATCTAAACAATA**TAA**AAGAAGACTAATGTGAAGTTC  
 TTCTGGTTGGTATTGAGAAGGTGGATGCGTTGTTGCGTTTGTATTTC

#### *ura4SJ\_3mer*

Integrado en *S. pombe* (Chr3:115780..116575)

TATCAAGAACCACTCTTCTGCCTGGACCCGAAACTAGCGATTCAAAAGCAAGGTT**ATG**AGCGACATTGCATTA  
 AAGACTTATACTGAAAGAGCTAATGTTTCATCCTAACGCCGTCGCTAAGAAGCTGCTGCTGATGGACGAAA  
 AGAAGTCCAACCTCTCAGTCGCGGTCGATCTGACAAAAAGAACCAAGTCTCGAGCTCGTTGATAAAATTG  
 GACCTTCTATATGTTTATAAAGACTCATATCGATATCGTGAAGACTTCGATGCGGACATGGTGCAGCAATTGG

TGGCTTTGGCAGAGAAGCATAAGTTCCTCATCTTCGAGGACCGAAAGTTCGCAGATATCGGAAATACAGTAA  
 AACTTCAGTATTCTGCTGGAGTATATAAGATTGCTTCTTGGGCTGATATCACGAACTGCCATACGGTGCCGGGT  
 GAAGGTATCATTTGGGTTTGAAGGAAGTCGGCTCCCACTCGGCCGAGGCCTGCTGCTCTGGCGGAGATG  
 AGCTCAAAGGAACATTAGCAACAGGATCTTATACTCAGGCTACTCTGGAGCTGGCTGAAAAGCACAAACGAC  
 TTCTGCATGGGTTTCATCGCTCGGAGGCGATTCCCAGGACTAAAATCGGATTTTCATCCACATGACACCCGGAG  
 TCGGACTCGATGTAAGGAGATGGATTAGGACAACAATATAGAACTCCTGAAGAGGTGATCTGCGAATCGC  
 AGTCGGACATCATCATCGTTGGTCGTGGTGTATGGTTCAGGACGAGACGCGGCCAGGAGGCCGAACGAT  
 ACCGAAAGGCGGGATGGGAAGCATATCAGAGAAGAATTTCTAAGCAAT**TAA**AAGAAAACAATGTTAAGTTC  
 TTCTGGCTGGTCATCGAAAAGGTGGACGCTTTGTTTCGCTTTCGTGTTTC

### Construcciones de la Figura 10

#### *ura4SJ\_2mer<sup>UTR</sup>*

Integrado en *S. pombe* (Chr3:115780..116575)

**ATG**TCAGATATAGCTTTAAAACTTATACTGAAAGAGCTAATGTGCATCCGAACGCGGTGGCGAAGAAGTTGT  
 TGCGTTTGATGGATGAGAAGAAAAGTAACCTATCAGTCGCAGTCGACCTCACGAAGAAGAACCAAGTCCTAG  
 AATTAGTAGATAAAAATAGGACCTTCTATATGTTTATTAATACTCATATTGATATTGTGAGGACTTCGACGCGG  
 ACATGGTTCAGCAGTTGGTTGCTCTGGCAGAAAAACACAAATTCCTCATCTTGAAGACCGCAAGTTCGCGG  
 ACATAGGAAAATACAGTAAAATTACAATATTCTGCTGGAGTTTATAAAATTGCTTCTTGGGCTGACATCACGAAC  
 TGTCACACTGTGCCTGGTGAGGGTATTATAAGCGGACTAAAAGAAGTAGGACTCCCACTCGGCCGCGGCCTC  
 CTCCTCTAGCAGAAATGAGTAGTAAAGGAACTTTAGCTACTGGATCTTATACTCAAGCTACTCTTGAGCTCGC  
 CGAGAAGCACAATGATTTTTGTATGGTTTTATTGCTAGAAGAAGATTCCCAGGACTCAAAGCGACTTCATC  
 CACATGACGCCGGGAGTAGGATTAGATGTAAGGAGATGGATTAGGTCAACAATATAGAACTCCTGAAGAG  
 GTGATCTGCGAGTCGCAGTCGGACATTATTATTGTTGGTCGTGGTGTCTACGGATCAGGAAGAGACGCAGCAC  
 AAGAAGCCGAGCGCTACCGAAAAGCAGGATGGGAAGCATATCAAAGAAGAATATCTAAACA**TAA**

#### *ura4SJ\_2mer<sup>UTR.CU</sup>*

Integrado en *S. pombe* (Chr3:115780..116575)

**ATG**AGTGATATAGCATTAAAACTTATACTGAAAGAGCTAATGTTTCATCCTAATGCTGTGGCTAAAAAGTTGTT  
 GCGTTTGATGGATGAAAAGAAATCAAATCTATCAGTCGCAGTCGACCTCACAAAAAAGAACCAAGTACTAGA  
 ATTAGTAGATAAAAATAGGACCATCAATATGTTTATTAATACTCATATTGATATTGTGGAGGATTTTGATGCTGAT  
 ATGGTGCAACAGTTGGTAGCTCTGGCAGAAAAACATAAATTCCTCATCTTGAAGACAGAAAGTTCGCAGAC  
 ATAGGAAATACAGTAAAATTACAATATTCTGCTGGAGTATATAAAATTGCTTCTTGGGCTGATATTACGAATTGT  
 CATACGGTGCCTGGTGAAGTATTATAAGCGGATTAAGAAGTTCGGACTCCCACTCGGCAGAGGACTGCTT  
 TTGCTTGCAGAAATGAGTAGTAAAGGAACATTAGCAACTGGATCTTATACTCAAGCTACTTTGGAGTTGGCTG  
 AGAAGCATAATGATTTTTGTATGGTTTTATTGCTAGAAGAAGATTCCCAGGACTCAAATCAGACTTCATCCAC  
 ATGACACCAGGAGTAGGATTAGATGTAAGGAGATGGATTAGGACAACAATATAGAACTCCTGAAGAAGTG  
 ATTTGCGAGTCGCAGTCGGATATTATTATTGTTGGTAGAGGTGTCTACGGATCAGGAAGAGACGCAGCACAAG  
 AAGCAGAACGATACAGAAAAGCAGGATGGGAAGCATATCAAAGAAGAATATCAAACA**TAA**

**Construcciones de la Figura 11*****ura4SP\_AT***

Integrado en *S. pombe* (Chr3:115780..116575)

**ATG**GATGCAAGAGTATTTC AATCTTATTCTGCTCGTGCTGAGGGTATGAAAAATCCTATTGCTAAGGAATTGTT  
GGCTTTGATGGAAGAAAAGCAAAGCAACTTATCAGTCGCAGTCGATTTAACAAAAAATCCGAAATCTTAGA  
ATTAGTAGATAAAATAGGACCCTATGTTTGTGTTATTAAGACTCATATTGATGTTGTTGAGGATTTTGATCAGGAT  
ATGGTTGAAAAATTAGTTGCTTTAGGAAAAAACATCGATTTTTAATCTTTGAAGATAGAAAATTCGCAGACAT  
AGGAAATACCGTCAAAC TACAATATGCTTCTGGTGTTTATAAAATTGCTTCTTGGGCTCATATTAATAATTGTCAT  
ACTGTTCTGGTGAGGGTATTATTCAAGGCTTAAAAGAAGTAGGATTACCATTAGGACGAGGATTATTATTATT  
AGCAGAAATGTCATCCAAAGGCTCATTAGCTACTGGTCTTATACTGAGAAACTTTAGAATGGTTTGAGAAG  
CATACTGATTTTTGTTTTGGTTTTATTGCTGGTAGACGATTTCCAAACTTACAAAGCGACTACATAACAATGCCC  
CCAGGAATAGGCTTAGATGTAAGGAGACGGATTAGGTCAGCAATATCGTACTCCTGAAGAAGTTATTGTTA  
ATTGTGTTCTGATATTATTATTGTTGGTCGTGGTGTCTATGGAGCAGGACGAAATCCAGTAGTCGAAGCCAAA  
AGATATAGAGAAGCAGGATGGAAAGCATATCAACAAAGACTTTCTCAGCATT**A**

***ura4SJ\_AT***

Integrado en *S. pombe* (Chr3:115780..116575)

**ATG**TCAGACATAGCATTGAAACTTATACTGAGCGTGCTAATGTTTCATCCTAATGCTGTTGCTAAGAAGTTGCT  
TCGTTTGATGGATGAGAAGAAATCAAACCTATCAGTAGCAGTAGACTTAACAAAAAACCAGTACTAGA  
ACTAGTAGACAAAATAGGACCATCTATTTGTTTGTGAAGACTCATATTGATATTGTTGAGGATTTTGATGCTGA  
TATGGTTCAACAGCTTGTGCTCTGGCAGAAAACACAAATTTTTAATATTTGAAGATAGAAAATTTGCAGAC  
ATAGGAAACACAGTAAACTACAATACTCTGCTGGTGTTTATAAGATTGCTTCTTGGGCTGATATTACTAATTGT  
CATACTGTTCTGGTGAGGGTATTATTAGCGGATTAAGAAGTAGGACTACCATTAGGAAGAGGATTATTACT  
ATTAGCAGAAATGTCATCAAAGGAACACTAGCTACTGGTCTTATACTCAAGCTACTTTGGAATTGGCTGAG  
AAGCATAATGATTTTTGTATGGTTTTATTGCTAGAAGAAGATTCCCAGGACTAAAAGCGACTTTATACACAT  
GACACCAGGAGTAGGACTAGACGTAAGGAGATGGACTTGGTCAACAATATCGTACTCCTGAGGAAGTTAT  
TTGTGAGTCTCAGTCTGATATTATTATTGTTGGTCGTGGTGTCTACGGAAGCGGAAGAGATGCAGCACAAGAA  
GCAGAAAGATACAGAAAAGCAGGATGGGAAGCATAACAAAGAAGAATTTCTAAGCAGT**A**

**Construcciones de la Figura 12*****Sec.artificial\_AT***

Integrado en *S. pombe* (Chr1:2753729)

CACATAATATATGGAATAGACACATACTTCTAATACCCACTATTATACATATTAACATCAAGAGTAAGAAAC  
ATTTTTCAGTATTATTGTTGGAAGAAGAAGCCTCGTCGTGTACTTTTATTCATGTTTCTGGTTTGTCTATACGTA  
GTAAGTGGAGATTCTATCCAGTACTAAAATACGTACCAACAACAATAGCATGCTCACTAAGCAAAGTATCAAG  
CCTACGTTCTTTGACTTTTGTGAGAAGTATGTTGGTGAGATACATCTTAAGTTGGGTCGTGTTACTCAGTTGAT  
TTTGTACGGACTACTATTAGAACCAAAAACACGACAGCATAAGACAATACTCATGCCAATACTACACAGACTGG  
GCAGCAGCAATTTCTTATCAGGTTGGTTTGTATCCTTATACTAGTACTATCTCTGTTTTTCTCATCGTTCTGGTAA

TTGTGCTTTGAGATATACATGCGCACTATGGCTAGTAACAACTGCAGAGACAGCTACCCAGGAATTACATGG  
 TTATTCAGCGTACCTTCTGCTGCTCAGCAGATTTGTAAGACTTGGACTTTGGTTGGTCATCCTGCTACTTTGCGT  
 ATTCATACTGCTGGAGGAAAAAGAATTTAAAATACGTAAGCTCATATGCAGACTGGGGAGAATGGCCAGGA  
 CCAATACCATTATACAGAATAAGGGTCTGATCATCATTCTCTTTCTCAGCTTCAGCTTACTATTTATGAGAATT  
 CTTTTTGAATCTTTTGACCAAGTAATAGTAGCATACTCATATTAGTAACAAGCCACGAATGGGTATGGTCAC  
 CACTAGTAAGCAGATCAACAAAGCGTAATGTTGATCGTTGTCGTTGACTTTGGATGGTTCTCCTTCTGAGCTT  
 GTTCGTCCTTCTTTGAGCGTCTG

### *Sec.artificial\_control*

Integrado en *S. pombe* (Chr1:2753729)

CATATAATCTATGGAATTGATACATACTTTCTGATACATCCTCTGTTGTACATATTGACTTCTCGCGTCCGAAATAT  
 TTCCAGTATTATTGCTGGAAGAAAAACCTCGTCGCGTTACTTTCATACATGTCTCAGGGCTAAGTATGCGTA  
 GTAACCTGGCGGTTCTATCCGGTACTTAAATACGTTCCAACAACCATCGCTTGTCACTCAGCAAAGTGTCCAGT  
 CTCAGATCGTTAACGTTCTGTTGGAGAAATATGTTGGAGAGATGCATCTCAAGTTAGGAAGGGTTACCCAATAA  
 TCTTATACGGTCTACTTTTGGAGCCGAAACATGATAGTATACGCCAGTACTCCTGCCAATACTACACTGACTGG  
 GCAGCGGCGATCTCATATCAGGTAGGCTTATACCCGTACACAAGCACTATGTCGGTATTTTCACATAGGTCTGG  
 CAACTGTGCTTTGAGATATACATGCGCTCTGTGGCTGGTAACTAACTGCCGCGACAGCTACCCCGGGATGACT  
 TGGTATTTTTCGGTCCCGTCTGCTGCGCAGCAGATATGCAAAACGTGGACTTTGGTCCGCCACCCGGCCACAT  
 TGCGGATTCATACCGCTGGAGGTAAGCGCATGCTAAAATACGTTTCTCGTATGCCGACTGGGGAGAATGGCC  
 GGGGCCATTCCATTGTACCGAACCAAAGGACCTGATCACCATTCTGTCTCAGCTGCAACTTACCATTATG  
 AGAATAGCTTCTGTAACCTCTTCGACCAGGTTATAGTAGCATACTCCTCGCTCGTTACTTCACACGAGTGGGTG  
 TGGTCACCTCTTGTGAGTCACTAAACGCAATGTTGATAGGTGCCGTTGTACGTTAGACGGTAGTCCATC  
 GGAACCTGGTCAGACCTAGCTTTGAACGTCTG

### **Construcciones de la Figura 13**

#### *ura4SP\_AT\_mov21*

Integrado en *S. pombe* (Chr3:115780..116575)

TTGTCTACCAAGAACCTCTTTTTGCTTGGATCGAAATTAAGGTTTAAAAGCAAAGTT**ATG**GACGCAAGAG  
 TATTCAAAGCTACTCAGCAAGAGCTGAGGGTATGAAAAATCCTATTGCTAAGGAATTGTTGGCTTTGATGGA  
 AGAAAAGCAAAGTAATTTGTCTGTTGCGGTAGACTTAACAAAAAATCAGAAATACTAGAATTAGTAGACAA  
 AATAGGACCATACGTATGCGTAATAAAAAACACATATTGATGTTGTTGAGGATTTTGTATCAGGATATGGTTGAAA  
 AACTTGTGCTTTAGGTA AAAAGCATCGTTTTCTTATATTGAAAGACCGAAAATTCGCAGACATAGGAAACAC  
 AGTAAACTACAATACGCATCAGGAGTATACAAAATAGCTTCTTGGGCTCATATTACTAATTGTCATACTGTCC  
 TGGTGAGGGTATTATTCAAGGTCTTAAAGAAGTTGGTTTACCATTAGGACGAGGACTTACTATTAGCAGAA  
 ATGTCATCAAAGGATCATTAGCAACAGGATCATAACAGAAAAAATTTAGAATGGTTTGAAGCATACTG  
 ATTTTGTCTTTGGCTTTATAGCTGGTCGTCGATTTCTAATCTTCAAAGCGACTACATAACAATGTCACCAGGAA  
 TAGGATTAGACGTA AAAAGGAGACGGACTAGGACAACAATACCGAACACCTGAAGAAGTTATTGTTAATTGTG  
 GTAGTGATATTATTATTGTTGGTCGTTGTTTATGGTGCTGGTCGTAATCCTGTAGTAGAAGCAAAAAGATACA

GAGAAGCAGGATGAAAAGCATACCAACAAAGACTATCACAACACT**TA**AAAAAAGACTAATGTAAAATTTTTT  
 GGTTGGTTATTGAAAAAGTCGATGCCTTGTTCGCTTTGTTTCCTAGGCGTTTTATGTCAGAAGGCATTAGA  
 ATTAGTATACAAGTACTCTTTGGTAAAATTTATGTAGCGACTAAAATATTAAGTATTATAGATAAACACCTGGG  
 AATAAAAAGTAATTTGCTATAGTAA

### *mug166SP\_AT\_mov21*

Integrado en *S. pombe* (Chr2:1247290..1248400)

ATTTGGTGCTAAAACCAACTACTACTAGGAACACTAAGAGGATCAAGCAACTCAACCGATTCTTCTTATTAC  
 ACACATACAAGGAGCATTAAAA**ATG**CCTAGGAGAGCTGTTCTTTTCGTGGTAATTCTTTTATTCTTTTCTTC  
 TATTCTCTGTAATGAGAGTAGCGCACACGGAGACGCATCACTAGGACAAGCAGACAAACCACTAACAGT  
 AGACGTAAGACGCATGGACAACACAATTTACTAAGAAGCAGACTAGGGTTTTTCATGAGGCTTATGTTAGT  
 GCTTATATTGATTTGCAGCAGAAGTATGCTCGACTAGAAAGACTACAAGAACAATACTAAACTTAGAAAGAT  
 CACTATTAAGTACAAATCACTAAAAAGCAACTCACTACTACTGTTGCTACTGTTCTTCTGAGAAGCGTTAT  
 TTTTTTCGTGTAAGCCTAAGACTTCTTCTACTACTAATCACTATCATCATCAAGCGTATCACACATGA  
 ACCAAGACGCAAACCCAAACAGCACAGTATTAACATCACTGCTGTTCTTATGATTCTTGTTCCTGCTCTCT  
 GATCGTCATACTCCTTCTAGTGCTTCTTAGGGCTAGTGAACAGGAACAACAAGCCACAAAGCATGAGA  
 AACCAAATAAGCCTACTATACTCAAAGTAGACCAAGTAAAAACAGAGATTGCTTCTGCTCAGGCTGTTATGC  
 ATGATTTGTTGCGTACTTATCCTGGTAGTGAGAGTTGCGAGCAGTCTACTCACACACAAT**TGAT**GCCCATAC  
 ACAACAAAAACTCACTATACTTCTAATTCTCAGTACTATCAGCATTCTCACTTGAGATGAATATTTTTGTTAG  
 GTTGTTTTTAGAGAATTCGTTGTCTATGGATGTTCAATTGATGAATTGAT

### Construcciones de la Figura 15

#### *ura4SP\_GC*

Integrado en *S. pombe* (Chr3:115780..116575)

**ATG**GACGCGCGCTGTTCCAGTCTACAGCGCGCGCGGAGGGCATGAAGAACCCGATCGCCAAGGAGC  
 TGCTGGCGCTGATGGAGGAGAAACAGTCCAACCTGAGCGTCGCGGTCGACCTGACCAAGAAGAGCGAGATC  
 CTGGAGCTGGTCGACAAGATCGGGCCCTACGTCTGCGTCATCAAGACGCACATCGACGTCGTGGAGGACTTC  
 GACCAGGACATGGTCGAGAAGCTGGTCGCCCTCGGCAAGAAGCACCGGTTCTCATCTTCGAGGACCGGAA  
 GTTCGCGGACATCGGCAACACCGTCAAGCTGCAGTACGCGAGCGGCGTCTACAAGATCGCGAGCTGGGCGC  
 ACATCACGAACTGCCACACCGTGCCGGGCGAGGGCATCATCCAGGGCCTGAAGGAGGTCGGCCTGCCGCTC  
 GGCCGGGGCCTGCTCCTGCTGGCCGAGATGAGCAGCAAGGGCAGCCTGGCGACGGGCGAGCTACACGGAG  
 AAGACCCTGGAGTGTTTCGAGAAGCACACCGACTTCTGCTTCGGGTTTCATCGCGGGCCGGCGGTTCCCGAA  
 CCTGCAGTCCGACTACATCACGATGAGCCCGGCATCGGGCTGGACGTCAAGGGCGACGGCCTGGGCCAGC  
 AGTACCGGACCGGAGGAGGTCATCGTCAACTGCGGCTCCGACATCATCATCGTCGGCCGGGGCGTATACG  
 GCGCCGGCCGAACCCGTCGTGGAGGCCAAGCGCTACCGCGAGGCCGGCTGGAAGGCGTACCAGCAGCG  
 CCTGAGCCAGCACT**A**





