



UNIVERSIDAD  
DE SALAMANCA

CAMPUS DE EXCELENCIA INTERNACIONAL



CSIC

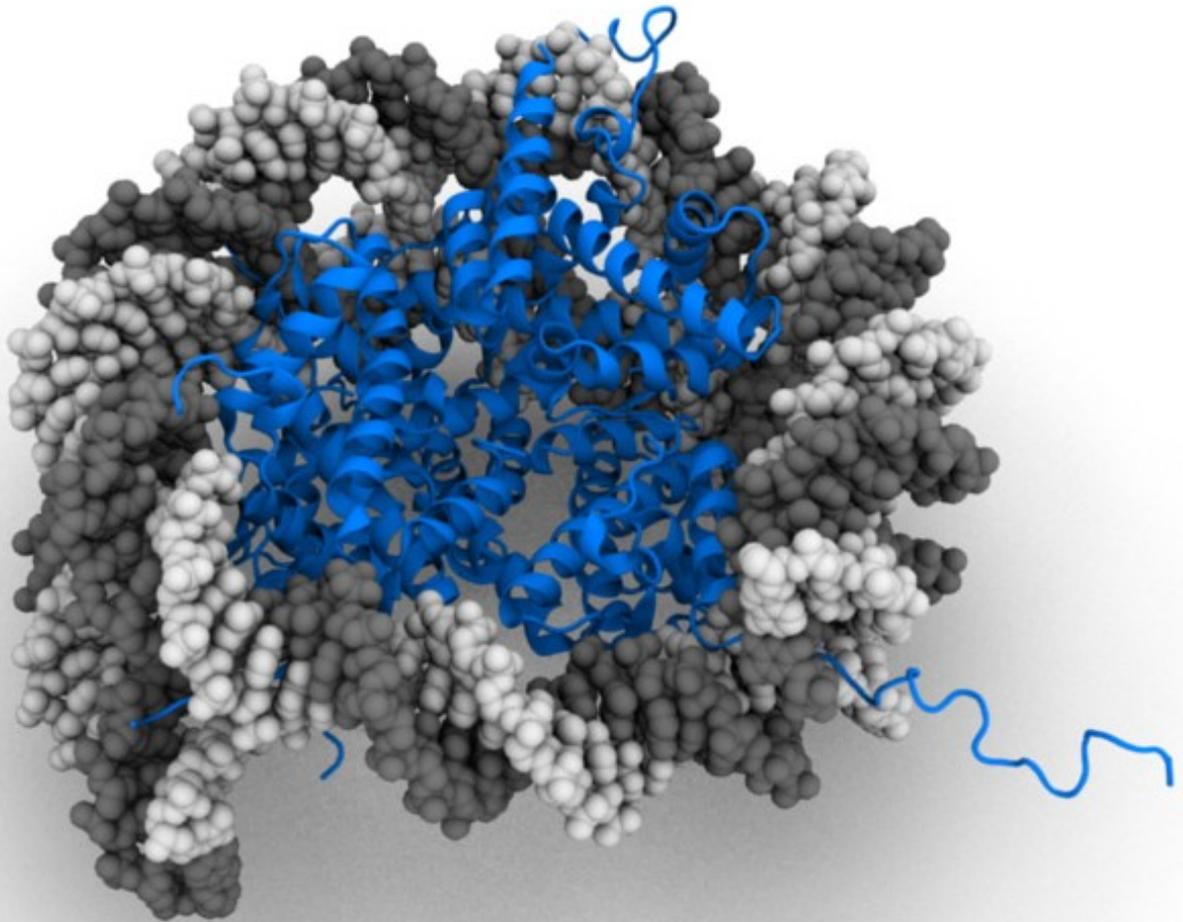
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

UNIVERSIDAD DE SALAMANCA

Instituto de Biología Funcional y Genómica



# Contribución de la secuencia del DNA al posicionamiento de los nucleosomas



Sara González Moreno

2017





**UNIVERSIDAD DE SALAMANCA**

**Instituto de Biología Funcional y Genómica**



## Contribución de la secuencia del DNA al posicionamiento de los nucleosomas

Sara González Moreno  
2017



**El Dr. Luis Antonio Miguel Quintales, Catedrático de Lenguajes y Sistemas Informáticos del Departamento de Informática y Automática de la Universidad de Salamanca**

CERTIFICA:

Que la memoria titulada “Contribución de la secuencia del DNA al posicionamiento de los nucleosomas” presentada por la licenciada Dña. Sara González Moreno para optar al grado de Doctor en Biología por la Universidad de Salamanca, ha sido realizada bajo la dirección del Dr. Francisco Antequera Márquez 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      de febrero de 2017.

Fdo. Luis Antonio Miguel Quintales

---

**El Dr. Francisco Antequera Márquez, Profesor de Investigación del Consejo Superior de Investigaciones Científicas**

CERTIFICA:

Que la memoria titulada “Contribución de la secuencia del DNA al posicionamiento de los nucleosomas” presentada por la licenciada Dña. Sara González Moreno para optar al grado de Doctor en Biología por la Universidad de Salamanca, 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, firma el siguiente certificado en Salamanca, a      de febrero de 2017.

Fdo. Francisco Antequera Márquez



Esta tesis doctoral ha sido financiada por una beca predoctoral del Programa Nacional de Formación de Personal Investigador (2012/2016) del Ministerio de Economía y Competitividad del Gobierno de España y por proyectos del Plan Nacional de I+D.



*“Así como los brotes, por crecimiento, dan origen a nuevos brotes, y éstos, si son vigorosos, se ramifican y sobrepujan por todos lados a muchas ramas más débiles, así también, a mi parecer, ha ocurrido, mediante generación, en el gran Árbol de la Vida, que con sus ramas muertas y rotas llena la corteza de la tierra, cuya superficie cubre con sus hermosas ramificaciones, siempre en una nueva división”*

*Charles Darwin*



## AGRADECIMIENTOS

Estas cien páginas recogen los resultados de varios años de trabajo y, aunque en la portada de esta tesis sólo aparezca mi nombre, nunca se habría hecho realidad sin la colaboración de muchas personas que durante esta etapa me han acogido, guiado, instruido, apoyado y animado.

En primer lugar, quisiera dar las gracias a Paco, quien no sólo ha contribuido notablemente a mi formación científica, sino también personal. Gracias por haberme revelado la enseñanza más importante que alguien puede recibir: que todo se puede aprender y que las palabras son las llaves que abren todas las puertas. Gracias por contagiarme cada día tu infinito entusiasmo por la ciencia, por escuchar y tener en cuenta todas mis sugerencias, incluso las más absurdas y en definitiva, por haber transmitido tanta confianza en aquella niña tímida e insegura que llegó a tu laboratorio hace ya cinco años sin saber muy bien qué decir.

Gracias también a Luis Quintales, no sólo porque sin sus matrices esta tesis sería mucho menos interesante, sino porque de manera totalmente desinteresada, me ha ilustrado y guiado en mis primeros pasos dentro del "misterioso mundo de la informática" que ahora, gracias a él se me antoja algo menos misterioso.

A todo el personal investigador porque sus comentarios han contribuido a enriquecer este trabajo. Gracias en especial a Mónica y a Rodrigo por todas las ideas que han aportado a lo largo de estos años en los seminarios de grupo.

A todo el personal técnico y administrativo del IBFG por su ayuda y su excelente labor que tanto me ha facilitado el trabajo.

Por supuesto, gracias a todos mis compañeros del laboratorio 2.5, a los que siguen aquí y a los que se marcharon. A Jose y Laura Marín, me siento afortunada por haber tenido la oportunidad de aprender de vosotros aunque fuese solo por un periodo breve de tiempo. Gracias a Nacho, que me acogió desde el primer día, según sus palabras "como me gustaría que me recibiesen a mí". A Rebeca, que me explicó pacientemente todo lo que necesitaba y respondió amablemente mis interminables preguntas. A Mar, cuya labor incansable evita que el laboratorio se convierta en una jungla, gracias por resolver todas mis dudas con precisión y eficacia, por abrirme las puertas de tu "despacho-confesionario" y por compartir risas y lamentos. A los que estuvieron temporalmente: a Alba, que me acompañó en mis primeros "tropiezos" con las transformaciones, a Pau, Dani, Raquel, JoseVi y a las nuevas incorporaciones, Aroa y Laura, gracias por vuestro entusiasmo y vuestras ganas. A Alberto, por animar las mañanas con tus chistes y por poner banda sonora al laboratorio con tus canturreos. A Quique, por estar siempre disponible para los tan necesarios análisis informáticos y por haber evitado que me perdiera en el caos durante la revisión del artículo.

Gracias a Alicia porque sin su ayuda esta tesis hubiese sido mucho más breve. Gracias por haberme permitido descubrir a tu lado todas las maravillas del trabajo en equipo, sin encontrar en el camino ningún inconveniente. Por compartir tu peculiar sentido del humor y por ser capaz de leer mi mente cuando no encuentro las palabras adecuadas. No puedo imaginar cómo hubiesen sido estos años si no hubieses estado a mi lado para compartir el peso de los experimentos fallidos y celebrar los acertados. Gracias por ser una de esas personas que se convierten irremediabilmente en ejemplos a seguir.

Al grupo de genética de la UNEX, con el que realicé el TFM que me abrió las puertas al mundo de la investigación. Gracias en especial a Elena, que me guió en mis primeros pasos dentro de un laboratorio. Gracias por haber confiado en mí. Gracias también a Encarna, Mari Carmen y a Belén por acogerme como a una más y por haber llenado de risas mis primeros contactos con el laboratorio. Emilia, José Emilio, Felipe, Helen... Muchas gracias por haber creado ese entorno de trabajo ideal que inspira y anima a seguir adelante incluso después de tanto tiempo.

Gracias a "Los elegidos +++" por estar siempre dispuestos a compartir unas cañas después del trabajo, por escuchar las penas, dudas y alegrías y por las maravillosas jornadas "Eno-gastro-culturales" que de culturales tenían más bien poco. A Parri, Rosana, Marta, Sandra, Álvaro, Alicia, Esther, Facundo y Sofi por convertir la sala de descanso en un refugio lleno de chistes y buenos consejos. A Esther, por haberse tomado el tiempo necesario para conocerme de verdad, que no es poco, por estar siempre dispuesta a escuchar y por sacarme de vez en cuando de mi introversión para despejar la mente con unas risas.

Gracias a todas esas personas que, a pesar de la distancia, siguen estando ahí cuando las necesito. Gracias en especial a Inma, Mari Jose y Ana, porque a pesar del tiempo, a vuestro lado todo sigue siendo igual. No os imagináis lo afortunada que me siento por poder disfrutar de vuestra compañía aunque sólo sea de vez en cuando. Gracias también a mis "amebas" que se iniciaron conmigo en el fascinante mundo de la biología, a pesar de que nuestros pasos nos hayan llevado por caminos distintos, sé que aún puedo contar con vosotras, ojalá sigamos organizando quedadas durante muchos años más.

Por último, gracias a toda mi familia por estar siempre ahí cuando los necesito. Gracias a Maribel, Leo, Juan, Jorge, Lorenzo y José, por seguir interesándose por lo que hago aunque nunca sea capaz de explicárselo bien. A mi abuela Cándida, por cuidar siempre de mi, esperar despierta mi llegada los viernes por la noche y "echarme de casa" los domingos por la tarde. A mi hermano Jorge, por compartir sus aficiones conmigo y por las interminables charlas literarias del viernes por la noche. A Jesús, por su cariño y por ese sándwich que siempre me espera en la mesa cuando vuelvo al pueblo. A Antonio, por conocerme mejor que nadie. Y por supuesto, a mis padres. Gracias por vuestros valiosísimos consejos para todos los aspectos de la vida, por confiar en mi y dejarme libertad para decidir qué quiero hacer y cómo quiero vivir mi vida. Gracias por apoyarme en todos mis propósitos, por secar mis lágrimas después de cada caída y compartir todas mis alegrías. Muchas gracias por enseñarme a vivir, a luchar por mis sueños y a defenderme del mundo. Sin vosotros, yo no sería quien soy.





## INTRODUCCIÓN

1. ORGANIZACIÓN DE LA CROMATINA EUCARIÓTICA.....	9
Estructura del nucleosoma .....	9
2. DISTRIBUCIÓN GENÓMICA DE LOS NUCLEOSOMAS .....	12
Ocupación y posicionamiento nucleosomal.....	12
Organización de los nucleosomas sobre regiones codificantes y reguladoras.....	13
Dinámica nucleosomal bajo diferentes situaciones fisiológicas.....	14
3. DETERMINANTES DEL POSICIONAMIENTO DE LOS NUCLEOSOMAS.....	15
Remodeladores de cromatina.....	17
Factores de transcripción.....	18
Secuencia del DNA.....	19
Secuencias que excluyen nucleosomas .....	19
Secuencias que favorecen el posicionamiento de los nucleosomas .....	20
<i>Signatures</i> nucleosomales.....	21
<b>OBJETIVOS.....</b>	<b>23</b>

## RESULTADOS

1. LA SECUENCIA DEL DNA CONTRIBUYE AL POSICIONAMIENTO DE LOS NUCLEOSOMAS .....	29
Cambios en la secuencia del gen <i>ura4</i> de <i>S. pombe</i> alteran el posicionamiento de los nucleosomas.....	29
Los cambios en el posicionamiento no son debidos a variaciones en el contenido total de A+T.....	33
La elección de codones sinónimos está optimizada para favorecer el posicionamiento nucleosomal .....	35
Los cambios en el posicionamiento del gen <i>ura4</i> no son debidos a un aumento en su expresión .....	36
La secuencia del DNA influye en el posicionamiento en regiones intergénicas.....	38
La información necesaria para posicionar nucleosomas es portátil.....	40
La información del posicionamiento está dispersa a lo largo de las 150 pb del DNA mononucleosomal.....	42
2. DISEÑO DE SECUENCIAS DE DNA CAPACES DE POSICIONAR NUCLEOSOMAS.....	44

Posicionamiento entre especies.....	44
Matrices de posicionamiento .....	47
Combinación del código genético y el código de posicionamiento.....	51
Remasterización de genes procariotas.....	53

## **DISCUSIÓN**

1. SENSIBILIDAD DEL POSICIONAMIENTO A CAMBIOS EN LA SECUENCIA DEL DNA.....	59
2. LA INFORMACIÓN RESPONSABLE DEL POSICIONAMIENTO VARÍA ENTRE ESPECIES .....	61
3. <i>SIGNATURES</i> NUCLEOSOMALES.....	62
4. GENERACIÓN DE SECUENCIAS PARA POSICIONAR NUCLEOSOMAS.....	64
6. INTERPRETACIÓN DE LAS <i>SIGNATURES</i> .....	65
7. FUNCIÓN DEL POSICIONAMIENTO DE NUCLEOSOMAS.....	67

<b>CONCLUSIONES .....</b>	<b>69</b>
---------------------------	-----------

## **MATERIALES Y MÉTODOS**

1. MICROORGANISMOS Y MEDIOS DE CULTIVO UTILIZADOS .....	75
2. GENERACIÓN DE SECUENCIAS QUE POSICIONAN NUCLEOSOMAS .....	75
2.1. Generación de la matriz de puntuación.....	75
2.2. Diseño de moléculas de DNA sintéticas no codificantes para el posicionamiento nucleosomal .....	76
2.3. Incorporación de las <i>signatures</i> nucleosomales en ORFs.....	76
2.4. Síntesis de moléculas de DNA.....	77
3. TÉCNICAS DE TRANSFORMACIÓN DE MICROORGANISMOS .....	77
3.1. Transformación de <i>E. coli</i> .....	77
3.2. Transformación de <i>S. pombe</i> .....	78
3.3. Transformación de <i>S. cerevisiae</i> .....	78
4. TÉCNICAS DE OBTENCIÓN Y MANIPULACIÓN DE DNA.....	78
4.1. Aislamiento de DNA plasmídico de <i>E. coli</i> .....	78
4.2. Aislamiento de DNA genómico de <i>S. pombe</i> .....	78
4.3. <i>Southern-blot</i> .....	79

---

4.4. Digestión con nucleasa micrococal y análisis de patrones de nucleosomas mediante <i>indirect end-labeling</i> .....	80
4.5. Preparación de DNA mononucleosomal para secuenciación masiva.....	82
4.6. Generación de mapas de nucleosomas de alta resolución mediante secuenciación masiva .....	83
4.7. Inmunoprecipitación de cromatina (ChIP) .....	84
4.8. Cuantificación mediante PCR (qPCR).....	85
5. TÉCNICAS DE MANIPULACIÓN Y OBTENCIÓN DE RNA.....	86
5.1. Extracción de RNA total .....	86
<b>BIBLIOGRAFÍA .....</b>	<b>87</b>
<b>ANEXO.....</b>	<b>99</b>
<b>ABREVIATURAS.....</b>	<b>111</b>



# INTRODUCCIÓN

---



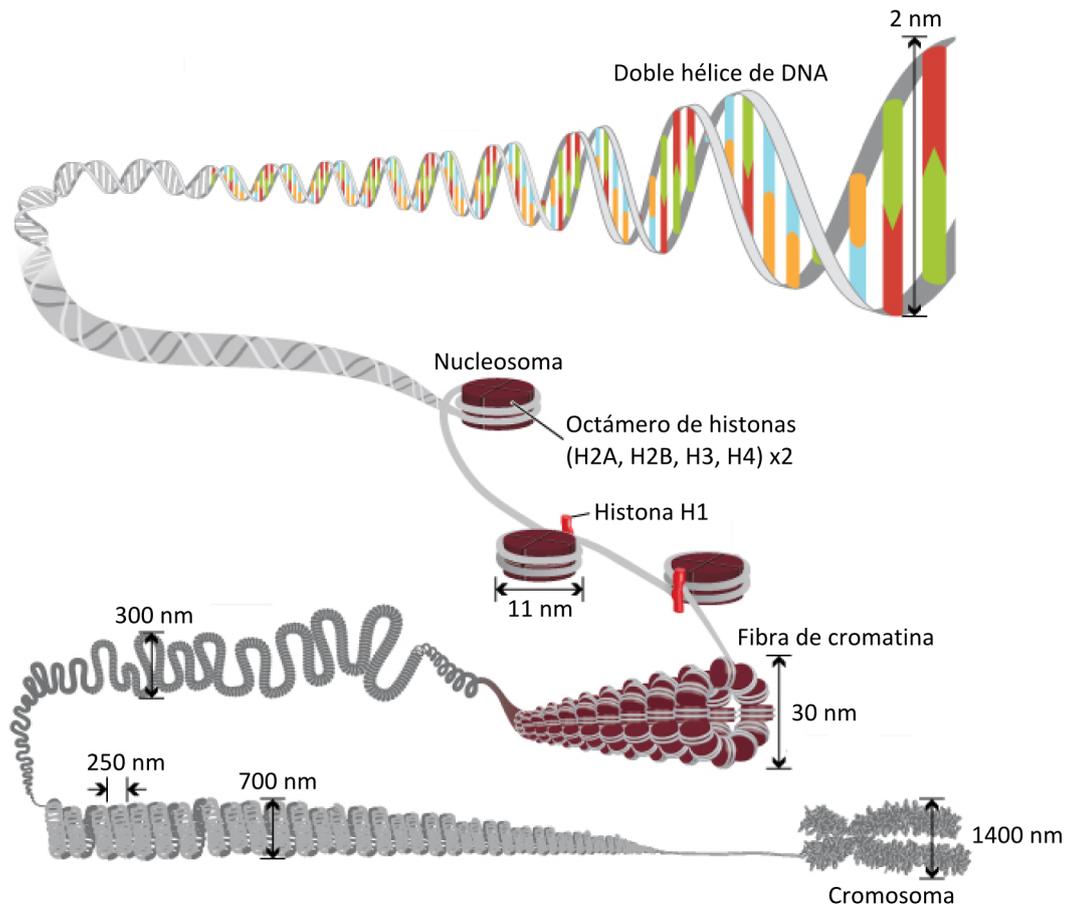
## 1. ORGANIZACIÓN DE LA CROMATINA EUCARIÓTICA

Los 3.000 millones de pares de bases que componen el genoma humano tienen una longitud de aproximadamente dos metros y deben empaquetarse en el núcleo celular, cuyo diámetro promedio es de unos 10  $\mu\text{m}$ . En este contexto, segmentos específicos del genoma deben activar o reprimir su transcripción de manera coordinada y realizar otras funciones para permitir a la célula reaccionar a un entorno en constante cambio. Para alcanzar este alto grado de compactación manteniendo la funcionalidad, los organismos eucariotas, organizan sus genomas en un complejo llamado cromatina.

La unidad fundamental de la cromatina es el nucleosoma, que se repite periódicamente cada 160 - 240 pb, según la especie, a lo largo de todo el genoma (McGhee & Felsenfeld 1980). El núcleo o *core* del nucleosoma se compone de un octámero de histonas alrededor del cual se enrollan 147 pb de DNA, dando aproximadamente 1.65 vueltas en una superhélice levógira. Cada *core* nucleosomal conecta con el siguiente a través de un segmento de DNA espaciador o *linker*, cuya longitud varía en función de la especie y que a menudo se asocia con la histona H1. Los nucleosomas se organizan entre ellos para dar lugar a estructuras de compactación superiores (**Figura 1**).

### Estructura del nucleosoma

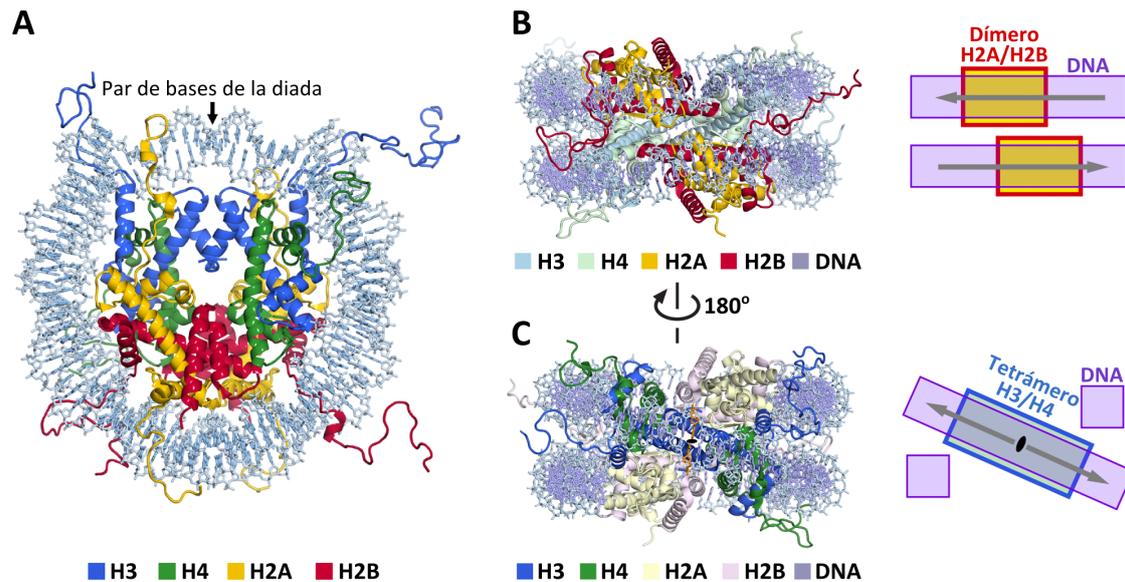
La estructura cristalográfica del nucleosoma fue determinada con resolución de 2.8 Å en 1997 (Luger et al. 1997). Cada una de las histonas canónicas (H2A, H2B, H3 y H4) consta de un motivo central llamado *histone-fold*, una extensión amino terminal y otra carboxilo terminal. El motivo *histone-fold* se compone de tres  $\alpha$ -hélices conectadas entre sí por dos bucles y su secuencia de aminoácidos está muy conservada en todos los eucariotas, mientras que la secuencia de las extensiones N-terminales de las histonas es más variable. La histona H2A se asocia a través del dominio *histone-fold* con la histona H2B y la histona H3 se asocia con la H4 de la misma forma. Dos heterodímeros H3/H4 se asocian formando un tetrámero, que constituye el centro de la parte proteica del nucleosoma. A cada lado de este tetrámero, se ensambla un dímero H2A/H2B (**Figura 2**). En conjunto, el octámero de histonas adquiere una forma de cilindro plano cuya superficie posee una fuerte carga positiva, la cual constituye la principal fuerza de unión entre éste y el DNA. Además de por interacciones iónicas, la unión entre el DNA y las histonas está mediada por puentes de hidrógeno y fuerzas hidrofóbicas (Revisado en McGinty & Tan 2015).



**Figura 1. Niveles de organización de la cromatina.** Modificado de Annunziato (2008).

En el recorrido del DNA alrededor del octámero de histonas, el surco menor de la doble hélice queda orientado hacia el octámero con una periodicidad de aproximadamente 10 pb. La unión entre el octámero de histonas y la molécula del DNA es muy estrecha, de tal forma que las histonas proyectan una arginina hacia el interior del surco menor a intervalos regulares (Revisado en McGinty & Tan 2015).

Las histonas proyectan diez colas flexibles hacia el exterior del *core* del nucleosoma: una cola amino-terminal por cada una de las histonas del octámero, y las colas carboxilo terminales de cada una de las dos histonas H2A. Estas colas salen del octámero a través del surco menor de la doble hélice en posiciones definidas y poseen numerosos residuos susceptibles de sufrir modificaciones post-traduccionales. Estas modificaciones modulan la unión de otras proteínas a la cromatina y juegan un papel clave en la regulación de procesos genómicos como la transcripción, la replicación, la recombinación y la reparación (Arya et al. 2010; Zentner & Henikoff 2013).



**Figura 2. Estructura del nucleosoma.** **A)** Visión frontal de la estructura cristalográfica del nucleosoma. Las histonas aparecen coloreadas como se indica en el diagrama. **B)** Vista del nucleosoma desde la posición opuesta a la diada. Los dos heterodímeros H2A/H2B interactúan con el DNA en dos planos paralelos. **C)** Vista del nucleosoma desde la diada. El tetrámero H3/H4 forma una rampa sobre la que cruza el DNA, conectando los dos planos representados en B. Las flechas apuntan en dirección opuesta a la diada. Modificado de McGinty & Tan (2015).

Además de las histonas canónicas existen otras variantes que se localizan en regiones específicas del genoma y cumplen diversas funciones. Estas variantes pueden introducir modificaciones en la estructura del nucleosoma, como por ejemplo la histona H2A.Z, que provoca una ligera desestabilización de la interacción entre el heterodímero H2A/H2B y el tetrámero H3/H4 (Suto et al. 2000). En *Schizosaccharomyces pombe*, no existe histona H1 y las únicas variantes de las histonas canónicas son la histona H3 centromérica (CENP-A en mamíferos y Cnp1 en *S. pombe*) y H2A.Z (codificada por el gen *pht1*). *Saccharomyces cerevisiae* posee, además de estas dos variantes, un homólogo de la histona linker H1 (codificada por el gen *HHO1*).

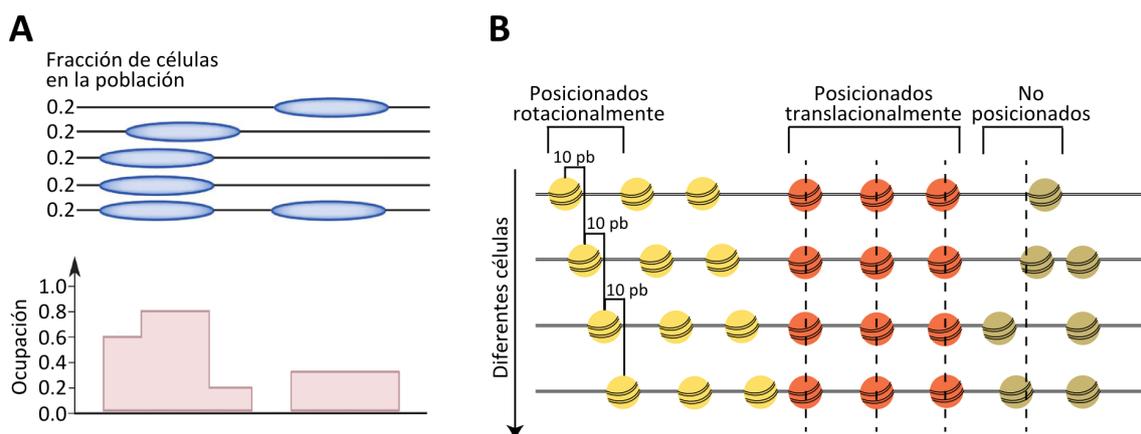
Los nucleosomas no son estructuras estáticas, sino que asistidos por chaperonas u otras proteínas como los factores remodeladores de cromatina, pueden desensamblarse para luego volverse a formar, permitiendo momentáneamente el acceso al DNA de otros factores. Como parte de este proceso, alguna de las histonas canónicas puede ser sustituida por alguna de sus variantes, proporcionando propiedades nuevas a la cromatina en regiones específicas (Bai et al. 2010). Las modificaciones post-traduccionales de las colas N-terminales de las histonas, la localización específica de las distintas variantes y la capacidad de los nucleosomas para ensamblarse y

desensamblarse en distintas situaciones confieren a la cromatina propiedades dinámicas y junto con las modificaciones covalentes que puede sufrir la propia molécula de DNA, constituyen la base de la epigenética de la célula.

## 2. DISTRIBUCIÓN GENÓMICA DE LOS NUCLEOSOMAS

### Ocupación y posicionamiento nucleosomal

Para estudiar la distribución de los nucleosomas a lo largo del genoma es importante tener en cuenta los conceptos de ocupación y posicionamiento. La ocupación nucleosomal hace referencia a la frecuencia con la que una determinada región del genoma se encuentra empaquetada en nucleosomas en una población de células. El término posicionamiento se refiere a la localización que ocupa cada nucleosoma con respecto a la secuencia del DNA. Cuando un nucleosoma se coloca sobre la misma secuencia de DNA en todas las células de una población, se dice que este nucleosoma está "posicionado". Este tipo de posicionamiento, conocido también como posicionamiento translacional, es el que vamos a estudiar en este trabajo. El término "posicionamiento rotacional" hace referencia a si un determinado par de bases está orientado hacia el interior o hacia el exterior del octámero. Dado que la doble hélice gira con una periodicidad de aproximadamente 10 pb, los nucleosomas pueden estar posicionados rotacionalmente sin estarlo translacionalmente (Revisado en Lieleg et al. 2015). Estos conceptos se ilustran en la **Figura 3**.

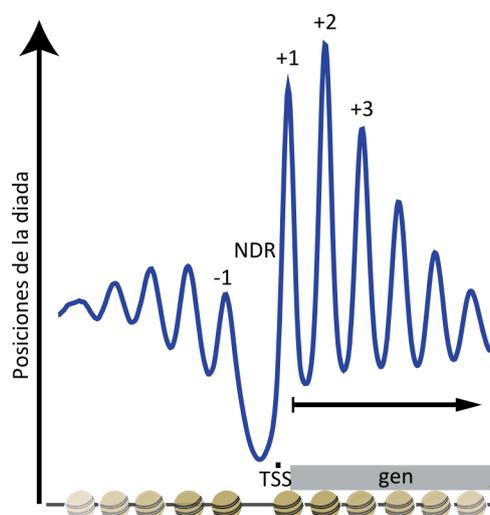


**Figura 3. Ocupación y posicionamiento nucleosomal.** **A)** En el diagrama superior aparece la distribución de nucleosomas en una población de células y en el inferior la ocupación nucleosomal por par de bases (Modificado de Struhl & Segal 2013). **B)** Diagrama de los distintos tipos de posicionamiento nucleosomal. Modificado de Lieleg et al. 2015.

## Organización de los nucleosomas sobre regiones codificantes y reguladoras

El desarrollo de las técnicas de secuenciación masiva ha permitido mapear la localización exacta de los nucleosomas en el genoma de múltiples especies, incluyendo levaduras (Lantermann et al. 2010; Tsankov et al. 2010; Givens et al. 2012), *Drosophila melanogaster* (Mavrigh et al. 2008a), *Caenorhabditis elegans* (Valouev et al. 2008) y *Homo sapiens* (Schones et al. 2008; Valouev et al. 2011). Estos estudios han identificado un patrón de organización común alrededor de los sitios de inicio de la transcripción (TSS). Por norma general, justo aguas arriba del TSS se localiza una región libre de nucleosomas (NDR), flanqueada por dos nucleosomas bien posicionados: el nucleosoma +1, que es el primer nucleosoma dentro de la región transcrita, y el nucleosoma -1, que se encuentra al otro lado de la NDR. En perfiles agregados de ocupación nucleosomal de las ORFs, los nucleosomas se colocan aguas abajo del nucleosoma +1 con un espaciamiento regular (Figura 4).

Aunque este patrón es muy similar en todas las especies estudiadas, existen algunas diferencias sutiles. Por ejemplo, en *S. cerevisiae* se ha descrito que el grado de posicionamiento de los nucleosomas decrece a medida que nos alejamos del TSS (Yuan et al. 2005; Lee et al. 2007; Mavrigh et al. 2008b), sin embargo, esto no ocurre en *S. pombe* (Soriano et al. 2013). En *S. cerevisiae* se ha descrito también que el posicionamiento del nucleosoma +1 está fuertemente influenciado por la ubicación del complejo de preiniciación (Zhang et al. 2009) y que la transcripción se inicia aproximadamente 12 pb dentro del nucleosoma +1 (Yuan et al. 2005; Lee et al. 2007; Mavrigh et al. 2008b), mientras que en metazoos el extremo de este nucleosoma se encuentra aproximadamente 60 pb aguas abajo del TSS (Mavrigh et al. 2008a; Valouev et al. 2008).

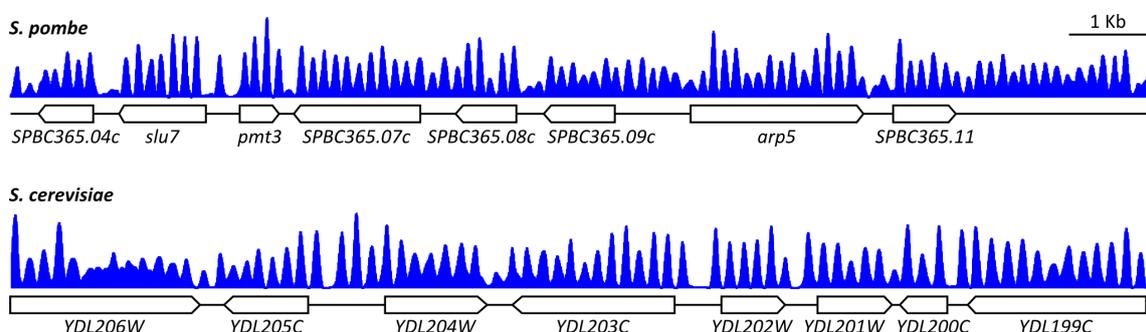


**Figura 4.** Organización de los nucleosomas alrededor del TSS en un cultivo de *S. cerevisiae* silvestre. Modificado de Lieleg et al. 2015.

La importancia de la NDR de la región 5' de los genes en su regulación transcripcional se ha puesto de manifiesto en diversos estudios. Sólo en algunos casos, la apertura o el cierre de esa región se corresponde con su activación o represión transcripcional, respectivamente. De hecho, se sabe que estas regiones están enriquecidas en motivos consenso reconocidos por factores de transcripción (Wunderlich & Mirny 2009; Hughes & de Boer 2013; Soriano et al. 2013; Weingarten-Gabbay & Segal 2014).

### Dinámica nucleosomal bajo diferentes situaciones fisiológicas

Los datos comentados en el apartado anterior han sido obtenidos a partir de perfiles agregados de nucleosomas, que permiten estudiar los niveles de ocupación promedio de todos los genes mediante la alineación a un punto concreto, que suele ser el TSS o el centro del nucleosoma +1. Cuando se estudia la distribución de los nucleosomas en regiones individuales, se puede observar que existe un alto grado de posicionamiento a lo largo de todo el genoma. La mayor parte del genoma de *S. pombe* está ocupado por nucleosomas muy bien posicionados, este posicionamiento se mantiene más allá de los sitios de terminación de la transcripción (TTS) y a lo largo de las regiones intergénicas (Soriano et al. 2013). En *S. cerevisiae*, el posicionamiento también está extendido a lo largo de todo el genoma, aunque se pierde en el extremo 3' de algunos genes (**Figura 5**). Soriano et al. (2013) observaron que, en *S. pombe*, el posicionamiento sobre las regiones transcritas se mantenía en un amplio rango de niveles de transcripción, de tal forma que sólo se perdía en los genes que se expresaban a niveles muy altos (al menos 180 veces por encima del fondo en microarrays de expresión), de los cuales sólo hay 59 en todo el genoma de *S. pombe*.



**Figura 5. Posicionamiento de nucleosomas en regiones individuales.** La figura muestra 15 Kb del genoma de *S. pombe* (arriba) y de *S. cerevisiae* (abajo). En *S. pombe*, el posicionamiento se mantiene tanto en ORFs (flechas) como en regiones intergénicas. En *S. cerevisiae*, el posicionamiento se pierde en la región 3' de algunos genes, como YDL206W.

Este alto grado de posicionamiento de nucleosomas, se mantiene constante bajo diferentes situaciones fisiológicas (Shivaswamy et al. 2008; Kaplan et al. 2009; Zawadzki et al. 2009; Cole et al. 2011; Zhang et al. 2011a; Huebert et al. 2012; Soriano et al. 2013). Por ejemplo, las diferencias entre mapas de nucleosomas de *S. pombe* realizados a distintos tiempos durante la meiosis se limitan a la aparición o desaparición de un nucleosoma en la región 5' de algunos genes, lo que en algunas ocasiones, aunque no siempre, se corresponde con represión o activación transcripcional, respectivamente (Soriano et al. 2013). Lo mismo ocurre en *S. cerevisiae*, donde la organización primaria de los nucleosomas sufre pocas variaciones durante todo el proceso de meiosis, aunque se ha observado un descenso gradual de la acetilación de la lisina 9 de la histona H3 (H3K9), así como de la trimetilación de H3K4, H2K36 y H3K79 (Zhang et al. 2011a).

En resumen, los nucleosomas se encuentran bien posicionados en la mayor parte del genoma de las levaduras. Esta organización tan compacta modula los procesos relacionados con el genoma mediante distintos mecanismos, como pueden ser la regulación del acceso a las regiones promotoras de los factores de transcripción y el complejo de preiniciación, la modificación covalente de las colas N-terminales de las histonas o la sustitución de las histonas canónicas por sus variantes histónicas.

### **3. DETERMINANTES DEL POSICIONAMIENTO DE LOS NUCLEOSOMAS**

El orden observado en la distribución de los nucleosomas a lo largo del genoma de las levaduras, suscitó la duda de cuáles serían los mecanismos encargados de dirigir y mantener el posicionamiento.

Un modelo muy aceptado para explicar el posicionamiento de los nucleosomas es el “modelo del posicionamiento estadístico” (Kornberg & Stryer 1988). Según este modelo, el posicionamiento podría explicarse con la existencia de “barreras de posicionamiento”, a partir de las cuales los nucleosomas se posicionarían de forma pasiva, manteniendo un espaciamiento regular. Si la distancia entre dos de estas barreras es igual a un número entero de nucleosomas, no habrá cabida para muchas variaciones en la posición de cada nucleosoma individual y por tanto, estos estarán bien posicionados. De esta forma, la conservación de las “barreras de posicionamiento” sería suficiente para que, en una población de células, los nucleosomas ocupasen siempre la misma posición con respecto a la secuencia del DNA.

Las “barreras de posicionamiento” podrían estar constituidas por regiones que excluyen nucleosomas, ya sea por propiedades intrínsecas de la secuencia del DNA o por la unión a éste de otras proteínas, como los factores de transcripción o el complejo de preiniciación de la RNA polimerasa. Secuencias de DNA que unen un nucleosoma con muy alta afinidad también podrían realizar esta función. Así, las principales “barreras de posicionamiento” propuestas son, o bien el NDR de la región 5’ de los genes, o bien el nucleosoma +1.

Una de las predicciones de este modelo es que el grado de posicionamiento podría disminuir a medida que nos alejamos de estas barreras. La principal objeción contra el “modelo del posicionamiento estadístico” es que, bajo condiciones de baja densidad de histonas, debería aumentar el espaciamiento entre nucleosomas y disminuir el grado de posicionamiento. Sin embargo, esto no ocurre cuando se reduce la concentración de histonas en experimentos de reconstitución de cromatina *in vitro* (Zhang et al. 2011b). También existen experimentos *in vivo* donde se reprime la expresión de las histonas (Gossett & Lieb 2012; van Bakel et al. 2013), en estos casos, la densidad de nucleosomas desciende incluso hasta el 50% con respecto a la situación en una cepa silvestre, sin embargo, el espaciamiento entre ellos no se altera. Este problema se podría resolver si se encontrase algún mecanismo capaz de agrupar nucleosomas, manteniendo un espaciamiento fijo independientemente de la densidad de estos. Se ha propuesto que esta acción agrupadora de nucleosomas podría ser realizada por los remodeladores de cromatina (Comentados en el apartado siguiente).

En paralelo al desarrollo del “modelo del posicionamiento estadístico”, se han desarrollado numerosos estudios que, en conjunto, llegan a la conclusión de que el posicionamiento de los nucleosomas depende de la acción combinada de varios factores. Particularmente, se sabe que juegan un papel esencial los remodeladores de cromatina, la unión al DNA de los factores de transcripción (TFs) y de otras proteínas reguladoras y la diferente afinidad del octámero de histonas por determinadas secuencias de nucleótidos. Algunos de estos determinantes contribuyen a completar el “modelo del posicionamiento estadístico”; otros estudios, defienden la existencia de un “código de posicionamiento de nucleosomas”, según el cual, la posición de cada nucleosoma estaría codificada en la secuencia de DNA. La contribución de todos estos factores al posicionamiento nucleosomal se detalla a continuación.

## Remodeladores de cromatina

Los remodeladores de cromatina son complejos multiproteicos que, mediante la hidrólisis de ATP, son capaces de debilitar la unión entre el DNA y el octámero de histonas (Hughes & Rando 2014). Como consecuencia de esta actividad y dependiendo del tipo de remodelador, pueden facilitar el desensamblaje de nucleosomas específicos, alterar la composición del octámero mediante el intercambio de histonas o desplazar los nucleosomas, sin necesidad de desensamblarlos, a distintas localizaciones a lo largo del DNA (Boeger et al. 2008; Hartley & Madhani 2009). Hasta la fecha se han descrito dos mecanismos implicados en dirigir a los remodeladores a *loci* específicos: o bien son reclutados directamente por modificaciones covalentes en las colas N-terminales de las histonas, o bien por la interacción física con factores de transcripción (Yadon et al. 2013).

La delección de algunos de estos remodeladores provoca la alteración global del patrón de posicionamiento de nucleosomas, como en el caso de *hrp3* en *S. pombe* (Hennig et al. 2012; Pointner et al. 2012; Shim et al. 2012) o *ISW1* y *CHD1* en *S. cerevisiae* (Gkikopoulos et al. 2011). Además, se ha observado que estos remodeladores son capaces de unirse a nucleosomas específicos y, mediante el desplazamiento direccional de estos, contribuir al mantenimiento de la organización de los nucleosomas sobre las regiones codificantes (**Figura 4**) (Yen et al. 2012). Sin embargo, es difícil determinar por experimentos *in vivo* si los remodeladores de cromatina contribuyen al posicionamiento de los nucleosomas de manera directa o indirecta, ya que mutaciones o delecciones de los genes que codifican para estos remodeladores suelen afectar a la expresión de otros factores que podrían a su vez afectar al posicionamiento.

En experimentos de reconstitución de cromatina *in vitro* se ha visto cómo la adición de extractos celulares totales de *S. cerevisiae* incrementa dramáticamente el grado de similitud del posicionamiento *in vitro* con el posicionamiento de las mismas secuencias *in vivo*. Además, este reposicionamiento depende estrictamente de la hidrólisis de ATP y es independiente de la transcripción (Kaplan et al. 2009; Zhang et al. 2009; Zhang et al. 2011b). Estos estudios confieren especial importancia a los remodeladores de cromatina en la generación de las NDRs y en el posicionamiento de los nucleosomas que los flanquean. Sin embargo, el grado de posicionamiento de los nucleosomas más alejados de la NDR, es sustancialmente menor en las reacciones de ensamblaje *in vitro* de lo que se observa *in vivo* (Struhl & Segal 2013).

## Factores de transcripción

Los motivos de unión para los factores de transcripción están dispersos por todo el genoma, pero tienden a agruparse formando *clusters* en las NDRs de las regiones 5' de los genes (Wunderlich & Mirny 2009; Hughes & de Boer 2013; Soriano et al. 2013; Weingarten-Gabbay & Segal 2014). Desde hace tiempo se sabe que los TFs pueden competir con las histonas por la unión al DNA (Workman & Kingston 1992) y por ello se ha propuesto que podrían jugar un papel crucial en la generación de las NDRs. Sin embargo, también los nucleosomas inhiben, de manera general, la unión de los TFs a sus motivos específicos y sólo algunos TFs son capaces de invadir el DNA que está empaquetado en un nucleosoma *in vivo* (Yu & Morse 1999).

En general, la delección de la mayoría de los TFs tiene un efecto limitado sobre las NDRs, como es el caso del mutante *leu3* de *S. cerevisiae* (Liu et al. 2006) y los mutantes *atf1Δ* y *pcr1Δ* de *S. pombe* (Soriano et al. 2013). No obstante, en *S. cerevisiae*, algunos TFs como Abf1, Reb1 y Rap1, conocidos como Factores Reguladores Generales (GRFs), son capaces de excluir nucleosomas (Yu & Morse 1999; Yarragudi et al. 2004) y su eliminación provoca un incremento en la ocupación nucleosomal en la mayor parte de las NDRs (Badis et al. 2008; Bai et al. 2011); lo mismo ocurre con la delección de *sap1* en *S. pombe* (Tsankov et al. 2011). De manera inversa, la integración de un motivo de unión para uno de estos GRFs, junto con un fragmento de poli(dA:dT) (Véase apartado siguiente) puede dirigir la formación de un NDR en *S. cerevisiae* (Raisner et al. 2005).

En conjunto, la unión combinada de múltiples factores de transcripción puede, mediante competición directa con las histonas, generar las NDRs y, junto con la formación del complejo de preiniciación de la transcripción, puede constituir barreras de posicionamiento como las que propone el modelo de posicionamiento estadístico. Además, los factores de transcripción también pueden contribuir al posicionamiento de nucleosomas mediante el reclutamiento de remodeladores de cromatina a regiones específicas (Yen et al. 2012).

## Secuencia del DNA

Existe cierta controversia en la literatura en relación al grado en que la secuencia del DNA puede determinar el posicionamiento de los nucleosomas. Según el modelo del posicionamiento estadístico, la secuencia del DNA se limita a especificar la localización de las barreras de posicionamiento, a partir de las cuales los nucleosomas se colocarán regularmente con poca o ninguna contribución por parte de ésta (Kornberg & Stryer 1988). En contraposición a este modelo, algunos autores defienden que existe un "código genómico del posicionamiento de nucleosomas" y que la posición de cada nucleosoma individual viene especificada por la secuencia de nucleótidos que forma parte de él o por secuencias inmediatamente adyacentes (Segal et al. 2006; Kaplan et al. 2009). En los apartados siguientes se describirán brevemente los aspectos conocidos hasta el momento en los que la secuencia del DNA contribuye al posicionamiento de los nucleosomas.

### Secuencias que excluyen nucleosomas

Para permitir la formación del nucleosoma, la molécula de DNA tiene que curvarse significativamente sobre su eje y por tanto su flexibilidad es esencial para permitir el empaquetamiento en nucleosomas. Se sabe que algunas secuencias de DNA ofrecen más resistencia a esta torsión que otras. Un ejemplo bien conocido es el caso de las secuencias homopoliméricas: los poli(dA:dT) y los poli(dG:dC).

Los poli(dA:dT) son secuencias particularmente rígidas y pueden excluir nucleosomas. De hecho, aunque secuencias largas de poli(dA:dT) pueden empaquetarse en nucleosomas *in vitro*, lo hacen con una afinidad menor que cualquier otra secuencia aleatoria (Anderson & Widom 2001). Además, se sabe que las NDRs de *S. cerevisiae* están enriquecidas en estos elementos (Yuan et al. 2005; Lee et al. 2007) y la integración de los mismos en el genoma de esta especie puede generar NDRs (Zhang & Reese 2007; Ravehsadka et al. 2012; Small et al. 2014). Los poli(dG:dC) parecen tener propiedades similares (Sekinger et al. 2005; Tsankov et al. 2011). No obstante, los efectos de estas secuencias de DNA sobre el posicionamiento de nucleosomas *in vivo* son variables. Los poli(dA:dT) no aparecen con más frecuencia en las NDRs de *S. pombe* (Lantermann et al. 2010) y de hecho, el contenido promedio de nucleótidos de A y T aumenta en el centro del DNA mononucleosomal (Quintales et al. 2015a).

### Secuencias que favorecen el posicionamiento de los nucleosomas

Las histonas deben encargarse de empaquetar todo el genoma y, como es de esperar, no tienen dominios que les permitan reconocer secuencias específicas de DNA. A pesar de ello, se han encontrado algunas secuencias capaces de dirigir por ellas mismas, sin necesidad de ningún factor adicional, el posicionamiento de un nucleosoma. Es el caso de la secuencia que codifica para el rRNA 5S del erizo de mar, que se asocia a un nucleosoma bien posicionado tanto *in vivo* como *in vitro* (Simpson & Stafford 1983). Además de esta secuencia natural, se han encontrado otras secuencias sintéticas capaces de posicionar nucleosomas *in vitro*. Lowary & Widom (1998) llevaron a cabo una serie de estudios de selección para aislar secuencias con una alta capacidad para posicionar nucleosomas *in vitro*. Estos estudios identificaron la secuencia sintética 601, que ha sido utilizada en numerosos estudios de posicionamiento de nucleosomas *in vitro*. Sin embargo, esta secuencia sintética no tiene especial afinidad para posicionar nucleosomas *in vivo* cuando se integra en el genoma de *S. cerevisiae* (Gaykalova et al. 2011; Perales et al. 2011).

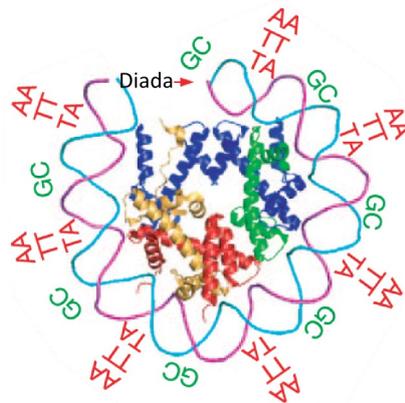
Tanto las secuencias naturales que posicionan nucleosomas como las sintéticas tienen una característica en común: los dinucleótidos AA/TA/TT aparecen preferentemente cuando el surco menor de la doble hélice queda orientado hacia el octámero de histonas, y los dinucleótidos GC son más frecuentes cuando el surco menor queda orientado hacia el exterior (**Figura 6**) (Drew & Travers 1985). Posiblemente debido a esta preferencia, en el DNA mononucleosomal de muchos organismos eucariotas puede encontrarse, en promedio, una periodicidad de dinucleótidos AA/TA/TT cada aproximadamente 10 pb y una periodicidad similar de dinucleótidos GC con un desfase de 5 pb con respecto a la periodicidad anterior (Satchwell et al. 1986; Brogaard et al. 2012). Esta periodicidad de dinucleótidos es considerada la principal característica de la secuencia del DNA que favorece la formación del nucleosoma.

La explicación biofísica de cómo estas periodicidades podrían favorecer la formación del nucleosoma aún no está clara, pero ha sido atribuida a que el surco menor de la doble hélice en regiones ricas en AT es más estrecho, mientras que en las secuencias ricas en GC, es más ancho. Esta disposición permitiría a la doble hélice sufrir la compresión requerida en el surco menor para curvarse alrededor del octámero de histonas (Drew & Travers 1985; Olson & Zhurkin 2011).

### Signatures nucleosomales

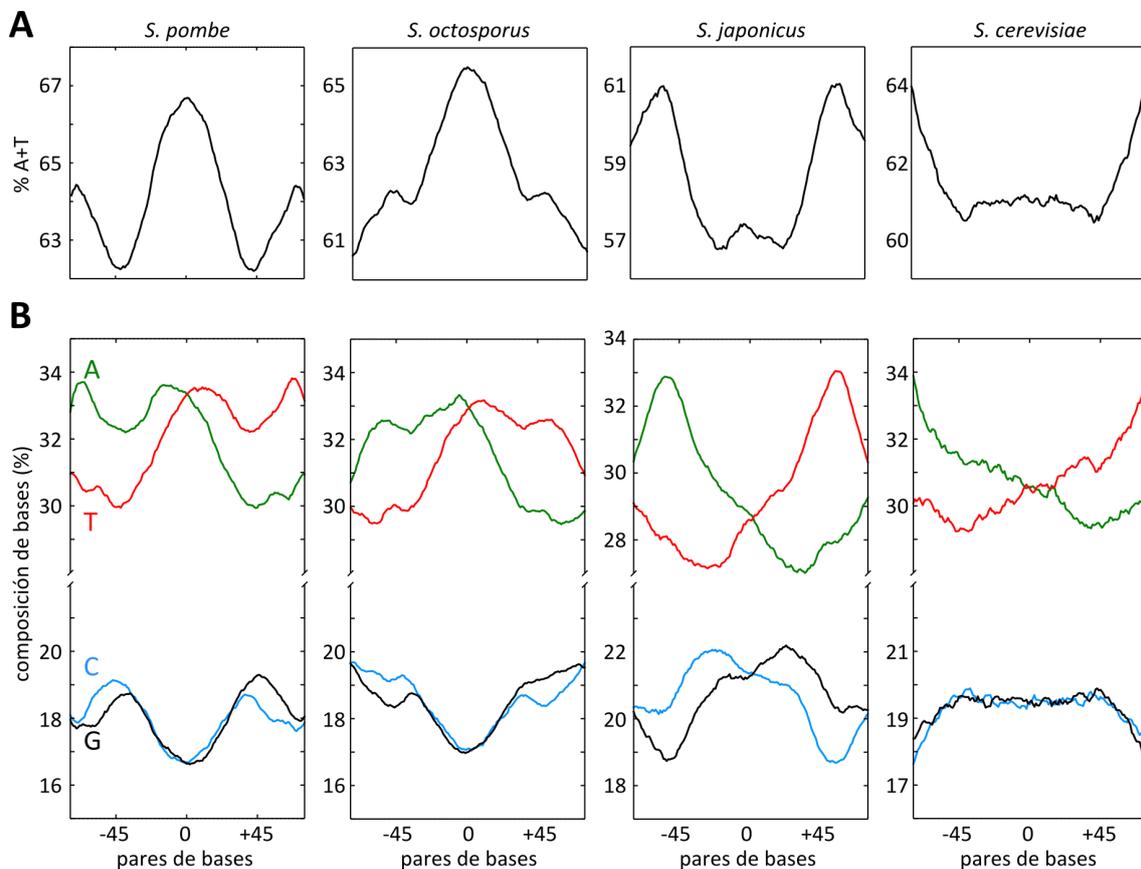
Además de la periodicidad de dinucleótidos comentada en el apartado anterior, existe otro tipo de señal en la secuencia del DNA mononucleosomal que podría contribuir al posicionamiento de los nucleosomas: las *signatures* nucleosomales. Éstas están definidas por un perfil característico que se genera cuando se representa la composición promedio de nucleótidos a lo largo de las 150 pb de las secuencias de DNA mononucleosomal asociadas a nucleosomas bien posicionados (Quintales et al. 2015a).

Una característica común en las *signatures* de todas las especies estudiadas es un mayor contenido en A en el extremo 5' del DNA mononucleosomal y en T en el extremo 3'. De hecho, los perfiles de A y T en la misma cadena del DNA forman imágenes especulares en todas las especies estudiadas, lo que indica que existe algún grado de palindromía a lo largo del DNA mononucleosomal de doble cadena. Lo mismo ocurre con los perfiles de G y C, aunque de un modo más sutil (**Figura 7B**). A pesar de estas características que las *signatures* de distintas especies tienen en común, el contenido promedio de nucleótidos a lo largo del DNA mononucleosomal varía incluso entre especies cercanas. De este modo, la *signature* de *S. pombe* se caracteriza por presentar zonas con un mayor contenido en A y T en la diada y el linker, entre las que se intercalan dos regiones con mayor riqueza en G y C, mientras que *S. japonicus* presenta un perfil inverso al de *S. pombe* y *S. octosporus* presenta un alto contenido en A y T en la diada pero no en el linker. Por su parte, *S. cerevisiae* presenta zonas ricas en A y T en el linker pero no en la diada (**Figura 7**).



**Figura 6.** Disposición de los dinucleótidos AA, TT, TA y GC en el DNA nucleosomal. Modificado de Segal et al. (2006).

No está claro si estas *signatures* son simplemente una huella generada por los nucleosomas como consecuencia de haber mantenido una posición estable durante largos periodos evolutivos, o si estas podrían estar contribuyendo al posicionamiento de los nucleosomas. En experimentos de ensamblaje de cromatina *in vitro*, la secuencia del DNA por sí sola es incapaz de reconstituir el patrón de posicionamiento que se observa *in vivo* (Zhang et al. 2011b). Por ello, su contribución al posicionamiento de los nucleosomas aún no está clara. En este trabajo, hemos estudiado la contribución de la secuencia del DNA al posicionamiento de los nucleosomas *in vivo* para intentar aclarar, en la medida de lo posible, la controversia existente a día de hoy en este campo.



**Figura 7. Patrones de distribución de nucleótidos a lo largo del DNA mononucleosomal.**

**A)** Contenido promedio en A+T de 38154, 46120, 27024 y 34526 secuencias de DNA mononucleosomal de *S. pombe*, *S. octosporus*, *S. japonicus* y *S. cerevisiae*, respectivamente, alineadas a la posición central. **B)** Distribución de los cuatro nucleótidos de las mismas secuencias de DNA que se muestran en A. El eje X indica la posición relativa al centro del DNA mononucleosomal. Modificado de Quintales et al. (2015a).

## OBJETIVOS





Los objetivos de este trabajo son:

1. Analizar la contribución de la secuencia del DNA al posicionamiento de los nucleosomas en el genoma.
2. Diseñar secuencias de DNA artificiales para analizar su capacidad de posicionar nucleosomas *in vivo*.



## RESULTADOS





## 1. LA SECUENCIA DEL DNA CONTRIBUYE AL POSICIONAMIENTO DE LOS NUCLEOSOMAS

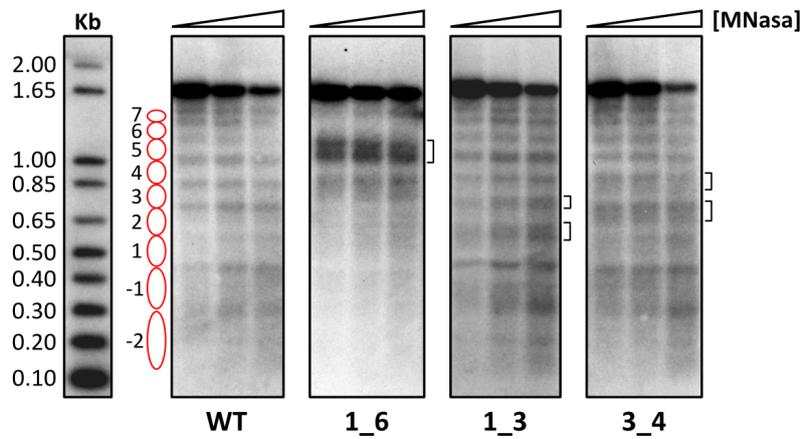
### Cambios en la secuencia del gen *ura4* de *S. pombe* alteran el posicionamiento de los nucleosomas

Varios estudios han establecido que la mayor parte de los nucleosomas en el genoma de las levaduras, ocupan siempre las mismas posiciones con respecto a la secuencia del DNA (Yuan et al. 2005; Lee et al. 2007; Zhang et al. 2011a; Soriano et al. 2013). Como se ha comentado en la introducción, existen múltiples factores que influyen en el posicionamiento de los nucleosomas, como por ejemplo los remodeladores de cromatina o los factores de transcripción. Sin embargo, el papel que juega la propia secuencia del DNA en el mantenimiento de dicho posicionamiento es aún un tema muy controvertido (Lieleg et al. 2015).

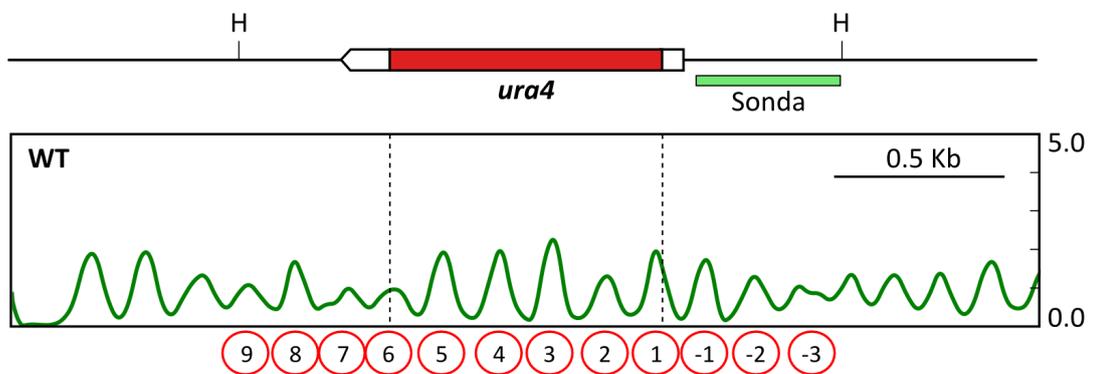
Para estudiar la influencia de la secuencia del DNA sobre el posicionamiento nucleosomal, se seleccionó la región del gen *ura4* de *S. pombe*, que está organizada en un patrón de nucleosomas regular, como se muestra en las **Figuras 8A (panel WT) y B**. Para ello, se sintetizaron distintas variantes de la ORF del gen *ura4* de *S. pombe* y se integraron en el genoma sustituyendo a la ORF endógena, dejando intactas las regiones promotora y terminadora. De este modo, se evita alterar en lo posible que distintos niveles de transcripción pudiesen afectar al posicionamiento de nucleosomas. A continuación, se estudió la posición de los nucleosomas sobre las ORFs modificadas mediante la digestión de la cromatina con nucleasa micrococcal (*MNase*) y el posterior análisis de los fragmentos obtenidos por *indirect end-labeling* o secuenciación masiva.

En primer lugar se diseñó una versión de la ORF de *ura4* que, manteniendo la codificación, fuese lo más distinta posible a la ORF silvestre (*ura4 WT*). Para ello, se sustituyó cada codón por su sinónimo más diferente, cambiando, siempre que el código genético lo permitiese, los nucleótidos A y T (W) por G y C (S) y viceversa. El resultado fue una ORF (*ura4 1\_6*) que, a nivel de la secuencia de nucleótidos, tenía una identidad del 61,51% con *ura4 WT*, pero que codificaba exactamente la misma proteína. Cuando se sustituyó la ORF silvestre por esta nueva secuencia, el posicionamiento de nucleosomas sobre la región se perdió completamente. En el panel WT de la **Figura 8A** puede observarse el patrón de bandas regular que presenta una cepa silvestre en la ORF del gen *ura4*, que indica la presencia de nucleosomas bien posicionados. En la **Figura 8B** aparece la misma región analizada por secuenciación. En el panel 1\_6 de la **Figura 8A** puede observarse que el posicionamiento de nucleosomas se ha perdido completamente en la

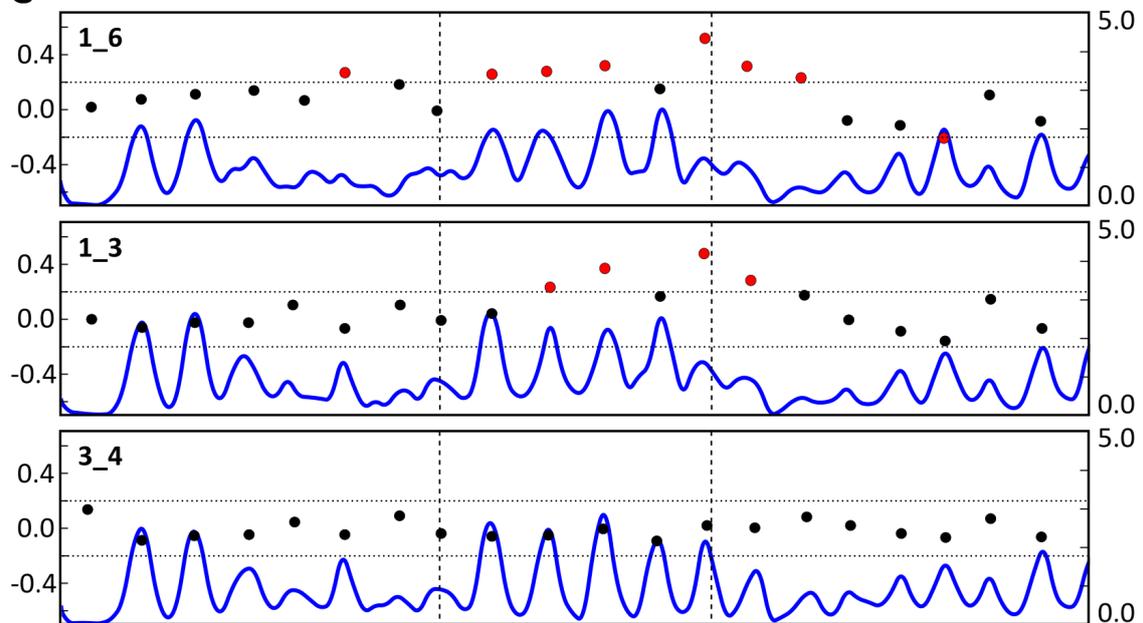
**A**



**B**



**C**



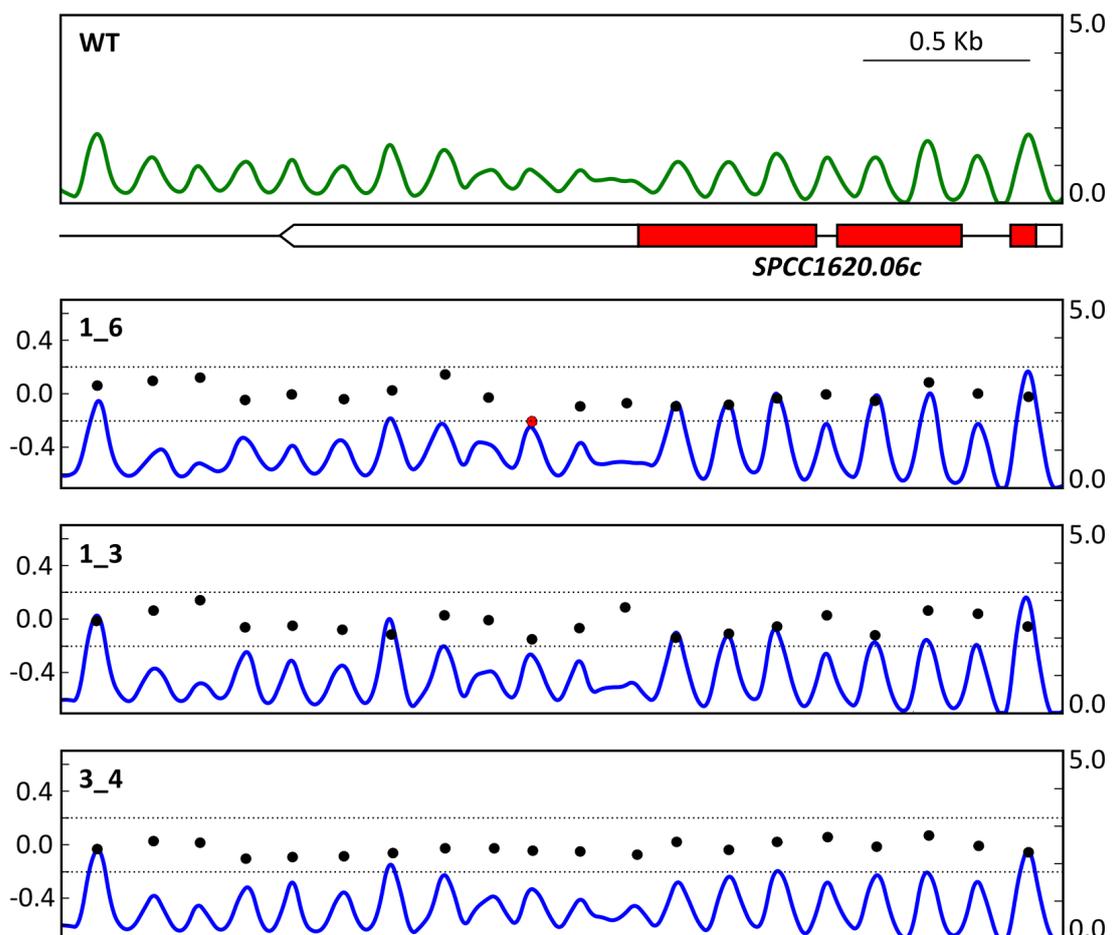
**Figura 8. Los cambios en la secuencia del DNA alteran el posicionamiento de nucleosomas del gen *ura4*.** **A)** *Indirect end-labeling* de la región *ura4* en las cepas WT, *ura4 1\_6*, *ura4 1\_3* y *ura4 3\_4*. La posición de los nucleosomas sobre la cepa WT se indica con óvalos rojos. **B)** Diagrama de la región del gen *ura4* analizada. La región codificante está representada en rojo. Se indica también la posición de la sonda (barra verde) y de los sitios de corte de HindIII (H). El mapa de ocupación de nucleosomas de células WT en esta región se muestra en verde. El eje Y indica la ocupación normalizada al promedio genómico y las líneas discontinuas verticales indican la posición de la ORF de *ura4*. Los nucleosomas están numerados como en A. **C)** Mapas de ocupación de nucleosomas generados por *MNase-seq* de las cepas 1\_6, 1\_3 y 3\_4 y análisis de DANPOS. El eje Y de la izquierda indica el  $\log_2$  de la diferencia de *fuzziness* ( $\log_2 \Delta \text{fuzziness}$ ) de cada mapa con respecto al WT (B). Los puntos rojos representan los valores mayores a  $2\sigma$  (líneas de puntos horizontales) de la distribución de esta diferencia entre cada versión modificada y el WT. El eje Y de la derecha representa la ocupación, igual que en B.

ORF modificada, donde no se observa ningún patrón reconocible salvo una zona de mayor sensibilidad a *MNase* (corchete). El efecto de estos cambios en la secuencia es tan drástico que la pérdida de posicionamiento se extiende más allá de la zona modificada, afectando también a los nucleosomas -1, -2 y 7.

Dado que en *S. cerevisiae* el nucleosoma +1 está siempre muy bien posicionado y en perfiles agregados de ocupación nucleosomal, el posicionamiento va decayendo paulatinamente a partir de él, se había especulado con la posibilidad de que este nucleosoma actuase como una barrera que, manteniendo una posición muy estable, impusiese el posicionamiento de los nucleosomas adyacentes a él (Mavrich et al. 2008b). Para comprobar si el efecto causado por la secuencia *ura4 1\_6* era debido a que ésta alterase la estabilidad del nucleosoma +1, se sintetizaron otras dos versiones del gen *ura4* modificando sólo los codones asociados a los nucleosomas 1, 2 y 3 (*ura4 1\_3*) ó 3 y 4 (*ura4 3\_4*). Los cambios realizados en estas nuevas secuencias son idénticos a los de *ura4 1\_6*, es decir, *ura4 1\_3* tiene la misma secuencia que *ura4 1\_6* en los nucleosomas 1, 2 y 3, mientras que el resto es igual a *ura4 WT*. Como puede observarse en la **Figura 8A** (panel 1\_3), la versión *ura4 1\_3* no provoca alteraciones tan drásticas como *ura4 1\_6*, lo que indica que el efecto observado a lo largo de la ORF no es debido únicamente a la alteración del nucleosoma +1. Sin embargo, puede apreciarse que las bandas entre los nucleosomas 1-2 y 2-3, están más dispersas, indicando que estos nucleosomas son menos estables. Asimismo, en el panel 3\_4 de la **Figura 8A** se observan las mismas diferencias entre los nucleosomas 2-3 y 3-4 (corchetes).

En la **Figura 8C** se muestran los mapas de nucleosomas de *MNase-seq* de las cepas analizadas en **A**. Esta técnica tiene mayor resolución que la de *indirect end-labeling* y además permite cuantificar las diferencias de posicionamiento entre distintos mapas

utilizando DANPOS (Chen et al. 2013). Esta herramienta bioinformática permite comparar el posicionamiento entre nucleosomas individuales de mapas distintos en una escala de desposicionamiento (*Δfuzziness*). Los puntos rojos en la **Figura 8C** indican que la diferencia de *fuzziness* de cada nucleosoma concreto entre las versiones modificadas de *ura4* y la cepa WT (**Figura 8B**) está por encima de 2 desviaciones estándar ( $\sigma = 0.1$ ) de la media de esta misma diferencia entre dos mapas de genoma completo diferentes. Los puntos negros indican que esta diferencia es menor a  $2\sigma$ . Las diferencias entre duplicados biológicos son muy pequeñas ( $\mu = 0.0$ ;  $\sigma = 0.1$ ). En la **Figura 9** se muestra el análisis de DANPOS en una región control no modificada de las cepas utilizadas en la **Figura 8**. Puede observarse que el posicionamiento en esta región es comparable en todas las cepas, lo que indica que las únicas diferencias significativas de *fuzziness* entre ellas se limitan a la zona modificada del gen *ura4*.



**Figura 9. Diferencias en la distribución nucleosomal en una región control no modificada de las cepas 1\_6, 1\_3 y 3\_4 respecto a la cepa WT.** La figura muestra 3 Kb de una región del cromosoma 3 de los mismos mapas utilizados en la Figura 8. Los símbolos y las escalas son iguales a los de la Figura 8.

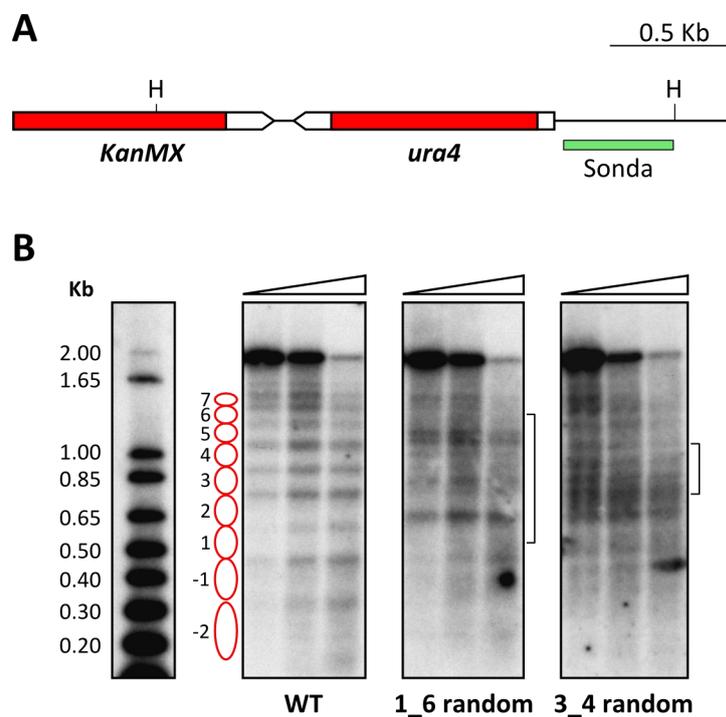
Los resultados de la cuantificación con DANPOS son comparables a los observados en los experimentos de *indirect end-labeling*. En el panel WT de la **Figura 8A** puede observarse que la posición de los nucleosomas obtenida por *indirect end-labeling* es equivalente a la determinada por *MNase-seq*, que aparece en la **Figura 8B** (óvalos rojos). En la **Figura 8C** Se observa que las modificaciones de *ura4 1\_6* extienden el desposicionamiento hasta los nucleosomas -2 y 8. En el panel 1\_3, se observa que el desposicionamiento no se limita a los tres nucleosomas modificados, sino que afecta también al -1 y al 4. En el caso de *ura4 3\_4*, no se detectaron diferencias significativas con DANPOS. Estas pequeñas variaciones entre los resultados de *indirect end-labeling* y secuenciación, pueden ser debidas a las diferencias en el grado de digestión con *MNase* existentes entre los dos métodos. En el caso del mapa de secuenciación, el tratamiento con nucleasa micrococcal se lleva a término, mientras que en el experimento de *indirect end-labeling*, la digestión es muy parcial, como indica el hecho de que la mayor parte de la señal se detecte en la banda de 1.7 Kb, que corresponde al fragmento de restricción intacto.

Estos resultados sugieren que los cambios en la secuencia del DNA afectan al posicionamiento de los nucleosomas de manera local. La pérdida de posicionamiento se limita a las regiones modificadas, afectando a veces a los nucleosomas inmediatamente adyacentes y el nucleosoma +1 no parece contribuir al posicionamiento global más que cualquier otro nucleosoma.

### **Los cambios en el posicionamiento no son debidos a variaciones en el contenido total de A+T**

La riqueza en A+T del genoma de *S. pombe* se ve reflejada en la preferencia de uso de codones, de tal manera que los más abundantes son los que poseen mayor contenido en A+T (Forsburg 1994). Una consecuencia directa de esto es que al intercambiar cada codón por aquél sinónimo más diferente, se incrementa considerablemente el contenido en G+C global de la ORF de *ura4*. Así, la versión *ura4 1\_6* descrita en el apartado anterior, posee un contenido en G+C total del 51.7%, frente al 42.9% de *ura4 WT*. Para comprobar si los cambios en el posicionamiento de nucleosomas observados en el apartado anterior se debían a esta variación en la composición de bases, se sintetizaron otras dos versiones de la ORF de *ura4* manteniendo exactamente los mismos nucleótidos, pero cambiando su orden al azar.

En la versión *ura4 1\_6 random*, se *randomizaron* todos los nucleótidos de la ORF de *ura4*, mientras que en *ura4 3\_4 random*, sólo aquellos que se encuentran asociados a los nucleosomas 3 y 4. Aunque estas dos versiones no mantienen la capacidad de codificación de la ORF, sí conservan el contenido global en A+T. Al igual que en el experimento anterior, se integraron las nuevas secuencias sustituyendo a la ORF de *ura4* sin alterar el promotor. Como estas versiones no son codificantes, las insertamos ligadas al gen marcador *KanMX*, pero dejamos intacta la región terminadora del gen *ura4* para evitar que la secuencia exógena de *KanMX* influya en el posicionamiento sobre él (**Figura 10A**). En el panel WT de la **Figura 10B**, se muestra el posicionamiento en una cepa WT que lleva *KanMX* integrado detrás del terminador de *ura4*, exactamente en el mismo lugar que en las cepas *ura4 3\_4 random* y *ura4 1\_6 random*. Como puede observarse, la presencia del gen marcador *KanMX* no influye en el posicionamiento del gen *ura4*, que es virtualmente idéntico al mostrado en la **Figura 8A**, dónde no se ha integrado el gen *KanMX*. Sin embargo, la pérdida de posicionamiento que producen las versiones modificadas es incluso más drástica que la observada en el apartado anterior (**Figura 10B**, paneles 3\_4 random y 1\_6 random), observándose un patrón de bandas totalmente

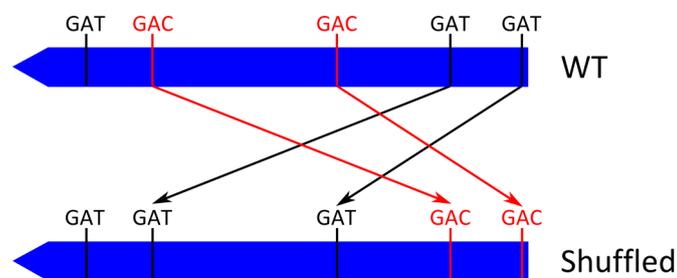


**Figura 10. Análisis de indirect end-labeling sobre versiones de *ura4* con la secuencia de la ORF randomizada. A)** Esquema de la región analizada; los símbolos son iguales que en la Figura 8. **B)** Análisis de indirect end-labeling de la región WT (panel WT) y de las versiones modificadas (3\_4 random y 1\_6 random). Los corchetes indican la región en la que la secuencia varía entre cada versión modificada y la WT.

irregular que indica una pérdida total del posicionamiento nucleosomal. Este efecto se extiende, de nuevo, más allá de las regiones modificadas, de modo que en la **Figura 10B** se puede apreciar que el nucleosoma -1 está alterado incluso en la versión *ura4\_3\_4 random*. Podemos concluir por tanto, que las modificaciones de la secuencia del DNA alteran el posicionamiento de los nucleosomas asociados a ellas independientemente del contenido global en A+T.

### La elección de codones sinónimos está optimizada para favorecer el posicionamiento nucleosomal

Como la mayoría de los organismos, *S. pombe* tiene un sesgo en el uso de codones y en su caso, como se ha mencionado anteriormente, tiende a utilizar con mayor frecuencia aquellos más ricos en A+T. Sin embargo, con baja frecuencia utiliza también algunos ricos en G+C. Se han propuesto muchas explicaciones para la existencia de este sesgo de codones, como la optimización de la traducción o incluso del plegamiento de las proteínas mediante la regulación de su velocidad de síntesis en diferentes tramos (Ikemura 1985; Cannarozzi et al. 2010; Tuller et al. 2010; Plotkin & Kudla 2011). Nosotros quisimos comprobar si la elección de determinados codones sinónimos en regiones específicas podría, además, contribuir al posicionamiento de los nucleosomas. Para ello se diseñó una nueva versión del gen *ura4* en la que se utilizaron exactamente los mismos codones de la ORF original, pero se cambió la posición de los codones sinónimos para cada aminoácido. De este modo, la nueva ORF, llamada *ura4 Shuffled*; mantiene la codificación y la composición de bases, pero los codones sinónimos de cada aminoácido ocupan posiciones diferentes. En la **Figura 11** se muestran los cambios introducidos en la posición de cinco de los codones que codifican ácido aspártico. El mismo proceso se aplicó a los codones sinónimos de los demás aminoácidos. *Ura4 Shuffled* tiene una identidad del 77.61% con la ORF original de *ura4*.



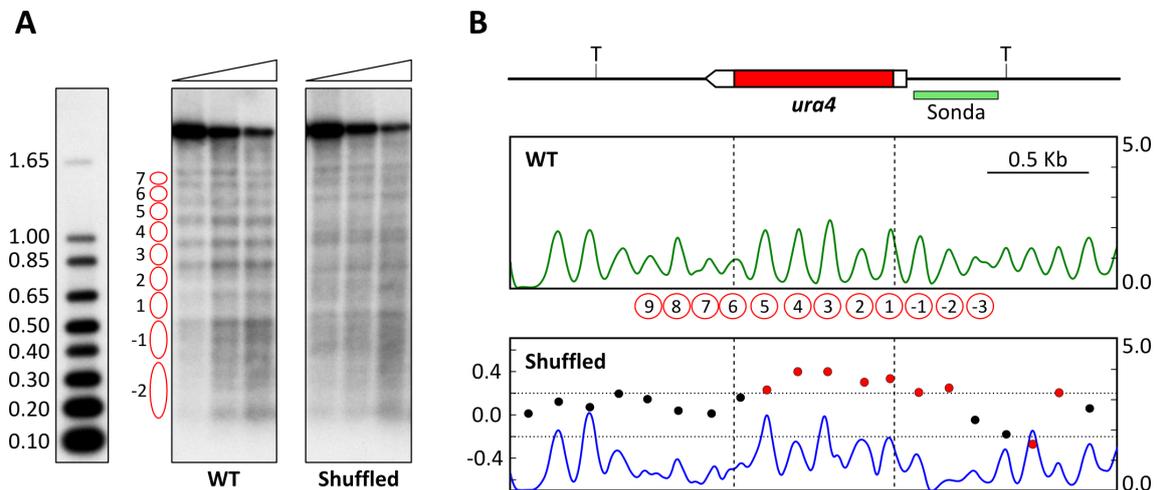
**Figura 11.** En la versión *ura4 Shuffled* se han permutado los codones sinónimos de todos los aminoácidos de la ORF WT de *ura4* de forma que no se altera su capacidad de codificación. La figura ilustra el caso concreto de cinco codones de Asp. Arriba aparece su localización en la ORF WT y abajo, después de la modificación (Shuffled).

En la **Figura 12A** aparece el análisis de *indirect end-labeling* de la versión *ura4 Shuffled*, en el que puede observarse que el patrón de nucleosomas silvestre se pierde por completo, ya que las bandas que aparecen están poco definidas y siguen un espaciamiento irregular. El apartado **B** de la misma figura representa el análisis de DANPOS a partir de los datos de *MNase-seq*. Los puntos rojos indican que el posicionamiento sobre esta ORF es menor que sobre la ORF WT (verde).

Estos resultados sugieren que la secuencia del DNA, además de contener información para la codificación de proteínas y la estabilidad de los mRNAs, también podría contribuir al posicionamiento de los nucleosomas dentro de las regiones codificantes mediante la selección de codones determinados en regiones concretas.

### Los cambios en el posicionamiento del gen *ura4* no son debidos a un aumento en su expresión

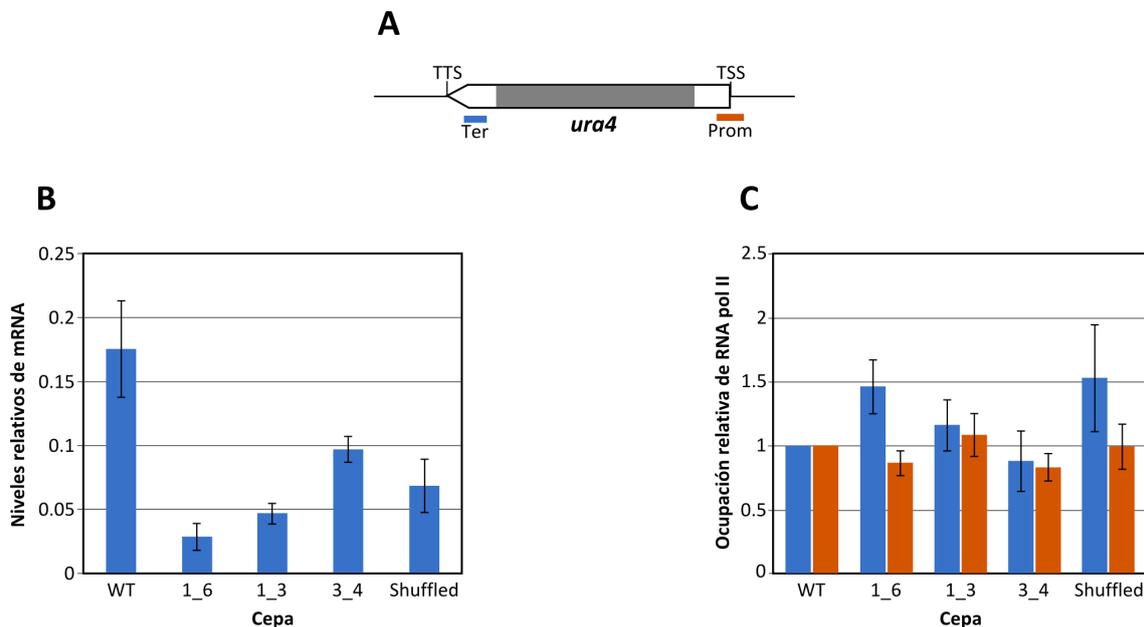
Se ha descrito que los genes que se expresan a niveles muy altos muestran un perfil de nucleosomas irregular (Koerber et al. 2009; Soriano et al. 2013). Aunque en ninguna de las construcciones descritas hasta ahora se ha modificado la región promotora del gen *ura4*, quisimos comprobar si los cambios observados en el posicionamiento eran debidos a un aumento en su expresión. Para ello, se extrajo RNA total de cultivos creciendo en las mismas condiciones que se utilizaron para realizar los mapas de nucleosomas y los



**Figura 12. La distribución de los codones sinónimos en una ORF contribuye al posicionamiento de los nucleosomas. A)** Perfiles de *indirect end-labeling* de la región del gen *ura4* en una cepa WT y en la cepa con la versión *Shuffled* de la ORF de este gen. Las posiciones de los nucleosomas en la cepa WT están indicadas con círculos rojos, como en la Figura 8. **B)** Mapas de nucleosomas y comparación por DANPOS sobre la región de *ura4* entre *ura4 Shuffled* (azul) y *ura4 WT* (verde). En la parte superior aparece el diagrama de la región analizada, donde se indican las posiciones de la sonda y de los sitios de restricción TfiI (T) utilizados en los experimentos de *indirect end-labeling*. Los símbolos y escalas son como los de la Figura 8.

experimentos de *indirect end-labeling*, es decir, medio rico y fase exponencial de crecimiento, y se cuantificaron los niveles de mRNA de cada versión del gen *ura4* por PCR cuantitativa (qPCR). Para dicho análisis, se utilizó una pareja de oligonucleótidos que anilla en la UTR 3' del RNA mensajero del gen *ura4*, que es común a todas las versiones (**Figura 13A**, barra azul) y se utilizaron los niveles de mRNA del gen *act1* para normalizar. En la **Figura 13B** se muestran los resultados obtenidos para las versiones *ura4 WT*, *ura4 1\_6*, *ura4 1\_3*, *ura4 3\_4* y *ura4 Shuffled*. En todos los casos, los niveles de mRNA son inferiores a los del WT, lo que podría ser debido a que la modificación de la secuencia de la ORF provocase que el mRNA fuese menos estable o a un descenso de los niveles de transcripción.

Para analizar esta posibilidad se cuantificó la ocupación total de la RNA Polimerasa II (RNAPII) sobre las regiones promotora y terminadora del gen *ura4*. Para normalizar utilizamos una región del gen *mat3-Mc*. En la **Figura 13C** se observa que los niveles de RNAPII de todas las versiones de *ura4* normalizadas respecto a *ura4 WT* son muy similares en todas las cepas analizadas, tanto en la región promotora de *ura4* (naranja) como en la terminadora (azul).



**Figura 13. Niveles de mRNA y ocupación de la RNA polimerasa II en el gen *ura4* y sus versiones modificadas.** **A)** Diagrama del gen *ura4*, la ORF aparece marcada en gris y se indican los sitios de inicio (TSS) y terminación (TTS) de la transcripción. Los amplicones utilizados para la cuantificación por qPCR se indican en azul y en naranja. **B)** Niveles de mRNA de distintas versiones del gen *ura4* normalizadas frente a los niveles de mRNA del gen *act1*. **C)** Ocupación relativa de la RNA Pol II en la zona del promotor (naranja) y del terminador (azul) en cepas con distintas versiones del gen *ura4*. Los datos están normalizados frente a una región de *mat3-Mc* y luego relativizados a los niveles de *ura4 WT*. En los análisis mostrados en B y C se indica la desviación estándar de tres experimentos independientes.

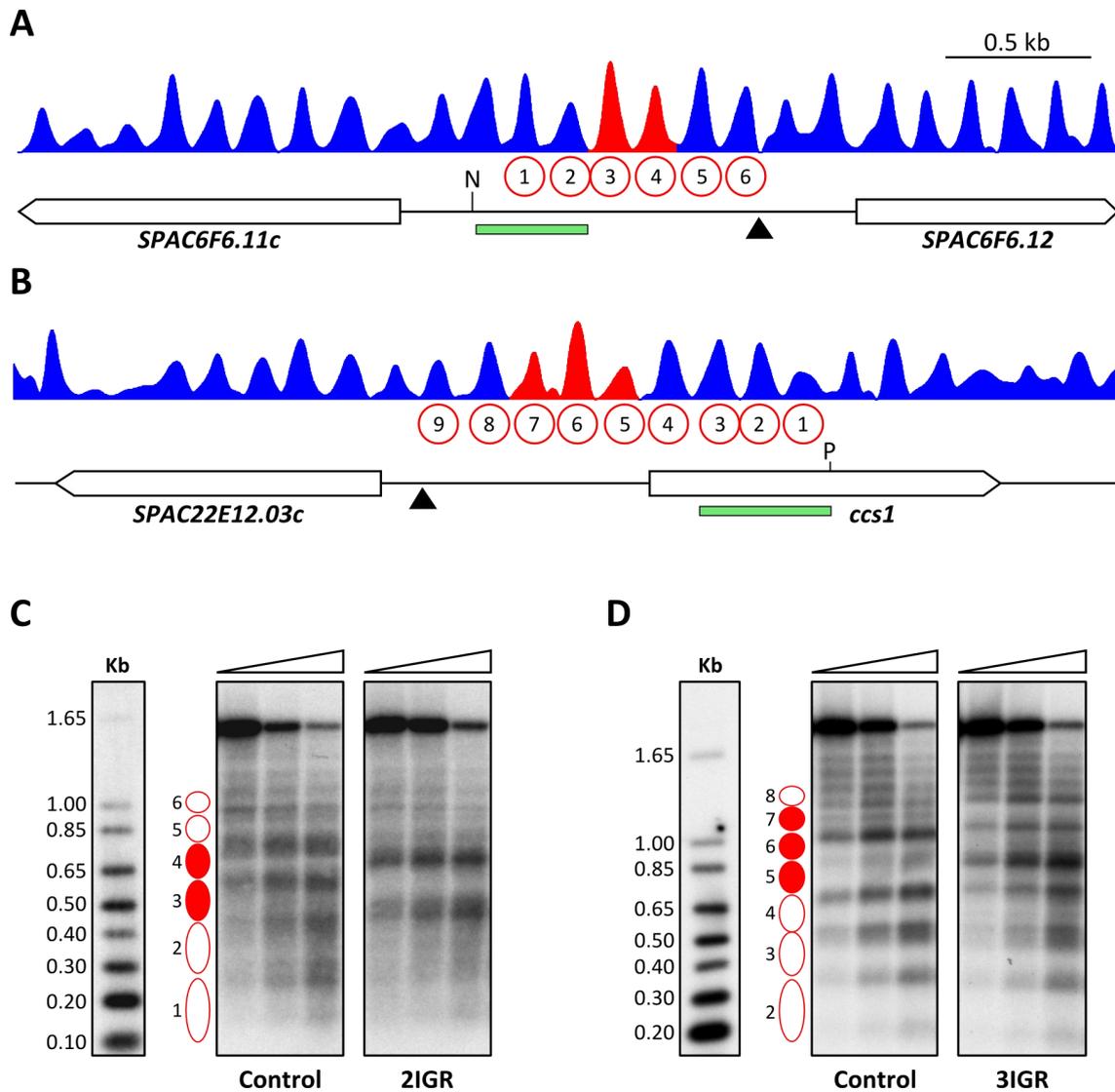
En las versiones *ura4 1\_6* y *ura4 Shuffled*, los niveles de RNAPII aumentan ligeramente en el terminador (1.5 veces), sin embargo, según estudios previos de nuestro laboratorio (Soriano et al. 2013), los niveles de expresión capaces de alterar el posicionamiento tendrían que ser al menos 20 veces superiores a los de *ura4 WT*. Se concluye por tanto que la pérdida de posicionamiento nucleosomal no es debida a alteraciones en el nivel de expresión del gen *ura4*.

### **La secuencia del DNA influye en el posicionamiento en regiones intergénicas**

Dado que los efectos de la secuencia del DNA sobre el posicionamiento parecían ser independientes de la transcripción, analizamos si el fenómeno también era detectable en regiones intergénicas no transcritas.

Para ello, seleccionamos dos regiones intergénicas grandes, con nucleosomas bien posicionados y sin ninguna transcripción detectable por *microarrays* de expresión o por RNA-seq (Rhind et al. 2011; Soriano et al. 2013) y *randomizamos* la secuencia del DNA correspondiente a dos o tres nucleosomas consecutivos (2IGR y 3IGR, respectivamente). En las **Figuras 14A** y **14B** se muestran los mapas de nucleosomas de una cepa silvestre de las dos regiones intergénicas seleccionadas. Los nucleosomas modificados en las cepas mutantes aparecen marcados en rojo. El triángulo negro indica el punto en el que está integrado el gen marcador *KanMX* tanto en las cepas modificadas como en sus respectivos controles. En el panel **control** de la **Figura 14C** aparece el perfil de *indirect end-labeling* correspondiente a la región intergénica mostrada en la **Figura 14A**, de la cepa control. La posición de los nucleosomas se indica con óvalos rojos numerados. En el panel 2IGR puede apreciarse que, a pesar de que la alteración a nivel de secuencia estaba limitada a 295 pb (nucleosomas 3 y 4), el cambio en el perfil nucleosomal es drástico en esa zona, de modo que en el lugar que en el control ocupan dos nucleosomas, en 2IGR tenemos sólo una región protegida, rodeada por dos regiones de hipersensibilidad a *MNase*. La pérdida de posicionamiento se extiende hasta los nucleosomas adyacentes 2 y 5, mientras que el resto se mantiene similar al control. En la **Figura 14D** aparecen los perfiles de *indirect end-labeling* de la cepa control (Panel control) y modificada (3IGR) correspondientes a la región intergénica mostrada en la **Figura 14B**. En este caso, la alteración del posicionamiento está restringida a la región modificada (nucleosomas 5, 6 y 7), mientras que las regiones adyacentes se mantienen invariables respecto a la cepa control.

Estos resultados son consistentes con los del gen *ura4* e indican que la modificación de secuencias de DNA mononucleosomal altera el posicionamiento de los nucleosomas asociados a ellas, independientemente de la transcripción.



**Figura 14. Cambios en la secuencia del DNA alteran el posicionamiento de nucleosomas en regiones no transcritas. A)** Mapa de nucleosomas de la región intergénica situada entre los genes *SPAC6F6.11c* y *SPAC6F6.12* en una cepa WT de *S. pombe*. El DNA mononucleosomal cuya secuencia ha sido modificada aparece marcado en rojo. El triángulo negro indica la región en la que se ha integrado el gen marcador *KanMX4* en la cepa control y en la cepa con la secuencia modificada. También se indican la posición de la sonda (barra verde) y el sitio de corte de *NsiI* (N) **B)** Mapa de nucleosomas de la región intergénica situada entre los genes *SPAC22E12.03c* y *ccs1* en una cepa WT. Se indica el sitio de corte de *PdmI* (P). Los símbolos son como en A. **C y D)** Perfiles de *indirect end-labeling* correspondientes a las regiones intergénicas mostradas en A y B de la cepa control (Control) y de las cepas en las que se han *randomizado* los nucleótidos correspondientes a los nucleosomas 3 y 4 (2IGR) y 5, 6 y 7 (3IGR).

### La información necesaria para posicionar nucleosomas es portátil

Hasta ahora hemos visto que los cambios en la secuencia del DNA afectan a la estabilidad de los nucleosomas incluso cuando estas modificaciones se limitan a regiones de unas 300 pb, que incluyen sólo dos nucleosomas (**Figuras 8 y 14**). Estos resultados sugieren la posibilidad de que la secuencia del DNA esté optimizada para contribuir al posicionamiento de los nucleosomas a lo largo de todo el genoma, tanto en regiones codificantes como intergénicas. Si esto fuese así, cada secuencia de 150 pb de DNA mononucleosomal en *S. pombe*, podría mantener la capacidad de posicionar un nucleosoma incluso aunque la cambiásemos de sitio, alterando su contexto genómico.

Para poner a prueba esta hipótesis, se eligieron 6 nucleosomas bien posicionados de distintas regiones del genoma (**Figura 15A**) y se generó una molécula de DNA compuesta por las seis secuencias de 150 pb asociadas a cada uno de ellos (**Figura 15B**, Puzle 1). A continuación, se sustituyó la ORF del gen *ura4* por esta secuencia, de manera que los fragmentos de DNA mononucleosomal de Puzle (1) mantuviesen la fase con los de las regiones adyacentes. Esta secuencia se ligó al gen marcador *KanMX*, ocupando éste la misma posición que la indicada en la **Figura 10**.

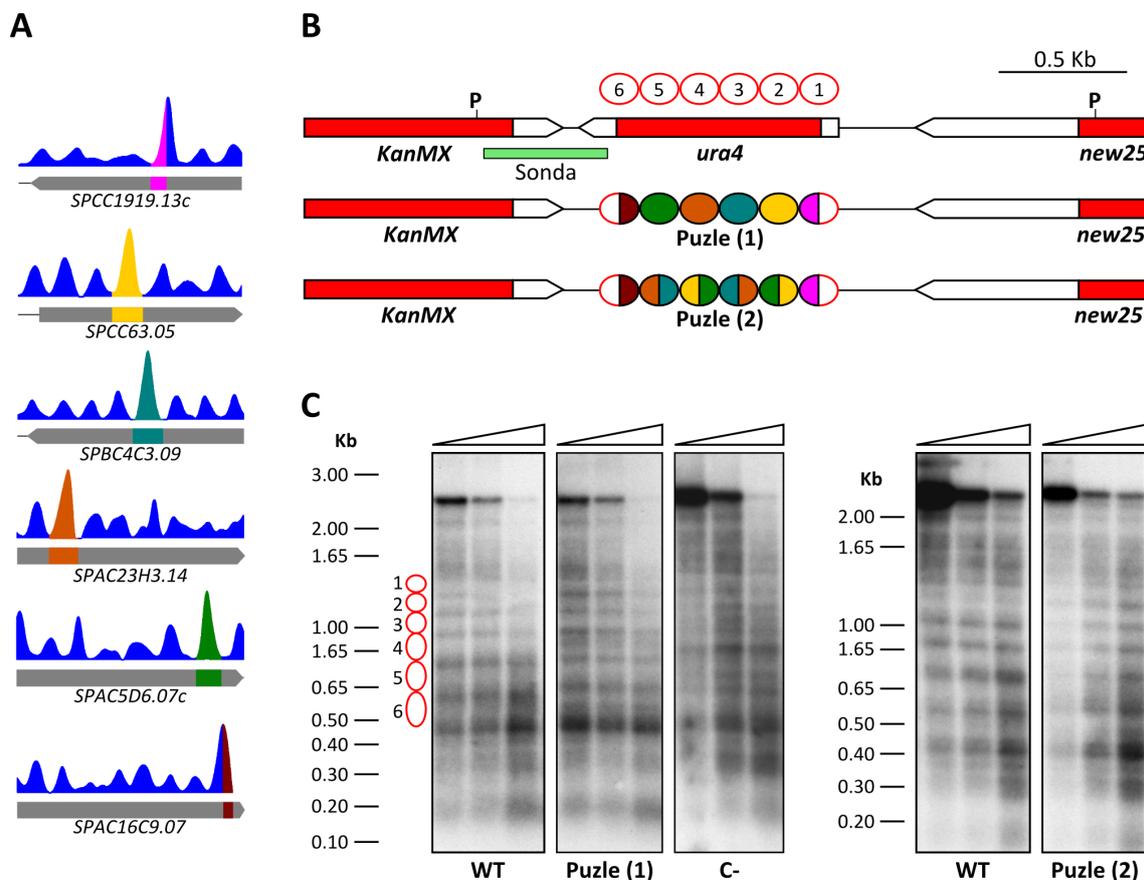
Como puede observarse en la **Figura 15C** (panel Puzle 1), estos seis fragmentos siguieron posicionando seis nucleosomas a pesar de encontrarse en otra región del genoma. El nucleosoma 4 parece ser menos estable, pero aún así los seis nucleosomas ocupan las mismas posiciones que ocupaban sobre la ORF de *ura4*. Este resultado es aún más llamativo si tenemos en cuenta que todos los cambios realizados hasta ahora en esta región generaron perfiles totalmente irregulares.

Como control negativo para este experimento, se *randomizó* individualmente cada secuencia de 150 pb y se generó una molécula de DNA compuesta por estas seis secuencias modificadas. En el panel C- de la **Figura 15C** puede observarse que en este caso el posicionamiento se pierde por completo.

Una característica notable de la secuencia Puzle (1), es que los nucleosomas 1 y 6 se mantienen en su sitio a pesar de que la secuencia asociada a ellos es mixta. En el caso del nucleosoma 1, la mitad de la secuencia sigue formando parte de la UTR 5' del gen *ura4* mientras que la otra mitad pertenece a un nucleosoma del gen *SPCC1919.13c*. Lo mismo ocurre en el caso del nucleosoma 6, donde una mitad corresponde a la UTR 3' del gen *ura4* y la otra a un nucleosoma del gen *SPAC16C9.07*. Esto podría indicar que la información contenida en cada mitad del DNA mononucleosomal puede combinarse con la otra mitad de otro DNA mononucleosomal, sin perder por ello la capacidad para

posicionar. Para analizar esta posibilidad, se compuso otro puzzle en el que se intercambiaron las mitades de estas mismas secuencias mononucleosomales (**Figura 15B**, Puzzle 2). Como puede observarse en el panel Puzzle (2) de la **Figura 15C**, esta nueva secuencia también reconstituye el perfil de nucleosomas de *ura4*.

En conjunto, estos resultados indican que la información para posicionar se encuentra contenida en las 150 pb de cada DNA mononucleosomal y que esta información es simétrica respecto a la posición central y se puede combinar entre distintos nucleosomas.

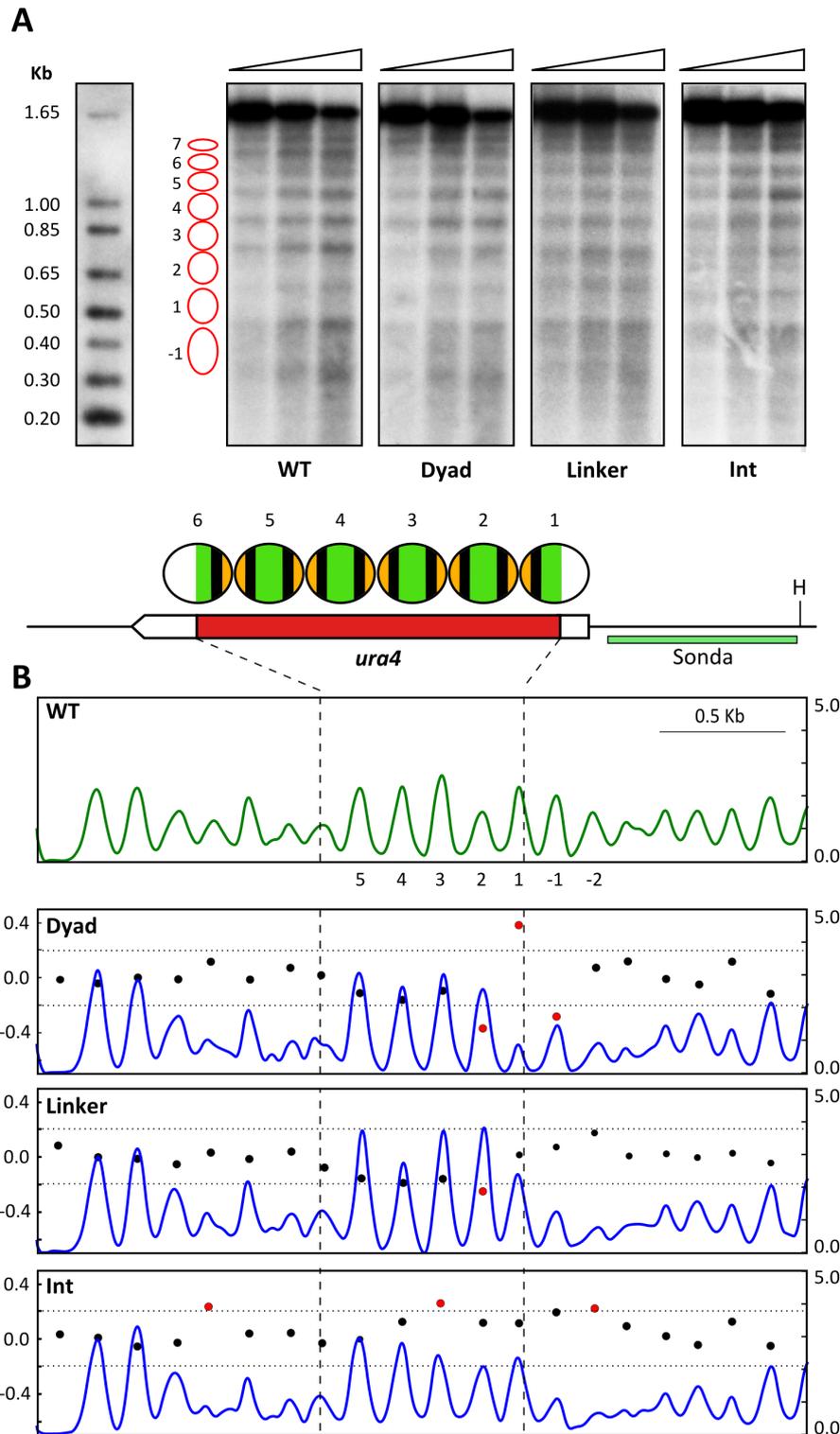


**Figura 15. El DNA mononucleosomal contiene información suficiente para posicionar nucleosomas.** **A)** Mapa de ocupación de nucleosomas en seis regiones diferentes del genoma de *S. pombe*. Los seis nucleosomas seleccionados para generar la molécula quimérica se resaltan en distintos colores. **B)** Diagrama de la región estudiada por *indirect end-labeling*. Se indican las posiciones del marcador *KanMX*, la ORF del gen *ura4*, los sitios de corte de *Psul* (P) y la sonda de hibridación utilizada (verde). Los colores de las construcciones Puzzle (1) y Puzzle (2) indican a qué DNA nucleosomal mostrado en A pertenece cada fragmento. **C)** Posicionamiento de nucleosomas analizado por *indirect end-labeling* de las tres construcciones del panel B. Los seis nucleosomas de la cepa WT sustituidos en las versiones quiméricas están representados por óvalos rojos. El panel C- se corresponde con una construcción en la que se han *randomizado* los nucleótidos de cada fragmento de DNA mononucleosomal utilizado en Puzzle (1) antes de ensamblarlos en el mismo orden que en Puzzle (1).

### **La información del posicionamiento está dispersa a lo largo de las 150 pb del DNA mononucleosomal**

Los resultados anteriores sugerían que muchos fragmentos de DNA mononucleosomal de *S. pombe* podrían contener la información necesaria para posicionar un nucleosoma. Sin embargo, dejaban abierta la posibilidad de que esta información estuviese restringida a determinadas regiones dentro de estas 150 pb, como podría ser el DNA *linker* o la zona de la diada.

Para determinar si alguna región del DNA mononucleosomal contribuía preferentemente al posicionamiento, se construyeron tres nuevas versiones de la ORF del gen *ura4* en las que se introdujeron exactamente los mismos cambios que en *ura4 1\_6* (indicados en la página 29), pero únicamente en un tercio de cada DNA mononucleosomal. En la versión *ura4 Dyad*, cambiamos 51 pb (correspondientes a 17 codones) centradas en la diada de cada nucleosoma de *ura4*; en *ura4 Linker*, 51 pb centradas en cada *linker* y en *ura4 Int*, las 51 pb restantes distribuidas en dos zonas simétricas respecto a la diada (secciones verdes, naranjas y negras de la **Figura 16A**, respectivamente). El análisis del posicionamiento de estas tres construcciones por *indirect end-labeling* mostró que las bandas internucleosomales eran ligeramente más difusas en los tres casos (**Figura 16A**). Sin embargo, el análisis de DANPOS no detectó diferencias significativas, con la excepción de algunos nucleosomas individuales en cada cepa. No obstante, estas diferencias no eran comparables a las alteraciones causadas por la versión *ura4 1\_6* (**Figura 8**) o *ura4 Shuffled* (**Figura 12**), lo que sugiere que ninguna de estas regiones contribuye preferentemente al posicionamiento de los nucleosomas, sino que la información está dispersa a lo largo de las 150 pb y todas las regiones contribuyen parcialmente a mantenerlo.



**Figura 16. La información para posicionar nucleosomas está dispersa a lo largo del DNA mononucleosomal. A)** Análisis de *indirect end-labeling* de una cepa WT y de las cepas con las versiones Dyad, Linker e Int de la ORF de *ura4*. Las secciones coloreadas indican qué región del DNA mononucleosomal ha sido modificada en Dyad (verde), Linker (naranja) e Int (negro). Se indica el sitio de restricción de HindIII (H) y la posición de la sonda de hibridación (verde). **B)** Mapas de ocupación de nucleosomas generados por *MNase-seq* de las cuatro cepas y análisis de DANPOS de los tres mutantes comparados con el WT. Símbolos como en la **Figura 8**.

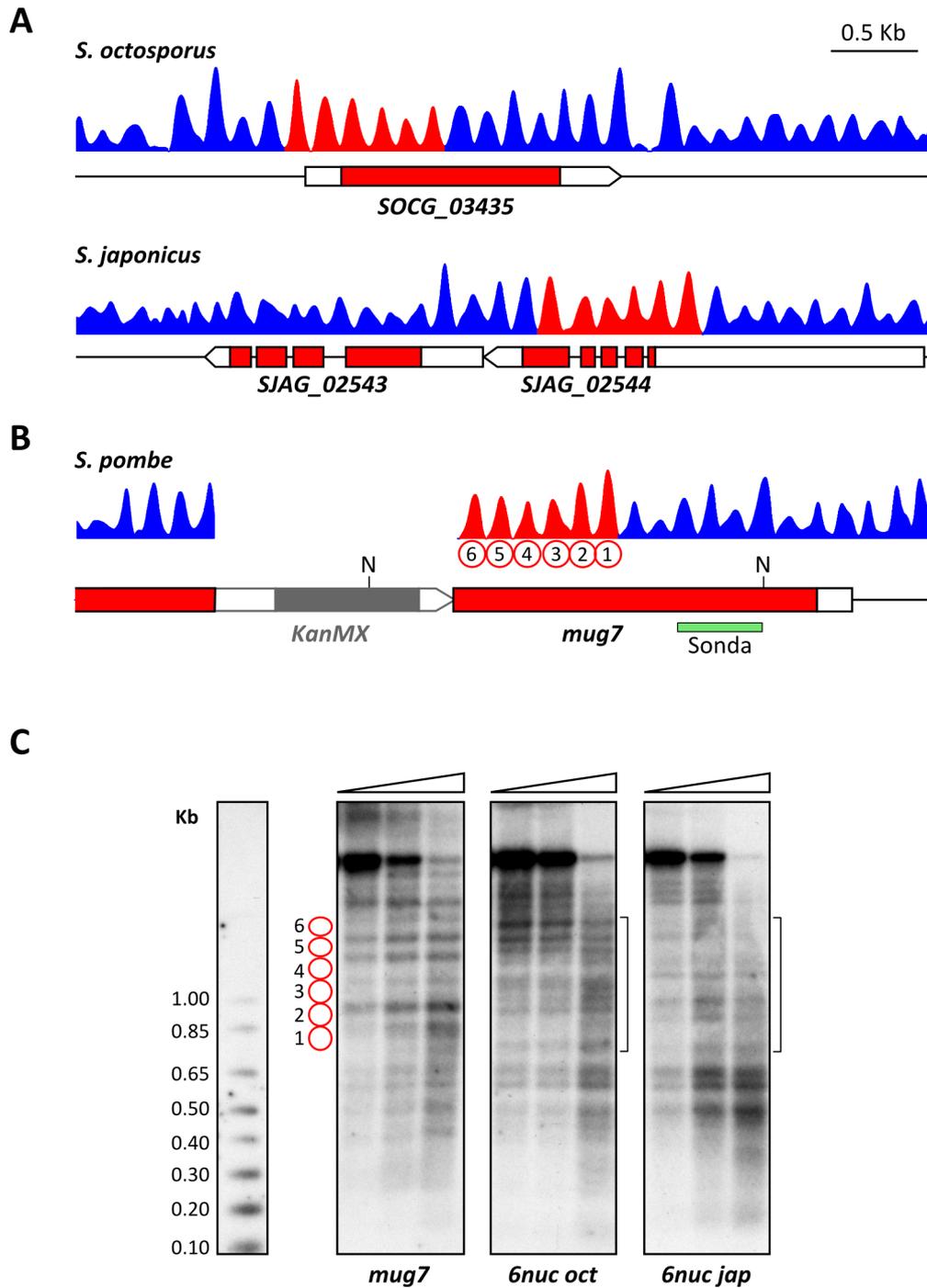
## 2. DISEÑO DE SECUENCIAS DE DNA CAPACES DE POSICIONAR NUCLEOSOMAS

### Posicionamiento entre especies

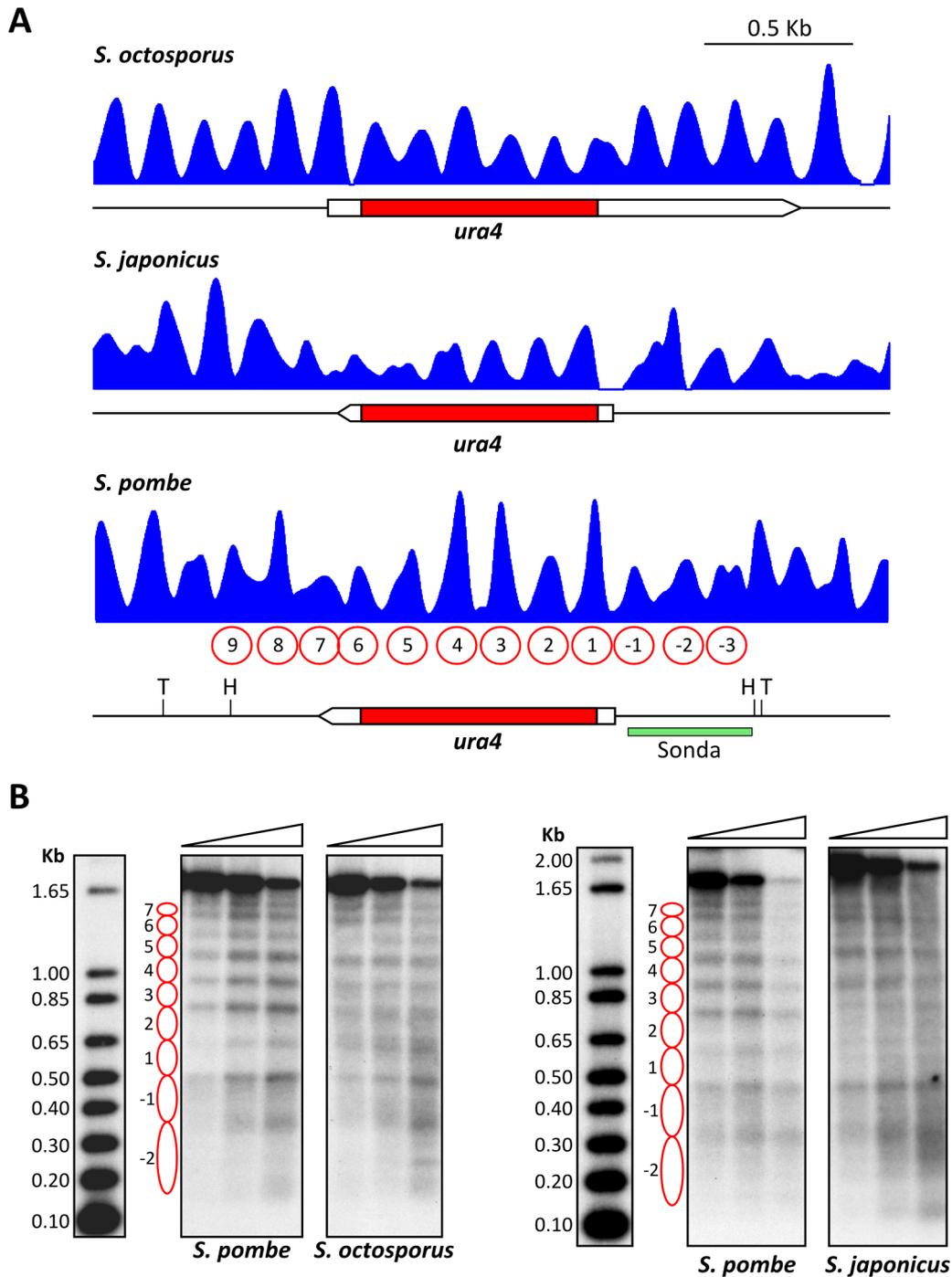
Los resultados descritos en los apartados anteriores indican que las secuencias de DNA mononucleosomal de *S. pombe* contienen la información necesaria para posicionar nucleosomas. Esta información podría ser universal o podría variar de unas especies a otras. Para probar si las secuencias mononucleosomales de otras especies también contienen información para posicionar nucleosomas que pueda ser interpretada por *S. pombe*, elegimos una secuencia asociada con 6 nucleosomas bien posicionados en *S. octosporus* y otra en *S. japonicus* (**Figura 17A**) y las integramos en *S. pombe*, sustituyendo a otra secuencia de 6 nucleosomas de aproximadamente el mismo tamaño, dentro del gen *mug7* (**Figura 17B**). En la **Figura 17C** aparece el resultado de *indirect end-labeling* para este experimento. En el panel *mug7*, se indica la posición de los 6 nucleosomas de *S. pombe* que son sustituidos en los otros dos paneles por los de *S. octosporus* (panel *6nuc oct*) y *S. japonicus* (panel *6nuc jap*). El patrón se altera completamente en esta región cuando se sustituye por las secuencias de estas dos especies. Más allá del nucleosoma 6 se encuentra el módulo *KanMX*, sin embargo, la región anterior al nucleosoma 1, que no ha sido modificada, presenta un patrón de bandas similar en todas las cepas.

Este resultado se confirmó con otros experimentos con el gen *ura4*. Las ORFs ortólogas de *ura4* de *S. octosporus* y *S. japonicus* tienen exactamente el mismo tamaño que la de *S. pombe* (795 pb) y, aunque codifican para proteínas ligeramente distintas (identidad del 82.6% para *S. octosporus* y del 74.6% para *S. japonicus*, con respecto a *S. pombe*), son capaces de sustituir funcionalmente a la de *S. pombe*. En cuanto a secuencia de DNA, la ORF de *ura4* de *S. octosporus* tiene una identidad del 74.97% con la de *S. pombe*, y la de *S. japonicus*, un 70.19%. En la **Figura 18A** se muestran los mapas de *MNase-seq* del gen *ura4* en sus sitios endógenos en cada una de estas tres especies. Las tres ORFs poseen seis nucleosomas en posiciones relativas similares en sus respectivos genomas.

El experimento consistió en reemplazar la ORF del gen *ura4* de *S. pombe* por la de *S. octosporus* o *S. japonicus* para analizar el posicionamiento de nucleosomas por *indirect end-labeling*. En el caso de la ORF de *S. octosporus*, se observa que los nucleosomas tienden a posicionarse en los mismos lugares que en *S. pombe*, aunque las bandas internucleosomales son más difusas (**Figura 18B**, panel *S. octosporus*). Esto puede indicar que, o bien el nucleosoma es menos estable y la nucleasa puede acceder mejor a



**Figura 17. La información de posicionamiento nucleosomal es específica de especie. A)** Mapa de ocupación de nucleosomas de una región del genoma de *S. octosporus* y otra de *S. japonicus*. **B)** Mapa de ocupación de nucleosomas en la región del gen *mug7* de *S. pombe*. La secuencia asociada a los seis nucleosomas marcados en rojo en fue sustituida por las de *S. octosporus* y *S. japonicus*, indicadas en rojo en A. Se indica la posición del marcador *KanMX* (gris) interrumpiendo al gen *mug7*, los sitios de corte de *NsiI* (N) y la posición de la sonda (verde). **C)** Análisis de *indirect end-labeling* de la cepa control (*mug7*) y de las construcciones *6nuc oct* y *6nuc jap*. Los corchetes señalan la región en la que la secuencia varía respecto a la cepa control.



**Figura 18. Posicionamiento de nucleosomas de los genes *ura4* de *S. octosporus* y *S. japonicus* en *S. pombe*.** **A)** Mapas de ocupación nucleosomal por MNase-seq de las regiones del gen *ura4* en los genomas de *S. octosporus*, *S. japonicus* y *S. pombe*. Se indican los sitios de restricción HindIII (H) y TfiI (T) en *S. pombe*, así como la posición de la sonda utilizada para los experimentos de *indirect end-labeling* (verde). **B)** Análisis del posicionamiento de nucleosomas sobre la ORF de *ura4* de *S. octosporus* y de *S. japonicus* cuando se integran sustituyendo a la de *S. pombe*. El análisis de *indirect end-labeling* de la ORF del gen *ura4* de *S. japonicus* se realizó con la enzima de restricción TfiI, a diferencia del resto de análisis mostrados en la figura, que se realizaron con HindIII. Dado que la distancia entre el sitio de corte de HindIII y TfiI es de tan solo 32 pb en el extremo en el que hibrida la sonda, los resultados son comparables aunque TfiI produzca un fragmento de restricción de mayor tamaño.

las regiones internucleosomales, o bien hay una ligera variación de la posición de los nucleosomas en las distintas células de la población.

En el panel *S. japonicus* de la **Figura 18B** se muestra el resultado para la ORF del gen *ura4* de *S. japonicus*. Se puede observar que el posicionamiento de los nucleosomas -1 y -2 es similar en los dos casos, sin embargo, en la región entre los nucleosomas 1 y 6, la cepa portadora de la ORF del gen *ura4* de *S. japonicus* ha perdido totalmente el posicionamiento. Las bandas observadas no tienen relación con las de la cepa WT y las distancias entre ellas son irregulares y no se corresponden con el tamaño de los nucleosomas.

El hecho de que la ORF de *ura4* de *S. octosporus* tenga un posicionamiento más próximo al de *S. pombe* que la de *S. japonicus*, puede explicarse porque evolutivamente, *S. octosporus* es más cercana a *S. pombe* (distancia evolutiva de 119 millones de años) que *S. japonicus* (distancia evolutiva de 221 millones de años) (Rhind et al. 2011). Estas diferencias filogenéticas se ven también reflejadas en las *signatures* nucleosomales descritas por Quintales et al. (2015), en las que puede observarse que tanto *S. pombe* como *S. octosporus* presentan una tendencia a acumular A+T en la diada, mientras que *S. japonicus* tiende a acumular G+C en esta región y muestra un perfil en su composición de bases inverso al de *S. pombe* (**Figura 7**).

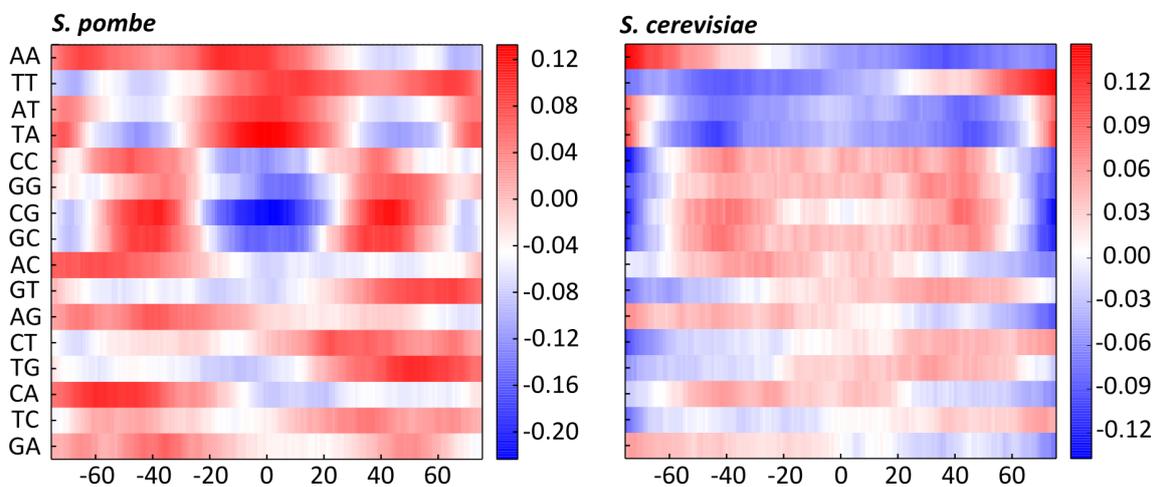
## Matrices de posicionamiento

Los resultados mostrados en los apartados anteriores indican que la secuencia del DNA juega un papel importante en el posicionamiento de los nucleosomas, ya que las secuencias de DNA mononucleosomal mantienen la capacidad para posicionar nucleosomas en regiones ectópicas del mismo genoma (**Figura 15**) y cualquier modificación de la secuencia del DNA afecta negativamente al posicionamiento de los nucleosomas (**Figuras 8, 10, 12 y 14**). Dado que los resultados mostrados en las **Figuras 17 y 18** sugieren que podría existir una relación entre las *signatures* nucleosomales y la capacidad de una secuencia de DNA para posicionar nucleosomas en una especie concreta, nos preguntamos si sería posible diseñar secuencias de DNA sintéticas capaces de posicionar nucleosomas a partir de la información contenida en estas *signatures*.

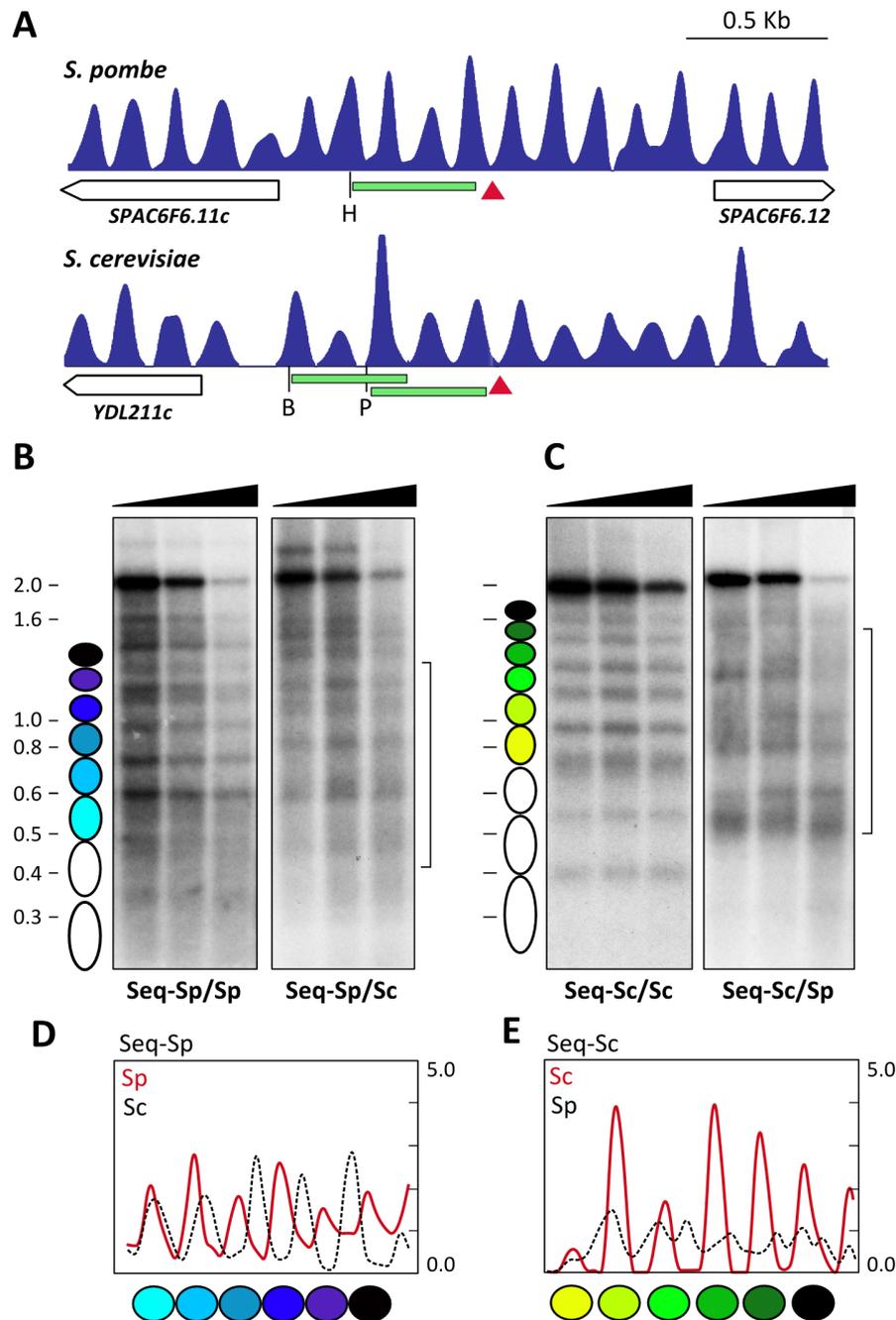
Para extraer esta información, Luis Quintales en nuestro laboratorio generó una "Matriz de puntuación de posición específica" (PSWM: *Position-Specific Weight Matrix*) para *S. pombe* y otra para *S. cerevisiae* (**Figura 19**). Como se explica en el **Apartado 2.1 de Materiales y Métodos**, estas matrices incorporan la frecuencia de cada dinucleótido a

lo largo de los perfiles agregados de 38.154 secuencias mononucleosomales en el caso de *S. pombe*, y 34.526 en el caso de *S. cerevisiae*. Estas matrices nos permiten asignar a secuencias individuales una puntuación, que será mayor cuanto más se aproxime su composición de nucleótidos a la *signature* de la especie correspondiente.

A partir de la información de las matrices de posicionamiento, se diseñó una secuencia de DNA de 918 pb (formada por seis secuencias de 153 pb —147 pb + 6 pb de linker) ajustada a la matriz de *S. pombe* y otra de 1008 pb (formada por seis secuencias de 168 pb —147 pb + 21 pb de linker) ajustada a la matriz de *S. cerevisiae*. El diseño de estas secuencias se explica más detalladamente en el **Apartado 2.2 de Materiales y Métodos**. Si las *signatures* nucleosomales contienen información para posicionar nucleosomas, cada una de estas secuencias debería ser capaz de posicionar seis nucleosomas en la especie correspondiente. Para analizar esta posibilidad, estas dos secuencias se ligaron al marcador de resistencia *KanMX* y se integraron entre dos nucleosomas bien posicionados de una región intergénica de *S. pombe* y otra de *S. cerevisiae* (**Figura 20A**).



**Figura 19. Matriz de puntuación de posición específica de las *signatures* nucleosomales.** Heat map de las matrices de *S. pombe* y *S. cerevisiae*. El eje X indica la posición de cada nucleótido relativa a la diada del nucleosoma. El eje Y indica la puntuación en log-odd de los 16 dinucleótidos a lo largo del DNA mononucleosomal. Esta puntuación se calcula como el cociente de la frecuencia de cada dinucleótido en cada posición con respecto a su frecuencia genómica. Las barras de la derecha representan la escala de color asociada con los valores numéricos de la matriz.



**Figura 20. Posicionamiento nucleosomal sobre secuencias artificiales *in vivo*.** **A)** Perfil de ocupación de nucleosomas de las regiones genómicas de *S. pombe* y *S. cerevisiae* donde se insertaron las secuencias artificiales (triángulos rojos). Se muestran los sitios de restricción de HindIII (H), BsmI (B) y PstI (P) y la localización de las sondas (barras verdes). **B)** Análisis de *indirect end-labeling* de la secuencia artificial de *S. pombe* integrada en *S. pombe* (Seq-Sp/Sp) y en *S. cerevisiae* (corchete, Seq-Sp/Sc). Los óvalos coloreados indican la distribución de seis nucleosomas en *S. pombe* sobre la secuencia artificial en las posiciones predichas por la matriz nucleosomal. **C)** Análisis de *indirect end-labeling* de la secuencia artificial de *S. cerevisiae* integrada en *S. cerevisiae* (Seq-Sc/Sc) y en *S. pombe* (corchete, Seq-Sc/Sp). Los óvalos coloreados indican lo mismo que en B, pero en el genoma de *S. cerevisiae*. **D)** Análisis de ocupación de nucleosomas por *MNase-seq* de la secuencia artificial de *S. pombe* integrada en *S. pombe* (rojo) y en *S. cerevisiae* (negro). **E)** Análisis de ocupación de nucleosomas por *MNase-seq* de la secuencia artificial de *S. cerevisiae* integrada en *S. cerevisiae* (rojo) y en *S. pombe* (negro).

Cuando se analizaron estas regiones por *indirect end-labeling* se vio que la secuencia diseñada según la *signature* de *S. pombe* era capaz de posicionar seis nucleosomas en esta especie en las posiciones predichas por la PSWM (**Figura 20B**, panel Seq-Sp/Sp, óvalos azules); sin embargo, cuando esta misma secuencia se insertó en *S. cerevisiae*, aunque los nucleosomas ocupaban algunas posiciones preferentes, estas no coincidían con las predichas en el diseño de la secuencia (**Figura 20B**, panel Seq-Sp/Sc). El análisis por secuenciación (*MNase-seq*) en ambas especies muestra el mismo resultado (**Figura 20D**). En este análisis se aprecia con más resolución que *S. cerevisiae* posiciona únicamente 5 nucleosomas (línea negra discontinua) en lugares distintos a los que lo hace en *S. pombe* (línea roja). Aún así, el hecho de que una secuencia artificial diseñada para posicionar nucleosomas en *S. pombe*, tenga capacidad de posicionar en *S. cerevisiae*, es llamativo y parece contradictorio con los resultados anteriores (**Figuras 17 y 18**). Sin embargo puede explicarse si tenemos en cuenta que en la *signature* nucleosomal, *S. cerevisiae* incrementa el contenido en A+T en el *linker* y en las zonas distales del DNA mononucleosomal, mientras que *S. pombe* lo hace en la diada. Esto puede explicar que *S. cerevisiae* coloque el *linker* donde *S. pombe* colocaría la diada, como de hecho ocurre en este experimento.

Cuando se realizó el mismo experimento con la secuencia artificial de *S. cerevisiae*, detectamos 6 nucleosomas perfectamente posicionados en las posiciones predichas en esta especie, como se muestra en el panel Seq-Sc/Sc de la **Figura 20C** y en el análisis de *MNase-seq* en la **Figura 20E** (línea roja). Sin embargo, cuando esta secuencia se integró en *S. pombe*, generó un perfil totalmente irregular (panel Seq-Sc/Sp de la **Figura 20C** y línea negra discontinua en **Figura 20E**) lo que indica que la información de posicionamiento derivada de la PSWM de *S. cerevisiae* incorporada en la secuencia, no es interpretada correctamente por *S. pombe*.

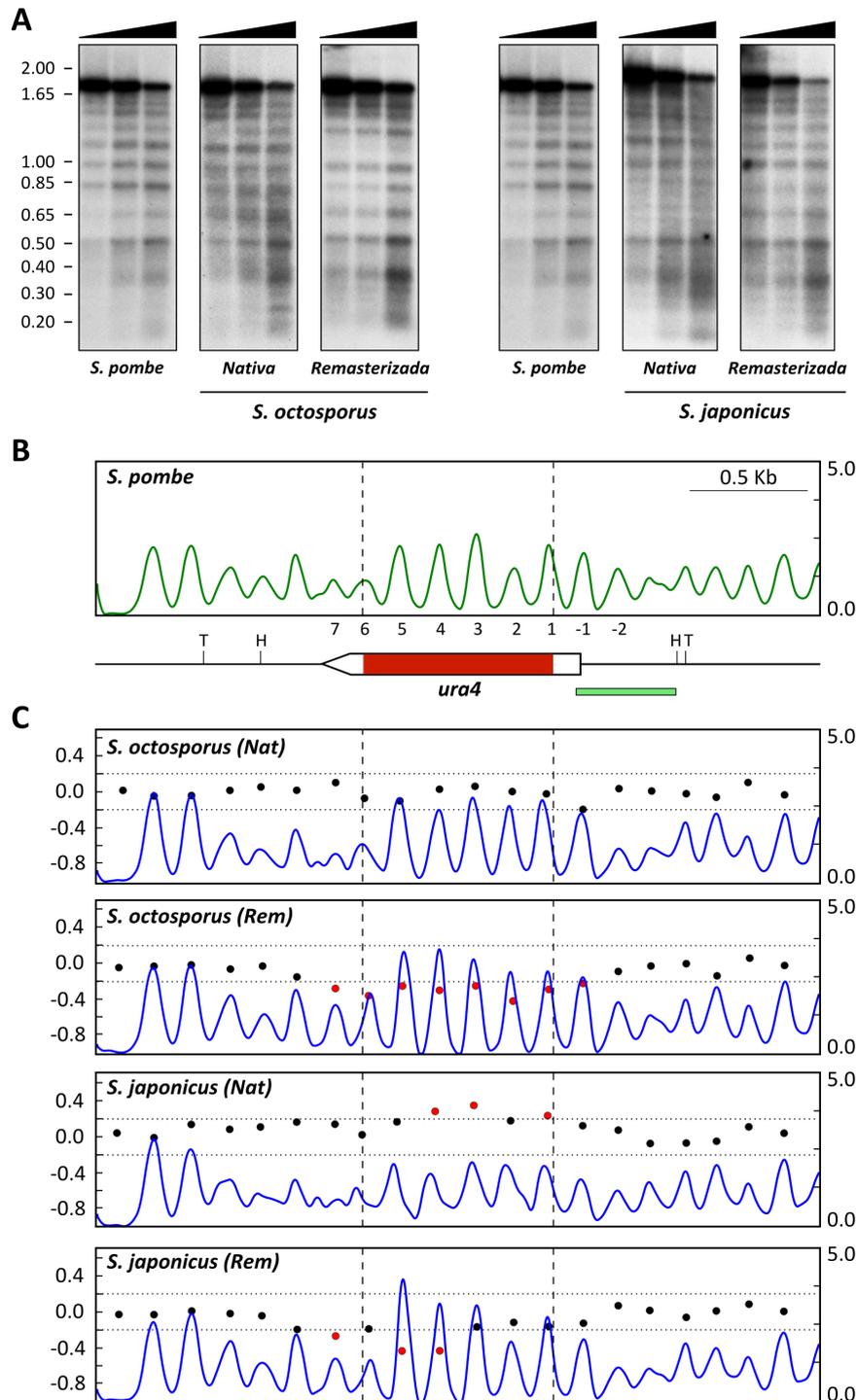
Estos resultados sugieren que la información contenida en las *signatures* nucleosomales es suficiente para posicionar nucleosomas en el contexto genómico, *in vivo* de forma específica de especie. Los requerimientos de la secuencia del DNA para posicionar nucleosomas, parecen ser más estrictos en *S. pombe* que en *S. cerevisiae*, ya que esta última es capaz de posicionar nucleosomas en secuencias diseñadas para *S. pombe*, aunque de forma mucho menos precisa que cuando se diseñan según su propia PSWM.

## Combinación del código genético y el código de posicionamiento

Los nucleosomas están posicionados a lo largo de todo el genoma, de modo que, si existe realmente un código para su posicionamiento, éste debe, necesariamente, solapar con el resto de códigos existentes, el mejor definido de los cuales es el código genético. Las *signatures* nucleosomales son muy degeneradas, por lo que en principio, existirían miles de secuencias distintas capaces de posicionar un nucleosoma en una especie concreta. Este hecho nos sugirió la posibilidad de que quizás pudiese combinarse la información de las *signatures* nucleosomales con el código genético a través del uso de codones sinónimos. Esto nos permitiría adaptar genes heterólogos para que pudieran posicionar nucleosomas en la especie hospedadora.

Para explorar esta posibilidad, tomamos como punto de partida las ORFs del gen *ura4* de *S. octosporus* y *S. japonicus*, que no eran capaces de posicionar nucleosomas en el genoma de *S. pombe* (**Figura 18**). Se modificaron estas dos ORFs seleccionando los codones sinónimos de cada aminoácido que más se aproximaban a la *signature* de *S. pombe* en función de su posición a lo largo del DNA mononucleosomal, de manera que pudieran posicionar 6 nucleosomas en los mismos lugares en los que lo hace la ORF endógena de *S. pombe*. En los paneles "Nativa" de la **Figura 21A**, se muestra el posicionamiento de las ORFs *ura4 WT* de *S. octosporus* y *S. japonicus* en el genoma de *S. pombe*, tal y como se analizó en la **Figura 18**. En los paneles "Remasterizada", puede observarse cómo las nuevas versiones con codones sinónimos restablecen el posicionamiento de la ORF original de *S. pombe*. Las ORFs remasterizadas codifican para las proteínas *ura4* de *S. octosporus* y de *S. japonicus*, y guardan una identidad con la ORF de *S. pombe*, a nivel de nucleótidos, del 73% y el 71%, respectivamente. Esta identidad es muy similar a la de las versiones nativas de *S. octosporus* y *S. japonicus* respecto a *S. pombe* (75% y 70%, respectivamente), lo que indica que el restablecimiento del posicionamiento de nucleosomas no depende directamente de la similitud global a nivel de secuencia.

Estas versiones de *ura4*, tanto las nativas de *S. octosporus* y *S. japonicus*, como las remasterizadas, fueron también analizadas por secuenciación. En la **Figura 21C** puede observarse que las diferencias de posicionamiento entre la ORF nativa de *ura4* de *S. octosporus* y la de *S. pombe* están por debajo del nivel de detección de DANPOS (panel *S. octosporus Nat*). Las de *S. japonicus*, sí son detectables, aunque las diferencias no son tan marcadas como las de *ura4 1\_6* (**Figura 8**) o *ura4 Shuffled* (**Figura 12**). Sin embargo, cuando se analizan sus versiones remasterizadas (*S. octosporus Rem* y *S. japonicus Rem*),



**Figura 21. Modificación de secuencias codificantes de *S. octosporus* y *S. japonicus* para que posicionen nucleosomas en *S. pombe*.** **A)** Análisis de *indirect end-labeling* de la región de *ura4* de una cepa WT de *S. pombe* (*S. pombe*), y de las cepas de *S. pombe* que contienen las versiones silvestres de *ura4* de *S. octosporus*, *S. japonicus*, o sus versiones remasterizadas. **B)** Análisis de ocupación de nucleosomas de *MNase-seq* de la región de *ura4* de una cepa WT de *S. pombe*. Se indican los sitios de restricción de HindIII (H) y TfiI (T) y la localización de la sonda (barra verde). **C)** Mapas de ocupación de nucleosomas de las cepas que contienen las versiones de *ura4* silvestres (Nat) y remasterizadas (Rem) de *S. octosporus* y *S. japonicus*. Análisis de DANPOS de cada una de estas cepas comparadas con el mapa WT de *S. pombe* (B). Símbolos iguales a los de la **Figura 8**.

se puede observar que el parámetro  $\Delta fuzziness$  adquiere valores negativos, indicando que el nivel de posicionamiento de esos nucleosomas es mayor incluso que el de los nucleosomas de la cepa WT de *S. pombe* que hemos utilizado como referencia.

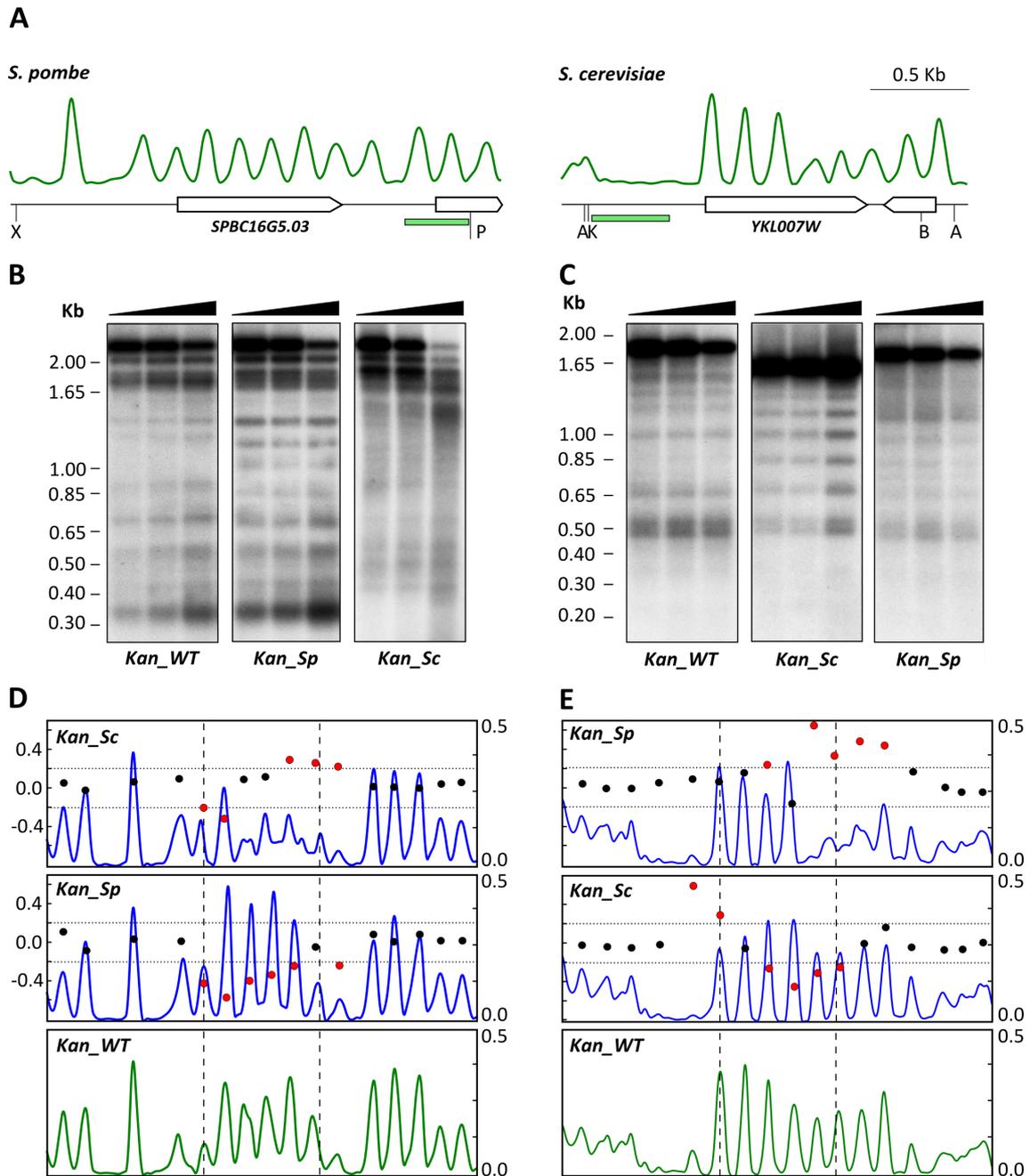
### Remasterización de genes procariotas

Para poner a prueba el potencial de nuestro método para posicionar nucleosomas, lo aplicamos al gen *Kan<sup>r</sup>*, de procedencia procariota, para analizar si incluso en este caso podría posicionar nucleosomas en *S. pombe* o en *S. cerevisiae*.

Dado que no hay ortólogos del gen *Kan<sup>r</sup>* en el genoma de la levadura, sustituimos las ORFs *SPBC16G5.03* de *S. pombe* y *YKL007W* de *S. cerevisiae* por la ORF del gen *Kan<sup>r</sup>*. Se seleccionaron estas dos ORFs porque tienen aproximadamente el mismo tamaño (807 pb) que la del gen *Kan<sup>r</sup>* (810 pb); porque no son esenciales para el crecimiento, y porque tienen nucleosomas bien posicionados en las dos especies (**Figura 22A**).

En contra de lo esperado en función de todos los resultados anteriores, la ORF del gen *Kan<sup>r</sup>* generó un perfil bastante regular tanto en *S. pombe* como en *S. cerevisiae* (Paneles *Kan\_WT* en **Figuras 22B y C**). Sin embargo, las versiones remasterizadas para cada una de estas especies generaron un perfil aún más regular y homogéneo cuando se integraron en el genoma correspondiente, incorporando seis nucleosomas en las posiciones predichas según el diseño de la secuencia (Paneles *Kan\_Sp* de la **Figura 22B** y *Kan\_Sc* de la **Figura 22C**). Como confirmación de este resultado, el análisis de DANPOS adjudicó valores negativos a las ORFs remasterizadas cuando se comparaban con la ORF del gen *Kan<sup>r</sup>* silvestre, tanto en *S. pombe* (*Kan\_Sp* en **Figura 22D**) como en *S. cerevisiae* (*Kan\_Sc* en **Figura 22E**).

Para analizar si la información del posicionamiento era específica de especie, se intercambiaron las dos versiones modificadas de *Kan<sup>r</sup>* entre las dos levaduras. En el panel *Kan\_Sc* de la **Figura 22B** puede observarse que el patrón regular generado por la ORF *Kan\_Sc* en *S. cerevisiae* (**Figura 22C**) se pierde por completo cuando se inserta en el genoma de *S. pombe*. Del mismo modo, la versión *Kan\_Sp* generó un perfil más irregular en *S. cerevisiae* (Panel *Kan\_Sp* en **Figura 22C**) que en *S. pombe* (Panel *Kan\_Sp* en **Figura 22B**). Estos mismos resultados se encuentran cuantificados por DANPOS en las **Figuras 22D** (versiones de *Kan<sup>r</sup>* en *S. pombe*) y **22E** (versiones de *Kan<sup>r</sup>* en *S. cerevisiae*). También puede observarse cómo las versiones remasterizadas para la otra especie posicionan incluso peor que la ORF procariota sin remasterizar (paneles superiores de las **Figuras 22D y E**).



**Figura 22. Modificación de genes procariotas para que posicionen nucleosomas en levaduras. A)** Distribución de nucleosomas de la región de *SPBC16G5.03* en *S. pombe* y *YKL007W* en *S. cerevisiae*. Se indican los sitios de restricción para *Xho*I (X), *Pst*I (P), *Ava*II (A), *Kpn*2I (K) y *Bst*XI (B) y la localización de las sondas (barras verdes). **B)** Análisis de *indirect end-labelling* de una cepa de *S. pombe* en la que la ORF del gen *SPBC16G5.03* ha sido sustituida por la ORF del gen *Kan<sup>r</sup>* (*Kan\_WT*) o por sus versiones remasterizadas para *S. pombe* (*Kan\_Sp*) y para *S. cerevisiae* (*Kan\_Sc*). **C)** Análisis de *indirect end-labelling* de una cepa de *S. cerevisiae* donde la ORF del gen *YKL007W* ha sido sustituida las mismas versiones del gen *Kan<sup>r</sup>*. **D)** Análisis de DANPOS de las versiones de *Kan<sup>r</sup>* en *S. pombe*. **E)** Análisis de DANPOS de las versiones de *Kan<sup>r</sup>* en *S. cerevisiae*. Las líneas de puntos verticales indican la posición de la ORF del gen *Kan<sup>r</sup>*. Símbolos como en **Figura 8**.

En conjunto, estos resultados refuerzan los obtenidos con el gen *ura4* de *S. octosporus* y *S. japonicus* (**Figura 21**) e indican que la información contenida en las *signatures* puede combinarse con ORFs de origen procariota para dirigir el posicionamiento de nucleosomas de manera específica de especie.



## DISCUSIÓN





## 1. SENSIBILIDAD DEL POSICIONAMIENTO A CAMBIOS EN LA SECUENCIA DEL DNA

Los mecanismos que dirigen el posicionamiento de los nucleosomas constituyen un campo de estudio muy activo en la actualidad. Como se ha comentado en la introducción, existen múltiples factores que influyen en el posicionamiento de los nucleosomas, entre los que se incluyen los remodeladores de cromatina, los factores de transcripción y la propia secuencia del DNA. En este trabajo nos hemos centrado en la contribución de esta última al posicionamiento de los nucleosomas en levaduras, principalmente en *S. pombe*.

La medida en la cual la secuencia del DNA afecta al posicionamiento de los nucleosomas es un tema muy controvertido. En la literatura pueden encontrarse distintas opiniones que oscilan entre dos extremos: por un lado, algunos autores defienden que existe un "código de posicionamiento", de tal forma que la secuencia del DNA determinaría individualmente la posición de la mayor parte de los nucleosomas (Segal et al. 2006; Kaplan et al. 2009; Kaplan et al. 2010). En el otro extremo está el "modelo de posicionamiento estadístico", que propone que el papel de la secuencia se limita a determinar la posición de "barreras de posicionamiento", a partir de las cuales los nucleosomas se posicionarían regularmente independientemente de la secuencia del DNA o con una influencia mínima por parte de la misma (Kornberg & Stryer 1988; Zhang et al. 2009; Möbius & Gerland 2010).

La mayoría de estos estudios llegan a dos conclusiones ampliamente aceptadas con respecto a la influencia de la secuencia del DNA en el posicionamiento. Por una parte, existen algunas secuencias de DNA que facilitan el plegamiento de la doble hélice alrededor del octámero de histonas. Concretamente, se ha propuesto que la disposición de los dinucleótidos dA:dA, dT:dA y dT:dT en el surco menor del DNA cuando éste interacciona con el octámero de histonas y dG:dC cuando está orientado hacia el exterior, favorecen la distorsión de la doble hélice necesaria para formar el nucleosoma (Drew & Travers 1985). Por otra parte, las secuencias enriquecidas en poli(dA:dT) y poli(dG:dC), confieren mayor rigidez a la molécula de DNA e impiden la formación del nucleosoma, por lo que podrían contribuir a la formación de las NDRs (Sekinger et al. 2005; Kaplan et al. 2009; Tsankov et al. 2011; Zhang et al. 2011b).

Existen numerosos estudios *in vitro* adicionales que han desvelado aspectos importantes de la interacción entre las histonas y el DNA, pero tienen la inevitable limitación de que no siempre son capaces de reproducir la situación *in vivo*, de forma que, a menudo, el posicionamiento sobre una misma secuencia *in vivo* e *in vitro* no es el mismo

(Zhang et al. 2009; Gaykalova et al. 2011; Perales et al. 2011).

En este trabajo hemos abordado esta cuestión mediante la modificación de distintas secuencias genómicas y la generación de secuencias artificiales para estudiar su efecto en el posicionamiento nucleosomal *in vivo*. El primer hallazgo llamativo ha sido la sensibilidad que los nucleosomas individuales muestran a los cambios en la secuencia, incluso cuando estos se limitan a regiones pequeñas que sólo abarcan dos o tres nucleosomas. Esta sensibilidad se manifiesta aunque no se alteren las secuencias asociadas al nucleosoma +1 y al promotor, que son las principales "barreras" propuestas por el "modelo de posicionamiento estadístico" (**Figura 8A, panel 3\_4**). Además, esta sensibilidad es independiente del contenido global en A+T y de la expresión, ya que los cambios en la secuencia afectan al posicionamiento de los nucleosomas tanto en regiones transcritas como en regiones intergénicas sin transcripción detectable (**Figura 14**). Esta independencia de la transcripción es consistente con el hecho de que el posicionamiento se mantenga más allá de los sitios de terminación de la transcripción y sobre genes inactivos durante la mitosis o la meiosis (Soriano et al. 2013).

Otra observación importante es que la pérdida de posicionamiento se restringe a la región de la secuencia modificada, afectando como mucho a los nucleosomas inmediatamente adyacentes. Este hecho, junto con los mencionados anteriormente, sugiere que el "posicionamiento estadístico" no juega un papel determinante en *S. pombe*. En vez de eso, alude a que la mayoría de las secuencias de DNA mononucleosomales podrían contener información que determinase dónde se coloca cada nucleosoma, y que fuese esta información la que hemos alterado mediante la modificación de la secuencia. Esta posibilidad estaría reforzada por el hecho de que las secuencias de DNA mononucleosomales mantienen la capacidad de posicionar cuando se integran en regiones del genoma distintas de su *locus* endógeno, propiedad que pierden cuando se *randomiza* su secuencia (**Figura 15, Puzzle 1 y C-**).

La información que determina el posicionamiento de cada nucleosoma podría estar dispersa a lo largo de todo el DNA mononucleosomal o restringida a regiones más pequeñas. En la **Figura 15 (Puzzle 2)**, puede observarse que distintas secuencias mononucleosomales pueden combinarse entre sí manteniendo la capacidad para posicionar. No obstante, cuando se ha intentado acotar más la región responsable del posicionamiento, no se ha encontrado ninguna que contribuya preferentemente a éste (**Figura 16**). Esto sugiere que la información del posicionamiento es redundante y se encuentra dispersa a lo largo de las 150 pb del DNA mononucleosomal. Así mismo, dado

que existen miles de secuencias diferentes de 150 pb asociadas con nucleosomas en *S. pombe*, esta información debe de ser también degenerada. La degeneración y la redundancia son dos propiedades importantes, ya que confieren robustez al posicionamiento, permitiendo además que la secuencia del DNA pueda sufrir variaciones sin afectar al orden global de los nucleosomas. Adicionalmente permite que exista un número enorme de combinaciones capaces de posicionar.

En resumen, sin excluir otros factores, nuestros resultados indican que la secuencia contribuye significativa-mente al posicionamiento de nucleosomas individuales a lo largo del genoma de *S. pombe* tanto en regiones codificantes como intergénicas y de forma independiente de la transcripción.

## 2. LA INFORMACIÓN RESPONSABLE DEL POSICIONAMIENTO VARÍA ENTRE ESPECIES

Como se ha comentado anteriormente, los nucleosomas se encuentran muy bien posicionados en el genoma de muchas especies, como por ejemplo *S. cerevisiae* (Kaplan et al. 2009), *Saccharomyces paradoxus* (Tirosch et al. 2010), *S. pombe* (Soriano et al. 2013), *S. octosporus*, *S. japonicus* (Quintales et al. 2015a) y *Dictyostelium discoideum* (Chang et al. 2012). Sin embargo, las características de la secuencia del DNA que influyen en el posicionamiento de los nucleosomas varían de unas especies a otras.

Se sabe que cuando se introducen fragmentos grandes de DNA de una especie en otra, las secuencias exógenas se empaquetan de manera diferente. Esto se observa tanto a nivel de cromosomas metafásicos (McManus et al. 1994) como a nivel de posicionamiento de nucleosomas (Hughes et al. 2012). En este último caso, más relacionado con nuestro trabajo, los autores introdujeron fragmentos de DNA de aproximadamente 140 Kb de *Kluyveromyces lactis*, *Kluyveromyces waltii* y *Debaryomyces hansenii* en *S. cerevisiae* a través de cromosomas artificiales de levaduras (YACs) y observaron que el posicionamiento de nucleosomas se perdía en *S. cerevisiae*. Además, aparecían NDRs en zonas que en la especie donadora se encontraban ocupadas por nucleosomas, incluso dentro de regiones codificantes. Este fenómeno se produce también entre especies filogenéticamente tan próximas como *Saccharomyces paradoxus* y *S. cerevisiae* (Tirosch et al. 2010) y cuando los fragmentos de DNA exógeno se limitan a unas 950 pb, como se muestra en la **Figura 17**. Incluso genes ortólogos pierden el posicionamiento cuando se integran en distintas especies del mismo género (**Figura 18**).

En resumen, por lo que sabemos hasta ahora, en la mayoría de las especies de levaduras los nucleosomas se mantienen posicionados a lo largo de todo el genoma, aunque la contribución de la secuencia a dicho posicionamiento parece ser diferente en cada especie. No obstante, existe una característica de la secuencia del DNA, relacionada con los nucleosomas, que sí parece estar conservada evolutivamente: la periodicidad de los dinucleótidos dA:dA, dT:dA y dT:dT cada 10 pb. Esta periodicidad ha sido observada en los fragmentos de DNA mononucleosomal de un gran número de organismos eucariotas, incluyendo *S. cerevisiae* y *S. pombe* (Brogaard et al. 2012; Moyle-Heyrman et al. 2013) y se ha sugerido que podría contribuir al posicionamiento rotacional de los nucleosomas (Mavrigh et al. 2008b).

### **3. SIGNATURES NUCLEOSOMALES**

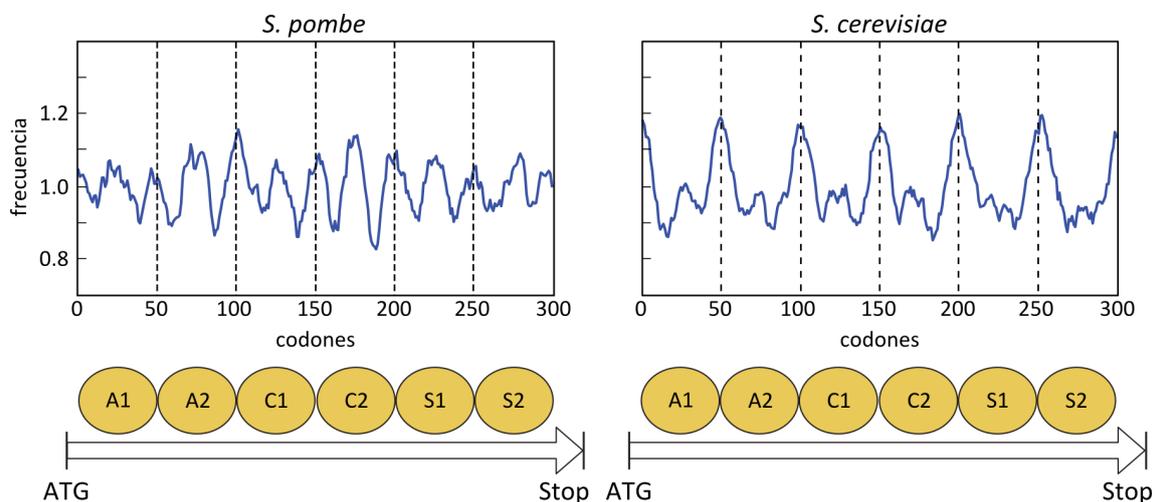
Además de la periodicidad de dinucleótidos comentada en el apartado anterior, existe una característica del DNA mononucleosomal que varía de unas especies a otras y que podría contribuir al posicionamiento translacional de los mismos: las *signatures* nucleosomales (Quintales et al. 2015a), que se han descrito en la introducción (pág. 21).

Se desconoce cómo se originaron estas *signatures* nucleosomales. Por un lado, podrían ser consecuencia de la diferente estabilidad del DNA a lo largo del octámero de histonas. Si la posición de los nucleosomas con respecto a la secuencia del DNA se hubiese mantenido estable durante largos periodos evolutivos, las *signatures* podrían haber aparecido como una huella dejada por el nucleosoma, como consecuencia de una diferencia en la frecuencia de mutación a lo largo del DNA mononucleosomal o de un sesgo en su reparación (Washietl et al. 2008; Sasaki et al. 2009). Por ejemplo, en *S. cerevisiae* se ha descrito que la frecuencia de sustitución de nucleótidos es mayor en la zona de la diada (Washietl et al. 2008).

Por otro lado, podría existir también un mecanismo de selección positiva de estas *signatures* por su contribución al posicionamiento de los nucleosomas. Prendergast & Semple (2011) realizaron estudios evolutivos con secuencias de DNA mononucleosomales procedentes de humanos y chimpancés. Estos autores encontraron una mayor frecuencia de sustituciones de los nucleótidos A y T hacia G y C en la zona de la diada, mientras que el cambio inverso ocurría en la región del linker. Además, encontraron que la frecuencia de sustitución de bases en determinadas zonas del DNA mononucleosomal variaba en función de la composición en G+C del contexto genómico, por ejemplo, en regiones ricas en G+C, la frecuencia de sustitución de bases era menor en

la diada y mayor en el linker, mientras que en regiones genómicas ricas en A+T, ocurría lo contrario. En base a estos estudios, Prendergast & Semple sugirieron la existencia de un mecanismo de selección positiva dirigido a mantener una composición óptima de G+C en la región de la diada para favorecer el posicionamiento de los nucleosomas.

Otra evidencia en favor del origen selectivo de las *signatures* es el impacto que estas causan sobre la composición de aminoácidos de las proteínas. Por ejemplo, el aminoácido Ile está codificado por codones ricos en A y T (ATT, ATA y ATC). Estos codones están enriquecidos en las regiones correspondientes al linker en *S. cerevisiae*, y al linker y la diada en *S. pombe*, lo que se refleja en la composición promedio de este aminoácido en las proteínas de estas dos especies, que sigue un perfil periódico (**Figura 23**). El resto de aminoácidos siguen periodicidades similares en función de la composición de sus codones. De hecho, las *signatures* nucleosomales pueden predecir correctamente la distribución relativa de los 20 aminoácidos a lo largo de las proteínas en base a la posición de los codones correspondientes a lo largo del DNA mononucleosomal, lo que sugiere una relación directa entre el posicionamiento nucleosomal y la evolución de las proteínas (Warnecke et al. 2008; Quintales et al. 2015a).



**Figura 23. Distribución periódica del aminoácido isoleucina a lo largo de las ORFs.** El eje Y indica la frecuencia relativa del aminoácido Ile, el eje X indica el número de codones a lo largo de los seis grupos de DNA mononucleosomal. El diagrama inferior representa la posición de los dos nucleosomas inmediatamente aguas abajo de cada codón ATG (A1 y A2), dos nucleosomas del centro de las ORFs (C1 y C2) y los dos nucleosomas inmediatamente aguas arriba de cada codón STOP (S1 y S2). Figura modificada de Quintales et al. (2015a).

#### 4. GENERACIÓN DE SECUENCIAS PARA POSICIONAR NUCLEOSOMAS

Si la secuencia del DNA contribuye de manera significativa al posicionamiento de los nucleosomas, debería ser posible diseñar secuencias artificiales capaces de colocar nucleosomas en regiones determinadas en el contexto genómico. La principal dificultad radica en que, en principio, esperaríamos que los requerimientos de la secuencia del DNA necesarios para posicionar un nucleosoma fuesen muy degenerados, ya que, como se ha mencionado anteriormente, existen miles de secuencias distintas asociadas a nucleosomas bien posicionados.

Las *signatures* nucleosomales cumplen la condición de ser degeneradas y, además, específicas de cada especie. Por ese motivo las tomamos como punto de partida para diseñar secuencias que podrían dirigir el posicionamiento de nucleosomas de manera específica en *S. pombe* y *S. cerevisiae* (**Figura 20**). Nuestros resultados sugieren que las *signatures* nucleosomales contienen información que es correctamente interpretada por las especies de las que proceden para posicionar los nucleosomas y constituyen la prueba más sólida que existe hasta ahora de la contribución de la secuencia del DNA al posicionamiento de los nucleosomas *in vivo*.

Una característica importante de estas secuencias artificiales es que, al igual que ocurre con las secuencias naturales, posicionan nucleosomas de manera específica de especie, es decir, las secuencias diseñadas para *S. pombe* posicionan en *S. pombe* y no en *S. cerevisiae* y viceversa. Estos resultados podrían contribuir a explicar las observaciones previas de que nucleosomas procedentes de especies filogenéticamente distantes empaqueten la misma secuencia de DNA de maneras distintas (Bernardi et al. 1992; McManus et al. 1994; Sekinger et al. 2005; Hughes et al. 2012).

Con anterioridad a este trabajo se habían publicado modelos probabilísticos y herramientas computacionales con la finalidad de predecir la posición de los nucleosomas a partir de la secuencia del DNA (Revisado en Teif 2016). No hemos probado la capacidad de predicción de nuestras PSWMs, pero el hecho de que la secuencia posicione de manera diferente en especies distintas, pone de manifiesto limitaciones importantes a la hora de aplicar cualquier predictor en especies distintas a las utilizadas para generarlo.

La degeneración que permiten las *signatures* nucleosomales es suficiente como para que esta información pueda combinarse con el código genético mediante el uso de

codones sinónimos. De este modo, hemos conseguido adaptar ORFs de especies tanto eucariotas como procariotas para que, manteniendo su codificación original, sean capaces de posicionar nucleosomas en *S. pombe* o en *S. cerevisiae*. (**Figuras 21 y 22**). Esta posibilidad es importante para las posibles aplicaciones de este método que se proponen más adelante.

## 5. INTERPRETACIÓN DE LAS SIGNATURES

Una de las cuestiones que queda por resolver es cuál es el mecanismo por el cual las *signatures* son interpretadas y por qué esta interpretación difiere entre especies. Generalmente, las propiedades de la secuencia del DNA que favorecen la formación de nucleosomas han sido atribuidas a una mayor facilidad para que la molécula de DNA se curve alrededor del octámero de histonas (Segal & Widom 2009). Sin embargo, las histonas son proteínas muy conservadas evolutivamente y la estructura del octámero es muy parecida en todas las especies. Esto hace difícil explicar que las propiedades de la secuencia del DNA que determinan el posicionamiento de los nucleosomas varíen incluso entre especies cercanas filogenéticamente. Aún así, se ha descrito recientemente que las histonas de distintos organismos muestran afinidades diferentes por las mismas secuencias de DNA (Allan et al. 2013). Estos autores sugieren que, a pesar de su similitud, los octámeros de histonas de distintos organismos pueden haberse adaptado durante la evolución para empaquetar genomas que difieren ampliamente en su composición de bases, tamaño, densidad de genes y otras propiedades estructurales y funcionales.

Las histonas de las levaduras utilizadas en este estudio varían en identidad de aminoácidos con respecto a *S. pombe* entre un 96.5% - 99.0% en *S. octosporus*, un 94.7% - 98.5% en *S. japonicus* y un 83.9% - 92.6% en *S. cerevisiae*. Actualmente no se sabe si estas pequeñas diferencias podrían ser responsables de la variación en la afinidad de los nucleosomas de distintos organismos por la misma secuencia de DNA (Bernardi et al. 1992; McManus et al. 1994; Sekinger et al. 2005; Hughes et al. 2012) o si podría haber otros factores implicados, como la concentración de histonas, el tamaño del linker u otros parámetros termodinámicos o biofísicos (Beshnova et al. 2014). Sería interesante estudiar con más detalle si las histonas son las responsables de esta interpretación de las *signatures*. Si así fuese, sería un factor clave a tener en cuenta para la realización de estudios estructurales que implican la reconstitución de cromatina *in vitro*. La mayoría de los experimentos de ensamblaje de cromatina *in vitro* descritos en la literatura se han realizado utilizando histonas de una especie y DNA de otra, en ocasiones muy alejada

filogenéticamente (Kaplan et al. 2009; Tsankov et al. 2011). Por ejemplo, la estructura tridimensional del nucleosoma, descrita por Luger et al. (1997) con resolución de 2.8 Å fue establecida a partir de histonas de *Xenopus laevis* y DNA humano (Luger et al. 1997).

Por otra parte, aunque las histonas tienen un contacto más íntimo con el DNA, los remodeladores de cromatina también podrían contribuir a la interpretación de las *signatures* nucleosomales. Existen experimentos de ensamblaje de cromatina *in vitro* que han conseguido imitar el posicionamiento *in vivo* de las regiones 5' de los genes de *S. cerevisiae* en presencia de remodeladores de cromatina dependientes de ATP (Zhang et al. 2011b). De manera inversa, se puede alterar el posicionamiento de nucleosomas *in vivo*, de forma global, mediante la delección de los genes que codifican para algunos de esos remodeladores, como *hrp3* en *S. pombe* (Pointner et al. 2012) o *ISW1* y *CHD1* en *S. cerevisiae* (Gkikopoulos et al. 2011).

En cualquier caso, al igual que las propiedades de la secuencia del DNA que determinan el posicionamiento varían entre especies, los distintos factores que contribuyen al posicionamiento (**véase Introducción**) podrían tener una influencia diferente en distintas especies e incluso en diferentes regiones del mismo genoma. En cuanto a las diferencias de posicionamiento entre *S. cerevisiae* y *S. pombe*, un parámetro que puede jugar un papel muy importante es la presencia de la histona H1 y el mayor tamaño del linker en *S. cerevisiae*. El hecho de que *S. cerevisiae* sea capaz de posicionar algunos nucleosomas en secuencias diseñadas para *S. pombe* (aunque en posiciones distintas a las esperadas según el diseño de la secuencia), podría explicarse por el modelo de posicionamiento estadístico (Kornberg & Stryer 1988). El mayor tamaño del linker en *S. cerevisiae* explicaría que los nucleosomas en la **Figura 20D** no ocupen las mismas posiciones, con respecto a la secuencia, que ocupan en *S. pombe*. Sin embargo, aunque la secuencia podría tener menos importancia para posicionar nucleosomas en *S. cerevisiae* que en *S. pombe*, sabemos que tiene alguna influencia, ya que el posicionamiento de nucleosomas sobre el gen *Kan<sup>r</sup>* remasterizado para *S. cerevisiae* es significativamente mayor que el del mismo gen sin remasterizar, y mucho mejor que el posicionamiento que se observa sobre el gen *Kan<sup>r</sup>* remasterizado para *S. pombe* (**Figuras 22C y E**).

## 6. FUNCIÓN DEL POSICIONAMIENTO DE NUCLEOSOMAS

En este trabajo hemos estudiado la influencia de la secuencia del DNA en el posicionamiento de nucleosomas, pero uno de los detalles más intrigantes, que aún está por resolver, es cuál es la función de este posicionamiento. El hecho de que todas las células de una población de levaduras de la misma especie tengan la mayor parte de los nucleosomas colocados en el mismo sitio con respecto a la secuencia del DNA, lleva a suponer que este orden deba tener alguna función, incluso aunque los mecanismos que contribuyen a mantenerlo no estén conservados evolutivamente.

De hecho, como se comentó en el apartado 4 de la Discusión, se han hallado indicios de selección positiva en las secuencias de DNA mononucleosomal humano (Prendergast & Semple 2011). En la **Figura 12** hemos mostrado cómo el simple hecho de cambiar el orden de los distintos codones sinónimos dentro de una ORF puede alterar el posicionamiento de esa región concreta. También es llamativo el hecho de que pueda encontrarse un sesgo en la distribución relativa de los 20 aminoácidos a lo largo de las proteínas, dependiendo de la posición de los codones correspondientes a lo largo del DNA mononucleosomal que, a su vez, está sesgada por la *signature* nucleosomal (Warnecke et al. 2008; Quintales et al. 2015a). Si, como parecen indicar estos estudios, la secuencia del DNA ha sido seleccionada a lo largo de todo el genoma e incluso en regiones codificantes, para favorecer el posicionamiento de los nucleosomas, es razonable suponer que este posicionamiento conlleve alguna ventaja evolutiva.

Una de las consecuencias de la pérdida de posicionamiento, ya sea por la delección de un remodelador concreto (Pointner et al. 2012) o por la inserción de fragmentos grandes de DNA exógeno (Tirosch et al. 2010; Hughes et al. 2012) es la activación, aparentemente descontrolada, de la transcripción antisentido, así como la aparición de promotores crípticos en lugares inusuales, como es en medio de regiones codificantes. Así pues, una de las funciones del posicionamiento de nucleosomas podría ser la regulación global del genoma mediante la modulación del acceso a la secuencia de los distintos factores que interaccionan con el DNA. De hecho, los efectos de los nucleosomas sobre la transcripción han sido estudiados ampliamente y se ha descrito que existe una compleja interacción entre los factores de transcripción, los nucleosomas y las enzimas remodeladoras de cromatina que, en conjunto, regula la expresión de los genes (Cairns 2009). Por ejemplo, se sabe que la mayor parte de las secuencias consenso de los factores de transcripción se encuentran agrupadas en *clusters* en las regiones libres de nucleosomas de los promotores de los genes (Badis et al. 2008; Soriano et al. 2013).

También se ha descrito que los sitios de unión de algunos factores de transcripción, como Rap1 y Reb1, tienen una orientación precisa sobre la superficie de un nucleosoma, lo que resalta la importancia del posicionamiento rotacional estricto de ese nucleosoma concreto, que colocado de otro modo, podría ocultar los motivos de unión de estos factores (Koerber et al. 2009).

Adicionalmente, se ha descrito que el posicionamiento está asociado con otros procesos genómicos como el *splicing* del mRNA, la replicación y la reparación. (Berbenetz et al. 2010; Duan & Smerdon 2010; Tilgner & Guigó 2010; Soriano et al. 2014). Los sitios de recombinación meiótica colocalizan con regiones libres de nucleosomas, algunas de las cuales se abren específicamente durante la meiosis (de Castro et al. 2012). También se ha descrito que los orígenes de replicación de *S. cerevisiae* presentan un NDR delimitado por dos nucleosomas bien posicionados que se desplazan ligeramente durante la activación del origen, ampliando la región libre de nucleosomas (Eaton et al. 2010; Soriano et al. 2014).

La capacidad para alterar el posicionamiento de los nucleosomas *in vivo* modificando la secuencia del DNA, podría ser una herramienta muy útil para estudiar, con más detalle, el papel de los nucleosomas en todos los procesos descritos anteriormente. Igualmente, sería interesante explorar si la modificación de secuencias exógenas, con la finalidad de imitar el posicionamiento del hospedador, podría mejorar la expresión o la estabilidad de genes y vectores de interés biotecnológico.

Por último, existen estudios que permiten modular la eficacia de los promotores basándose en su capacidad para posicionar o excluir nucleosomas (Curran et al. 2014). El efecto de las *signatures* en el posicionamiento descrito en este trabajo podría permitir ampliar el método descrito por Curran et al. (2014) a otras especies. El potencial de las *signatures* nucleosomales para modelar el posicionamiento de nucleosomas tanto en regiones codificantes como intergénicas, junto con el diseño de promotores de potencia variable, abre la posibilidad de incorporar esta información en el diseño de genomas sintéticos (Annaluru et al. 2014; Haimovich et al. 2015), permitiendo tal vez una mayor estabilidad de los cromosomas artificiales y un mayor control sobre la expresión génica.

## CONCLUSIONES





1. La modificación de la secuencia del DNA afecta al posicionamiento de los nucleosomas *in vivo* a nivel individual, tanto en regiones codificantes como no codificantes e independientemente de la transcripción.
2. La información que contribuye al posicionamiento de los nucleosomas es degenerada y se encuentra dispersa a lo largo de los 150 pares de bases del DNA mononucleosomal. Esta información se mantiene en secuencias integradas en posiciones ectópicas del genoma y es específica de especie.
3. La distribución de los codones sinónimos a lo largo de las regiones codificantes de los genes está optimizada para favorecer el posicionamiento de los nucleosomas.
4. Es posible diseñar secuencias artificiales capaces de posicionar nucleosomas de manera específica de especie mediante la incorporación de la información de las *signatures* nucleosomales, a través de la generación de Matrices de Puntuación Específica de Posición (PSWM).
5. La información contenida en las *signatures* nucleosomales es lo suficientemente degenerada como para permitir modificar las ORFs de una especie, sin alterar la secuencia original de aminoácidos de las proteínas codificadas, para optimizar el posicionamiento de nucleosomas cuando se incorporan en el genoma de otra especie diferente.



## MATERIALES Y MÉTODOS

---



## 1. MICROORGANISMOS Y MEDIOS DE CULTIVO UTILIZADOS

Las cepas descritas en este trabajo se han construido a partir de la estirpe *h-972* de *S. pombe*, salvo las cepas que contienen variantes de la ORF del gen *ura4*, que se han construido a partir de un mutante *ura4-DS/E (h-leu1-32 ura4-DS/E)*, el cual tiene una delección de aproximadamente 300 pb dentro de la ORF del gen *ura4*. Las cepas de *S. cerevisiae* se construyeron a partir de la estirpe *aw303-1a (MATa ade2-1 ura3-1 his3-11,15 trp1-1 leu2-3,112 can1-100)*. Además, se ha utilizado *S. octosporus* (CBS 1804), *S. japonicus var. japonicus* (FY53 *ade12<sup>-</sup>*) y la cepa DH5 $\alpha$  de *Escherichia coli* (*F- endA1 gyr96 hsdR17  $\Delta$ lacU169  $\beta$ 80lacZ $\Delta$ M15 recA1 relA1 supE44-thi-1*) en experimentos de clonación y amplificación de plásmidos.

Las cepas del género *Schizosaccharomyces* se cultivaron a 32 °C en los medios descritos por Moreno et al. (1991), concretamente en medio rico (YES) o en medio mínimo (MM) con 225 mg/L de los suplementos necesarios. Las cepas de *S. cerevisiae* se cultivaron a 28 °C en YEPD (1% de extracto de levadura, 2% de bacto-peptona y 2% de glucosa). *E. coli* se cultivó en medio rico LB (0,5% extracto de levadura, 1% bacto-triptona y 1% NaCl). El crecimiento en medio líquido se determinó midiendo la turbidez del cultivo por su absorbancia a 595 nm ( $OD_{595}$ ).

## 2. GENERACIÓN DE SECUENCIAS QUE POSICIONAN NUCLEOSOMAS

### 2.1. Generación de la matriz de puntuación

Las *signatures* nucleosomales se definieron por la composición asimétrica y palindrómica de los cuatro nucleótidos a lo largo de 38154, 46120, 27024 y 34526 secuencias de DNA mononucleosomal de *S. pombe*, *S. octosporus*, *S. japonicus* y *S. cerevisiae*, respectivamente, alineadas con respecto a la posición de la diada (Quintales et al. 2015a). Para extraer la información contenida en las *signatures* nucleosomales, se generó una matriz de puntuación de posición específica (PSWM) que incorporaba la frecuencia de cada uno de los 16 dinucleótidos a lo largo de las 149 pb de las secuencias de DNA asociadas a nucleosomas bien posicionados indicadas anteriormente. Por tanto, las dimensiones de la PSWM fueron 16 (dinucleótidos) x 149 (posiciones a lo largo del DNA mononucleosomal). Los valores para cada posición (*i*) y para cada dinucleótido (*NN*) se calcularon de acuerdo a la siguiente expresión:

$$Puntuación_i^{NN} = \log_2 \frac{Freq_i^{NN}}{Freq_{genómica}^{NN}}$$

donde  $Freq_i^{NN}$  es la frecuencia del dinucleótido NN en la posición  $i$  de todas las secuencias alineadas y  $Freq_{genómica}^{NN}$  es la frecuencia promedio del mismo dinucleótido en todo el genoma. Cada uno de los 2384 (16 x 149) elementos de la matriz representa la puntuación para cada dinucleótido dependiendo de su posición a lo largo del DNA mononucleosomal. Cada especie genera una matriz diferente dependiendo de su *signature* nucleosomal.

## 2.2. Diseño de moléculas de DNA sintéticas no codificantes para el posicionamiento nucleosomal

Para incorporar la información de las *signatures* nucleosomales de *S. pombe* en las moléculas sintéticas mostradas en la **Figura 20**, generamos seis secuencias aleatorias de 153pb (147 pb + 6 pb de linker) (Lantermann et al. 2010), con un contenido promedio de G+C del 36%. En el caso de *S. cerevisiae*, las seis secuencias tenían una longitud de 168 pb (147 pb + 21 pb de linker) (Lantermann et al. 2010) y un contenido promedio del 38% en G+C. Cada una de estas secuencias individuales se sometió, *in silico*, a rondas sucesivas en las que se incorporaron mutaciones puntuales al azar y se seleccionaron aquellas secuencias que, después de cada ronda, mejoraban su puntuación en la PSWM. Someter a estas secuencias a un número ilimitado de ciclos produciría seis secuencias mononucleosomales idénticas. Para evitar esto, detuvimos los ciclos de mutación antes de que el porcentaje de identidad entre ellas fuese significativamente mayor del 25% de identidad promedio que hay entre distintas secuencias de DNA de 150 pb seleccionadas al azar en el genoma.

## 2.3. Incorporación de las *signatures* nucleosomales en ORFs

Para reproducir el patrón de posicionamiento de una ORF endógena de *S. pombe* con otra ORF del mismo tamaño procedente de otra especie, identificamos en primer lugar el nucleótido central de un nucleosoma bien posicionado en la ORF endógena. Consideramos que el resto de nucleosomas se encuentran a distancias equivalentes a múltiplos de 153 pb. A continuación, para cada codón a lo largo de la ORF exógena, se seleccionó el codón sinónimo con la mayor puntuación en la PSWM de *S. pombe* en la posición correspondiente. Para adaptar las ORFs exógenas a la matriz de *S. cerevisiae*, se utilizó el mismo protocolo, pero en ese caso, se tomaron 168 pb como distancia

internucleosomal y se utilizó como referencia la PSWM de *S. cerevisiae*. Dado que ninguna de las ORFs remplazadas en *S. pombe* o en *S. cerevisiae* contenía un número entero de nucleosomas, mantuvimos los codones de ATG y STOP y las secuencias externas a la ORF, quedaron sin modificar.

## 2.4. Síntesis de moléculas de DNA

Todas las moléculas de DNA sintético utilizadas en este trabajo se encargaron a GENEWIZ® o GeneArt™ (Life Technologies). En el diseño de las secuencias se incluyeron los sitios de recombinación y de restricción necesarios para todos los pasos siguientes. Estas empresas envían las secuencias solicitadas integradas en plásmidos, en los que introducimos un marcador junto a la secuencia sintetizada cuando es necesario. Por último, extraemos el fragmento de interés mediante digestión con enzimas de restricción y lo utilizamos para transformar las levaduras. 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.

## 3. TÉCNICAS DE TRANSFORMACIÓN DE MICROORGANISMOS

Los distintos mutantes, tanto de *S. pombe* como de *S. cerevisiae* se generaron por recombinación homóloga. Las variantes del gen *ura4* se integraron directamente en el *locus* del gen *ura4* de *S. pombe* y las variantes del gen *Kan<sup>r</sup>* se integraron en el *locus* del gen *SPAC16G5.03* de *S. pombe* o *YKL007W* de *S. cerevisiae*. En todos los casos, se sustituyó la ORF del gen endógeno por la ORF modificada correspondiente y se dejaron intactas las regiones promotoras y terminadoras. El resto de secuencias intergénicas y no codificantes se ligaron al gen marcador *KanMX4* antes de ser utilizadas para transformar. La comprobación final de los transformantes se llevó a cabo mediante PCR, *Southern-blot* o secuenciación.

### 3.1. Transformación de *E. coli*

La transformación de *E. coli* se realizó siguiendo el procedimiento descrito por Kushner (1978). La cepa de *E. coli* utilizada para la clonación de fragmentos de DNA y la amplificación de los plásmidos fue la estirpe DH5 $\alpha$ .

### **3.2. Transformación de *S. pombe***

La transformación de *S. pombe* se llevó a cabo por la técnica de electroporación (Forsburg & Rhind 2006). Cuando se integró el gen *ura4*, las células se seleccionaron directamente en placas de MM. Cuando se utilizó el gen marcador *KanMX4* o alguna variante de éste, las células se recuperaron en placas de YES durante 24h; para seleccionar los transformantes, se hicieron réplicas en placas de YES con sulfato de geneticina (G418; Gibco) a una concentración de 25 mg/L para los casos en los que el gen *KanMX* está bajo el promotor de *SPAC16G5.03* de *S. pombe*, o 100 mg/L para el módulo *KanMX4*, que posee el promotor TEF de *Ashbya gossypii*.

### **3.3. Transformación de *S. cerevisiae***

La transformación de *S. cerevisiae* se realizó por la técnica del acetato de litio (Amberg et al. 2005). Para seleccionar los transformantes, las células se incubaron durante 2h en YEPD a 28°C en agitación y después se sembraron en placas de YEPD con G418 a una concentración final de 100 mg/L.

## **4. TÉCNICAS DE OBTENCIÓN Y MANIPULACIÓN DE DNA**

### **4.1. Aislamiento de DNA plasmídico de *E. coli***

Para aislar DNA plasmídico de *E. coli* a pequeña escala se utilizó el método de la lisis alcalina (Birnboim & Doly 1979). Para la obtención de DNA plasmídico a gran escala se usó el kit Wizard Plus Midipreps de Promega, siguiendo las especificaciones del fabricante.

### **4.2. Aislamiento de DNA genómico de *S. pombe***

La extracción de DNA genómico de *S. pombe* se realizó siguiendo el protocolo descrito por Moreno et al. (1991), exceptuando las preparaciones de DNA a partir de cromatina digerida con nucleasa micrococcal, que se detallan en los **Apartados 4.4 y 4.5**.

### 4.3. Southern-blot

Los fragmentos de DNA, generados con enzimas de restricción, se separaron mediante electroforesis en geles de agarosa. Dependiendo de su tamaño, se usaron geles de agarosa de concentración del 0.8% al 2% en tampón TAE a pH 8.0 (Tris-acetato 40 mM, EDTA 1 mM). La electroforesis se realizó a voltaje constante (50 - 120 V) en TAE. A continuación se incubó el gel durante 10 min en HCl 0,125 M para depurinar el DNA; seguidamente, se desnaturalizó en una solución de NaCl 1,5 M y NaOH 0,5 M durante 20 min. La transferencia del DNA a una membrana de nylon Hybond-XL (GE Healthcare) se realizó por capilaridad y en la misma solución de desnaturalización, tras lo cual se neutralizó la membrana incubándola durante 30 min en 2x SSC (0,3 M NaCl, 30 mM citrato sódico).

Para su hibridación con una sonda radiactiva, se prehibridó la membrana durante al menos 2 horas a 65°C en 8 ml de Solución de Prehibridación (NaCl 360 mM; Na<sub>2</sub>HPO<sub>4</sub> · 7 H<sub>2</sub>O 20 mM; EDTA 2 mM, SDS 1%, leche en polvo 0,5%, azida sódica 0,01% y DNA de esperma de salmón sonicado 0,5 mg/ml (Roche) previamente desnaturalizado). A continuación se sustituyó la solución anterior por 8 ml de la solución de hibridación, con la misma composición que la solución de prehibridación pero incluyendo Dextran Sulfato 10% (Sigma) y la sonda marcada radiactivamente.

Para preparar las sondas, en primer lugar se amplificó por PCR la región de interés y se purificó el fragmento resultante mediante electroforesis en gel de agarosa, utilizando el kit *GFX™ PCR DNA and Gel Band Purification* (GE Healthcare). 50 ng del producto de la PCR, previamente desnaturalizados, se incubaron durante 1 hora a 37°C en una mezcla de 20 µl totales que contenía 75 µM de dATP, dTTP y dGTP, 2 µl de *Hexanucleotide mix 10x* (Roche), 2 U de enzima Klenow (Roche) y 50 µCi de [<sup>32</sup>P]-α-dCTP. Para eliminar los nucleótidos no incorporados, la sonda se purificó mediante columnas *ProbeQuant G-50* (GE Healthcare). La hibridación se llevó a cabo durante 12 - 14 horas a 65 °C en la solución de hibridación descrita anteriormente. A continuación, se eliminó la solución de hibridación y se lavó dos veces la membrana: primero con una solución que contiene SSPE 2x (NaCl 360 mM; Na<sub>2</sub>HPO<sub>4</sub> · 7 H<sub>2</sub>O 20 mM; EDTA 2 mM) y SDS 0,1% durante 10 minutos a 65°C y a continuación durante 30 minutos a 65°C con SSPE 0,1x y SDS 0,1%. Finalmente, las membranas se envolvieron en plástico y se utilizaron para exponer películas de autorradiografía *BioMax XAR* (Kodak) a una temperatura de -80°C durante el tiempo necesario en función de la intensidad de la señal.

#### 4.4. Digestión con nucleasa micrococcal y análisis de patrones de nucleosomas mediante *indirect end-labeling*

Para estudiar el posicionamiento de nucleosomas en regiones específicas utilizamos la técnica de *indirect end-labeling* (Lantermann et al. 2009).

##### *Indirect end-labeling en S. pombe*

Para realizar la técnica de *indirect end-labeling* en *S. pombe*, se centrifugaron 400 ml de cultivo a  $OD_{595} = 0.8$  en un rotor F14S-6X250 a 6000 rpm durante 10 min y las células sedimentadas se lavaron con H<sub>2</sub>O destilada estéril. A continuación se resuspendieron en 20 ml de solución de preincubación (ácido cítrico 20 mM, Na<sub>2</sub>HPO<sub>4</sub> 20 mM, EDTA pH 8.0 40 mM) y se incubaron a 30°C durante 10 minutos en un baño de agua con agitación (200 rpm). Posteriormente se recogieron las células por centrifugación a 3200 g durante 5 min, se resuspendieron en 10 ml de tampón Sorbitol/Tris (sorbitol 0.5 M, Tris/HCl 50 mM pH 7.5, β-mercaptoetanol 10 mM) con 40 mg de Zimoliasa 20T y se incubaron durante 30 minutos a 30°C con agitación (200 rpm). Seguidamente, se centrifugaron las células a 3200 g, se lavaron los protoplastos con 40 ml de tampón Sorbitol/Tris sin β-mercaptoetanol y, tras volver a centrifugar en las mismas condiciones, se resuspendieron en 7 ml de tampón NP (sorbitol 1 M, NaCl 50 mM, Tris/HCl 10 mM pH 7.5, MgCl<sub>2</sub> 5 mM, CaCl<sub>2</sub> 1 mM, β-mercaptoetanol 1 mM, espermidina 0.5 mM, NP-40 0.07%). A continuación se dividió la muestra en siete alícuotas de 1 ml, se añadió nucleasa micrococcal a seis de ellas, a concentraciones finales de 0, 1.5, 3, 7, 15 y 30 U/ml, y se incubaron durante 10 min a 37°C en un baño de agua sin agitación; inmediatamente después, se paró la reacción con 125 μl de una solución de SDS 5% y EDTA 50 mM y se incubaron todas las alícuotas con 1 mg/ml de proteinasa K durante 12 - 14 h a 65°C.

El DNA de todas las alícuotas se extrajo con 1 ml de fenol:cloroformo (1:1) y se precipitó añadiendo 1/25 volúmenes de NaCl 5 M y 0.7 volúmenes de isopropanol 100%, incubando durante 1 h a -20°C y centrifugando durante 45 min a 4°C a 16000 g. 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 durante 1 h a 37°C.

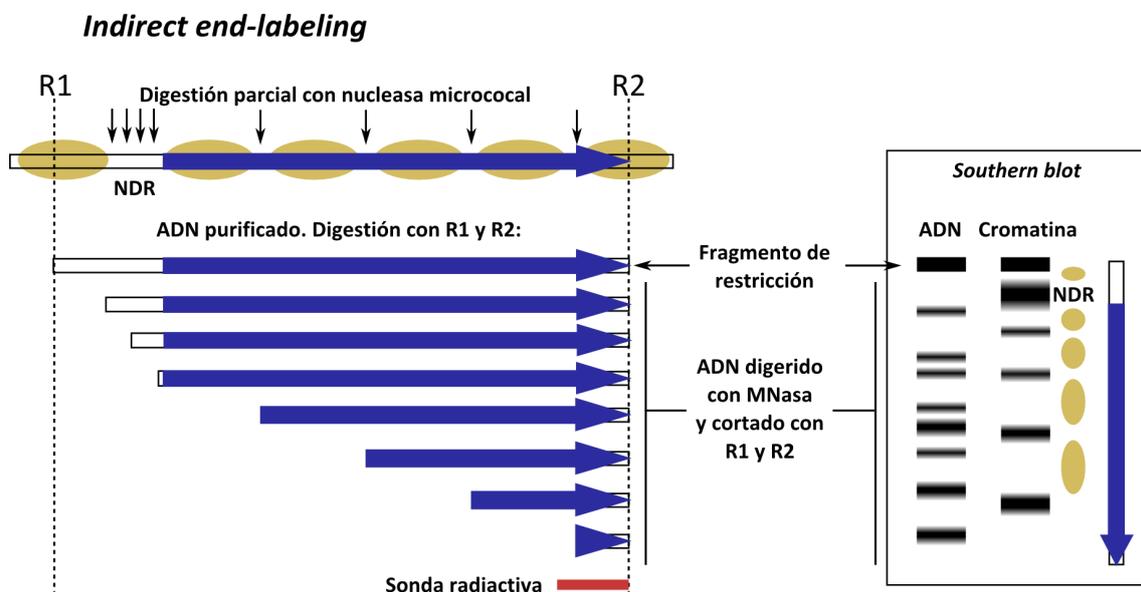
Tras la purificación del DNA, la alícuota que no había sido tratada con *MNase* anteriormente, se trató con 0.15 - 0.3 U/ml de *MNase* durante 10 min a 37°C. Esta muestra constituirá el control de DNA desnudo tratado con *MNase*, que sirve para descartar que el patrón de bandas observado en el experimento se deba a preferencias de la *MNase* por determinadas secuencias.

Para terminar, se digirieron independientemente cantidades equivalentes de DNA de las 7 muestras con la enzima de restricción adecuada, se realizó una electroforesis en gel de agarosa al 1.5% durante 210 min a 120 V y se detectaron los fragmentos de DNA nucleosomal de diferentes tamaños por *Southern blot*, hibridando con una sonda radiactiva correspondiente al extremo del fragmento de restricción (**Figura 24**).

#### Indirect end-labeling en *S. cerevisiae*

Para realizar estos experimentos en *S. cerevisiae*, introdujimos algunas modificaciones. Partimos de un cultivo de 200 ml, eliminamos el paso con la solución de preincubación y aumentamos la concentración de sorbitol del tampón Sorbitol/Tris, de modo que la composición final de esta solución fue: sorbitol 1M, Tris 50 mM y  $\beta$ -mercaptoetanol 10 mM.

El tratamiento con zimoliasa se realizó en un volumen de 20 ml de tampón Sorbitol/Tris con 10 mg de zimoliasa durante 10 minutos a 30 °C bajo agitación constante. Antes del tratamiento con nucleasa micrococcal, se resuspendió el pellet celular en 3.5 ml de tampón NP y se dividió la muestra en 7 alícuotas de 500  $\mu$ l. Al igual que en *S. pombe*, guardamos una de las alícuotas para obtener DNA genómico y tratarlo con *MNase*, como control del posible sesgo de la propia enzima. El resto de las alícuotas fueron



**Figura 24. Análisis del posicionamiento de nucleosomas por indirect end-labeling.** A la izquierda se muestra el mapa de una región a analizar, flanqueada por dos sitios de restricción (R1 y R2). La nucleasa micrococcal corta preferentemente en las regiones internucleosomales (flechas). A la derecha se muestra un resultado esquemático del *Southern blot*, donde el ADN linker estaría representado por bandas oscuras discretas y los NDRs por bandas más anchas, debido a la acumulación de fragmentos de tamaño similar. Modificado de Clark (2010).

tratadas con las siguientes concentraciones de *MNase*: 0, 0.5, 1, 1.5, 3 y 6 U/ml. Para todo lo demás, el protocolo fue idéntico al de *S. pombe*.

#### **4.5. Preparación de DNA mononucleosomal para secuenciación masiva**

##### Preparación de DNA mononucleosomal de *S. pombe*

El protocolo utilizado para preparar DNA mononucleosomal para secuenciación masiva fue similar al utilizado para el análisis por *indirect end-labeling*, aunque con algunas modificaciones. Se partió de un cultivo de *S. pombe* de 200 ml y, antes de recoger las células, estas se trataron con formaldehído al 0.5% a 25 °C durante 20 minutos con agitación suave (60 rpm) para fijar los nucleosomas al DNA. Tras esta incubación, se inactivó el formaldehído añadiendo glicina hasta una concentración de 125 mM e incubando durante 10 minutos a 25°C con agitación continua (100 rpm). Los pasos de centrifugación, lavado de las células y tratamiento con solución de preincubación se realizaron del mismo modo que en los experimentos de *indirect end-labeling*.

La cantidad de zimoliasa utilizada fue de 8 mg en un volumen de 5 ml de tampón Sorbitol/Tris. Tras el tratamiento con zimoliasa, los protoplastos se centrifugaron a 3200 g y se lavaron con Sorbitol/Tris sin  $\beta$ -mercaptoetanol. A continuación, los protoplastos se centrifugaron de nuevo a 3200 g y se resuspendieron en 3 ml de tampón NP. En el protocolo descrito por Lantermann et al. (2009) se recomienda digerir la cromatina con nucleasa micrococcal hasta alcanzar una proporción de un 80% - 90% de mononucleosoma con respecto a dinucleosoma. Esta proporción permite recuperar la mayor parte de la muestra en forma de mononucleosoma; una digestión mayor produciría fragmentos menores a 150 pb, lo que indicaría que la nucleasa micrococcal ha empezado a digerir incluso el DNA protegido por el nucleosoma. Para alcanzar esta proporción, se trataron los protoplastos con 150 U/ml de *MNase* durante 45 minutos a 37 °C.

##### Preparación de DNA mononucleosomal de *S. cerevisiae*

En el caso de *S. cerevisiae*, el tratamiento con formaldehído se realizó con una concentración final del 1% y tras la parada con glicina, se lavaron dos veces las células con PBS (NaCl 137 mM, KCl 2,7 mM, Na<sub>2</sub>HPO<sub>4</sub> 8.1 mM, KH<sub>2</sub>PO<sub>4</sub> 1 mM). Al igual que con los experimentos de *indirect end-labeling* de *S. cerevisiae*, eliminamos el paso de la solución de preincubación y aumentamos la concentración de sorbitol del tampón Sorbitol/Tris hasta 1M.

Las células se trataron con 10 mg de zimoliasa en un volumen de 20 ml de tampón Sorbitol/Tris durante 5 minutos a 30 °C con agitación suave (60 rpm). Para obtener la proporción 80:20 de mononucleosoma con respecto a dinucleosoma, los protoplastos se incubaron con 150 U/ml de *MNase* durante 20 minutos a 37 °C. Para todo lo demás, el protocolo fue idéntico al de *S. pombe*.

Tras la digestión con *MNase*, tanto en el caso de *S. pombe* como en el caso de *S. cerevisiae*, se extrajo el DNA con fenol-cloroformo, se precipitó y se trató con RNasa de la misma forma que en el protocolo de *indirect end-labeling*. A continuación, el DNA se cargó en un gel de agarosa al 1.5% en TAE y se corrió hasta que las bandas de DNA mononucleosomal quedaron bien separadas de los otros fragmentos. Por último, se cortaron esas bandas y se extrajo el DNA utilizando las columnas de BIO-RAD (*Freeze 'N Squeeze DNA Gel Extraction Spin Columns*).

Para generar las librerías necesarias para secuenciación, se utilizó el kit "*NEXTflex™ Illumina ChIP-Seq Library Prep Kit*" (Bioo Scientific), omitiendo los pasos de selección de tamaño, que no son necesarios puesto que el material de partida tiene un tamaño homogéneo y adecuado para el protocolo del kit.

#### **4.6. Generación de mapas de nucleosomas de alta resolución mediante secuenciación masiva**

Las muestras de DNA mononucleosomal se secuenciaron en el ultrasecuenciador de Illumina NextSeq 500 utilizando el protocolo de *paired-end*. La cobertura genómica de los diferentes experimentos osciló entre 186 y 582 veces. Las lecturas se alinearon utilizando la herramienta Bowtie (Langmead et al. 2009) frente al genoma de *S. pombe* (ASM294v2.20) o de *S. cerevisiae* (SacCer 3). Para la representación de los datos, se utilizó la herramienta NUCwave (Quintales et al. 2015b). Las comparaciones de posicionamiento de nucleosomas entre distintos mapas se realizaron con la herramienta DANPOS 2 (Chen et al. 2013).

#### 4.7. Inmunoprecipitación de cromatina (ChIP)

Los experimentos de inmunoprecipitación de cromatina (ChIP) se realizaron siguiendo el protocolo descrito por (Keogh et al. 2003).

Cultivos de *S. pombe* de 200 ml y  $DO_{595} = 0.5$  se fijaron con formaldehído al 1% durante 20 min a 25 °C, a continuación se añadió glicina a una concentración final de 360 mM y se incubaron 5 min más. Seguidamente se centrifugaron a 3200 g durante 10 min a 4°C y se lavaron las células dos veces con tampón TBS frío (Tris-HCl 20 mM pH 7.6 y NaCl 150 mM) y una vez con tampón de lisis (HEPES-KOH 50 mM, NaCl 150 mM, EDTA 1 mM, Tritón x-100 1%, deoxicolato sódico 0.1%, SDS 0.1%). En este punto se congeló el precipitado celular a -80°C.

Para preparar el extracto celular, se descongelaron las células y se resuspendieron en 500 µl de tampón de lisis con inhibidores de proteasas y fosfatasas (cóctel de inhibidores Roche, PMSF 10 mM, NaF 10 mM, ortovanadato sódico 20 mM). A continuación se añadió el mismo volumen de *glass beads* de 450 µm y se lisaron las células con tres pulsos de FAST-PREP a velocidad 4.5 durante 45 seg. Se eliminaron las *glass beads* por centrifugación. Para sonicar la cromatina, se ajustó el volumen de las muestras a 1.5 ml con tampón de lisis con inhibidores. A cada muestra se le aplicaron cinco ciclos de sonicación de 10 min (con pulsos alternativos de 30 seg encendido y 30 seg apagado) con el sonicador Bioruptor™ (Diagenode), hasta obtener fragmentos de DNA de un tamaño de 200 - 600 pb. A continuación, se centrifugó cada muestra a 13000 rpm durante 10 min a 4°C para eliminar los restos celulares, y se transfirió el sobrenadante a un tubo nuevo. Se añadió sarcosil a una concentración final 1% y se cuantificó la concentración de proteínas en Nanodrop (Thermo Scientific, ND-1000). Se utilizaron 20 mg de proteína por inmunoprecipitación.

Para la inmunoprecipitación de la RNA polimerasa II (RNAPII) se utilizó un anticuerpo monoclonal anti-Rpb1 (8WG16, Abcam), que reconoce la subunidad mayor de la polimerasa. Se tomó 1/10 del volumen a inmunoprecipitar como muestra control de DNA total (INPUT) sin anticuerpo. Las inmunoprecipitaciones se realizaron en primer lugar incubando los extractos durante 3 - 4 h a 4°C en agitación suave con el anticuerpo y añadiendo después 20 µl de Proteína G (Dynabeads ProteinG, Invitrogen) equilibrada con tampón de lisis y bloqueada con ssDNA de salmón (2 mg/ml). Posteriormente, las muestras se incubaron durante toda la noche a 4°C con agitación constante. A continuación se realizaron cinco lavados sucesivos: dos con 1.4 ml de tampón de lisis, uno con 1.4 ml de tampón de lisis con NaCl a 500 mM, uno con 1.4 ml de tampón LiCl

(Tris-HCl 10 mM pH 8.0, LiCl 250 mM, NP-40 0.5%, deoxicolato sódico 0.5%, EDTA 1 mM) y uno con 1.4 ml de TE. Después del último lavado, se añadieron 100  $\mu$ l de TE con 20  $\mu$ g de RNasa A. En este punto, las muestras INPUT se descongelaron y se trataron también con RNasa A. Posteriormente, las proteínas se eluyeron mediante dos incubaciones sucesivas con 250  $\mu$ l de solución de elución ( $\text{NaHCO}_3$  0.1 M, SDS 1%), durante 15 min y después se incubaron durante 6 h a 65°C. Al INPUT se le añadieron directamente 400  $\mu$ l de solución de elución y se incubó a 65°C con el resto de las muestras. Tras esta incubación, se digirieron las proteínas de todas las muestras añadiendo 25  $\mu$ l de Tris-HCl, 10  $\mu$ l de EDTA 0.5 M pH 8.0 y 20 mg de proteinasa K e incubando a 37°C durante 1 h.

Por último, para purificar el DNA de las muestras inmunoprecipitadas, se realizó una extracción con 500  $\mu$ l de fenol:cloroformo (1:1) y el DNA se precipitó durante toda la noche a -20°C con 10  $\mu$ g de glicógeno, 1/10 del volumen de acetato sódico 3 M y 2.5 volúmenes de etanol 100%. Las muestras se centrifugaron a 13000 rpm durante 30 - 40 min a 4°C y el precipitado se lavó con etanol 70%, se secó por centrifugación en vacío y se resuspendió en 100  $\mu$ l de TE. Finalmente, se purificó el DNA con columnas *GFX™ PCR DNA and Gel Band Purification Kit (GE-Healthcare)*, eluyendo en 70  $\mu$ l de TE.

Para la PCR cuantitativa (qPCR) se utilizó 1  $\mu$ l de una dilución 1/50 del INPUT y 1  $\mu$ l del resto de las muestras sin diluir.

#### **4.8. Cuantificación mediante PCR (qPCR)**

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).

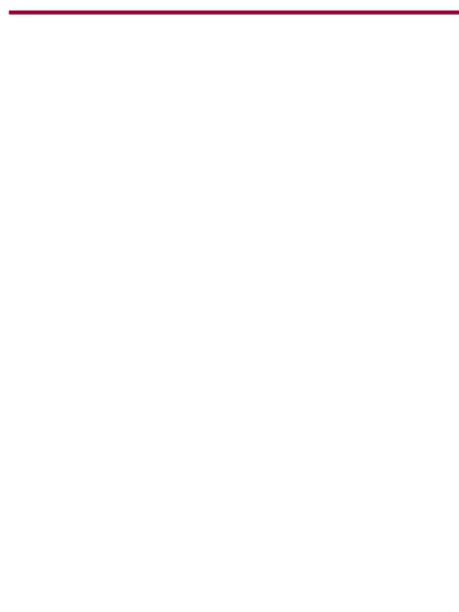
## 5. TÉCNICAS DE MANIPULACIÓN Y OBTENCIÓN DE RNA

### 5.1. Extracción de RNA total

Se recogieron 25 ml de un cultivo celular de *S. pombe* en fase exponencial y se lavaron las células con agua. Se resuspendió el precipitado de células en 20  $\mu$ l de solución de extracción (EDTA 100 mM pH8.0, NaCl 100 mM, Tris 50 mM pH8.0), se añadieron 20  $\mu$ l de una mezcla de fenol:cloroformo (1:1), 5  $\mu$ l de SDS 10% y el equivalente a 200  $\mu$ l de *glass beads* de 450  $\mu$ m de diámetro. Se rompieron las células en dos ciclos de FAST-PREP a velocidad 5.5 durante 15 seg, parando durante dos minutos para enfriar los tubos en hielo entre cada ciclo. Se añadieron 200  $\mu$ l más de solución de extracción y 200  $\mu$ l más de fenol:cloroformo y se repitió el paso de FAST-PREP. Se centrifugó a 16000 g durante 5 min a 4°C, se recogió el sobrenadante y se extrajo el RNA tratando una vez con fenol, 4 veces con fenol:cloroformo (1:1) y una vez con cloroformo:alcohol isoamílico (24:1). Por último, se precipitó el RNA con acetato sódico 0.3 M y 2.5 volúmenes de etanol 100%, dejándolo a -20°C durante 1 h y centrifugando después a 16000 g durante 45 min. Finalmente, se eliminó el sobrenadante, se lavó el precipitado con etanol 70% y se resuspendió el RNA en 30  $\mu$ l de H<sub>2</sub>O tratada con DEPC al 0.2%.

Para la síntesis de cDNA, se trató 1  $\mu$ g de RNA total con DNAsaI (1 U/ml, Invitrogen), siguiendo las recomendaciones del fabricante. Posteriormente, se utilizaron 400 ng del RNA tratado con DNAsaI para la reacción de transcripción reversa con el kit comercial *iScript Reverse Transcription Supermix for RT-qPCR* (Bio-Rad). Los niveles de cDNA se analizaron a continuación por qPCR (**Apartado 4.8**).

## BIBLIOGRAFÍA





- Allan, J., Fraser, R.M., Owen-Hughes, T., Docherty, K., and Singh, V. (2013). A comparison of in vitro nucleosome positioning mapped with chicken, frog and a variety of yeast core histones. *J. Mol. Biol.* *425*, 4206–4222.
- Amberg, D.C., Burke, D., and Strathern, J.N. (2005). *Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual* (CSHL Press).
- Anderson, J.D., and Widom, J. (2001). Poly(dA-dT) promoter elements increase the equilibrium accessibility of nucleosomal DNA target sites. *Mol. Cell. Biol.* *21*, 3830–3839.
- Annaluru, N., Muller, H., Mitchell, L.A., Ramalingam, S., Stracquadiano, G., Richardson, S.M., Dymond, J.S., Kuang, Z., Scheifele, L.Z., Cooper, E.M., et al. (2014). Total synthesis of a functional designer eukaryotic chromosome. *Science* (80-. ). *344*, 55–58.
- Annunziato, A.T. (2008). DNA packaging: Nucleosomes and chromatin. *Nature Education*, *1*(1), 26.
- Arya, G., Maitra, A., and Grigoryev, S.A. (2010). A structural perspective on the where, how, why, and what of nucleosome positioning. *J. Biomol. Struct. Dyn.* *27*, 803–820.
- 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., et al. (2008). A Library of Yeast Transcription Factor Motifs Reveals a Widespread Function for Rsc3 in Targeting Nucleosome Exclusion at Promoters. *Mol. Cell* *32*, 878–887.
- Bai, L., Charvin, G., Siggia, E.D., and Cross, F.R. (2010). Nucleosome-Depleted Regions in Cell-Cycle-Regulated Promoters Ensure Reliable Gene Expression in Every Cell Cycle. *Dev. Cell* *18*, 544–555.
- Bai, L., Ondracka, A., and Cross, F.R. (2011). Multiple Sequence-Specific Factors Generate the Nucleosome-Depleted Region on CLN2 Promoter. *Mol. Cell* *42*, 465–476.
- van Bakel, H., Tsui, K., Gebbia, M., Mnaimneh, S., Hughes, T.R., and Nislow, C. (2013). A compendium of nucleosome and transcript profiles reveals determinants of chromatin architecture and transcription. *PLoS Genet.* *9*, e1003479.
- Berbenetz, N.M., Nislow, C., and Brown, G.W. (2010). Diversity of Eukaryotic DNA Replication Origins Revealed by Genome-Wide Analysis of Chromatin Structure. *PLoS Genet.* *6*, e1001092.

- Bernardi, F., Zatchej, M., and Thoma, F. (1992). Species specific protein-DNA interactions may determine the chromatin units of genes in *S.cerevisiae* and in *S.pombe*. *EMBO J.* *11*, 1177–1185
- Beshnova, D.A., Cherstvy, A.G., Vainshtein, Y., and Teif, V.B. (2014). Regulation of the Nucleosome Repeat Length In Vivo by the DNA Sequence, Protein Concentrations and Long-Range Interactions. *PLoS Comput. Biol.* *10*, e1003698.
- Birnboim, H.C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* *7*, 1513–1523.
- Boeger, H., Griesenbeck, J., and Kornberg, R.D. (2008). Nucleosome Retention and the Stochastic Nature of Promoter Chromatin Remodeling for Transcription. *Cell* *133*, 716–726.
- Brogaard, K., Xi, L., Wang, J.-P., and Widom, J. (2012). A map of nucleosome positions in yeast at base-pair resolution. *Nature* *486*, 496–501.
- Cairns, B.R. (2009). The logic of chromatin architecture and remodelling at promoters. *Nature* *461*, 193–198.
- Cannarozzi, G., Schraudolph, N.N., Faty, M., von Rohr, P., Friberg, M.T., Roth, A.C., Gonnet, P., Gonnet, G., and Barral, Y. (2010). A Role for Codon Order in Translation Dynamics. *Cell* *141*, 355–367.
- de Castro, E., Soriano, I., Marín, L., Serrano, R., Quintales, L., and Antequera, F. (2012). Nucleosomal organization of replication origins and meiotic recombination hotspots in fission yeast. *EMBO J.* *31*, 124–137.
- Chang, G.S., Noegel, A.A., Mavrich, T.N., Müller, R., Tomsho, L., Ward, E., Felder, M., Jiang, C., Eichinger, L., Glöckner, G., et al. (2012). Unusual combinatorial involvement of poly-A/T tracts in organizing genes and chromatin in *Dictyostelium*. *Genome Res.* *22*, 1098–1106.
- Chen, K., Xi, Y., Pan, X., Li, Z., Kaestner, K., Tyler, J., Dent, S., He, X., and Li, W. (2013). DANPOS: Dynamic analysis of nucleosome position and occupancy by sequencing. *Genome Res.* *23*, 341–351.
- Clark, D.J. (2010). Nucleosome positioning, nucleosome spacing and the nucleosome code. *J. Biomol. Struct. Dyn.* *27*, 781–793.

- Cole, H.A., Howard, B.H., and Clark, D.J. (2011). Activation-induced disruption of nucleosome position clusters on the coding regions of Gcn4-dependent genes extends into neighbouring genes. *Nucleic Acids Res.* *39*, 9521–9535.
- Curran, K.A., Crook, N.C., Karim, A.S., Gupta, A., Wagman, A.M., and Alper, H.S. (2014). Design of synthetic yeast promoters via tuning of nucleosome architecture. *Nat. Commun.* *5*, 4002.
- Drew, H.R., and Travers, A.A. (1985). DNA bending and its relation to nucleosome positioning. *J. Mol. Biol.* *186*, 773–790.
- Duan, M.R., and Smerdon, M.J. (2010). UV damage in DNA promotes nucleosome unwrapping. *J. Biol. Chem.* *285*, 26295–26303.
- Eaton, M.L., Galani, K., Kang, S., Bell, S.P., and MacAlpine, D.M. (2010). Conserved nucleosome positioning defines replication origins. *Genes Dev.* *24*, 748–753.
- Forsburg, S.L. (1994). Codon usage table for *Schizosaccharomyces pombe*. *Yeast* *10*, 1045–1047.
- Forsburg, S.L., and Rhind, N. (2006). Basic methods for fission yeast. *Yeast* *23*, 173–183.
- Gaykalova, D.A., Nagarajavel, V., Bondarenko, V.A., Bartholomew, B., Clark, D.J., and Studitsky, V.M. (2011). A polar barrier to transcription can be circumvented by remodeler-induced nucleosome translocation. *Nucleic Acids Res.* *39*, 3520–3528.
- Givens, R.M., Lai, W.K.M., Rizzo, J.M., Bard, J.E., Mieczkowski, P.A., Leatherwood, J., Huberman, J.A., and Buck, M.J. (2012). Chromatin architectures at fission yeast transcriptional promoters and replication origins. *Nucleic Acids Res.* *40*, 7176–7189.
- Gkikopoulos, T., Schofield, P., Singh, V., Pinskaya, M., Mellor, J., Smolle, M., Workman, J.L., Barton, G.J., and Owen-Hughes, T. (2011). A Role for Snf2-Related Nucleosome-Spacing Enzymes in Genome-Wide Nucleosome Organization. *Science* (80-. ). *333*, 1758–1760.
- Gossett, A.J., and Lieb, J.D. (2012). In vivo effects of histone H3 depletion on nucleosome occupancy and position in *Saccharomyces cerevisiae*. *PLoS Genet.* *8*, e1002771.
- Haimovich, A.D., Muir, P., and Isaacs, F.J. (2015). Genomes by design. *Nat. Rev. Genet.* *16*, 501–516.
- Hartley, P.D., and Madhani, H.D. (2009). Mechanisms that Specify Promoter Nucleosome Location and Identity. *Cell* *137*, 445–458.

- Hennig, B.P., Bendrin, K., Zhou, Y., and Fischer, T. (2012). Chd1 chromatin remodelers maintain nucleosome organization and repress cryptic transcription. *EMBO Rep.* *13*, 997–1003.
- Huebert, D.J., Kuan, P.-F., Keleş, S., and Gasch, A.P. (2012). Dynamic changes in nucleosome occupancy are not predictive of gene expression dynamics but are linked to transcription and chromatin regulators. *Mol. Cell. Biol.* *32*, 1645–1653.
- Hughes, A.L., and Rando, O.J. (2014). Mechanisms Underlying Nucleosome Positioning In Vivo. *Annu. Rev. Biophys.* *43*, 41–63.
- Hughes, T.R., and de Boer, C.G. (2013). Mapping yeast transcriptional networks. *Genetics* *195*, 9–36.
- Hughes, A.L., Jin, Y., Rando, O.J., and Struhl, K. (2012). A functional evolutionary approach to identify determinants of nucleosome positioning: a unifying model for establishing the genome-wide pattern. *Mol. Cell* *48*, 5–15.
- Ikemura, T. (1985). Codon usage and tRNA content in unicellular and multicellular organisms. *Mol. Biol. Evol.* *2*, 13–34.
- Iyer, V., and Struhl, K. (1995). Poly(dA:dT), a ubiquitous promoter element that stimulates transcription via its intrinsic DNA structure. *EMBO J.* *14*, 2570–2579.
- 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., et al. (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* *458*, 362–366.
- Kaplan, N., Hughes, T.R., Lieb, J.D., Widom, J., and Segal, E. (2010). Contribution of histone sequence preferences to nucleosome organization: proposed definitions and methodology. *Genome Biol.* *11*, 140.
- Keogh, M.-C., Podolny, V., and Buratowski, S. (2003). Bur1 Kinase Is Required for Efficient Transcription Elongation by RNA Polymerase II. *Mol. Cell. Biol.* *23*, 7005–7018.
- Koerber, R.T., Rhee, H.S., Jiang, C., and Pugh, B.F. (2009). Interaction of Transcriptional Regulators with Specific Nucleosomes across the *Saccharomyces* Genome. *Mol. Cell* *35*, 889–902.
- Kornberg, R.D., and Stryer, L. (1988). Statistical distributions of nucleosomes: Nonrandom locations by a stochastic mechanism. *Nucleic Acids Res.* *16*, 6677–6690.

- Kushner, S. (1978). An improved method for transformation *Escherichia coli* with ColE1 derived plasmids. In Genetic engineering, H.W.B.a.S. Nicosia, ed. (Amsterdam: Elsevier/North-Holland Biomedical Press), pp. 17-23.
- Lantermann, A., Strålfors, A., Fagerström-Billai, F., Korber, P., and Ekwall, K. (2009). Genome-wide mapping of nucleosome positions in *Schizosaccharomyces pombe*. *Methods* 48, 218–225.
- Lantermann, A.B., Straub, T., Strålfors, A., Yuan, G.-C., Ekwall, K., and Korber, P. (2010). *Schizosaccharomyces pombe* genome-wide nucleosome mapping reveals positioning mechanisms distinct from those of *Saccharomyces cerevisiae*. *Nat. Struct. Mol. Biol.* 17, 251–257.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25.
- Lee, W., Tillo, D., Bray, N., Morse, R.H., Davis, R.W., Hughes, T.R., and Nislow, C. (2007). A high-resolution atlas of nucleosome occupancy in yeast. *Nat. Genet.* Publ. Online 16 Sept. 2007; | doi10.1038/ng2117 39, 1235.
- Lieleg, C., Krietenstein, N., Walker, M., and Korber, P. (2015). Nucleosome positioning in yeasts: methods, maps, and mechanisms. *Chromosoma* 124, 131–151.
- Liu, X., Lee, C.-K., Granek, J.A., Clarke, N.D., and Lieb, J.D. (2006). Whole-genome comparison of Leu3 binding in vitro and in vivo reveals the importance of nucleosome occupancy in target site selection. *Genome Res.* 16, 1517–1528.
- Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* 276, 19–42.
- Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251–260.
- 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., et al. (2008a). Nucleosome organization in the *Drosophila* genome. *Nature* 453, 358–362.
- Mavrich, T.N., Ioshikhes, I.P., Venters, B.J., Jiang, C., Tomsho, L.P., Qi, J., Schuster, S.C., Albert, I., and Pugh, B.F. (2008b). A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res.* 18, 1073–1083.

- McGhee, J.D., and Felsenfeld, G. (1980). Nucleosome Structure. *Annu. Rev. Biochem.* *49*, 1115–1156.
- McGinty, R.K., and Tan, S. (2015). Nucleosome structure and function. *Chem. Rev.* *115*, 2255–2273.
- McManus, J., Perry, P., Sumner, A.T., Wright, D.M., Thomson, E.J., Allshire, R.C., Hastie, N.D., and Bickmore, W.A. (1994). Unusual chromosome structure of fission yeast DNA in mouse cells. *J. Cell Sci.* *107 (Pt 3)*, 469–486.
- Möbius, W., and Gerland, U. (2010). Quantitative test of the barrier nucleosome model for statistical positioning of nucleosomes up- and downstream of transcription start sites. *PLoS Comput. Biol.* *6*, e1000891.
- Moreno, S., Klar, A., and Nurse, P. (1991). Molecular genetic analysis of fission yeast *Schizosaccharomyces pombe*. *Methods Enzymol.* *194*, 795–823.
- Moyle-Heyrman, G., Zaichuk, T., Xi, L., Zhang, Q., Uhlenbeck, O.C., Holmgren, R., Widom, J., and Wang, J.-P. (2013). Chemical map of *Schizosaccharomyces pombe* reveals species-specific features in nucleosome positioning. *Proc. Natl. Acad. Sci. U. S. A.* *110*, 20158–20163.
- Olson, W.K., and Zhurkin, V.B. (2011). Working the kinks out of nucleosomal DNA. *Curr. Opin. Struct. Biol.* *21*, 348–357.
- Perales, R., Zhang, L., and Bentley, D. (2011). Histone occupancy in vivo at the 601 nucleosome binding element is determined by transcriptional history. *Mol. Cell. Biol.* *31*, 3485–3496.
- Plotkin, J.B., and Kudla, G. (2011). Synonymous but not the same: the causes and consequences of codon bias. *Nat. Rev. Genet.* *12*, 32–42.
- Pointner, J., Persson, J., Prasad, P., Norman-Axelsson, U., Strålfors, A., Khorosjutina, O., Krietenstein, N., Svensson, J.P., Ekwall, K., and Korber, P. (2012). CHD1 remodelers regulate nucleosome spacing in vitro and align nucleosomal arrays over gene coding regions in *S. pombe*. *EMBO J.* *31*, 4388–4403.
- Prendergast, J.G.D., and Semple, C.A.M. (2011). Widespread signatures of recent selection linked to nucleosome positioning in the human lineage. *Genome Res.* *21*, 1777–1787.

- Quintales, L., Soriano, I., Vázquez, E., Segurado, M., and Antequera, F. (2015a). A species-specific nucleosomal signature defines a periodic distribution of amino acids in proteins. *Open Biol.* *5*, 140218.
- Quintales, L., Vázquez, E., and Antequera, F. (2015b). Comparative analysis of methods for genome-wide nucleosome cartography. *Brief. Bioinform.* *16*, 576–587.
- Raisner, R.M., Hartley, P.D., Meneghini, M.D., Bao, M.Z., Liu, C.L., Schreiber, S.L., Rando, O.J., and Madhani, H.D. (2005). Histone variant H2A.Z Marks the 5' ends of both active and inactive genes in euchromatin. *Cell* *123*, 233–248.
- Raveh-sadka, T., Levo, M., Shabi, U., Shany, B., Keren, L., Lotan-Pompan, M., Zeevi, D., Sharon, E., Weinberger, A., and Segal, E. (2012). Manipulating nucleosome disfavoring sequences allows fine-tune regulation of gene expression in yeast. *Nat. Genet.* *44*, 743–750.
- Rhind, N., Chen, Z., Yassour, M., Thompson, D.A., Haas, B.J., Habib, N., Wapinski, I., Roy, S., Lin, M.F., Heiman, D.I., et al. (2011). Comparative functional genomics of the fission yeasts. *Science* *332*, 930–936.
- Sasaki, S., Mello, C.C., Shimada, A., Nakatani, Y., Hashimoto, S., Ogawa, M., Matsushima, K., Gu, S.G., Kasahara, M., Ahsan, B., et al. (2009). Chromatin-associated periodicity in genetic variation downstream of transcriptional start sites. *Science* *323*, 401–404.
- Satchwell, S.C., Drew, H.R., and Travers, A.A. (1986). Sequence periodicities in chicken nucleosome core DNA. *J. Mol. Biol.* *191*, 659–675.
- Schones, D.E., Cui, K., Cuddapah, S., Roh, T.-Y., Barski, A., Wang, Z., Wei, G., and Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. *Cell* *132*, 887–898.
- Segal, E., and Widom, J. (2009). What controls nucleosome positions? *Trends Genet.* *25*, 335–343.
- Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thåström, A., Field, Y., Moore, I.K., Wang, J.-P.Z., and Widom, J. (2006). A genomic code for nucleosome positioning. *Nature* *442*, 772–778.
- Sekinger, E.A., Moqtaderi, Z., and Struhl, K. (2005). Intrinsic histone-DNA interactions and low nucleosome density are important for preferential accessibility of promoter regions in yeast. *Mol. Cell* *18*, 735–748.

- Shim, Y.S., Choi, Y., Kang, K., Cho, K., Oh, S., Lee, J., Grewal, S.I.S., and Lee, D. (2012). Hrp3 controls nucleosome positioning to suppress non-coding transcription in eu- and heterochromatin. *EMBO J.* *31*, 4375–4387.
- Shivaswamy, S., Bhinge, A., Zhao, Y., Jones, S., Hirst, M., and Iyer, V.R. (2008). Dynamic remodeling of individual nucleosomes across a eukaryotic genome in response to transcriptional perturbation. *PLoS Biol.* *6*, 0618–0630.
- Simpson, R.T., and Stafford, D.W. (1983). Structural features of a phased nucleosome core particle. *Proc. Natl. Acad. Sci. U. S. A.* *80*, 51–55.
- Small, E., Xi, L., and Wang, J. (2014). Single-cell nucleosome mapping reveals the molecular basis of gene expression heterogeneity. *Proc. Natl. Acad. Sci. U. S. A.* *111*, E2462–71.
- Soriano, I., Quintales, L., and Antequera, F. (2013). Clustered regulatory elements at nucleosome-depleted regions punctuate a constant nucleosomal landscape in *Schizosaccharomyces pombe*. *BMC Genomics* *14*, 813.
- Soriano, I., Morafraila, E.C., Vázquez, E., Antequera, F., and Segurado, M. (2014). Different nucleosomal architectures at early and late replicating origins in *Saccharomyces cerevisiae*. *BMC Genomics* *15*, 791.
- Struhl, K., and Segal, E. (2013). Determinants of nucleosome positioning. *Nat. Struct. Mol. Biol.* *20*, 267–273.
- Suto, R.K., Clarkson, M.J., Tremethick, D.J., and Luger, K. (2000). Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat. Struct. Biol.* *7*, 1121–1124.
- Teif, V.B. (2016). Nucleosome positioning: resources and tools online. *Brief. Bioinform.* *17*, 745–757.
- Tilgner, H., and Guigó, R. (2010). From chromatin to splicing: RNA-processing as a total artwork. *Epigenetics* *5*, 180–184.
- Tirosh, I., Sigal, N., and Barkai, N. (2010). Divergence of nucleosome positioning between two closely related yeast species: genetic basis and functional consequences. *Mol. Syst. Biol.* *6*, 365.
- Tsankov, A., Yanagisawa, Y., Rhind, N., Regev, A., and Rando, O.J. (2011). Evolutionary divergence of intrinsic and trans-regulated nucleosome positioning sequences reveals plastic rules for chromatin organization. *Genome Res.* *21*, 1851–1862.

- Tsankov, A.M., Thompson, D.A., Socha, A., Regev, A., and Rando, O.J. (2010). The role of nucleosome positioning in the evolution of gene regulation. *PLoS Biol.* 8, e1000414.
- Tuller, T., Carmi, A., Vestsigian, K., Navon, S., Dorfan, Y., Zaborske, J., Pan, T., Dahan, O., Furman, I., and Pilpel, Y. (2010). An Evolutionarily Conserved Mechanism for Controlling the Efficiency of Protein Translation. *Cell* 141, 344–354.
- Valouev, A., Ichikawa, J., Tonthat, T., Stuart, J., Ranade, S., Peckham, H., Zeng, K., Malek, J.A., Costa, G., McKernan, K., et al. (2008). A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res.* 18, 1051–1063.
- Valouev, A., Johnson, S.M., Boyd, S.D., Smith, C.L., Fire, A.Z., and Sidow, A. (2011). Determinants of nucleosome organization in primary human cells. *Nature* 474, 516–520.
- Warnecke, T., Batada, N.N., and Hurst, L.D. (2008). The impact of the nucleosome code on protein-coding sequence evolution in yeast. *PLoS Genet.* 4, e1000250.
- Washietl, S., Machné, R., and Goldman, N. (2008). Evolutionary footprints of nucleosome positions in yeast. *Trends Genet.* 24, 583–587.
- Weingarten-Gabbay, S., and Segal, E. (2014). The grammar of transcriptional regulation. *Hum. Genet.* 133, 701–711.
- Workman, J.L., and Kingston, R.E. (1992). Nucleosome core displacement in vitro via a metastable transcription factor-nucleosome complex. *Science* 258, 1780–1784.
- Wunderlich, Z., and Mirny, L.A. (2009). Different gene regulation strategies revealed by analysis of binding motifs. *Trends Genet* 25, 434–440.
- Yadon, A.N., Singh, B.N., Hampsey, M., and Tsukiyama, T. (2013). DNA Looping Facilitates Targeting of a Chromatin Remodeling Enzyme. *Mol. Cell* 50, 93–103.
- Yarragudi, A., Miyake, T., Li, R., and Morse, R.H. (2004). Comparison of ABF1 and RAP1 in chromatin opening and transactivator potentiation in the budding yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 24, 9152–9164.
- Yen, K., Vinayachandran, V., Batta, K., Koerber, R.T., and Pugh, B.F. (2012). Genome-wide nucleosome specificity and directionality of chromatin remodelers. *Cell* 149, 1461–1473.
- Yu, L., and Morse, R.H. (1999). Chromatin opening and transactivator potentiation by RAP1 in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 19, 5279–5288.

- Yuan, G.-C., Liu, Y.-J., Dion, M.F., Slack, M.D., Wu, L.F., Altschuler, S.J., and Rando, O.J. (2005). Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* *309*, 626–630.
- Zawadzki, K.A., Morozov, A. V, and Broach, J.R. (2009). Chromatin-dependent transcription factor accessibility rather than nucleosome remodeling predominates during global transcriptional restructuring in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* *20*, 3503–3513.
- Zentner, G.E., and Henikoff, S. (2013). Regulation of nucleosome dynamics by histone modifications. *Nat. Struct. Mol. Biol.* *20*, 259–266.
- Zhang, H., and Reese, J.C. (2007). Exposing the core promoter is sufficient to activate transcription and alter coactivator requirement at RNR3. *Proc. Natl. Acad. Sci. U. S. A.* *104*, 8833–8838.
- Zhang, L., Ma, H., and Pugh, B.F. (2011a). Stable and dynamic nucleosome states during a meiotic developmental process. *Genome Res.* *21*, 875–884.
- Zhang, Y., Moqtaderi, Z., Rattner, B.P., Euskirchen, G., Snyder, M., Kadonaga, J.T., Liu, X.S., and Struhl, K. (2009). Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. *Nat. Struct. Mol. Biol.* *16*, 1–7.
- Zhang, Z., Wippo, C.J., Wal, M., Ward, E., Korber, P., and Pugh, B.F. (2011b). A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. *Science* *332*, 977–980.

## **ANEXO**





## Construcciones de la Figura 8

### *ura4 1\_6*

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

TTAGTGTGGCTCAAGCGCTGTTGGTACGCTTCCAGCCGGCCTCGCGGTAGCGTTTAGCCTCT  
ACGACCGGGTTCCTGCCGGCGCCGTATACGCCCTGCCGACGATAATAATGTCTGAGCCACAAT  
TGACGATTACCTCCTCCGGCGTCTGTACTGTTGGCCTAATCCATCGCCCTTGACGTCTAGTCC  
TATGCCCGGACTCATCGTGATATAATCTGACTGCAAATTCGGGAACCTCCTGCCCGGATGAAT  
CCGAAACAGAAGTCAGTGTGTTTTTCGAACCACTCGAGTGTCTTTCCGTATAACTGCCCGTCC  
CTAGGCTTCCCTTACTGCTCATCTCTGCTAGCAATAGTAAGCCCCTGCCTAGCGGCAGGCCGAC  
CTCCTTTAATCCCTGGATGATGCCTTACCCGGTACGGTGTGACAGTTCGTTATGTGCGCCAG  
CTCGCGATCTTATATACGCCGCTCGCGTACTGCAATTTAACAGTGTGCGGATATCCGCAAAT  
TTCTGTCTTCGAATATCAAGAACCTGTGTTTTCTTGCCGAGTGCTACTAACTTCTCGACCATGTC  
TTGATCAAAGTCTTCTACGACATCGATGTGCGTTTTAATGACGCATACGTATGGCCCGATCTT  
GTGCTAGCTCGAGTATCTCACTCTTTTTAGTTAGGTCTACAGCTACGCTTAGATTTGACTG  
TTTCTCCTCCATTAGCGCTAGTAGCTCTTTAGCGATCGGGTTCTTCATACCTTCCGCGCGCGC  
CTGTATGACTGGAACACGCGCGCTCCAT

### *ura4 1\_3*

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

TTAATGCTGAGAAAGTCTTTGCTGATATGCCTTCCAACCAGCTTCTCTATATCTCTTGGCTTCG  
ACAACAGGATTACGACCAGCTCCATAGACTCCACGACCAACAATGATGATATCGCTACCGCAG  
TTTACAATCACTTCTTCAGGAGTACGATATTGCTGTCCCAGCCCGTCTCCTTTAACATCCAAGC  
CGATAACCAGGGGACATAGTTATGTAGTCGCTTTGAAGGTTAGGAAATCGACGACCAGCTATAA  
AGCCAAAGCAAAAATCGGTATGCTTCTCAAACCATTCTAAGGTTTTCTCTGTGTAGGAACCAG  
TAGCCAAAGAGCCTTTGGAAGACATTTACGCCAAAGCAAGAGACCACGTCCCAAAGGTAAC  
CAACTTCTTTGAGGCCTTGTATGATGCCTTACCCGGTACGGTGTGACAGTTCGTTATGTGCGC  
CCAGCTCGCGATCTTATATACGCCGCTCGCGTACTGCAATTTAACAGTGTGCGGATATCCGCA  
AACTTTCTGTCTTCGAATATCAAGAACCTGTGTTTTCTTGCCGAGTGCTACTAACTTCTCGACCA  
TGTCTTGATCAAAGTCTTCTACGACATCGATGTGCGTTTTAATGACGCATACGTATGGCCGA  
TCTTGTCGACTAGCTCGAGTATCTCACTCTTTTTAGTTAGGTCTACAGCTACGCTTAGATTTGA  
CTGTTTTCTCCTCCATTAGCGCTAGTAGCTCTTTAGCGATCGGGTTCTTCATACCTTCCGCGCGC  
GCGCTGTATGACTGGAACACGCGCGCTCCAT

### *ura4 3\_4*

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

TTAATGCTGAGAAAGTCTTTGCTGATATGCCTTCCAACCAGCTTCTCTATATCTCTTGGCTTCG  
ACAACAGGATTACGACCAGCTCCATAGACTCCACGACCAACAATGATGATATCGCTACCGCAG  
TTTACAATCACTTCTTCAGGAGTACGATATTGCTGTCCCAGCCCGTCTCCTTTAACATCCAAGC  
CGATAACCAGGGGACATAGTTATGTAGTCGCTTTGAAGGTTAGGAAATCGACGACCAGCGATGA  
ATCCGAAACAGAAGTCAGTGTGTTTTTCGAACCACTCGAGTGTCTTTCCGTATAACTGCCCG  
TCGCTAGGCTTCCCTTACTGCTCATCTCTGCTAGCAATAGTAAGCCCCTGCCTAGCGGCAGGCC  
GACCTCCTTTAATCCCTGGATGATGCCTTACCCGGTACGGTGTGACAGTTCGTTATGTGCGCC  
CAGCTCGCGATCTTATATACGCCGCTCGCGTACTGCAATTTAACAGTGTGCGGATATCCGCAA  
ACTTTCTGTCTTCGAATATCAAGAACCTGTGTTTTTACCTAAGGCCACCAGTTTTTCTACCAT  
ATCCTGGTCGAAATCCTCGACAACGTCAATATGTGTCTTGATAACACAGACATAGGGTCCAAT  
TTTATCTACCAATTCTAAGATTTCCGATTTCTTCGTCAAATCGACCGCGACTGACAAGTTGCT  
TTGCTTTTCTTCCATCAAAGCCAACAATTCCTTGGAATGGGATTTTTTCATCCCCTCAGCTCTA  
GCTGAATAGCTTTGAAATACTCTAGCATCCAT

## Construcciones de la Figura 10

### *ura4 1\_6 random*

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

TCACATCCTTGCGCCTGATGTCACCTTAATTCTAACATACATTAGAAATCCCGGTTACGAATCT  
 AATCGCGATAGCATTGATCCATCGCGATACTATTCAATCGAGGGAATGATCGTCTATTACAAA  
 CTATACGTCCTTAAAATGCCTTAATAGGAAAACTGTTTCCGCTACCCGTATGACCTCCTGGTC  
 GATCCGTTAATATAATAGAAGGACACGTGGAACACTAAAGGATGGAGTGAGGGAAGCGAATC  
 CCTTTAATATCAGCATCCGACATATCAACATACAAGGGTTCCGAAACTATATCCTCTAAGCGC  
 TCAAGATACTCGTACACTACGTCTTTTAGTCTAGCCGGACCGCGCCCTGAACCACTAGGTTCC  
 TCAATTTTTATTATCATATCAGCAATGTGAGACTCGCCCGGTATCATTAGTCTGAAAGTCAA  
 CGCTACCTACATGCGCTCTCATCGTGCTTAGTACCCCGTTGCCTATATCCCAACGCCTTTG  
 GGTCTTGGCAAATATTTTATGTAGAATAACATATAACAGTGGATTGAGCTCTTAGATTAAGA  
 ATTCTAGCCTATTTGGCGTCAAGACTAAGGATACTAAGGCAATTACTTAATTCGATGCGTTGG  
 CGCAAGCCTGGTGTACCTAAATATACTATTGGGACAGTTCCGCTATCCTAGCCCTAATGCCAT  
 CTCCTTTCCCAATCATGATCTCCGATAGTTTCGAACCTAATGATTGAGACGGTCAACAGCCTTTC  
 ATGACTAACTGTATCTCCGATTATTAGCGGAACAT

### *ura4 3\_4 random*

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

TTAATGCTGAGAAAGTCTTTGCTGATATGCCTTCCAACCAGCTTCTCTATATCTCTTGGCTTCG  
 ACAACAGGATTACGACCAGCTCCATAGACTCCACGACCAACAATGATGATATCGCTACCGCAG  
 TTTACAATCACTTCTTCAGGAGTACGATATTGCTGTCCCAGCCGTCTCCTTTAACATCCAAGC  
 CGATAACCAGGGGACATAGTTATGTAGTCGCTTGAAGGTTAGGAAATCGACGACCAGCGCACG  
 TTGCCGTGGAAGTGTACGGTCTAACGTACTTTCCCTCCAACCTTACTGACATTATTCGCATGGGG  
 ATCTATCTAGAGCAAAAATACGGTATTCTTTAATGAGTCAAAAAAGAGGCTTGTACGGCA  
 ATCAAGCCACAGTGTGACCACCCGATCGAAAGGCATAAGTAGTGGAATTCAGTTCTAAATTC  
 TATGTAAGCCAGTGAAGATCAAATTTTACAAGGACGATACATATATAAACCAGTCGTTAGGCA  
 AGTTATGCTTACCATACTTAACGAGCCCTCACTAGTTTTACCTAAGGCCACCAGTTTTTCTACC  
 ATATCCTGGTCGAAATCCTCGACAACGTCAATATGTGTCTTGATAACACAGACATAGGGTCCA  
 ATTTTATCTACCAATTCTAAGATTTCCGATTTCTTCGTCAAATCGACCGGACTGACAAGTTG  
 CTTTGCTTTTCTTCCATCAAAGCCAACAATTCCTTGGCAATGGGATTTTTTCATCCCCTCAGCTC  
 TAGCTGAATAGCTTTGAAATACTCTAGCATCCAT

## Construcciones de la Figura 12

### *ura4 Shuffled*

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

TTAATGTTGAGACAAACGTTGTTGATAAGCTTTCCAACCAGCTTCTCTATAACGTTTAGCTTC  
 GACGACAGGATTTCTACCAGCACCATAGACACCTTACCGACAATGATAATATCAGAACC  
 ATTTACGATAACTTCTCAGGTGTACGATATTGTTGACCCAAACCATCACCTTTACATCCAAA  
 CCAATGCCAGGAGACATTGTTATATAATCAGACTGCAAATTGGGAAAGCGTCTTCCAGCAATA  
 AAGCCAAAGCAAAAATCTGTATGCTTCTCAAACCATCCAATGTTTTTTTCAGTATAGGAACCA  
 GTAGCCAAGGATCCTTTGCTGGACATTTAGCCAACAACAACAACCCTCTTCCCTAAAGGTAAT  
 CCTACTTCTTAAGTCCCTGGATTATCCCTTCTCCAGGGACGGTATGGCAGTTGGTAATATGAG  
 CCCAGCTAGCAATTTTATAAACGCCGCTTGCCTACTGTAATTTACAGTGTTCGGATATCAGC  
 AAATTTACGATCCTCAAAGATAAGGAAACGATGTTTCTTACCCAATGCAACAAGTTTTTCTAC  
 CATATCTTGATCAAAGTCTTCCACGACGTCTATATGGGTCTTAATAACACAAACGTAGGGGCC  
 AATTTTGTCAACAAGCTCTAAGATTTCTGACTTCTTCGTCAGGTCCACAGCCACTGAGAGATT  
 GCTTTGCTTTTCTCCATTAGGGCCAGGAGCTCTTTGGCGATTGGGTTCTTCATGCCTTCTGCT  
 CGGGCAGAGTAGGACTGGAAGACACGCGCTCCAT

## Construcciones de la Figura 14

### 2IGR

Integrado en *S. pombe* (Chr1:2753584..2753879)

GGTACTAATATCAACATTTAAATGTGATTTGGTCTTTATCTTATGTAACCTTAGAAATAAT  
 TTTTATACGACAAATAACTTGTGGCCTTGTAACATCGGTATAACAATTCATTAATATTA  
 ATATTAATTATACAGAGTACAGTTCAAATCTAAAGTAGTTCGCAACTCTACTCTCGCATCCTC  
 TCGATAAACAAACATAGAAGTAATACAAAAGCCAGAAATTGCTCTAAAAATGAATTTTGTACC  
 GTAGGTATCCTTTGTGTTACCGCATATAAATACTGATACGAAG

### 3IGR

Integrado en *S. pombe* (Chr1:5024357..5024797)

TTCATTCTATCGTAATTATCAGAGACTCTATAATCGTTTATAATCAAGAACTAATACTACAAA  
 TACAGAAGAAATATCAATTATTTGAAGTATATATTTAATATCTTTCTAGCCCACTCCATC  
 AAAAAAGTTAGCTAAATTTACTAAGTATTAACCTAAATACACATCTTCAAATATTTAACTTTA  
 TGGAAAAATATACACAAATATAAAAACAGATAATTTTAAAAAAAGTAGAAATTTTAATAAAT  
 TATACTGACCTCAATTTATTTGACCTAAGTAAAGACTCCGAACTATAATATACACATAAATTA  
 CCAACTCATTATCTTTTAAAGTACAAATTCACGTTTCCCAATAATAAACTGTCAACAAAA  
 TATGGCACTTGATGTAAGTAAACCTTGAAAACCTGAGTATAAAATATCCATCTAGATTAGTTAA

## Construcciones de la Figura 15

### **Puzle (1)**

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

TTAATCCCATTTCAAACATTTTGAAGAAAATCAAGGAAAGAAGCCTGCTTTGTGTTTAGAAG  
 CTGACTAAATGTTCTAGTAGATGGAACATCACTGGGAAGAAAAAAGGAATGACAGGAATTT  
 GAGAATACGCATCAATTGCAGCTGTTTTAAAGTTGCAGCACTATTTTCCGACGGAGTTTGAG  
 AATCCTTCGTTGCTTTTGGAAATTGAAAAAGACCTTTGGTGTGACTTATTGGTATTTTTTAGTA  
 GAACAAGAGCATTCTGTCTCCATTTCGCACAAAAAACTTTTCAAGTTGATGTTAAGAGCAATGT  
 GGACATCTTGCTCCTTCCTAGTGAGTACATGAAAAAATCTCGTAGTACATCTAAGTTCGAAA  
 AATCGCCTTGAGAAAAATAGTTTTTCAATACTAGGAGATCGCTATCCAGAATGCAGACCTTAT  
 CAAACTGTGTTTGTCAAATACACTGAGTTTTGTGAACATGTATTGATAACGATAACTGCCCA  
 TGCCAATAGACAAACCCTTTCAATTAAGTCTTCAGTTTTAATTTGGTCCAATGTATCTTCAA  
 CAGCAATTTGTAATTTAGTGAGCCAAAATTTCTCTCCTGCTCTTCATTGGTTTTCCGAGTCCTT  
 AGACAAAGCCTTTAACTCCTTCTCAAGGTTTTGACGCATTCTGTATCTGGCGATTTTTAGAGT  
 CCTTGATCTATAGTATCAAATTTTTGTGTCTTAAGCTTTTTATGCAGATGCAGTTTCGTTGC  
 ATCAATTTTCATGCATAACCTTTCCCCCATTTATTTCCAT

### **C-**

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

ACGTTGCGGGGTTGTCGAAAAGAACAAGCAATACATTACAGAACTAATAATATATGAGTATAT  
 CATTTATCATCGCATTAACCTTTGGTGAGGAACAAAACCTGTTCAATGAGTCCAATTATGATT  
 CAAAGGAGTATCATTAAACGTCCTTTGGCCTGATCAATTAATGGCAATTGGGTAAGAGGTCCCT  
 GCGGTGTTATATACTGATAGTATTGGAGTGAAATGATGATTTTCGATGATTTAGTATGGTCAGG  
 TAGTCATTGATTGAAGAGTATACTACTAGAATCCTTGACCAAACCTCTGATTTTTTCAAGATAC  
 AGGGAGTCAAATCACGCAAAACTCTTAAGCAAACACTTTCATAGATTAGACCGTAATAATGAT  
 CCTTTTTATAAGGAACCGTAAGTGTATTTACTCCGAGATTCATCAGTCGACACTAACATCCGA  
 TGCTCTTTTGCAATAGGTTCAATGACGCGATTCTATACATAACATTTAAATGAATATGGTCCG  
 GTAACAACTCCATTTTTTATAAAAAACTGTCGGCGGTATTACTATCAATTCTCATGCTTCTG  
 GCTCCCTTTACACGAGTTTCATCGAGATGTCAGCACTAATTCTTCTAAGTGTCCGTCGATCCC  
 TCTGATGTAGCTTGACATAATGTTTTTATACCGCATAGTTTTTTGAATTGATACTCACCATGTT  
 ATGTCTCAGTTAAGTTTCATTTCATCATTTCTTTTTAAAATAGATTTCTCCATGATAATATATT  
 GCCTCCTTCACGTTTACATTTCCGGTCTCCATGGCTAGT

### **Puzle (2)**

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

TTAATCCCATTTCAAACATTTTGAAGAAAATCAAGGAAAGAAGCCTGCTTTGTGTTTAGAAG  
 CTGACTAAATGTTCTAGTAGATGGAACATTTTTTCTCAAGGCGATTTTTTCGAACTTAGATGTA  
 CTACGAGATTTTTTTCATGTACTCACTAGGAAGGAGCAAGATGAACATGTATTGATAACGATA  
 ACTGCCCATGCCAATAGACAAACCCTTTCAATTAAGTCTTCAGTTTTAATTTGGATAACAAG  
 ACTCTAAAAATCGCCAGATACAGAATGCGTCAAAACCTTGAGAAGGAGTTAAAGGCTTTGTCT  
 AAGGGTGTGCAACTTTAAAAACAAGCTGCAATTGATGCGTATTCTCAAATTCCTGTCATTCTT  
 TTTTTTCTTCCCAGTGATGTTTTTCAATACTAGGAGATCGCTATCCAGAATGCAGACCTTATC  
 AAAGTGTGTTTGGTCAAATACACTGAGTTTTGTGTCCACATTGCTCTTAACATCAACTTGAAA  
 AGTTTTTTGTGCGAATGGAGACAGAATGCTCTGTTCTACTAAAAAATACCAATAAGTCACAC  
 CAAAGGTCTTTTTCAATTCAAAAGCAACGAAGGATTCTCAAACCTCCGTCGGAAAATAACTCG  
 GAAACCAATGAAGAGCAGGAGAGGAAATTTGGCTCACTAAATTACAAATGCTGTTGAAGAT  
 ACATTGGACTATAGTATCAAATTTTTGTGTCTTAAGCTTTTTATGCAGATGCAGTTTCGTTGC  
 ATCAATTTTCATGCATAACCTTTCCCCCATTTATTTCCAT

## Construcciones de la Figura 16

### *ura4 Dyad*

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

TTAGTGTTGGCTCAATCTTTGCTGATATGCCTTCCAACCAGCTTCTCTATATCTCTTGGCTTCG  
 ACAACAGGATTACGACCAGCTCCATAGACTCCACGACCAACAATGATGATATCGCTACCGCAG  
 TTTACAATCACTTCTCCGGCGTCTGTACTGTTGGCCTAATCCATCGCCCTTGACGTCTAGTC  
 CGATACCAGGGGACATAGTTATGTAGTCGCTTTGAAGGTTAGGAAATCGACGACCAGCTATAA  
 AGCCAAAGCAAAAATCGGTATGCTTCTCAAACCATTCTAAGGTTTTTTCCGTATAACTGCCCG  
 TCGCTAGGCTTCCCTTACTGCTCATCTCTGCTAGAAGCAAGAGACCACGTCCCAAAGGTAAACC  
 AACTTCTTTGAGGCCTTGTATAATACCCTCGCCTGGCACTGTATGGCAATTTGTGATGTGCGCC  
 CAGCTCGCGATCTTATATACGCCGCTCGCGTACTGCAATTTAACGGTATTTCCAATGTCTGCGA  
 ATTTGCGATCCTCAAAGATAAGAAAACGATGCTTTTTACCTAAGGCCACCAGTTTTTCTACCA  
 TATCCTGGTGC AAAATCCTCGACAACGTCAATATGCGTTTTAATGACGCATACGTATGGCCGA  
 TCTTGTGCTAGCTCGAGTATTTCCGATTTCTTCGTCAAATCGACCGCGACTGACAAGTTGCT  
 TTGCTTTTCTCCATCAAAGCCAACAATTCCTTGGCAATGGGATTTTTTCATCCCTTCCGCGCGC  
 GCGCTGTATGACTGGAACACGCGCGCTCCAT

### *ura4 Linker*

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

TTAATGCTGAGAAAGTCTTTGCTGATATGCCTTCCAGCCGGCCTCGCGGTAGCGTTTAGCCTCT  
 ACGACCGGGTTCTCGCGGCTCCATAGACTCCACGACCAACAATGATGATATCGCTACCGCAGT  
 TTACAATCACTTCTTCAGGAGTACGATATTGCTGTCCAGCCCGTCTCCTTTAACATCCAAGCC  
 GATACCAGGGCTCATCGTGATATAATCTGACTGCAAATTCGGGAACCTCCTGCCCGCGATAAA  
 GCCAAAGCAAAAATCGGTATGCTTCTCAAACCATTCTAAGGTTTTTCTCTGTGTAGGAACCAGT  
 AGCCAAAGAGCCTTTGGAAGACATTTTCAGCCAAAAGCAAGAGACCACGTCCCAACGGCAGGCC  
 GACCTCCTTTAATCCCTGGATGATGCCTTCACCCGGTACTGTATGGCAATTTGTGATATGAGCC  
 CAAGAAGCAATTTTGTACACACCAGATGCATATTGTAGCTTGACGGTATTTCCAATGTCTGCG  
 AATTTGCGATCTTCGAATATCAAGAACCTGTGTTTCTTGCCGAGTGCTACTAACTTCTCTACCA  
 TATCCTGGTGC AAAATCCTCGACAACGTCAATATGTGTCTTGATAACACAGACATAGGGTCCAA  
 TTTTATCTACCAATTCTAAGATTTCCGATTTCTTCGTTAGGTCTACAGCTACGCTTAGATTTG  
 ACTGTTTCTCCTCCATTAGCGCCAACAATTCCTTGGCAATGGGATTTTTTCATCCCTCAGCTCT  
 AGCTGAATAGCTTTGAAATACTCTAGCATCCAT

### *ura4 Int*

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

TTAATGCTGAGAAAGGCGCTGTTGGTACGCTTTCAGCCAGCTTCTCTATATCTCTTGGCTTCG  
 ACAACAGGATTACGACCAGCTCCATAGACTCCACGACCGACGATAATAATGTCTGAGCCACAG  
 TTTACAATCACTTCTTCAGGAGTACGATATTGCTGTCCAGCCCGTCTCCTTTAACGTCTAGTC  
 CTATGCCCGACTCATAGTTATGTAGTCGCTTTGAAGGTTAGGAAATCGACGACCAGCTATAA  
 AGCCAAAGCAGAAGTCAGTGTGTTTTTTCGAACCATTCTAAGGTTTTTCTCTGTGTAGGAACCAG  
 TAGCCAAAGAGCCTTTGGAAGACATTTTCAGCTAGCAATAGTAAGCCCTGCCTAGAGGTAAC  
 CAACTTCTTTGAGGCCTTGTATAATACCCTCGCCTGGTACGGTGTGACAGTTCGTTATGTGAG  
 CCCAAGAAGCAATTTTGTACACACCAGATGCATATTGTAGCTTGACGGTGTGCGGATATCCG  
 CAACTTTCTATCCTCAAAGATAAGAAAACGATGCTTTTTACCTAAGGCCACCAGTTTTTCTA  
 CCATGTCTTGATCAAAGTCTTCTACGACGTCAATATGTGTCTTGATAACACAGACATAGGGTC  
 CAATTTTATCTACCAATTCGAGTATCTCACTTTTTTAGTTAGATCGACCGCGACTGACAAGT  
 TGCTTTGCTTTTTCTCCATCAAAGCCAATAGCTCTTTAGCGATCGGGTTCTTCATCCCTCAGC  
 TCTAGCTGAATAGCTTTGAAATACTCTAGCATCCAT

## Construcciones de la Figura 17

### *6nuc oct*

Integrado en *S. pombe* (Chr1:5075934..5076896)

TCATGTACTAGAATTTTCATACTGCGTGTCTCGAATTCGACAAGCGTATTCAAGTATGTGCACC  
 GTTTTTCTAATAAAAAAATTTGCTATAAAGTGTTTTAAAAACTATTCACATCTACTCAGTCTTT  
 CAGCCTTGTTTTATTTCATTATATGCCGCCTGTAAGGAACATATATAGAAGTGCTGTGACTCTT  
 TATTTTATTTCAAATTTTCTTTTCTTTATTTTTTTTTATAAATTTGCATAAAAAAAGAAAGAA  
 CCTGCTCGTGAATTCTGTCTTTATTTATGCCTCACTACTGATATACTATATACATAATAGAAA  
 ACCTCGTTCCTTCCAGCATTCCCTCAAGGTGTTAGATTATCCCATGAATTATAAAGGAATATTG  
 GAAATAGTATTTATGAATTATCAGTAACATCAATTGATTTTTTTTTTTGAATAAATTTTTTATAC  
 TTTTTACCGTTGTGTTGGTTTTCCCTCTGCATTGAACTAACCATGTCCAAATCTGACGCAGTTT  
 CTCGCTTGTTGGCAGAGATTTCTGCTCAGCAAAGCCAAAAACAAGTTACTCCAGAATTCCCTT  
 ATCCTACGTCTAGTGGCATAATAAACTTGAACGATCCTAATCCATCATTTGGGAACATTCGTCC  
 GTCTTTATGATCCTACTTTCTCTTCCCAGATCTATCAAGACACCTCGACGGGGACCGTCCCTGT  
 AAACCCGTCTGCCTATCATCCACCGAACCCATCTGTCTCCTTAAAACTTCGTATTGTTTCCCT  
 GTCAAAGCTGCTATGTGGGCATGCTCATTGGAAAACGCGGCGAAAAGCTCCAAAAAATGGAG  
 TCTGCCACAAAGACAAGAGTACAATTTATGCCTGAAACCACTTCCAAACCAGGCGAACGAACA  
 GCTATCATTACCGGCCAAACCG

### *6nuc jap*

Integrado en *S. pombe* (Chr1:5075934..5076896)

CTCAATAAACCACAGAAAACCTCGACGAAAAGCAATAATCGGGATTGACGTCTCGCGATTCCCG  
 AAAAGAGACGAGACAAACAATTGCACTAATATTATCATTTTTAGCTTTTTAAAATTTGCCGCG  
 ATTTTTTCTGTTTTTTATTTATTTTTTCTGTTATTTTGGAAATAATAAAAAAATAAAAAAAT  
 CATCATTCACCAATTATCTACGAACTTCCTCAACTGTTTATGGAATTTTCTCCAGATATTTTAT  
 AAAAAATGCCTCTAACTTATTTAAAGCACAAGAGCTGGAATGGTATGGGACCAATAGAATAC  
 AGGTGATAACTTTACAGTCTACAGTGAAAAAACCGTGAACGTGTGGAAAGGGATGAAGAGT  
 TGGCAGCTTAGAAGCTGTGGAAGCAGCAAAAAGCCGCTCCTGGAGTGTACGAAACGAATCAT  
 TAGAGTGCCTGTTGCCCTTGCTAACTCTCGAAACTATAAAGCGACAAGGAGAGAAATTAACAGA  
 ACTACCTAATGTGCGAACATCATCACATGCCAACTTTTGGGAAAAAGAGGAACGCGAAGTATG  
 TCAGTGCATTCTAACTATACATGCTAATAGGCCAATTCAGCATCTGGAACATGAACGCTTTC  
 AGCAGGAGACAAGCGAGAGCTGGATGCGATGAAAGAAAGACACGGACTGGGTAAGAAACAC  
 GTTTCAACTTTTGAACGATGCCGCTTTCGCCGACTCTTACAGATTTTATGTACTAACTACCTG  
 TAGGGCCGCTTCCATGGTATGCGACATCGCCATCAAAGAAGACAATAATCTCGATCCACAAT  
 TACCCATCAAGTCTGTCAAAGTTCCAGCCTAGAAGACCCTCTGGCGTTCATGAATCGTGTGTT  
 GAGCACAAAACAATCACAAGATTCTAAAACAACGGCCTCTGCTTCCATAAACGTT

## Construcciones de la Figura 18

### *ura4 S. octosporus*

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

TTATTTTTTGGTAAGACGTTCTTGGTAGGCATCCCAGCCTGCTTTTTTGTAAACGTTGGGCTTC  
 AACTACGGGATCCCGGCCAGCACCATAGATTCCACGGCCAACGATGATAATATCGCTGCCACA  
 ATTCACAATAACTTCATGAGGTGTACGGTATTGCTGGCCCATACCGTCGCCTTTAACATCCAA  
 TCCAATACCGGGAGACATAGTGATAAAGTCATGCTGAAGATTGGGGAAACGACGGCCGGCAAT  
 GAAGCCAATGCAAAAGTCGCTATGTTTTTCAGCCCATTCTAAAGTAGTTTCTGTGTAAGAACC  
 AGTGGCCAGACTTCCTTTAGAAGACATTTCCGCCAAGAGAAGCAAGCCGCGACCAAGAGGTAA  
 ACCAACTTCTTTCAAACCCTGAACAATGCCTTCTCCAGGGACAGTATGACAATTAGTAATGTG  
 AGCCCAAGAAGCAATTTTGTAAACACCAGAAGAGTATTGGAGCTTGACGGTGTTCCTCAATATC  
 GGCAAACCTTACGGTCTCAAAGATGAAGAATCGATGCTTCTCGGAAAGGGCGGTCAATTGCTG  
 AACAAAGTCTTGATCAAAATCTTCAATGACATCGACATGCGTCTTAATGACACAAACAAAAGG  
 ACCAATTTTTTCGATGAGAGCAAGGATGTCAGCTTTTTTTGTCAAGTCGACGGCAACAGACAA  
 ATTCGATTGCTTTTGCTCCATTAAGCAAGCAAAGCCTTTGCTAGAGGGTTACTTAACCCATC  
 TGCTCTTTGTGAATATTGGTAATTGAGTCTAGCATCCAT

### *ura4 S. japonicus*

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

TTACTGCTTGAAATGCGACGCTGGTAGGCCTCCCAGCCGGCCTTTCTGTAGCGCTCAGCTTCT  
 TGAGCAGCATCACGGCCGCTGCCGTAGACACCGCGACCGACGATAATGATATCGCTCTGGCTCT  
 CACAGATCACTTCCTCTGGTGTACGGTATTGTTGACCAAGGCCATCGCCCTTAACGTCCAAACC  
 GACACCGGGTGTGATGTGAATAAAGTCGCTCTTGAGACCGGGGAAGCGGCGTCTGGCGATAAA  
 ACCCATAACAGAAGTCGTTGTGCTTCTCAGCCAATTCCAATGTGGCTTGTGTGTAGCTTCCAGTG  
 GCCAGGGTTCCCTTCGAAGACATCTCAGCCAAAAGCAACAAACCACGACCCAAGGGAAGACCA  
 ACTTCCTTCAAACCGCTAATAATACCCTCACCGGGAACAGTGTGGCAGTTGGTGATGTCAGCCC  
 AAGAAGCAATCTTGTACACACCAGCAGAGTACTGAAGCTTGACAGTGTGCCAATGTCGGCAA  
 ACTTGCGATCCTCAAAAATCAAAAACCTTGTGCTTCTCAGCCAGAGCAACCAGCTGTTGGACCA  
 TGTCTGCGTCAAAGTCTCCACAATGTCAATGTGAGTCTTCAACAAGCAGATCGAGGGACCAA  
 TCTTGTCAACGAGTTCAGGACTTGGTTCTTCTTCGTCAAGTCCACAGCGACCGAGAGGTTTG  
 ACTTCTTCTCGTCCATCAAACGCAGCAACTTCTTGGCGACTGCATTAGGATGCACATTGGCGCG  
 CTCCGTGTACGTTTTCAAGGCGATGTCCGACAT



## Construcciones de la Figura 21

### *ura4 S. octosporus Remasterizada*

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

TTATTTTTTAGTTAATCTTTCTTGATAAGCATCCCAGCCTGCTTTCTTGTATCTCTGTGCTTCT  
 ACTACTGGATCTCTTCTGCTCCATAGATACCACGACCAACAATAATAATATCAGAACCACAA  
 TTCACAATCACCTCATGAGGAGTTCTATATTGTTGACCCATTCCATCTCCTTTTACATCTAGTC  
 CTATTCCTGGGCTCATTGTGATGAAGTCATGTTGTAGATTTGGGAATCTTCTGCCAGCAATAA  
 AACCAATACAAAAATCAGAATGCTTCTCAGCCCACTCAAGAGTAGTCTCAGTATAAGAACCAG  
 TAGCTAAAGATCCTTTTGAGCTCATCTCTGCGAGGAGGAGGAGTCTCTTCCTAGTGGTAATC  
 CTACTTCTTTTAGTCCCTGTACAATACCCTCACCAGGAACAGTATGACAATTCGTGATGTGAG  
 CCCAAGAAGCAATTTTATAAACTCCAGAAGAATATTGAAGTTTTACTGTGTTTCCTATGTCTG  
 CGAATTTTCTGTCTTCGAATATGAAGAATCTATGTTTTTCTGAAAGAGCAGTCAATTGTTGAA  
 CAAAATCCTGATCAAAATCCTCAATCACATCAACATGAGTTTTAATAACACAAAACAAAAGGTC  
 CTATTTTTTCTATGAGAGCTAGGATGTCTGCTTTCTTTGTGAGGTCTACTGCTACTGATAGAT  
 TTGATTGCTTCTGTTCCATAAGAGCAAGAAGAGCCTTAGCAAGAGGATTCGAGAGGCCATCAG  
 CTCTTTGAGAATATTGATAATTTAATCTAGCATCCAT

### *ura4 S. japonicus Remasterizada*

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

TTATTGTTTAGAAAATCTTCTTTGATAAGCTTCCCAGCCTGCTTTTCTGTATCTCTCTGCTTCT  
 TGTGCTGCATCTCTTCTGATCCATAGACACCACGACCAACAATAATAATATCAGACTGAGAC  
 TCACAGATCACCTCCTCAGGAGTTCTATATTGTTGACCTAATCCATCTCCTTTTACATCTAGTC  
 CTACTCCTGGCGTCATGTGGATGAAGTCTGATTTTAATCCTGGGAATCTTCTTCTAGCAATAA  
 AACCATACAAAAATCATTATGCTTCTCAGCAAGCTCAAGAGTAGCTTGAGTATAAGAACCAG  
 TAGCTAAAGTTCTTTTGAGCTCATCTCTGCGAGGAGGAGGAGTCTCTTCCTAGTGGTAATC  
 CTACTTCTTTTAGTCCGCTTATAATACCCTCACCAGGAACAGTATGACAATTCGTGATGTCAGC  
 CCAAGAAGCAATTTTATAAACTCCAGCAGAATATTGAAGTTTTACTGTGTTTCCTATGTCTGC  
 GAATTTTCTGTCTTCGAATATTAGGAATTTATGTTTTTCTGCCAGAGCAACCAATTGTTGAAC  
 CATATCAGCATCAAAATCCTCAACAATGTCAATATGAGTTTTTAAAAGACAAATAGAAGGTCC  
 TATTTTATCTACTAGCTCTAGGACTTGGTTCTTCTTTGTGAGGTCTACTGCTACTGATAGATT  
 TGATTTCTTCTCATCCATAAGACGAAGAAGTTTCTTAGCAACAGCATTCGGATGCACATTAGC  
 TCTTTTCAGTATAAGTTTTTAAAGCTATATCACTCAT

## Construcciones de la Figura 22

### ***Kan WT***

Integrado en *S. pombe* (Chr2:4218674..4219480) y en *S. cerevisiae* (Chr:429302..430108)

ATGGGTAAGGAAAAGACTCACGTTTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTA  
TATGGGTATAAATGGGCTCGCGATAATGTGCGGCAATCAGGTGCGACAATCTATCGATTGTAT  
GGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTT  
ACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCAT  
TTTATCCGTA CTCTGATGATGCATGGTTACTCACCCTGCGATCCCCGGCAAAAACAGCATTCC  
AGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGC  
GCCGTTGCATTCGATTCCTGTTTGTAATTGTCTTTAACAGCGATCGCGTATTTTCGTCTCGC  
TCAGGCGCAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAA  
TGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAGCTTTTGCCATTCTCACCGGATTC  
AGTCGTCATCATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGG  
TTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAA  
CTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAA  
TCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAA

### ***Kan Sp***

Integrado en *S. pombe* (Chr2:4218674..4219480) y en *S. cerevisiae* (Chr:429302..430108)

ATGGGAAAAGAAAAA ACTCATGTCTCTCGTCCTCGTCTGAACTCGAACATGGATGCTGATCTT  
TATGGTTATAAATGGGCTCGTGATAATGTAGGACAATCAGGAGCAACAATATACAGACTCTAC  
GGAAAGCCAGACGCTCCAGAATTATTTTTAAAACATGGAAAAGGTTCTGTTGCTAATGATGTG  
ACTGATGAGATGGTTCGTTTGAATTGGCTTACTGAATTTATGCCTTTACCTACTATTAACAT  
TTCATAAGAACACCAGATGACGCATGGCTCCTCACAACAGCGATCCCAGGAAAAACAGCTTTT  
CAAGTTTTAGAAGAATATCCTGATTCTGGAGAGAACATTGTTGATGCTCTTGCTGTTTTTCTT  
CGTCGTTTACATTCTATACCTGTTTGTAACTGTCCATTCAATAGTGACAGAGTATTCAGACTC  
GCACAAGCGCAGAGCAGAATGAACAATGGATTAGTAGATGCTTCTGATTTTGATGATGAGAGA  
AATGGCTGGCCTGTTGAGCAGGTTTGGAAAGGAGATGCATAAATTATTACCTTTTTCTCCTGAT  
AGTGTAGTAACACATGGAGACTTCAGCCTAGACAACCTCATCTTCGACGAAGGAAAATTAATA  
GGATGTATAGATGTTGGAAGAGTTGGAATTGCTGATCGCTATCAGGATCTTGCGATTCTTTGG  
AATTGCTTGGTGAGTTTAGTCCTTCTTTACAAAAAAGATTATTCCAAAAATATGGAATAGAC  
AACCCAGACATGAACAAGCTTCAATTTCAATTTAATGTTAGATGAATTTTTTTTAA

### ***Kan Sc***

Integrado en *S. pombe* (Chr2:4218674..4219480) y en *S. cerevisiae* (Chr:429302..430108)

ATGGGCAAGGAGAAGACCCACGTGTCCC GCCCCCGCCTGAACTCCAACATGGACGCGGACCTTT  
ACGTTTATAAGTGGGCTCGTGATAATGTGCGGCAATCAGGAGCAACAATCTACAGGCTCTACG  
GCAAGCCCGACGCGCCGAGCTGTTCCCTGAAGCACGGCAAGGGCAGTGTGGCCAACGACGTGA  
CCGACGAGATGGTGCGGCTGAACTGGCTGACGGAGTTCATGCCTCTTCTACTATTAACATT  
TTATCCGTA CTCCAGATGATGCATGGCTGCTCACAACCGCCATCCCCGGCAAGACGGCGTTCCA  
GGTGTGGAGGAGTACCCGGACAGTGGGGAGAACATCGTGGACGCCCTGGCGGTGTTCTCTGCG  
CCGGCTGCACTCGATTCTGTTTGTAATTGTCTTTAACAGCGATCGCGTATTTAGATTAGCA  
CAAGCGCAAAGCCGCATGAACAACGGCCTCGTCGACGCCAGCGACTTCGACGACGAGCGCAACG  
GGTGGCCGGTGGAGCAGGTGTGGAAGGAGATGCACAAGCTGCTGCCGTTTTCTCCTGATTCTG  
TTGTTACTCATGGTGATTTCTCATTAGATAATCTAATCTTCGACGAAGGCAAGCTCATCGGCT  
GCATCGACGTCGGCAGGTTGGGCATCGCGGACAGGTACCAGGACCTGGCCATCCTGTGGAAC  
GCCTGGGGGAGTTCTCGCCGTCGCTTCAGAAGCGTCTTTTTCAAAAATATGGTATTGATAATC  
CAGATATGAACAACTGCAATTCACCTCATGCTCGACGAGTCTTCTGA





---

<b>A</b>	Adenina
<b>ATP</b>	Adenosín trifosfato
<b>C</b>	Citosina
<b>ChIP</b>	Inmunoprecipitación de cromatina
<b>dA</b>	Desoxiadenosina
<b>dATP</b>	Desoxiadenosina trifosfato
<b>dC</b>	Desoxicitidina
<b>dCTP</b>	Desoxicitidina trifosfato
<b>DEPC</b>	Dietil Pirocarbonato
<b>dG</b>	Desoxiguanosina
<b>dGTP</b>	Desoxiguanosina trifosfato
<b>DNA</b>	Ácido Desoxirribonucleico
<b>dNTP</b>	Desoxinucleótido trifosfato
<b>dT</b>	Desoxitimidina
<b>dTTP</b>	Desoxitimidina trifosfato
<b>EDTA</b>	Ácido etilen diamino tetra acético
<b>G</b>	Guanina
<b>GRF</b>	Factor regulador general
<b>h</b>	Horas
<b>IAA</b>	Alcohol isoamílico
<b>IGR</b>	Región intergénica
<b>Kb</b>	Kilobase
<b>L</b>	Litro
<b>log</b>	Logaritmo
<b>M</b>	Molar
<b>mg</b>	Miligramo
<b>min</b>	Minutos
<b>ml</b>	Mililitros
<b>MM</b>	Medio mínimo
<b>mM</b>	Milimolar
<b>MNase</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>OD</b>	Densidad óptica

<b>ORF</b>	Marco abierto de lectura
<b>pb</b>	Pares de bases
<b>Pol</b>	Polimerasa
<b>PSWM</b>	Matriz de puntuación específica de posición
<b>qPCR</b>	PCR cuantitativa
<b>RNA</b>	Ácido Ribonucleico
<b>RNAPII</b>	RNA polimerasa II
<b>rpm</b>	Revoluciones por minuto
<b>SDS</b>	Dodecil sulfato sódico
<b>seg</b>	Segundos
<b>seq</b>	Secuenciación
<b>SSC</b>	Citrato de sodio salino
<b>T</b>	Timina
<b>TAE</b>	Tris/Acetato/EDTA
<b>TE</b>	Tris/EDTA
<b>TF</b>	Factor de transcripción
<b>TSS</b>	Sitio de inicio de la transcripción
<b>TTS</b>	Sitio de terminación de la transcripción
<b>U</b>	Unidades
<b>μCi</b>	Microcurios
<b>μg</b>	Microgramos
<b>μl</b>	Microlitros

