

OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA



1 Número de publicación: **2 523 845**

51 Int. CI.:	
C12Q 1/68	(2006.01)

(12)	TRADUCCIÓN DE PATENTE EUROPEA			Т3
 96) Fecha de presentación 97) Fecha y número de put 	y número de la solicitud europea: blicación de la concesión europea:	02.06.2009 03.09.2014	E 09757758 (9) EP 2297340	
54 Título: Métodos de de	ección de interacciones cromos	ómicas de larg	go alcance	

30 Prioridad:	Titular/es:	
 02.06.2008 GB 0810051 (45) Fecha de publicación y mención en BOPI de la traducción de la patente: 02.12.2014 	OXFORD BIODYNAMICS LIMITED (100.0%) 26 Beaumont Street Oxford OX1 2NP, GB 72 Inventor/es:	
	RAMADASS, AROUL SELVAM; AKOULITCHEV, ALEXANDRE y MELLOR, ELIZABETH JANE	
	(74) Agente/Representante:	
	LAZCANO GAINZA, Jesús	

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Métodos de detección de interacciones cromosómicas de largo alcance

Campo de la invención

La presente invención se refiere a métodos para controlar los cambios epigenéticos.

10 Antecedentes

5

El genoma eucariota está organizado en estructuras complejas de orden superior; de hecho, las primeras micrografías electrónicas de la cromatina extraída revelaron una estructura sin histona que forma bucles radiales (Earnshaw y Laemmli, 1983, J Cell Bio 96, 84-93). El superenrollamiento negativo sin restricciones en los genomas de mamíferos sugiere tamaños

- 15 de "dominio" del cromosoma del orden de decenas de kilobases (Kramer y Sinden, 1997, Biochem 36, 3151-3158). Más recientemente, el uso de la tecnología 3C (captura de la conformación del cromosoma) se ha demostrado que existen interacciones de largo alcance de orden superior en una amplia variedad de eucariotas, incluyendo células de *S. cerevisiae* (Dekker et al., 2002, Science 295, 1306-1311), mosca (Blanton et al., 2003 Genes Dev 17, 664-675), ratón (Tolhuis et al., 2002, Mol Cell 10, 1453-1465) y humano (Carroll et al., 2005, Cell 122, 33-43), generalmente de un tamaño de kilobases. En
- 20 células de mamíferos, se ha demostrado la asociación de regiones de control del locus o reforzadoras con los genes expresados de forma activa en los loci de β-globina (Tolhuis et al., 2002, Mol Cell 10, 1453-1465) y de la proteína C reactiva (Choi et al., 2007, Nucleic Acids Res. 35, 5511-5519), así como la formación de bucles reunidos de genes recombinantes de inmunoglobulina (Skok et al., 2007, Nat Immunol 8, 378-387). En células de mamíferos, se propone una transcripción continua para dirigir la organización del genoma y la formación de bucles de genes (Chakalova et al., 2005 Nat Rev Genet 6,
- 25 669-677; Marenduzzo et al., 2007, Trends Genet 23, 126-133 y las referencias allí), pero análisis recientes sugieren que, al menos para el locus de β-globina, se mantienen algunas interacciones de ADN de largo alcance después de inhibir la transcripción (Palstra et al., 2008 PLoS ONE 3, e1661) argumentando en contra de los modelos que sugieren que las funciones comprometidas de la ARN polimerasa como las ataduras de los bucles de cromatina.
- 30 En el genoma de levadura, la transcripción activa parece ser importante para la formación de "bucles de genes", interacciones de largo alcance que enlazan las regiones 5' y 3' de los genes activos (Ansari y Hampsey, 2005 Genes Dev 19, 2969-2978, O'Sullivan et al., 2004, Nat Genet 36, 1014-1018; Singh y Hampsey, 2007, Mol Cell 27, 806-816). La inmunoprecipitación de la cromatina (ChIP) demuestra la presencia de TFIIB y la forma fosforilada de RNAPII en ambos promotores y terminadores. Además, se requiere de TFIIB funcional para formar estas interacciones de largo alcance (Singh y Hampsey, 2007, Mol Cell 27, 806-816). TFIIB es capaz de interactuar con el ARN no codificante y la pérdida de TFIB
- 35 y Hampsey, 2007, Mol Cell 27, 806-816). TFIIB es capaz de interactuar con el ARN no codificante y la pérdida de TFIIB promovida por el ARN no codificante conduce a la pérdida de la interacción de largo alcance en DFHR (Martianov et al., 2007, Nature 445, 666-670).
- Los altos niveles de ARN no codificante transcritos en todo el genoma incluyen transcriptos antisentido para marcos de lectura abiertos. Existen tanto en genomas eucariotas como procariotas (Johnson et al., 2005, Trends Genet 21, 93-102, Kapranov et al., 2007, Nat Rev Genet 8, 413-423; Selinger y otros, 2000, Proc Natl Acad Sci EE.UU. 103, 4.192-4197). En eucariotas, muchos de estos transcriptos no están destinados nunca para su traducción en proteína y algunos están destinados para degradación mediada por el exosoma por parte del complejo TRAMP (Bickel y Morris, 2006 Cell Mol 22, 309-316). En la levadura estos transcriptos inestables crípticos (los CUT) se detectan en promotores (Berretta et al., 2008, Genes Dev 22, 615-626; Davis y Ares, 2006, Proc Natl Acad Sci EE.UU. 103, 3262-3267), en las regiones entre los genes
- (Wyers et al., 2005, Cell 121, 725-737) y antisentido de genes (Camblong et al., 2007 Cell 131, 706-717; Uhler et al., 2007, Proc Natl Acad Sci EE.UU. 104, 8011-8016).
- 50 Experimentos detallados que perfilan los ARN de *S. cerevisiae* utilizando microchips han demostrado que la transcripción se produce en casi todas partes del genoma de la levadura (Perocchi et al., 2007, Nucleic Acids Res 35, e128; Samanta et al., 2006, Proc Natl Acad Sci EE.UU. 103, 4192-4197; Miura et al y otros, 2006, PNAS 103, 17846-17851; Hongay et al., 2006, Cell 127, 735-745; David y otros, 2006, Proc Natl Acad Sci EE.UU. 103, 5320-5325; Havilio et al., 2005, BMC Genomics 6, 93). Estos análisis también indican que se transcriben al menos parcialmente entre 100 y 370 genes en ambas direcciones que producen transcriptos sentido y antisentido poliadenilados estables. Muchos de estos genes se transcriben activamente.
- 55

Los presentes inventores han utilizado el locus *GAL* como un sistema modelo en el cual comparar los estados inducidos y reprimidos y observar diferencias en el transcripto antisentido y la regulación epigenética. Los inventores muestran que hay transcriptos antisentido que controlan tanto los estados inducidos como los reprimidos en el locus *GAL*, y que estos transcriptos difieren en tamaño, posición y abundancia. Transcriptos antisentido muy abundantes en el locus inducido están asociados con la producción del transcripto sentido a partir del promotor *GAL10*. Además, los niveles de transcriptos antisentido se correlacionan fuertemente con los niveles de Hda1 asociados con el locus pero no con la acetilación misma

de las histonas. Los inventores también identifican que Hda1 parece ser necesario para las interacciones de largo alcance en el locus *GAL* reprimido, lo que sugiere un vínculo entre los transcriptos antisentido, las modificaciones epigenéticas y las estructuras de cromatina de orden superior en el estado reprimido. Los cambios en la conformación del locus después de la conmutación del estado reprimido al estado inducido han sido identificados con implicación del ARN antisentido en el control de este. La presente invención se basa en el descubrimiento de que la represión de genes es un estado proactivo de la regulación que implica cambios epigenéticos específicos del locus.

Resumen

5

- 10 De acuerdo con un primer aspecto de la invención, se proporciona un método para controlar los cambios epigenéticos que comprenden el seguimiento de los cambios en las interacciones cromosómicas condicionales de largo alcance en al menos un locus cromosómico donde diferentes conformaciones de interacciones de largo alcance están asociadas con una condición asociada con un cambio en el nivel de expresión de uno o más genes o un cambio en uno o más productos génicos en donde la condición se selecciona entre cáncer, trastornos cardiovasculares, condiciones inflamatorias,
- 15 incluyendo trastornos autoinmunes y respuestas inflamatorias a enfermedades inflamatorias y trastornos genéticos heredados modulados por mecanismos epigenéticos dicho método comprendiendo las etapas de:

(i) entrecruzamiento *in vitro* de dichas interacciones cromosómicas de largo alcance presentes en al menos un locus cromosómico;

- 20 (ii) aislamiento el ADN entrecruzado de dicho locus cromosómico;
 - (iii) sometimiento de dicho ADN entrecruzado a una digestión de restricción con una enzima que corta al menos una vez dentro de al menos un locus cromosómico;
 - (iv) ligar dichos extremos del ADN escindido entrecruzado para formar bucles de ADN;
 - (v) la identificación de la presencia de bucles de ADN;
- 25

en donde la presencia de bucles de ADN indica la presencia de una interacción cromosómica específica de largo alcance y en donde las interacciones cromosómicas de largo alcance no son interacciones de largo alcance entre genes y sus elementos reguladores.

- 30 Se entenderá que las interacciones cromosómicas condicionales de largo alcance siempre estarán presentes en la cromatina. Se entenderá además que estas interacciones son dinámicas y cambiarán dependiendo del estado de la región del cromosoma, es decir, si se está transcribiendo o reprimiendo en respuesta al cambio de las condiciones fisiológicas.
- Tal como se utiliza aquí, el término interacciones condicionales de largo alcance se refiere a las interacciones entre las regiones distales de un locus en un cromosoma, siendo dichas interacciones dinámicas y que producen alteración dependiendo del estado de la región del cromosoma.
- Como se usa en este documento, el término espectro de interacción de largo alcance se refiere a las diferentes conformaciones de las interacciones cromosómicas de largo alcance que pueden estar presentes en un locus cromosómico dado. Se entenderá que como se describió anteriormente, estas interacciones son dinámicas, con diversas interacciones de conformación o de rompimiento de largo alcance dependiendo del estado del locus.

Se entenderá además que las interacciones cromosómicas de largo alcance se pueden entrecruzar mediante cualquier medio adecuado. En una realización preferida, las interacciones cromosómicas de largo alcance se entrecruzan utilizando formaldehído.

Se entenderá además que los bucles de ADN presentes pueden ser indicativos de la transcripción o la represión de dicho locus cromosómico, o alternativamente, de la expresión de un producto alterado a partir de dicho locus cromosómico.

- 50 La presencia de los bucles de ADN puede ser identificada como se describe en el presente aquí a continuación en relación con el locus GAL. Será fácilmente evidente para el experto en la materia que el método descrito en relación con este locus puede ser adaptado para ser utilizado en cualquier otro locus donde se cree que ocurren las interacciones de largo alcance. Estos bucles pueden ser detectados usando técnicas conocidas en la técnica tales como el ensayo 3C (Conformación de cromosoma de captura) ensayo (Dekker, 2006, Nat Methods Nat 3, 17-21; Dekker et al., 2002, Science 295, 1306-1311; 0'Sullivan et al., 2004, Nat Genet 36, 1014-1018).
 - La persona experta será consciente de numerosos enzimas de restricción que pueden ser usadas para cortar el ADN en el locus cromosómico de interés. Será evidente que la enzima particular utilizada dependerá del locus estudiado y de la secuencia del ADN localizado allí.
- 60

La presente invención se basa en el sorprendente descubrimiento por parte de los inventores de que las interacciones

cromosómicas condicionales de largo alcance están siempre presentes en un locus dado en el cromosoma y que el perfil de las interacciones cromosómicas condicionales de largo alcance cambia en función del estado real de la región, su actividad y las condiciones fisiológicas, es decir, la presencia o ausencia de una interacción particular de largo alcance, proporcionará una indicación del estado de esa región.

5

10

entre los genes.

Además, los inventores han descubierto que concordante con los datos genéticos anteriores, estas interacciones cromosómicas condicionales de largo alcance pueden solaparse e incluir las regiones de los cromosomas que se muestra que codifican genes relevantes o no descritos, pero igualmente pueden estar en las regiones entre los genes. Además, debe tenerse en cuenta que los inventores han descubierto que las interacciones de largo alcance en todas las regiones son igualmente importantes en la determinación del estado del locus cromosómico. Estas interacciones de largo alcance no están necesariamente en la región de codificación de un gen particular localizado en el locus y pueden estar en las regiones

Se entenderá además por parte de la persona experta que el término epigenética se refiere a cambios heredables en la función del gen dentro de una célula que son causados por cambios distintos de los cambios en la secuencia de ADN subyacente; estos cambios pueden ser causados, por ejemplo, por factores ambientales, metilación del ADN, transcriptos de ARN antisentido no codificantes, carcinógenos no mutagénicos, modificaciones de las histonas, remodelación de la cromatina y las interacciones locales específicas del ADN de largo alcance todos los cuales han sido implicados en la creación del entorno específico para la actividad transcripcional definida en los genes o las unidades transcripcionales de ARN no codificante dentro de la región de interés.

Se entenderá que los cambios epigenéticos pueden ser causados por cambios en la secuencia de ácido nucleico subyacente, que de por sí no afectan directamente un producto génico o el modo de expresión del gen, tales cambios pueden ser, por ejemplo, del SNP dentro y/o fuera de los genes, y fusiones de genes y/o supresiones de ADN entre los genes.

25 genes.

Además, será evidente que el término condición fisiológica específica se refiere a cualquier condición en la que hay un cambio en el estado fisiológico definido de la célula. Esto puede ser por un cambio en el nivel de expresión de uno o más genes, o un cambio en uno o más productos génicos. De acuerdo con la presente invención tales condiciones incluyen

30 trastornos cardiovasculares, condiciones inflamatorias, incluyendo trastornos autoinmunes y las respuestas inflamatorias a las enfermedades infecciosas en desarrollo, trastornos genéticos heredados modulados por mecanismos epigenéticos.

Preferiblemente, la presencia de los bucles de ADN se identifica usando técnicas de PCR. Se entenderá que la presencia de un bucle puede estar indicada por la presencia de un producto de la PCR que está ausente en la ausencia de un bucle de ADN o viceversa. También se entenderá que el tamaño del producto de la PCR producido puede ser indicativo del bucle de ADN específico presente y por lo tanto se puede utilizar para identificar el estado del locus.

En una realización preferida, la presencia de un bucle de ADN indica un estado alterado de la transcripción indicativo de una condición fisiológica específica.

40

35

En una segunda forma de realización preferida, la ausencia de un bucle de ADN indica un estado alterado de la transcripción indicativo de una condición fisiológica específica.

45 Será evidente para el experto que el método de acuerdo con el primer aspecto puede ser utilizado no sólo para controlar la presencia de una interacción cromosómica específica de largo alcance en un locus cromosómico, sino igualmente para controlar la ausencia de una interacción cromosómica específica de largo alcance en dicho locus cromosómico.

Se entenderá que en cualquier aspecto de la presente invención, los cambios en las interacciones cromosómicas condicionales de largo alcance de una muestra pueden ser controlados mediante la comparación de la conformación de las interacciones cromosómicas de largo alcance en un locus en diferentes instantes de tiempo en el mismo tejido o tipo de célula o mediante comparación con una muestra correspondiente a un estado fisiológico conocido.

Además, se debe entender que las interacciones cromosómicas de largo alcance de la presente invención no se refieren a las interacciones de largo alcance entre los genes y sus elementos reguladores tal como fue descrito anteriormente por Chambeyron et al., (2004), Curr Opin Biol. 16, 256-262; de Laat et al., (20030 Chromosome res 11, 447-459; y Dekker, (2003), J. Trends Biochem. Sci. 28, 277-280. En vez de eso, la presente invención se refiere a los cambios condicionales en las interacciones de largo alcance dentro de un locus particular como una indicación de un cambio en la actividad de un gen.

También se describe aquí un método para controlar los cambios epigenéticos que comprende el seguimiento de los cambios en las interacciones cromosómicas condicionales de largo alcance en al menos un locus cromosómico donde el espectro de la interacción de largo alcance se asocia con una condición fisiológica específica, dicho método comprendiendo la etapa de

identificar un cambio en el perfil de ARN antisentido expresado a partir de al menos un locus cromosómico.

Será evidente para el experto que el cambio en el perfil de ARN antisentido puede ser un cambio en el tamaño, la posición de partida, y/o el número de transcriptos de ARN antisentido.

5

Se entenderá por la persona experta que se conoce el fenómeno de la producción de transcriptos de ARN antisentido en los loci reprimidos en los cromosomas. Sin embargo, los inventores han descubierto, sorprendentemente, que el perfil de ARN antisentido transcrito a partir de cambios en un locus cromosómico dependiendo de si el locus es inducido o reprimido y que los transcriptos de ARN antisentido producidos desempeñan un papel central en el control de la transcripción del transcripto de la tr

10 de RNA sentido de un locus particular y las condiciones epigenéticas en ese locus, incluyendo las interacciones de largo alcance.

También se describe aquí un método para diagnosticar un trastorno asociado con al menos un cambio epigenético en un sujeto, comprendiendo dicho método la identificación de un cambio en una o más interacciones cromosómicas de largo alcance en al menos un locus cromosómico asociado con dicho trastorno en una muestra aislada del sujeto; en donde dicho método comprende el método de cualquiera de los aspectos uno o dos.

El cambio epigenético puede dar como resultado la transcripción alterada del locus cromosómico.

20 Se entenderá que la transcripción alterada puede ser sobrerregulación, represión, o producción de un transcripto alternativo con un sitio de inicio y/o de terminación cambiado, o una variante de empalme de los mismos.

Será evidente que el cambio epigenético provoca un cambio en la expresión de al menos un gen y/o unidad transcripcional dentro de la parte no codificante del genoma.

25

15

El trastorno se puede seleccionar de entre cáncer, trastornos cardiovasculares, condiciones inflamatorias, incluyendo trastornos autoinmunes y respuestas inflamatorias a enfermedades infecciosas y trastornos genéticos heredados modulados por mecanismos epigenéticos. Cualquier otra condición que resulta en un cambio en al menos una interacción cromosómica de largo alcance también puede ser diagnosticada por los métodos actuales.

30

También se divulga aquí un método de regulación de la transcripción de al menos un gen en un paciente que sufre de un trastorno asociado con expresión génica alterada, comprendiendo dicho método la administración a dicho paciente de un ARN antisentido en una cantidad eficaz para alterar la transcripción de dicho al menos un gen.

35 En un ejemplo, el trastorno resulta de la sobreexpresión de dicho al menos un gen.

Será evidente para el experto que el trastorno puede igualmente resultar de la represión de dicho al menos un gen, o de la producción de un producto génico alterado a partir de dicho al menos un gen.

- 40 El trastorno se puede seleccionar de entre cáncer, trastornos cardiovasculares, condiciones inflamatorias, incluyendo trastornos autoinmunes y las respuestas inflamatorias a enfermedades infecciosas y trastornos genéticos heredados moduladas por mecanismos epigenéticos.
- En un ejemplo preferido, dicho ARN antisentido elige como objetivo al menos un sitio de enlazamiento del CTCF. 45

CTCF es un factor multifuncional, que, como se discute más adelante está implicado en el establecimiento y mantenimiento de estructuras de cromatina de orden superior.

En un ejemplo adicional, la administración de dicho ARN antisentido resulta en la modulación de las enzimas HDAC. 50

Se sabe que la acetilación de histonas está involucrada con la modulación de la transcripción. Se ha sugerido que esta modulación también se controla a través del ARN antisentido.

Las enzimas HDAC se clasifican en cuatro clases dependiendo de la identidad de la secuencia y de la organización del dominio. En un ejemplo, dicha enzima HDAC se selecciona de una enzima HDAC de clase i-iv.

Las enzimas HDAC de clase I incluyen HDAC1, HDAC2, HDAC3, HDAC8; Las enzimas HDAC de clase II incluyen HDAC4, HDAC5, HDAC6, HDAC7A, HDAC9, HDAC10; Las enzimas HDAC de clase III incluyen homólogos de Sir2 en la levadura *Saccharomyces cerevisiae*, y las sirtuinas en mamíferos (SIRT1, SIRT2, SIRT3, SIRT4, SIRT5, SIRT6, SIRT7); Las enzimas HDAC de clase IV incluyen HDAC11.

IDAG de clase iv incluyen indag i

También se divulga aquí un método de regulación de la transcripción de al menos un gen en un paciente que sufre de un trastorno asociado con la expresión de genes alterados, comprendiendo dicho método la administración a dicho paciente de ARN de interferencia complementario a una molécula de ARN antisentido implicado en la modulación de dicho gen.

5

Será evidente para el experto en la materia que el trastorno puede resultar de la sobreexpresión o represión de dicho al menos un gen, o de la producción de un producto génico alterado a partir de dicho al menos un gen.

El trastorno se puede seleccionar de entre cáncer, trastornos cardiovasculares, condiciones inflamatorias, incluyendo 10 trastornos autoinmunes y respuestas inflamatorias a las enfermedades infecciosas, y trastornos genéticos heredados modulados por mecanismos epigenéticos.

También se divulga aquí un ARN antisentido para el tratamiento de un trastorno asociado con expresión de genes alterados, en donde dicho ARN antisentido regula la transcripción de dicho gen.

15

También se divulga aquí el uso de ARN antisentido en la fabricación de un medicamento para el tratamiento de un trastorno asociado con expresión de genes alterados, en donde dicho RNA antisentido regula la transcripción de dicho gen.

El trastorno se selecciona de entre cáncer, trastornos cardiovasculares, condiciones inflamatorias, incluyendo trastornos 20 autoinmunes y respuestas inflamatorias a enfermedades infecciosas y trastornos genéticos heredados modulados por mecanismos epigenéticos.

En un primer ejemplo, dicho ARN reprime la transcripción de dicho gen.

25 En un segundo ejemplo, dicho ARN induce la transcripción de dicho gen.

En un ejemplo preferido, dicho ARN antisentido elige como objetivo al menos un sitio de enlazamiento al CTCF.

En un ejemplo preferido adicional, dicho ARN antisentido modula enzimas HDAC. 30

Será evidente para el experto en la materia que los medicamentos anteriores se pueden formular en formas de dosificación farmacéuticas, junto con vehículos farmacéuticamente aceptables adecuados, tales como diluyentes, rellenos, sales, amortiguadores, estabilizantes, solubilizantes, etc. La forma de dosificación puede contener otros excipientes farmacéuticamente aceptables para modificar condiciones tales como pH, osmolaridad, sabor, viscosidad, esterilidad, lipofilicidad, solubilidad etc.

35

Las formas de dosificación adecuadas incluyen formas sólidas de dosificación, por ejemplo, comprimidos, cápsulas, polvos, gránulos dispersables, tabletas y supositorios, incluyendo formulaciones de liberación sostenida y de liberación retardada. Los polvos y comprimidos comprenderán generalmente de aproximadamente 5% hasta aproximadamente 70% de

- 40 ingrediente activo. Los vehículos y excipientes sólidos adecuados se conocen generalmente en el arte e incluyen, por ejemplo, carbonato de magnesio, estearato de magnesio, talco, azúcar, lactosa, etc. Los comprimidos, polvos, tabletas y cápsulas son todos formas de dosificación adecuadas para administración oral.
- Las formas de dosificación líquidas incluyen soluciones, suspensiones y emulsiones. Las preparaciones en forma líquida se 45 puede administrar por vía intravenosa, intracerebral, intraperitoneal, parenteral o inyección o infusión intramuscular. Las formulaciones inyectables estériles pueden comprender una solución o suspensión estéril del agente activo en un diluyente o solvente farmacéuticamente aceptable no tóxico. Los diluyentes y disolventes adecuados incluyen agua estéril, solución de Ringer, y solución isotónica de cloruro de sodio, etc. Las formas de dosificación líquidas también incluyen soluciones o aerosoles para administración intranasal.
- 50

60

Las preparaciones en aerosol adecuadas para inhalación pueden incluir soluciones y sólidos en polvo, que pueden combinarse con un vehículo farmacéuticamente aceptable, tal como un gas comprimido inerte.

También se incluyen formas de dosificación para administración transdérmica, incluyendo cremas, lociones, aerosoles y/o 55 emulsiones. Estas formas de dosificación se pueden incluir en parches transdérmicos del tipo matriz o depósito, que se conocen generalmente en la técnica.

Las preparaciones farmacéuticas se pueden preparar convenientemente en forma de dosificación unitaria, de acuerdo con procedimientos convencionales de formulación farmacéutica. La cantidad de compuesto activo por dosis unitaria se puede variar de acuerdo con la naturaleza del compuesto activo y el régimen de dosificación deseado.

Los agentes activos se deben administrar a sujetos humanos en "cantidades terapéuticamente efectivas", es decir una dosificación suficiente para provocar un resultado médicamente deseable en el paciente. La dosis exacta y la frecuencia de administración de una cantidad terapéuticamente efectiva del agente activo variarán, dependiendo de factores tales como la naturaleza de la sustancia activa, la forma de dosificación y la vía de administración.

5

De acuerdo con un octavo aspecto de la presente invención, se proporciona un método para identificar el estado de la transcripción de un locus cromosómico, dicho método comprendiendo las etapas de: identificar el perfil del transcripto de ARN antisentido expresado a partir de dicho locus cromosómico; y comparar dicho perfil con el perfil del transcripto de ARN antisentido de dicho locus cromosómico en un estado conocido.

10

Preferentemente, dicho locus cromosómico comprende al menos un gen.

Preferiblemente, dicho gen es un gen conocido o que se sospecha que está involucrado en una condición fisiológica específica.

15

Se entenderá que dicha condición puede ser cualquier condición asociada con un cambio epigenético, por ejemplo, cáncer, trastornos cardiovasculares, condiciones inflamatorias, incluyendo trastornos autoinmunes y respuestas inflamatorias a enfermedades infecciosas, y trastornos genéticos heredados modulados por mecanismos epigenéticos.

- 20 Preferiblemente, dicho método se realiza in vitro. Cuando una cantidad de características se indican como opciones, se contempla cada característica individual como aplicable de forma individual o en combinación con cualquier otra característica descrita en la solicitud.
- Las divulgaciones hechas aquí se describirán ahora adicionalmente con referencia a los siguientes ejemplos y figuras en las que:

Figura 1. Transcriptos antisentido condicionales en GAL10.

Transferencias tipo Northern de ARN total probadas con sondas sentido y antisentido específicas en GAL10. A, B. Se cultivó

- 30 la cepa BY4741 en galactosa (carril 1), se lavó (carril 2) y se la transfirió a un medio que contiene glucosa durante 15 minutos (carril 3), 60 minutos (carril 4), 120 minutos (carril 5), 180 minutos (carril 6) o 360 minutos (carril 7). Se muestran dos exposiciones de la señal antisentido en B. Se indica la posición de las bandas de ARNr. Se expusieron el panel de la izquierda en B y el panel A durante el mismo tiempo. A los 15 minutos, los AS de GAL10 de 2,25kb y de 2,4 kb (véase la Fig. 3B) son ambos detectables a medida que se produce el cambio de la del estado activo al estado reprimido. C. Se lavaron los cultivos de levadura durante la noche en galactosa y se transfirieron a un medio fresco que contiene el azúcar
- 35 indicado. En este experimento, son evidentes altos niveles del GAL10 y del transcripto de fusión GAL10-7 en galactosa con los transcriptos antisentido equivalentes. La exposición sentido es el 20% de la antisentido. SE carga el doble de ARN en el carril 2*.

Figura 2. Transcriptos del mapeo por (RT)-PCR de transcripción inversa en el locus de GAL.

A. Posición de los cebadores de RT para la transcripción reversa (RT) del transcripto sentido (S) o antisentido (AS) y los cebadores anidados para la amplificación por PCR. Cada conjunto de cebadores está diseñado para amplificar una región de aproximadamente 200-300 pb en los sitios mostrados en el locus *GAL*. B. Mapeo de transcriptos sentido y antisentido en ARN total en tres fuentes de carbono a través de *GAL7* y *GAL10*. El control que carece de RTasa para el ARN preparado en glucosa se muestra en el carril 7. C. Mapeo de transcriptos antisentido en *GAL7* y *GAL10* en el ARN total preparado a partir de células cultivadas en glucosa. Los controles incluyen la omisión de la etapa de RT y un control positivo para la eficiencia del cebador de PCR sobre el ADN total (no mostrado).

Figura 3. Transcriptos de mapeo en el locus *GAL* con sondas específicas de la hebra.

- A. Autorradiografías de ARN total preparado a partir de BY4741 (carril 1), W303-1a (carril 2) y YMH147 (carril 3) cultivado en glucosa, rafinosa o galactosa e hibridado con sondas específicas monocatenarias 1-7 (todas diseñadas para detectar transcriptos antisentido con respecto a *GAL10*). Las posiciones de los ARNr (*RDN25* y *RDN18*) y ARNt están marcadas con líneas negras a través de las autorradiografías, que fueron expuestas durante 24 horas. B. Resumen de los datos en la Fig. 2 y la Fig. 3A que muestran transcriptos en locus reprimido (glucosa) o inducido (galactosa).
- Figura 4. Se requieren secuencias en el extremo 3' de *GAL10* para inducción del transcripto de sentido. A. Diagrama esquemático que muestra la posición de los transcriptos principales en GAL10 reprimido o inducido (WT) y un derivado de GAL10 que contiene una inserción de *pTEF: KanMX: TEFter* (Mut) que crea una supresión de 554 pb en el extremo 3' de GAL10. B y C. Transferencias tipo Northern de ARN total a partir de las cepas mostradas cultivadas en las fuentes de carbono indicadas. En B se indica la posición de los ARNr con líneas punteadas y se utilizaron sondas específicas sentido y antisentido. Se expusieron las autorradiografías durante tiempos equivalentes. En C se indujeron cultivos y se prepararon muestras a intervalos de 10 minutos después de la transferencia de rafinosa (carril 1) a galactosa (carriles 2 6).

Figura 5. La asociación de Hda1 y Eaf3 con GAL10 está relacionada con los niveles de transcripto antisentido. Inmunoprecipitación de la cromatina preparada a partir de células cultivadas durante la noche en galactosa y transferida a medio fresco que contenía galactosa o glucosa durante 20 minutos y se detecta usando PCR en tiempo real en *GAL10*. A. Hda1-myc normalizado con el control sin etiquetar. B. Eaf3 normalizado con el control sin etiquetar. C. H3K18ac normalizado con H3K18R y luego con histona H3.

Figura 6. 3C con inmunoprecipitación en el locus *GAL* reprimido. A. Mapa de sitios de restricción DpnII (líneas verticales) que muestra el número, la posición y orientación de los cebadores (flechas 3'OH) y la posición aproximada de los transcriptos sentido inducidos y los transcriptos antisentido en el locus reprimido. B. 3C con IP (tres paneles superiores Rbp1, panel inferior Myc) en las cepas mostradas (BY4741 en el fondo) en los sitios indicados. C. Controles para la reacción estándar para no incluir entrecruzamiento de formaldehído (-Form), no digestión (-dig), no ADN ligasa, no ATP (-lig). En el gel es un producto 3C (etapa -IP) y el 3C estándar con inmunoprecipitación (Expt). Se muestran los productos de PCR para la interacción de largo alcance en *GAL10*. Todas las demás interacciones mostraron dependencias similares (no mostradas). D. 3C con IP (Rbp1) en las cepas mostradas (BY4741 en el fondo) en los sitios indicados.

Figura 7. Interacciones condicionales de largo alcance en el locus GAL. Mapa de sitios de restricción DpnII (líneas verticales) que muestra el número, posición y orientación de los cebadores (flechas 3'OH) y la posición aproximada de los transcriptos sentido inducidos y el transcripto antisentido en el locus reprimido. B. Productos 3CIP (Rbp1) en los sitios indicados de la cromatina aislada de células BY4741 cultivadas en las fuentes de carbono indicadas. C. 3C IP (Rbp1) sobre GAL10-7 y FMP27 en la cepa WT (BY4741 en el fondo) cultivado en tres diferentes fuentes de carbono. D. Un modelo para las interacciones dinámicas de largo alcance en el locus reprimido. Se proponen los transcriptos antisentido de GAL7 (mostrados), GAL10 y otros transcriptos no codificantes en las regiones entre los genes, junto con las lisina desacetilasas

- Hda1 y Eaf3 para crear un entorno adecuado para las interacciones de largo alcance. En la glucosa, las interacciones a través de *GAL10*, *GAL7* y *GAL-7* pueden ser detectadas y se muestra un modelo para acomodarla. Dada su asociación con RNAPII y TFIIB, las interacciones de largo alcance parecen representar un estado en equilibrio pero reprimido sin transcripción activa en el locus. E. Un modelo para las interacciones condicionales de largo alcance en el locus *GAL*. Como en el locus reprimido, las interacciones de largo alcance parecen representar un estado no activo en equilibrio sobre los genes. Un aumento en la conmutación dinámica que implica interacciones de largo alcance en *GAL10* y *GAL7* cuando se
- 30 quita la represión de la glucosa (es decir, en rafinosa) daría lugar a la pérdida de la interacción de largo alcance sobre GAL-7. El inicio de la transcripción activa contribuiría a la conmutación dinámica entre los estados expresados y reprimidos. Se prevé que la producción de los transcriptos antisentido en el locus inducido juega dos papeles: para facilitar la producción del transcripto sentido y a través de Eaf3 y Hda1 para cebar la región para la represión y la formación de interacciones de largo alcance cuando no se expresa activamente el gen.
- 35

5

Figura 8. Sistema transcripcional que depende de CTCF *in vivo*. a, El gen reportero integrado de luciferasa contiene el sitio de enlazamiento N-Myc de CTCF de 90 pb (pN-MycLuc de tipo silvestre) o un mutante deficiente para el enlazamiento de CTCF (pN-MycLuc de tipo silvestre), en lugar del promotor ¹³. b, El mapa ampliado del constructo integrado con las posiciones de los cebadores indicados. Los cebadores para ensayos de ChIP en los sitios de N-Myc se muestran en rojo y se describieron previamente ¹³. Los cebadores en el extremo 5' se representan en color marrón, en el extremo 3' en azul; los

- 40 se describieron previamente ¹³. Los cebadores en el extremo 5' se representan en color marrón, en el extremo 3' en azul; los cebadores utilizados en los ensayos 3C y 4C se indican en verde. Estos cebadores se describen en Materiales y métodos en el Suplemento en línea. Los sitios de restricción relevantes *Taql* se muestran como triángulos de color azul pálido; sus posiciones están marcadas en relación con el sitio de inicio de la transcripción de luciferasa. c, La presencia del sitio de enlazamiento de CTCF de tipo silvestre es suficiente para conducir la expresión de luciferasa en pN-MycLuc de tipo
- 45 silvestre; el tratamiento con alfa-amanitina reduce drásticamente la actividad de la luciferasa. d, CTCF y Pol II se enlazan al sitio de N-Myc en PN-MycLuc de tipo silvestre, pero no en PN-MycLuc mutante, como se muestra por medio del ensayo de ChIP con los anticuerpos indicados. e, El tratamiento con alfa-amanitina conduce a la disociación de Pol II, pero no de CTCF del sitio de N-Myc, como se muestra por el ensayo de ChIP con los anticuerpos indicados. En los paneles "d" y "e", se muestran los cebadores utilizados para el análisis en rojo en la Figura 1b y descritos anteriormente ¹³.
- 50
- Figura 9. Ensayos de captura de conformación de la cromatina (3C) y ChIP. a, Los extremos 5' y 3' (flechas verdes en la Figura 1b) están yuxtapuestos en pN-MycLuc de tipo silvestre (carril 3), pero no en N-MycLuc mutante (carril 4) como se revela por el ensayo 3C. El carril 1 es el control positivo para la plantilla ligada, el carril 2 es el control negativo de ligación. Los cebadores utilizados para el análisis se describen en Materiales y métodos en el Suplemento en línea. b, Los sitios 5'
- (carril 3) y 3' (carril 5) de pN-MycLuc de tipo silvestre, pero no de pN-MycLuc mutante (carriles 4 y 6), están ocupados por Pol II, según lo revelado por el ensayo ChIP con el anticuerpo anti-Pol II. Carril 1 - entrada, carril 2 - suero preinmune; c, Los sitios 5' (carril 3) y 3' (carril 5) de pN-MycLuc de tipo silvestre, pero no de pN-MycLuc mutante (carriles 4 y 6), están ocupados por CTCF. Carril 1 - entrada, Carril 2 - suero preinmune. En los paneles de "b" y "c" los iniciadores específicos del extremo 5' se muestran en color marrón y los cebadores específicos del extremo 3' se muestran en color azul.
- 60

Figura 10. El ensayo 4C (combinación del ensayo ChIP y el ensayo 3C) con los anticuerpos anti-Pol II y CTCF. Los

cebadores utilizados para el análisis se muestran en color verde en la Figura 1b y se describen en Materiales y métodos en el Suplemento en línea. a, Pol II está presente en la yuxtaposición de los sitios 5' y 3' sitios del pN-MycLuc de tipo silvestre (carril 3), pero no del pN-MycLuc mutante (carril 4). El tratamiento con alfa-amanitina conduce a la desaparición de Pol II de las estructuras de cromatina de orden superior en pN-MycLuc de tipo silvestre (carril 5). b, CTCF está presente en la yuxtaposición de los sitios 5' y 3' del pN-MycLuc de tipo silvestre (carril 3), pero no del pN-MycLuc mutante (carril 4). CTCF aún está asociado con la estructura de cromatina de orden superior en pN-MycLuc de tipo silvestre después del tratamiento con alfa-amanitina (carril 5).

Figura 11. Un modelo que muestra cómo CTCF puede enlazar la transcripción con las estructuras de cromatina de orden superior. a, CTCF se enlaza a dos sitios en los extremos 5' y 3' del pN-MycLuc de tipo silvestre. b, La estructura de orden superior se establece entre los extremos 5' y 3' del PN-MycLuc de tipo silvestre a través de CTCF. c, Pol II se enlaza a CTCF e inicia la transcripción; d, Después de la inhibición de la transcripción y de la remoción de Pol II aún se puede detectar la estructura restante debido a su asociación con CTCF.

15 Ejemplo 1. Control epigenético del locus de GAL

Transcriptos antisentido condicionales en GAL10.

- La adición de glucosa a un cultivo de células que crecen en galactosa da como resultado la inhibición rápida de la transcripción. Como era de esperar, dentro de los 15 minutos después de la adición de glucosa, caen dramáticamente los niveles tanto del transcripto *GAL10* de 2,25 kb y del transcripto de fusión GAL10-7 más largo de 4,1 kb (Greger y Proudfoot, 1998, EMBO J. 17, 4771-4779; St. John y Davis, 1981, J Mol Biol 152, 285-315) (Fig. 1A). Se ha reportado un transcripto antisentido en el gen *GAL10* reprimido (Perocchi et al., 2007, Nucleic Acids Res 35, e128; Samanta et al., 2006, Proc Natl Acad Sci EE.UU. 103, 4192-4197; David et al., 2006, Proc Natl Acad Sci EE.UU. 103, 5320-5325; Miura et al., 2006, PNAS 103, 17846-17851). Los inventores han investigado si los transcriptos antisentido son una característica general del gen
- 25 103, 17846-17851). Los inventores han investigado si los transcriptos antisentido son una característica general del gen GAL10 y, si es así, cuando aparecen estos transcriptos.

En cultivos inducidos, se observaron tres transcriptos antisentido (Fig. 1B), dos de los cuales son similares en tamaño a los transcriptos sentido de 2,25 kb y de 4,1 kb en el locus de *GAL* inducido (Fig. 1A). El tercer transcripto antisentido detectable es más pequeño, de 1,5 kb aproximadamente. Las autorradiografías en la figura 1A y 1B se hibridaron con sondas de actividad específica similar y se expusieron durante períodos de tiempo similares lo que sugiere que en el gen *GAL10* activo, los transcriptos antisentido representan una proporción significativa del ARN total. Además, incluso el transcripto GAL10-7 endógeno de 4,1 kb, que se extiende sobre ambos genes (Fig. 1A, C) coincide con un transcripto antisentido equivalente (Fig. 1B, C).

35

40

5

Durante la represión de la glucosa, la abundancia de los transcriptos antisentido *GAL10* baja considerablemente (Fig. 1B). Como resultado, se muestran para mayor claridad dos exposiciones de la transferencia tipo Northern en la Figura 1B. Los datos en la Figura 1 demuestra que el tamaño de los transcriptos antisentido en el locus reprimido también cambia. Los estudios cinéticos indican que 15 minutos después de la adición de glucosa al cultivo, se observa un cambio de los transcriptos antisentido de 4,1 kb, 2,25 kb y 1,5 kb a tres transcritos, de aproximadamente 3,5 kb, 2,4 kb y 1,5 kb. Tanto el transcripto de 2,25 kb como el de 2,4 kb se pueden detectar 15 minutos después de la adición de glucosa. Por lo tanto, la abundancia y el tamaño de los transcriptos antisentido en *GAL10* refleja si se induce o se reprime el gen aumentando la interesante posibilidad de que los transcriptos antisentido puedan desempeñar papeles reguladores en el locus *GAL* en ambas condiciones.

45

Mapeo de los transcriptos alrededor de GAL10 utilizando RT-PCR.

- Se utilizó transcripción inversa (RT) con cebadores específicos de cadena acoplados a PCR para determinar la posición de los transcriptos sentido y antisentido alrededor de *GAL10* en células cultivadas en glucosa, rafinosa y galactosa (Fig. 2). La estrategia utilizada aquí empleó cebadores RT, específicos para la detección ya sea de transcriptos sentido (S) o antisentido (AS) y cebadores de PCR anidada para producir productos de aproximadamente 200 pb (Fig. 2A). En primer lugar se analizaron el transcrito sentido de *GAL10* de 2,25 kb y su equivalente antisentido. Estos son transcriptos relativamente abundantes susceptibles al mapeo por RT-PCR (Figura 2B). El transcripto sentido puede ser detectado con conjuntos de cebadores 6S a 8S situados en toda la región de codificación de *GAL10* sólo en las muestras preparadas a partir de células
- 55 cultivadas en galactosa (Fig. 2B, carril 2). Una señal de conjuntos de cebadores 6AS a 8AS, diseñados para transcripción inversa y amplificación de transcriptos antisentido, también es evidente sólo cuando las células se cultivan en galactosa (Fig. 2B, carril 5). Esto sugiere que los transcriptos sentido y antisentido transcripciones en *GAL10* inducido se extienden sobre las mismas regiones.
- 60 En glucosa y rafinosa, sin embargo, se detectaron señales con los conjuntos de cebadores específicos antisentido 7AS y 8AS pero no con 6AS o cualquiera de cebadores equivalentes para el transcripto sentido. Esto está de acuerdo con el

mapeo del microarreglo global, que muestra al transcripto que surge dentro de la región de codificación de *GAL10* aproximadamente 500 pb del extremo 3' en condiciones de represión (Perocchi et al., 2007, Nucleic Acids Res 35, e128; David et al., 2006, Proc Natl Acad Sci EE.UU. 103, 5320-5325). Esto sugiere que hay dos transcriptos antisentido distintos sobre *GAL10* correspondientes al estado inducido y reprimido. Si el transcripto antisentido de 2,4 kb en el locus reprimido comienza dentro de la región de codificación de *GAL10* es probable que se extienda en la región entre genes *GAL10-1*, confirmado con señales de conjuntos de cebadores 9 y 10 (Fig. 2C).

Las señales para conjuntos de cebadores 5S y 5AS, localizados sobre la región entre genes *GAL10-7*, son mucho más débiles lo que sugiere que la mayoría de los transcriptos sentido y antisentido terminan o inician, respectivamente, dentro de esta región. Se observó que conjuntos de cebadores sentido y antisentido específicos para la región *GAL7* también revelaron evidencia de transcriptos sentido y antisentido (Fig. 2C). En este caso, conjuntos de cebadores 3S y 3AS, que abarcan el extremo del ORF y la región 3', fueron capaces de detectar el transcripto sentido pero no el antisentido, lo que sugiere que el transcripto antisentido es promovido además secuencia arriba. Los datos indican que los transcriptos pueden compensarse entre sí. Se concluye que *GAL10* y *GAL7* producen ambos transcriptos sentido y antisentido cuando se inducen las células con galactosa y que los transcriptos son similares en tamaño.

Mapeo de los transcriptos antisentido alrededor de GAL10 usando transferencias tipo Northern.

5

- Se utilizaron transferencias tipo Northern con sondas específicas de hebra para identificar la posición aproximada de los transcriptos antisentido en células cultivadas en glucosa, rafinosa y galactosa alrededor de *GAL10* a partir de tres diferentes ambientes de cepas (Fig. 3). Las transferencias tipo Northern tienen la ventaja de ser capaces de correlacionar el tamaño de un transcripto con su posición en el locus *GAL*. Sondas 2 y 3 mostraron rutinariamente una pobre hibridación con señales de fondo altas haciendo la hibridación con transcriptos largos difíciles de discernir.
- En cultivos inducidos son evidentes altos niveles de ARN antisentido con *GAL7* y *GAL10* (sondas 2 y 4), confirmando los datos de RT-PCR en la Figura 2. El antisentido de 2,25 kb en *GAL10* inducido muestra una ligera hibridación con la sonda 3, pero ninguno con la sonda de 5 lo que sugiere que está confinado principalmente a la región ORF de *GAL10*. La sonda 6 hibrida con el transcripto sentido *GAL1* ya que *GAL1* es en la cadena de Watson. Los otros dos transcriptos sólo se producen en células inducidas. Estos dos surgen dentro de *GAL1* y se extienden en la cadena sentido a través del terminador de *GAL1* y en FUR4. Una transcripto es aproximadamente > 5 kb y el segundo es mucho mayor.

En cultivos reprimidos, el transcripto antisentido de 2,4 kb (*GAL10*) mostró una fuerte hibridación con una sonda corta en la región entre genes *GAL10-1* y con menos fuertemente con una sonda *GAL1*. No se observa hibridación con estas sondas con el ARN preparado a partir de células cultivadas en galactosa, de conformidad con el mapeo por PCR en la Fig. 2. Esto sugiere que el transcripto antisentido de *GAL10* en el locus reprimido surge aproximadamente 500 pb desde el extremo 3' del ORF de *GAL10* y se extiende sobre la región entre los genes *GAL10-1*. De una manera similar, el transcripto más largo de *GAL10* antisentido de 3,5 kb también muestra una fuerte hibridación con estas sondas lo que sugiere que también se extiende a través de la región entre genes *GAL10-1* y en *GAL1* (Fig. 3A).

- 40 Cuatro pequeños transcriptos no codificantes son evidentes en células cultivadas tanto en condiciones de represión como de inducción (Fig. 3A). Notablemente, estos transcriptos se presentan predominantemente en las regiones entre los genes en el locus. El primero es un transcripto de 10 kb detectable en la región entre los genes *GAL7-10* con una sonda en la orientación antisentido con respecto a los transcriptos *GAL10* y *GAL7* (Sonda 3). El segundo es un transcripto de aproximadamente 600 pb que hibrida con una sonda de orientación sentido al terminador *GAL1* (Sonda 7). El tercer transcripto es de aproximadamente 1,7 kb y se extiende antisentido desde la región 5' de *KAP104* y a través de la región del
- terminador *GAL7* (Sonda 1). El cuarto es aproximadamente de 1,5 kb e inicia dentro de *GAL10* y se extiende en la región entre los genes *GAL10-1* (Sonda 5).
- En este análisis, se preparó ARN a partir de tres diferentes ambientes de cepas. Las cepas BY4741 (carril 1) y W303-1a (carril 2) produjeron perfiles similares en las tres fuentes de carbono. En la cepa YMH147 (carril 3), sin embargo, la expresión del locus *GAL* parece estar no reprimida en rafinosa. En algunas regiones, el perfil del transcripto en esta cepa es diferente de aquel en glucosa o galactosa (comparar, por ejemplo, los carriles 1 y 3 en las tres fuentes de carbono en hibridados con sondas de 3, 6 y 7).
- 55 Un resumen de los datos de mapeo (de las Figuras 1-3) se muestra en la Fig. 3B. Hay tres puntos notables. El primero es el nivel relativamente alto de transcriptos antisentido para cada uno de los transcriptos sentido cuando se inducen las células con galactosa (*GAL10, GAL7 y GAL10-7*). Los transcriptos sentido y antisentido en el locus inducido parecen estar apareados y sinergísticos. El segundo es el cambio en el tamaño de los transcriptos antisentido en represión (en glucosa) acoplados a los cambios en los sitios predichos de iniciación y terminación en *GAL10*. En tercer lugar, está la presencia de los transcriptos sobre las regiones entre los genes (terminador GAL7; región entre los genes *GAL10-7*; región entre los genes *GAL10-1*; terminador GAL1) tanto en el locus reprimido como en el inducido.

Se requieren secuencias en la región 3' de GAL10 para la inducción del transcripto sentido.

Los datos sugieren que la posición de los transcriptos antisentido sobre *GAL10* refleja si el gen está reprimido o inducido. Estos transcriptos son probablemente iniciados en diferentes sitios (Fig. 4A). Los inventores diseñaron un experimento para identificar secuencias requeridas para la transcripción antisentido inducida (probablemente para surgir en o cerca de las secuencias de flanqueo 3' de *GAL10*) separando la región de flanqueo 3' del gen principal. Este constructo permitió analizar un promotor putativo y el examen de su efecto sobre la inducción del transcripto sentido. Se suprimieron 554 pb entre las bases 2453 y 3007 dejando 92 pb en el extremo 3' del ORF de *GAL10* y las secuencias de flanqueo 3' intactas. La región suprimida también fue reemplazada por un casete de expresión en la misma dirección que la expresión de *GAL10*. Se

- Suprimida también fue reemplazada por un casere de expresión en la misma dirección que la expresión de GAL10. Se preparó ARN a partir de células cultivadas en rafinosa y transferidas a galactosa durante 15 y 60 minutos, suficiente para fuerte inducción del transcripto sentido y la aparición del transcripto antisentido inducido de 2,3 kb en el tipo silvestre (Fig. 4B carril 5 y 7). La inserción en el extremo 3' de GAL10 es suficiente para comprometer la expresión tanto de los transcriptos antisentido GAL10 de 3,5 kb como de 2,4 kb (Fig. 4B, carriles 6 y 8). Es importante destacar que, se observaron niveles muy
- 15 bajos de transcriptos sentido en estas muestras (Fig. 4B, carriles 6 y 8). Por lo tanto, a pesar de que la región promotora de la transcripción sentido esté intacta, la producción de ARN sentido se ve comprometida en la cepa con la inserción en el extremo 3' del *GAL10*. Esto eleva la posibilidad de que se requiera el transcripto antisentido para la transcripción a partir del promotor *GAL10*. Se examinaron también los perfiles de inducción en *GAL7* y *GAL1* (Fig. 4C) en la cepa mutante. No se observó ningún efecto evidente en los niveles reprimidos en la agrupación *GAL*. Además, hubo un efecto pequeño de esta inserción en la inducción de los transcriptos en *GAL7* o *GAL1*.

Por lo tanto se requieren secuencias en el extremo 3' de *GAL10* para la producción del transcripto antisentido inducido, que es a su vez está implicado en la transcripción eficiente del transcripto sentido inducido. Esto sugiere la coordinación de eventos entre los extremos 5' y 3' del gen inducido.

25

30

Las secuencias en la región 3' de GAL10 no se requieren para la transcripción antisentido en el locus reprimido.

En el locus reprimido, es probable que los transcriptos antisentido inicien en una posición diferente dentro del ORF de *GAL10* (Fig. 4A). La inserción en el extremo 3' del *GAL10* nos permitió preguntarnos si se requieren secuencias similares para la transcripción antisentido en el locus reprimido como en el locus inducido (Fig. 4B). Sorprendentemente, se observó

- sólo una pequeña reducción en el tamaño de los dos transcriptos antisentido en el mutante en comparación con el WT. Los niveles y la diferencia de tamaño relativa entre el antisentido largo y el corto permanecieron muy similares a los de WT. Esto sugiere que las secuencias en las que normalmente inician estas transcripciones han sido suprimidas o interrumpidas por la inserción en el extremo 3' y que ambos transcriptos inician en un nuevo sitio, un poco más cerca de la región 5' de *GAL10*.
 35
- Una explicación son las secuencias en el casete de expresión insertado. Sin embargo, como estos nuevos transcriptos no están presentes en las condiciones inducidas, es poco probable que sean promovidos a partir de secuencias dentro del casete insertado. Por lo tanto, hay diferentes requisitos de secuencia para el transcripto antisentido inducido y el antisentido en el locus reprimido.
- 40

El enganche de Hda1 para GAL10 refleja los niveles de transcriptos antisentido pero no de H3K18ac.

Ya que hay significativamente más transcripto antisentido producido cuando *GAL10* está en el estado inducido en comparación con el estado reprimido, se llevaron a cabo experimentos para verificar si la asociación de Hda1 en *GAL10* refleja ya sea el nivel de ARN antisentido o si la expresión del gen es inducida o reprimida. La asociación de Hda1-myc con *GAL10* fue evaluada por ChIP usando cromatina preparada a partir de células inducidas o reprimidas (Fig. 5A). Se usó como control Eaf3, un componente tanto de la lisina desacetilasa Rpd3S lisina como de la lisina acetiltransferasa NuA4 (Allard et al., 1999, EMBO J 18, 5108-5119). Las cepas que carecen de Eaf3 muestran altos niveles de acetilación de histonas en el promotor y dentro de la región de codificación de genes en lugar del perfil normal en el que los niveles son elevados en el promotor y caen en la región codificante, lo que sugiere un efecto importante de la pérdida de Eaf3 en el complejo Rpd3S en particular (Reid et al., 2004, Mol Cell Biol 24, 757-764).

Hda1-myc se asocia tanto con la región 5' como con la región 3' del *GAL10* en células inducidas. La señal para Hda1-myc cae cerca de 5 veces en la cromatina reprimida. Esta diferencia no es debida a diferencias en los niveles de Hda1- myc en las células cultivadas en condiciones reprimidas o de inducción (datos no presentados). Estos datos son consistentes con la presencia de altos niveles de transcripto antisentido en células inducidas y altos niveles de Hda1 a través de *GAL10*. Eaf3 muestra una tendencia similar pero menos pronunciada que muestra niveles más bajos de la asociación, tanto en el promotor como en el extremo 3' del *GAL10* en represión (Fig. 5B).

60 Para identificar si existe una relación entre la asociación de Hda1 y acetilación de histonas en GAL10 H3K18ac reprimida e inducida, se examinó un sustrato conocido para Hda1. Los niveles de H3K18ac en el gen reprimido son bajos, similares a

los niveles en una cepa H3K18R, en consonancia con la desacetilación activa por Hda1. En células inducidas, los niveles de H3K18ac son significativamente mayores que en las células reprimidas, a pesar de los altos niveles de Hda1 de la cepa inducida (Fig. 5C). Esta diferencia está presente incluso cuando la pérdida de nucleosoma que acompaña a la transcripción activa en GAL10 es tenida en cuenta y la señal de H3K18ac se normaliza a niveles de H3. Esto sugiere que no hay una correlación directa entre la asociación de Hda1 y H3K18ac en GAL10. Si la actividad de Hda1 está relacionada con transcriptos antisentido como lo proponen Camblong et al., 2007 Cell 131, 706-717, entonces, los altos niveles de H3K18ac en el gen inducido, sobre todo en la región 5', pueden ser explicados por la acetilación de histonas que acompaña la activación transcripcional y el alargamiento del transcripto sentido. El equilibrio de acetilación y desacetilación se desplazaría hasta resultar en una ganancia neta de H3K18ac en el gen inducido.

10

5

Interacciones de la cromatina de largo alcance en el locus GAL.

Las interacciones de la cromatina de largo alcance, también conocidas como bucles del gen, han sido descritas en un número limitado de genes de levadura activos (Ansari y Hampsey, 2005, Genes Dev 19, 2969-2978; O'Sullivan et al., 2004, 15 Nat Genet 36, 1014-1018; Singh y Hampsey, 2007, Mol Cell 27, 806-816). Estas interacciones representan la yuxtaposición de las regiones 5' y 3' de los genes de levadura y se asocian con RNAPII, TFIIB y la maquinaria de terminación de la transcripción del CPF. Teniendo en cuenta los datos que muestran la transcripción antisentido activa en el locus GAL reprimido y la comunicación del extremo 3' al 5' en el locus GAL inducido, se llevó a cabo una investigación para determinar si las interacciones de largo alcance están implicadas en la regulación antisentido de la expresión de GAL. Se analizó el 20 locus GAL por la presencia de interacciones de largo alcance y, mediante la inclusión de una etapa de inmunoprecipitación, si estas interacciones están asociadas con RNAPII y TFIIB.

Las interacciones entre las regiones 5' y 3' de GAL10, la región 5' y 3' de GAL7 y la región 5' de GAL10 con la región 3' de GAL7 a partir de células cultivadas en glucosa, rafinosa y galactosa fueron controladas usando una técnica 3C modificada (conformación del cromosoma de captura) (Dekker, 2006, Nat Methods 3, 17-21; Dekker et al., 2002, Science 295, 1306-

- 25 1311; O'Sullivan et al., 2004, Nat Genet 36, 1014-1018) (Fig. 6A). Como control, se utilizó la interacción de largo alcance detectada previamente en FMP27, que se asocia con RNAPII y está presente en todas las tres condiciones de crecimiento (véase la Fig. 7C)
- 30 Interacciones de largo alcance en el locus GAL reprimido.

Las interacciones de largo alcance pueden ser detectadas sobre GAL10, GAL7 y entre la región 5' de GAL10 y la región 3' de GAL7 (GAL10-7) en células cultivadas en glucosa (Fig. 6B). Además, estas interacciones están asociadas con RNAPII ya que las interacciones 3C pueden ser inmunoprecipitadas con anticuerpos para Rbpl. Los controles para las interacciones de

- 35 GAL10 y GAL10-7 muestran que los productos de la PCR son específicos y dependen del entrelazamiento de formaldehído, la digestión de la cromatina con DpnII, la reacción de ligación y adición de plantilla con la reacción PCR (Fig. 6C). En algunas reacciones se observó más de un producto de la PCR. Estos productos de la PCR se aislaron de los geles v se analizaron las secuencias sobre las uniones de DpnII. Esto puso de manifiesto, por ejemplo, que el sitio de DpnII colindante con el conjunto cebador 6 en la región entre los genes GAL10-1 está parcialmente protegida en la preparación de la
- 40 cromatina entrecruzada, resultando en un producto de aproximadamente 550 pb además del producto de 260 pb esperado para la interacción de largo alcance sobre GAL10. Análisis similares en los tres sitios de interacción de largo alcance confirmaron que cada uno de los productos de la PCR observados depende del entrecruzamiento de formaldehído y es consistente con la interacción válida de largo alcance mostrando yuxtaposición de secuencias distantes.
- 45 El número limitado de interacciones de largo alcance descritas hasta la fecha, han sido observadas en genes transcritos activamente. Las interacciones descritas aquí se producen en un locus que está reprimido para la expresión de GAL aunque la presencia de transcriptos antisentido sugiere que el locus es transcripcionalmente activo. El siguiente paso fue examinar si el mecanismo que dirige la formación del bucle en loci reprimidos es similar al de los genes activos. Las interacciones de largo alcance se reducen en una cepa que porta el alelo sua7-1, que expresa una versión de TFIIB con una sustitución de
- 50 E62K (ácido glutámico 62 por lisina) (Singh y Hampsey, 2007, Mol Cell 27, 806-816). Esta mutación es defectuosa en las interacciones en la región 3' pero no en la región 5' de los genes activos. En condiciones reprimidas, las interacciones de largo alcance de GAL10, GAL7, GAL10-7 y el FMP27 de control todos mostraron dependencia de TFIIB funcional lo que sugiere que los bucles en loci reprimidos tienen los mismos requisitos que aquellos en genes activos (Fig. 6D). Además, las interacciones de largo alcance de GAL10 y GAL7, pero no las de GAL10-7, se pueden enriquecer en inmunoprecipitados a
- 55 partir de una cepa que expresa TFIIB-myc en comparación con una cepa sin etiquetar (Fig. 6B). Esto sugiere que TFIIB-myc está estrechamente asociada con las regiones de la cromatina que muestran interacciones de largo alcance en GAL10 y GAL7. La razón de la incapacidad para detectar TFIIB asociada con la interacción de largo alcance a través de GAL10 y GAL7 no está clara, ya que se requiere TFIIB funcional para la interacción. Una posibilidad es que la etiqueta del epítopo no es accesible en esta interacción. Por lo tanto, las interacciones de largo alcance pueden ser detectadas en el locus GAL
- 60 reprimido y estas interacciones muestran los mismos requisitos que los bucles de genes en los genes activos.

Una interacción de largo alcance condicional en el locus GAL.

Las interacciones de largo alcance para *GAL10* y *GAL7* se observan también en la rafinosa, un medio de represión del crecimiento, y en galactosa cuando se expresan los genes (Figs. 7B). En ambos casos, la interacción se asocia con RNAPII. La interacción de largo alcance entre la región 5' de *GAL10* y le región 3' de *GAL7* (*GAL10-7*), sin embargo, depende de la fuente de carbono en el medio de cultivo (Fig. 7B, C). Más importante aún, esta interacción de largo alcance se pierde en condiciones inducidas (galactosa) e incluso cuando se remueve la represión de glucosa (en rafinosa). Por lo tanto, es poco probable que la interrupción de la interacción de largo alcance se pueda explicar simplemente por los altos niveles de transcripción sentido que se observan en condiciones de inducción. Más bien, parece que la condicionalidad de la interacción de largo alcance de *GAL10-7* está relacionada con la pérdida de la represión de glucosa.

Las interacciones de largo alcance en el locus GAL están influenciadas por lisina desacetilasas.

Los transcriptos no codificantes y las interacciones de largo alcance están presentes en el locus *GAL* tanto en condiciones de represión como de inducción. Se investigó si la presencia de transcriptos no codificantes está relacionada con las interacciones de largo alcance. Los transcriptos antisentido están asociados con las lisina desacetilasas Hda1 (Fig. 5) (Camblong et al., 2007 Cell 131, 706-717). En el locus *GAL10* reprimido e inducido, el nivel de cromatina asociada con Hda1 se correlaciona con el nivel de transcripto antisentido. Se preguntó si se requieren Hda1 y Eaf3 para las interacciones de largo alcance en el locus *GAL* y en *FMP27*.

20

La pérdida de Hda1 tiene un efecto dramático en las interacciones de largo alcance sobre *GAL10*, *GAL7* y *GAL10-7* en células reprimidas (Fig. 6B). El efecto sobre *FMP27* es mucho menos dramático. Por el contrario, se requiere Eaf3 para la interacción de largo alcance en *FMP27* y *GAL10*, pero es un requisito menor para este factor en *GAL7* o *GAL10-7*. Algunos loci, tales como *FMP27* y *GAL7* muestran especificidad por uno o el otro complejo, mientras que loci tales como *GAL10* parecen requerir ambas actividades. Estos datos implican a las lisina desacetilasas, Hda1 y Rpd3S, directa o

25 parecen requerir ambas actividades. Estos datos implican a las lisina desacetilasas, Hda1 y Rpd3S, directa o indirectamente, en la formación de interacciones de largo alcance y sugieren que esto puede estar relacionado con los transcriptos antisentido.

Discusión

Los inventores han utilizado el locus *GAL* como un sistema modelo en el cual comparar los estados inducido y reprimido y observar las diferencias en el transcripto antisentido y la regulación epigenética. Ellos muestran que la represión de genes es un estado proactivo que implica la producción regulada de transcriptos antisentido en asocio con los cambios epigenéticos para el locus.

35

30

El mapa del transcripto no codificante en el locus *GAL* es complejo y condicional. La presencia de transcriptos relativamente cortos sobre las regiones entre los genes es una reminiscencia de transcriptos asociados al promotor asociados con genes de levadura tales como *SER3* (Martens et al., 2004, Trends Genet 23, 126-133), *IMD2*, *LEU4* (Davis y Ares, 2006, Proc Natl Acad Sci EE.UU. 103, 3262-3267) y Ty1 (Berretta et al., 2008 Genes Dev 22, 615-626) o los genes de mamíferos tales como *DHFR* (Martianov et al., 2007, Nature 445, 666-670). Ya que estos están presentes tanto en el locus *GAL* inducido como reprimido con porte probable que tangente provincian estante en el locus *GAL* inducido como representes como probable que tangente provincian estante estante están presentes tanto en el locus *GAL* inducido estante estant

- 40 como *DHFR* (Martianov et al., 2007, Nature 445, 666-670). Ya que estos están presentes tanto en el locus *GAL* inducido como reprimido es poco probable que tengan funciones reguladoras relacionadas con activación y represión directamente, pero pueden influir en otros aspectos de la topología del locus.
- El tamaño y la posición de los transcriptos antisentido no codificantes en *GAL10* cambian con las condiciones de crecimiento. En el locus inducido, hay un transcripto antisentido abundante cuyos niveles suben y bajan de forma sinergísticamente con el transcripto sentido. Los datos sugieren que se requieren secuencias en la región de flanqueo 3' de *GAL10* para promover la expresión de este transcripto. Además de los sitios de enlazamiento para el regulador Gal4 (esta región también contiene el promotor para *GAL7*) también hay sitios de enlazamiento para Reb1 en esta región. Ya que este es un transcripto condicional, es posible un doble papel para Gal4 en la activación tanto del transcripto de *GAL7* y como del
- 50 antisentido de GAL10. Sin embargo, los inventores también han mapeado un transcripto antisentido de alto nivel que surge en la región 3' de GAL7 inducida y no existen sitios de enlazamiento con Gal4 en esta región. Dado que la pérdida del transcripto antisentido de GAL10 está asociada con niveles bajos del transcripto sentido de GAL10, y la demostración de las interacciones de largo alcance entre la región 3' y 5' de GAL10 y GAL7, una posibilidad alternativa es que las secuencias promotoras juegan un papel en la activación *in trans* de los transcriptos antisentido.
- 55

Los transcriptos antisentido en *GAL10* inducido comparten propiedades con dos genes que previamente demostraron estar asociados con ARN antisentido. Al igual que *PHO5* (Uhler et al., 2007, Proc Natl Acad Sci EE.UU. 104, 8011-8016), el transcrito antisentido en *GAL10* inducido está enlazado a la transcripción del transcripto sentido. El transcripto de *GAL10* es diferente de aquel en *PHO5*, ya que no parece extenderse en la región del promotor y por lo tanto es poco probable que funcione en la forma propuesta para *PHO5* mediante la remodelación de la cromatina del promotor. Al igual que *PHO84* (Camblong et al., 2007 Cell 131, 706-717), la presencia del transcripto antisentido se correlaciona con la asociación de la

lisina desacetilasa Hda1 (KDAC) con la cromatina. Aunque las lisina desacetilasas están asociadas tanto con la activación como con la represión de la expresión génica (Bernstein et al., 2000, Proc Natl Acad Sci EE.UU. 91, 13708-13713), en funciones PHO84 Hda1 con el transcripto antisentido para reprimir al promotor de sentido (Camblong et al., 2007 Cell 131, 706-717). Paradójicamente, los altos niveles de H3K18ac se mantienen sobre el gen activo *GAL10* lo que sugiere que la

- 5 acetilación asociada con la transcripción sentido desplaza el equilibrio dinámico hacia la acetilación. Subyacente a esto está un estado fundamental de represión activa a través de los KDAC y los transcriptos antisentido. Los inventores sugieren que este estado fundamental está representado en parte por las interacciones de largo alcance sobre *GAL10* y *GAL7* (Fig. 7E). En este modelo las interacciones de largo alcance, aunque asociadas con RNAPII y TFIIB, representarían una transcripción no activa en equilibrio según lo previsto anteriormente (Ansari y Hampsey, 2005, Genes Dev 19, 2969-2978; O'Sullivan et
- 10 al., 2004, Nat Genet 36, 1014- 1018; Singh y Hampsey, 2007, Mol Cell 27, 806-816). En la levadura, moscas y mamíferos existen muchos genes en un estado estable en equilibrio con RNAPII acoplado y las modificaciones epigenéticas asociadas en el promotor (Guenther et al., 2007, Cell 130, 77-88; Radonjic et al., 2005, Mol Cell 18, 171-18; Zeitlinger et al., 2007, Nat Genet 39, 1512-1516).
- 15 En el locus *GAL* inducido, la transcripción antisentido se extiende sobre la misma región general que la transcripción sentido. En la represión hay un interruptor en el sitio de iniciación y un cambio en la naturaleza de las secuencias requeridas para promover la transcripción antisentido. Además, el transcripto antisentido se vuelve dominante. Las exposiciones largas de transferencias tipo Northern, sin embargo, revelan niveles muy bajos de transcriptos sentido de tamaño equivalente lo que sugiere que la relación entre la transcripción sentido y antisentido se mantiene incluso en un gen reprimido. Esto
- 20 refuerza la idea de la represión "activa" y apoya el estado reprimido como una variación de los eventos que ocurren en la activación. Una consecuencia natural de los transcriptos antisentido dominantes es la represión activa a través de los KDAC como Hda1 y Eaf3. Es interesante que en el locus reprimido los transcriptos antisentido se extienden sobre la región entre los genes *GAL10-1* de una manera similar al transcripto antisentido regulado por el exosoma en *PHO84* (Camblong et al., 2007, Cell 131, 706-717769). La desacetilación sobre esta región y otra región entre los genes puede promover o estabilizar
- 25 las interacciones de largo alcance en el locus, por ejemplo la interacción sobre GAL10-7.

También destacados en el mapa están los transcriptos largos que se extienden desde un gen al otro. Estos transcriptos son condicionales, por ejemplo el transcripto largo *GAL1:FUR4* en el locus inducido o el transcripto antisentido *GAL10-1* en el locus reprimido. Los transcriptos no codificantes largos se observan en el locus de β-globina en células de mamífero y pueden estar involucrados en interruptores condicionales (Gribnau et al., 2000, Mol Cell 5, 377-386). En la levadura, estos

- 30 pueden estar involucrados en interruptores condicionales (Gribnau et al., 2000, Mol Cell 5, 377-386). En la levadura, estos transcriptos pueden simplemente reflejar un pobre procesamiento del transcripto. Alternativamente, pueden ser la primera indicación de dos tipos diferentes de eventos de transcripción (represivo o de activación) que dirigen o rompen órdenes superiores de organización de genes de levadura.
- 35 Procedimientos experimentales

Cepas

- Se usaron tres ambientes de cepas diferentes en este estudio: BY4741 (MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0), W303-1a (MATa, ura3-52, leu2-3-112, his3-11, ade2-1, can1-100, trp1Δ2) y YMH14 (MATα, cyc1-5000, cyc7-67, ura3-52, leu2-3-112, cyh2). El alelo sua7-1 está en el ambiente de YMH14 (Pinto et al., 1994, J Biol Chem 269, 30569-30573). Las cepas, incluyendo los derivados etiquetados con epítopo, truncamientos y supresiones de genes se construyeron mediante el reemplazo de genes en un solo paso utilizando fragmentos de ADN generados por PCR (Longtine et al., 1998, Yeast 14, 953-961). Se insertó un pTEF: KanMX:TEFter en la región 3' de *GAL10* resultando en la pérdida de los residuos 2453 y 3007 con respecto a la ATG. La transcripción del marcador seleccionable es en la misma dirección que la transcripción sentido de *GAL10*. Se etiquetó SUA7 en el terminal C con el epítopo myc en BY4741. Se construyeron supresiones de hda1Δ y eaf3Δ en el ambiente de BY4741.
 - Medios y condiciones de cultivo.

50

60

Se prepararon los medios de cultivo utilizando métodos estándar en YE suplementado con 2% de glucosa, rafinosa o galactosa según se requiera. Se tomaron las levaduras de placas frescas y se cultivaron hasta una DO_{600} de 0,6 a 0,8 en 50 a 100 ml. Se recolectó la levadura por centrifugación y se lavó en H₂O antes de la transferencia a un medio fresco.

55 Protocolo de inmunoprecipitación de la cromatina

Se llevó a cabo la inmunoprecipitación de la cromatina como se describió (Meluh y Broach, 1999, Nature 445, 666-670; Morillon et al., 2005, Mol Cell 18, 723-734). En resumen, se elaboró ChIP utilizando 50 ml de cultivos fijados con formaldehído al 1% durante 15 minutos seguido por la adición de glicina hasta una concentración final de 0,25 mM. Se rompieron las células de levadura usando perlas de vidrio en un MagnaLyser (Roche) y se cortó la cromatina fijada mediante sonicación utilizando un Bioruptor (Diagenode). Las longitudes promedio de los fragmentos de ADN fueron de 150 a 300 pb.

Después de la centrifugación (30 min, 10K, 4 °C), se incubó la cromatina soluble con el anticuerpo con los siguientes epítopos; 5 µl de H3 (Abcam), 5 µl de H3K18ac (Upstate), 20 µl de Eaf3 (Abcam), 10 µl de Y80 (Santa Cruz) y 10 µl de myc (Sigma) en 1,5 ml de tubos Eppendorf recubiertos de silicona a 4 °C durante 15 a 20 horas y se inmunoprecipitaron con proteína A - Sefarosa durante 90 minutos a temperatura ambiente. Después del lavado, se eluyó la cromatina de las perlas a 65 °C durante 30 minutos. Se revirtieron los enlaces cruzados por incubación a 65 °C durante 6 a 20 horas y se trataron con proteína y APNaga A. Se avuíficó en 100 vil de cruza Se

- proteasa y ARNasa A. Se purificó el ADN utilizando mini columnas de PCR Qiagen y se eluyó en 100 µl de agua. Se utilizaron muestras de IP y controles, por ejemplo, sin anticuerpo, sin etiqueta, puros, mientras que los ADN de control (entrada) se diluyeron en forma correspondiente. Se sometieron las muestras a PCR en tiempo real utilizando una mezcla Corbett Rotorgene y Sybr Green (Sensymix, Quantace). Se utilizó PCR en tiempo real para amplificar las regiones correspondientes a aquellas mostradas en *GAL10*. Se calcularon los datos (IP sin anticuerpo) / TOT y se expresaron como un porcentaje de la entrada. Las barras de error reflejan la desviación estándar de la señal promedia entrada entra
- 10 correspondientes a aquellas mostradas en GAL10. Se calcularon los datos (IP sin anticuerpo) / TOT y se expresaron como un porcentaje de la entrada. Las barras de error reflejan la desviación estándar de la señal promedio obtenida entre diferentes experimentos (n = 2 a 4).

Generación de sondas específicas de la cadena

Se incorporó un promotor T7 en el extremo de la región específica del ADN utilizando PCR. Se uso la ARN polimerasa T7 para generar sondas monocatenarias con especificidad para la cadena sentido o antisentido de ADN utilizando ³²P αUTP y el kit Ambion MAXIscript® (Cat # AM1308-AM1326).

20 Transferencia tipo Northern

Se separaron 15 µg de ARN total, preparado a partir de células utilizando fenol caliente:cloroformo y perlas de vidrio, en geles de formaldehído al 1,1% y se hizo transferencia a membranas de nailon Magna y se hizo la cocción a 80 °C durante 2 horas y después se hibridó durante la noche en PerfectHyb Plus (Sigma) a 64 °C, se lavaron dos veces en 1XSSC / SDS al 0.1% durante 20 minutes aceda lavada. Se avaguairem triagmenta las membranas

- 25 0,1%, dos veces en 0,2XSSC/SDS al 0,1% durante 20 minutos cada lavado. Se expusieron típicamente las membranas durante 24 horas a menos que se indique lo contrario. Los niveles de ARN total cargado se controlaron mediante la especie ARNr, que son muestras iguales a menos que se indique otra cosa.
 - PCR de transcripción inversa.

Para cada una de las diez posiciones a través de *GAL10-7* en la figura 2, hay un cebador de RT para el transcripto sentido (con respecto a la dirección de la expresión de *GAL10*), un cebador de RT para el transcripto antisentido, dos cebadores alternativos para la primera ronda de PCR y dos cebadores anidados alternativos para la amplificación. Los detalles de los cebadores se dan en la tabla complementaria 1. Se efectuó la síntesis de la primera hebra utilizando el kit de ADNc Verso de ABgene con cebadores específicos para los transcriptos sentido o antisentido. Se preparó una reacción estándar con 2

- 35 de ABgene con cebadores específicos para los transcriptos sentido o antisentido. Se preparó una reacción estándar con 2 µL de ARN; cebador (25 µM) 1 µl; H₂O 9 µl, calentando a 70 °C durante 5 min para eliminar cualquier estructura secundaria y colocando en hielo inmediatamente. Luego se añadieron 4 µl de amortiguador, 2 ml de dNTP, 1 µl de RT y 1ml de reforzador de RT y se incubó la reacción a 52 °C durante 50 min, a continuación, durante 2 min a 95 °C. En el control de RT, no se añadió transcriptasa inversa a las reacciones. La amplificación por PCR fue una reacción anidada utilizando ADN
- 40 polimerasa TakaRa en las siguientes reacciones: 2X amortiguador I, 25 μl; dNTP (2,5 mM), 8 μl; ADNc, 2 μl; cebadores (25 μM), hacia adelante 1 μl, inverso 1 μl; Taq polimerasa TakaRa LA, 0,5 μl; H₂O 12, 5 μl a 94 °C 5 min, 24 ciclos de 94 °C 1 min, 55 °C 45 segundos, 72 °C 30 segundos y 72 °C 5 min. Para la segunda ronda se utilizaron 2 μl del producto de la primera ronda con los cebadores anidados en una reacción como la anterior: 94 °C 5 min, los 18 ciclos a 94 °C 45 segundos, 54 °C 30 segundos, 72 °C 20 segundos y finalmente una incubación a 72 °C 5 min.
- 45

5

15

30

Conformación del cromosoma de captura con inmunoprecipitación (3CIP)

Se extrajeron los núcleos de 100 ml de un cultivo de Saccharomyces cerevisiae que se desarrolló en un medio apropiado hasta una densidad óptica A600 = 0,2. Se añadió formaldehído al 1% (2,44 ml del 41%) y se agitó durante 10 minutos. Se inactivó el formaldehído mediante la adición de glicina a 0,125 M (5 ml de 2,5 M). Se lavó el sedimento de células dos veces en amortiguador de Mg/K (K₂HPO₄ / KH₂PO₄ 0,1 M (relación 35:65), MgCl₂ 5 mM, pH 6,5) y se resuspendió en amortiguador de formación de esferoplastos (sorbitol 1,2 M, 500 U de enzima lítica de levadura y DTT 25 mM en amortiguador de Mg/K) durante 15 minutos a temperatura ambiente. Se lavaron una vez los esferoplastos en amortiguador MES (MES 0,1 M, sorbitol 1,2 M, EDTA 1 mM, MgCl₂ 0,5 mM pH 6,4). Se lisaron los esferoplastos usando 10 golpes con un homogenizador de mano. Se formó una capa del lisado sobre un gradiente de sacarosa (5 ml de sacarosa 1,8 M, 10 ml de sacarosa 1,1 M en lisis MES en un tubo Corex) y se separó por centrifugación durante 10 min a 10.000 RPM en un rotor Beckman 14.17. El sedimento de núcleos se ubicó en la interfaz en la pared de vidrio. El sedimento de núcleos se ubicó en la interfaz en la pared de vidrio.

rotor Beckman JA-17. El sedimento de núcleos se ubicó en la interfaz en la pared de vidrio. El sedimento en la parte inferior del tubo se remueve mediante un lavado con agua y se desecha. Se lavó el sedimento de núcleos del vidrio con amortiguador CSK (NaCl 100 mM, sacarosa 300 mM, PIPES 10 mM, MgCl₂ 3 mM, EGTA 1 mM, Triton X-100 al 0,5%, leupeptina 10 M, 1: 1000 AEBSF) a 4 °C, se lavo de nuevo y se resuspendió en 1 ml de amortiguador CSK y se dejó durante

20 minutos en hielo. Se sedimentaron los núcleos y se removió todo menos 100 pl del sobrenadante. Se añadieron 40 µl de NaCl 5 M y se incubó durante 10 min sobre hielo. Se diluyó la mezcla viscosa con 1,2 ml de H₂O. Se añadió anticuerpo y se puso en rotación la mezcla a 4 ºC durante la noche. Se prepararon ~ 40 µl de suspensión de proteína G - Sefarosa (20-30 µl de perlas) mediante lavado dos veces en H2O y una vez con 1 ml de amortiguador de lavado de restricción (Tris-HCI 50 mM 5 (pH = 8,1), NaCl 100 mM, MgCl₂ 10 mM) y centrifugación a 2000 rpm durante 3 min para recoger las perlas. Se incubó la mezcla de cromatina con rotación durante 60 minutos y se recogieron las perlas por centrifugación a 1000 rpm durante 3 min, se lavó 3 veces con 1 ml de amortiguador de lavado de restricción mediante rotación a 4 ºC durante 5 min y centrifugación a 2000 rpm durante 3 min. Se usaron 10 µl de 10X amortiguador 3, 50 U de enzima de restricción y agua hasta 100 µl y se digirió la cromatina durante la noche a 37 °C durante la noche con DpnII C, o 25 °C con CviQI, luego a 65 10 °C durante 10 min para matar la enzima de restricción. Se removieron las enzimas insensibles al calor tales como CviQI mediante el lavado de las perlas dos veces con amortiguador de lavado de restricción. Se diluyó la mezcla y se ligó con 410 µl de H2O, 60 µl de 10X amortiguador de ligación, 30 µl de ligasa T4 (12000 U) y se incubó a 16 °C durante 4 horas. Se incubó la mezcla durante la noche a 65 ºC para el desentrecruzamiento. Se añadió 1 µl de 1 mg/ml de ARNasa A y se incubó a 37 °C durante 30 min seguido por 60 µl de 20 mg/ml de proteinasa K y se incubó a 42 °C durante una hora. Se 15 extrajo el ADN utilizando 660 µl de fenol:cloroformo:alcohol isoamílico y se precipitó con 30 µl de NaCl 5 M, 0,5 µl de 10 mg/ml de glucógeno y 1 ml de etanol frío y se incubó a -80 °C durante una hora. Se recogió el sedimento, se lavó y se resuspendió en 20 µl de H₂O. Se incluyeron los siguientes controles en el protocolo según lo recomendado (Dekker, 2006, Nat Methods 3, 17-21): Se excluyó el paso de inmunoprecipitación para hacer un procedimiento 3C directo. Tanto para el 3C y el 3C con IP, se incluyeron tratamientos con ARNasa o proteinasa K antes de la etapa de ligación para demostrar la 20 dependencia con el ARN o la proteína. Se llevó a cabo el protocolo en núcleos aislados sin la etapa de tratamiento con formaldehído. Los pasos de inmunoprecipitación se realizaron después de la etapa de restricción y ligación utilizando un protocolo estándar ChIP (véase arriba). Se detectaron los productos de la reacción utilizando PCR anidada usando polimerasa TakaRa en 50 µl de reacción. Los patrones de cebadores fueron de 25 µM. La primera reacción contenía 25 µl de amortiguador GC I, 8 µl de solución dNTP (1,25 mM cada uno), 1 ml de plantilla, 1 ml de cada cebador, 13,5 µl de H₂O, 25 0.5 ul de ADN polimerasa TakaRa (5U/ul) durante 25 ciclos (94 °C 5 min - 194 °C - 45 s, 60 °C 30 s, 72 °C 20 s] - 72 °C 5 min). La segunda reacción contenía 25 µl de amortiguador GC I, 8 µl de solución dNTP (1,25 mM cada uno), 2 µl de plantilla de la primera reacción, 1 µl de cada cebador, 12,5 µl de H2O, 0,5 µl de ADN polimerasa TakaRa (5U/µl) durante 18 ciclos (94 °C 5min - [94 °C - 45 s, 61 °C 30 s, 72 °C 20 s] - 72 °C 5 min). Como hay cuatro productos posibles para cada interacción de largo alcance, se incluyeron todas las combinaciones de cebadores, excepto hacia adelante:hacia adelante, en el análisis 30 inicial. Sólo se muestran los datos con una combinación de cebadores. La orientación del cebador (hacia adelante F; inverso R) se da con respecto a la dirección del ORF. Cada conjunto se anida con un cebador interior (i) y exterior (o) (ver la Tabla 2). Se controlaron las reacciones PCR mediante la omisión de la plantilla o la ADN polimerasa. Las plantillas de control para la eficiencia del cebador se prepararon mediante ligación de ADN genómico restringido por DpnII. Se titularon las plantillas de control y las plantillas experimentales para determinar el intervalo lineal de la amplificación; únicamente se muestra un 35 producto equivalente en este intervalo para cada muestra.

Tabla 1			
Los cebadores para el análisis por RT-PCR			
Ceba	dores específicos o	del gen sentido para RT	
		5' 3'	
1.	S1KAP104	CTGTAAAAGAGTTGC	
2.	INF1GAL7	CTGCAACATCCAAT	
3.	INF2GAL7	AAGGACCACTCTTAC	
4.	S1GAL7	CATGTGAAACCAAC	
5.	INFIGAL10	CTACTTTACCAAACG	
6.	S1GAL10	CAAGGTTACACAATC	
7.	S2GAL10	CTTCACCAGCAGTC	
8.	S3GAL10	GCAAGATAGCAAAC	
9.	S4GAL10	TTAGCTCTACCACAG	
10.	INF1GAL1	TGGTTATGAAGAGG	
Ceba	dores específicos o	del gen antisentido para RT	
1.	INR1GAL7	CAAGGCTCATTGTC	
2.	INR2GAL7	ACGGAGTGACAATA	
3.	AS2GAL7	CTTGGTTGGTTTTG	
4.	AS1GAL7	TGGTGCTTAGAATC	
5.	INR1GAL10	CTACAGATTTTCCTG	
6.	AS1GAL10	AGGTGATCTATTGGT	
7.	AS2GAL10	GCAAGATTTGTGAC	
8.	AS3GAL10	GTTTTGGTTACAGG	

9.	INR1GAL1	
10.	S1GAL1	

GTGAAGACGAGGAC CAATCACTTCTTCTG

Primera ronda de cebadores para PCR

RKAP1041	CGATGGAAATCCTGCACCTA
FGAL71	CGTAATAAACTTCAACAGAGCCTAAA
RGAL71	TTCTAGTTATGTAAGAGTGGTCCTTTC
FGAL73	TTGTCACTCCGTTCAAGTCG
RGAL73	GCCTCAAAGAGATTTAACTTCG
FGAL75	ACCAGTCGCATTCAAAGGAG
RGAL75	TGAAGTTTCGCAAGAATTGAAA
FGAL101	GCGCTTCGCAATAGTTGT
RGAL101	TTGCCAGCTTACTATCCTTCTTG
FGAL103	CATCAATGTATCTACCAGGCTCA
RGAL103	AAATTGACTGCTGGTGAAGC
FG10AS2	ATTTTGAATGATGGGTCCC
RG10AS2	AGATTTCAAGCCACGTTTGC
FGAL105	TGGCGTATTTCGTATGACCA
RGAL105	TGTTGCTGATAACCTGTCGAA
FGAL107	TGGATGGACGCAAAGAAGTT
RGAL107	GCTCGGCGGCTTCTAATC
FGAL11	CGAATCAAATTAACAACCATAGGA
RGAL11	AATACAAACTGAAAATGTTGAAAGT
R anidados	
R anidados FKAP1042	AAAACCAAAGACTGCGGAAT
R anidados FKAP1042 RKAP1042	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTTGCTT
R anidados FKAP1042 RKAP1042 FGAL72	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTTGCTT TGTCAATAAAGTGGAAATGTGTCA
R anidados FKAP1042 RKAP1042 FGAL72 RGAL72	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTTGCTT TGTCAATAAAGTGGAAATGTGTCA GAATTTTAGGAATACAATGCAGCTT
R anidados FKAP1042 RKAP1042 FGAL72 RGAL72 FGAL74	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTTGCTT TGTCAATAAAGTGGAAATGTGTCA GAATTTTAGGAATACAATGCAGCTT GAAACCAGGCAGTTAATAGAAAAA
R anidados FKAP1042 RKAP1042 FGAL72 RGAL72 FGAL74 RGAL74	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTTGCTT TGTCAATAAAGTGGAAATGTGTCA GAATTTTAGGAATACAATGCAGCTT GAAACCAGGCAGTTAATAGAAAA GCTGCTGAAAAACTAAGAAA
R anidados FKAP1042 RKAP1042 FGAL72 RGAL72 FGAL74 RGAL74 FGAL76	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTTGCTT TGTCAATAAAGTGGAAATGTGTCA GAATTTTAGGAATACAATGCAGCTT GAAACCAGGCAGTTAATAGAAAA GCTGCTGAAAAACTAAGAAA TTAAAATCGAGGCGAGG
R anidados FKAP1042 RKAP1042 FGAL72 RGAL72 FGAL74 RGAL74 FGAL76 RGAL76	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTTGCTT TGTCAATAAAGTGGAAATGTGTCA GAATTTTAGGAATACAATGCAGCTT GAAACCAGGCAGTTAATAGAAAA GCTGCTGAAAAACTAAGAAA TTAAAATCGAGGCGAGG
R anidados FKAP1042 RKAP1042 FGAL72 RGAL72 FGAL74 RGAL74 FGAL76 RGAL76 FGAL102	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTTGCTT TGTCAATAAAGTGGAAATGTGTCA GAATTTTAGGAATACAATGCAGCTT GAAACCAGGCAGTTAATAGAAAAA GCTGCTGAAAAACTAAGAAA TTAAAATCGAGGCGAGG
R anidados FKAP1042 RKAP1042 FGAL72 RGAL72 FGAL74 RGAL74 FGAL76 RGAL76 FGAL102 RGAL102	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTTGCTT TGTCAATAAAGTGGAAATGTGTCA GAATTTTAGGAATACAATGCAGCTT GAAACCAGGCAGTTAATAGAAAAA GCTGCTGAAAAACTAAGAAA TTAAAATCGAGGCGAGG
R anidados FKAP1042 RKAP1042 FGAL72 RGAL72 FGAL74 RGAL74 FGAL76 RGAL76 FGAL102 RGAL102 FGAL104	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTGCTT TGTCAATAAAGTGGAAATGTGTCA GAATTTTAGGAATACAATGCAGCTT GAAACCAGGCAGTTAATAGAAAAA GCTGCTGAAAAACTAAGAAA TTAAAATCGAGGCGAGG
R anidados FKAP1042 RKAP1042 FGAL72 RGAL72 FGAL74 RGAL74 FGAL76 RGAL76 FGAL102 RGAL102 FGAL104 RGAL104	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTGCTT TGTCAATAAAGTGGAAATGTGTCA GAATTTTAGGAATACAATGCAGCTT GAAACCAGGCAGTTAATAGAAAAA GCTGCTGAAAAACTAAGAAA TTAAAATCGAGGCGAGG
R anidados FKAP1042 RKAP1042 FGAL72 RGAL72 FGAL74 RGAL74 FGAL76 RGAL76 FGAL102 RGAL102 FGAL104 RGAL104 FG10AS1	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTGCTT TGTCAATAAAGTGGAAATGTGTCA GAATTTTAGGAATACAATGCAGCTT GAAACCAGGCAGTTAATAGAAAA GCTGCTGAAAAACTAAGAAA TTAAAATCGAGGCGAGG
R anidados FKAP1042 RKAP1042 FGAL72 RGAL72 FGAL74 RGAL74 FGAL76 RGAL76 FGAL102 RGAL102 FGAL104 RGAL104 FG10AS1 RG10AS1	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTGCTT TGTCAATAAAGTGGAAATGTGTCA GAATTTTAGGAATACAATGCAGCTT GAAACCAGGCAGTTAATAGAAAAA GCTGCTGAAAAACTAAGAAA TTAAAATCGAGGCGAGG
R anidados FKAP1042 RKAP1042 FGAL72 FGAL72 FGAL74 RGAL74 FGAL76 FGAL102 RGAL102 FGAL104 RGAL104 FG10AS1 FGAL106	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTGCTT TGTCAATAAAGTGGAAATGTGTCA GAATTTTAGGAATACAATGCAGCTT GAAACCAGGCAGTTAATAGAAAAA GCTGCTGAAAAACTAAGAAA TTAAAATCGAGGCGAGG
R anidados FKAP1042 RKAP1042 FGAL72 FGAL72 FGAL74 RGAL74 FGAL76 RGAL76 FGAL102 RGAL102 FGAL104 RGAL104 FG10AS1 FGAL106 RGAL106 RGAL106	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTGCTT TGTCAATAAAGTGGAAATGTGTCA GAATTTTAGGAATACAATGCAGCTT GAAACCAGGCAGTTAATAGAAAAA GCTGCTGAAAAACTAAGAAAA TTAAAATCGAGGCGAGG
R anidados FKAP1042 RKAP1042 FGAL72 FGAL72 FGAL74 RGAL74 FGAL76 FGAL102 RGAL102 FGAL104 RGAL104 FG10AS1 FGAL106 RGAL106 FGAL108	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTTGCTT TGTCAATAAAGTGGAAATGTGTCA GAATTTTAGGAATACAATGCAGCTT GAAACCAGGCAGTTAATAGAAAAA GCTGCTGAAAAACTAAGAAA TTAAAATCGAGGCGAGG
R anidados FKAP1042 RKAP1042 FGAL72 FGAL72 FGAL74 RGAL74 FGAL76 FGAL102 RGAL102 FGAL104 RGAL104 FG10AS1 FGAL106 RGAL106 FGAL108 RGAL108 RGAL108	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTGCTT TGTCAATAAAGTGGAAATGTGTCA GAATTTTAGGAATACAATGCAGCTT GAAACCAGGCAGTTAATAGAAAAA GCTGCTGAAAAACTAAGAAA TTAAAATCGAGGCGAGG
R anidados FKAP1042 RKAP1042 FGAL72 RGAL72 RGAL74 RGAL74 RGAL76 RGAL76 FGAL102 RGAL102 RGAL104 RGAL104 RGAL104 FG10AS1 RG10AS1 FGAL106 RGAL106 RGAL108 RGAL108 RGAL12	AAAACCAAAGACTGCGGAAT TCCGGGTTATAGAGTTTGCTT TGTCAATAAAGTGGAAATGTGTCA GAATTTTAGGAATACAATGCAGCTT GAAACCAGGCAGTTAATAGAAAAA GCTGCTGAAAAACTAAGAAA TTAAAATCGAGGCGAGG
	RGAL71 RGAL73 FGAL73 FGAL75 FGAL101 RGAL101 FGAL103 RGAL103 FG10AS2 FGAL105 FGAL107 FGAL107 FGAL107 FGAL107 FGAL11

Tabla 2

l abla 2
Secuencia (5' - 3')
GCTCATTGTCGGTGTCGTTA
CGATGGAAATCCTGCACCTA
TTGTCACTCCGTTCAAGTCG
TCCGAAGTTAAATCTCTTTGAGG
TTGCTTTGCCTCTCCTTTTG
CGTTTGGTAAAGTAGAGGGGGTA
CGCACCATAATCTCCGTACC
CGCTTCACCAGCAGTCAAT
GGGCCTACTAATCCGTATGGT
TCCCAGAAGAATGTCCCTTAG
GAGGAAAAATTGGCAGTAACCT

6 GAL 10 promotor F(i) 7 FMP27 promotor F(o)	GCCCCACAAACCTTCAAAT ATCAAAGCCACGCCAAAC
	(continuación)
Designación del cebador	Secuencia (5' - 3')
7 FMP27 promotor F(i)	CCTACACGCAAAGGAACTAGAGA
8 FMP27 terminador F(o)	AGCAAACCGAACATCAAACC

- 5 DpnII: Los pares cebadores 4+6, 5+3 y 3+6 detectarán las interacciones de largo alcance a través de *GAL10*. Los pares cebadores 2+4, 3+1 y 4+1 detectarán las interacciones a través de *GAL7*. Los pares cebadores 6+2 y 5+1 y 1+6 detectarán las interacciones a través de *GAL10-7*. Para los análisis mostrados en la Figura 5 sólo se muestra la última de las combinaciones de cebadores para cada interacción.
- 10 La Tabla 3 muestra las posiciones cromosómicas potenciales a través del locus Gal donde pueden producirse las interacciones de largo alcance. Para cada región del cromosoma, se diseña un conjunto de cebadores hacia adelante e inversos. La interacción de largo alcance en el locus *GAL* se controla mediante análisis 3C entre los cebadores diseñados para cada región del cromosoma. Por ejemplo, para controlar la interacción entre las regiones de Gal 7 y de Gal 10, se utilizarán los cebadores de la fila 3 (274081-87) y de la fila 5 (278016-19). Si se van a controlar las interacciones en otras regiones, se utilizarán otras combinaciones de cebadores.

Tabla 3

Organismo	Número del cromosoma	Posición de arranque del cebador	Posición de parada del cebador	Hebra
Levadura	II	273036	273043	+
Levadura	II	273126	273130	+
Levadura	II	274081	274087	+
Levadura	II	274838	274852	+
Levadura	II	278017	278019	+
Levadura	II	278022	278022	+
Levadura	II	278025	278026	+
Levadura	II	279321	279331	+
Levadura	II	279952	279962	+
Levadura	II	281268	281284	+
Levadura	II	279941	279959	-
Levadura	II	279595	279595	-
Levadura	I	274080	274084	-
Levadura	II	273684	273694	-

- 20 Este sistema es igualmente aplicable a cualquier otro locus cromosómico donde se cree que se produzcan las interacciones cromosómicas de largo alcance. Una vez que se identifica la región, se pueden diseñar los cebadores para identificar la presencia o ausencia de una interacción especificada de largo alcance que indica una condición fisiológica particular.
 - Ejemplo II Control de las interacciones de largo alcance por CTCF
- 25

30

Se ha demostrado que varios factores de transcripción juegan un papel tanto en las interacciones de ADN de largo alcance como en la transcripción y, por tanto, pueden ser buenos candidatos para proporcionar un enlace entre estos procesos. Los ejemplos incluyen el factor de transcripción básico, TFIIB que se demostró que organizan la formación de bucles de varios genes en la levadura Mol Cell, 27, 806-16, y factores de transcripción EKLF, GATA-1 y FOG-1 responsables por las interacciones de ADN de largo alcance en el gen para la β-globina (Drissen et al., 2004, Genes Dev, 18, 2485-90; Vakoc, et al., 2005, Mol Cell, 17, 453-62).

Los inventores han identificado proteínas de enlazamiento de CCCTC (CTCF) como otro candidato para llevar a cabo estas funciones de todo el genoma. CTCF está implicado tanto en la regulación transcripcional como la formación de estructuras cromosómicas intra e inter conformacionales de alto orden (Klenova, 2002, Semin Cancer Biol, 12, 399-414; Kurukuti et al., 2006, Proc Natl Acad Sci EE.UU., 103, 10684-9; Zhao et al., 2006, Nat Genet, 38, 1341-7; Splinter et al., 2006, Genes Dev, 20, 2349-54). Se estima que existen > 15000 de los sitios de enlazamiento de CTCF en el genoma (Kim et al., 2007, Cell, 128, 1231-1245), sin embargo, es probable que el número real de tales sitios sea > 30.000 (Vetchinova et al., 2006, Anal Biochem, 354, 85-93). En la transcripción, CTCF puede actuar como un factor de transcripción clásico; un informe reciente

II) (Chernukhin et al., 2007, Mol Cell Biol, 27, 1631-1648). La colocalización de CTCF - Pol II en los sitios de inicio de la transcripción (TSS) de genes activos en todo el genoma refuerza aún más esta posibilidad (Birney et al., 2007, Nature, 447, 799-816). Por último, CTCF puede formar dímeros que pueden ser importantes para la organización de los bucles de ADN (Pant et al., 2004, Mol Cell Biol, 24, 3497-504).

5

10

15

Las propiedades únicas de CTCF han llevado a los inventores a investigar si se puede vincular mecánicamente la formación de estructuras cromosómicas de orden superior y la transcripción, posiblemente a través de su asociación con Pol II. Para investigar esta transcripción mínima in vivo se han utilizado sistemas de células basados en de dos líneas de células NIH3T3 genéticamente modificadas. Estas líneas portan vectores de expresión integrados en forma estable que contienen el sitio de enlazamiento de CTCF y su variante mutada deficiente para el enlazamiento de CTCF, fusionada al gen reportero de luciferasa sin promotor (pN-MycLuc de tipo silvestre y pN-MycLuc mutante, Figuras 8a, b).

El sitio único de tipo silvestre, pero no su variante mutante, fue suficiente para conducir la expresión del gen reportero de la luciferasa (Figura 8c) (Chernukhin et al., 2007, Mol Cell Biol, 27, 1631-1648). Ambos, CTCF y Pol II, estuvieron presentes en el tipo silvestre, pero no el sitio de enlazamiento de CTCF mutante (Fig. 8d), apoyando así el modelo anterior sobre que CTCF ayuda a enganchar Pol II en ausencia del promotor endógeno.

En este sistema, los procesos de transcripción pueden estar vinculados con la formación de estructuras de ADN de orden superior, en particular, entre las regiones 5' y 3' de pN-MycLuc de tipo silvestre y pN-MycLuc de tipo silvestre del ADN integrado. Las estructuras conformacionales de orden superior pueden ser controladas mediante el ensayo de captura de conformación cromosómica (3C), que detecta la cercana proximidad de los sitios distantes en el ADN cromosómico *in vivo*. Los inventores han aplicado el análisis 3C a los loci de pN-MycLuc de tipo silvestre y pN-MycLuc mutante integrados. Se identificaron dos sitios en la posición 5' y en la posición 3' (Figura 8b) como yuxtapuestos en el ensayo 3C en pN-MycLuc de tipo silvestre, pero no en pN-MycLuc mutante (Figura 9a).

25

30

35

Con base en el trabajo anterior sobre la interacción de CTCF y la colocalización con Pol II, los inventores plantearon la hipótesis de que CTCF y Pol II pueden estar vinculados a la formación de estructuras de orden superior en el gen de pN-MycLuc de tipo silvestre transcrito. Para investigar esto, se ha trabajado para identificar si ambos factores están presentes en los sitios yuxtapuestos recientemente identificados (Fig. 9a). El ensayo de inmunoprecipitación de cromatina (ChIP) reveló que, efectivamente, tanto CTCF como Pol II están presentes en los sitios 5' y 3', identificados como yuxtapuestos, en PN-MycLuc de tipo silvestre, pero no en pN-MycLuc mutante (Figura 9b, c).

El N-Myc es un sitio objetivo conocido de CTCF y se investigaron previamente los rasgos característicos de las secuencias dentro de N-Myc implicado en el enlazamiento de CTCF (Chernukhin et al., 2007, Mol Cell Biol, 27, 1631-1648; Lutz et al., 2003, EMBO J., 22, 1579-1587). La alta frecuencia de ocurrencia de sitios de enlazamiento de CTCF en los genomas nos permiten formular la hipótesis de que puede haber otro sitio potencial objetivo de CTCF en el extremo 3' del pN-MycLuc de tipo silvestre integrado en el gen de la luciferasa (Fig. 9c). Como es difícil predecir el enlazamiento de CTCF de la secuencia, se utilizó el ensayo de cambio de la movilidad electroforética (EMSA) y el análisis de la huella para investigar si

- 40 CTCF puede enlazar directamente a las secuencias identificadas. De hecho, se demostró el enlazamiento por EMSA con la sonda marcada que contenía las secuencias 3' mediante CTCF recombinante. El análisis de la huella de esta secuencia reveló la región protegida de la digestión con ADNasa I por CTCF recombinante. Por lo tanto, en conjunto con ChIP, EMSA y los datos de huellas confirmaron que CTCF de hecho se puede enlazar con el sitio 3'. Por lo tanto, es concebible que dos moléculas de CTCF enlazadas a los dos sitios pueden estar implicadas en la formación de la yuxtaposición.
- 45 Para confirmar aún más la participación de CTCF y Pol II en el establecimiento de estructuras de orden superior, se llevó a cabo el ensayo 4C (ensayos ChIP, ya sea con anticuerpos anti-Pol II o anti-CTCF seguidos por el 3C). Los análisis 4C de Pol II 4C y 4C de CTCF demostró la presencia de Pol II y CTCF en los sitios 5' y 3' durante la yuxtaposición en pN-MycLuc de tipo silvestre, pero no en pN-MycLuc mutante (Fig. 10a, b). A partir de estos experimentos se concluyó que el establecimiento de las yuxtaposiciones entre las regiones 5' y 3' del constructo activo de tipo silvestre, pN-MycLuc de tipo silvestre, estaba asociado con la presencia de CTCF y Pol II en los sitios identificados y también con el estado de tipo silvestre del constructo.

Se emprendió el análisis adicional de la dependencia del fenómeno observado en la transcripción. Para este propósito se trataron las células con el inhibidor de la transcripción, alfa-amanitina. El tratamiento abolió la actividad de pN-MycLuc de tipo silvestre y pN-MycLuc mutante (Fig. 8c) y condujo a la desaparición de Pol II, pero no de CTCF del sitio N-myc (Fig. 8e). El ensayo 4C de Pol II realizado con el anticuerpo anti-Pol II en las células tratadas no reveló la presencia de Pol II en la yuxtaposición entre los sitios 5' y 3'. Sin embargo se detectó esta yuxtaposición con el anticuerpo anti-CTCF (Fig. 10a, b). El análisis 3C de las células tratadas con alfa-amanitina confirmó la existencia de la estructura (Fig. 9a).

60 Los inventores también analizaron si el CTCF recombinante mezclado in vitro con el ADN del plásmido linealizado desnudo, pN-MycLuc de tipo silvestre, podría ser suficiente para formar yuxtaposiciones detectadas *in vivo*. Utilizando un ensayo 4C

(ensayo ChIP en combinación con 3C), se detectaron las estructuras similares a las estructuras formadas *in vivo* en este sistema básico. Significativamente se observaron las señales más débiles con el constructo pN-MycLuc mutante utilizado como control indicando así que las estructuras detectadas eran dependientes de dos sitios intactos de CTCF. La presencia del enlazamiento de CTCF en el pN-MycLuc de tipo silvestre, fue confirmado por el ensayo ChIP para ambos sitios 5' y 3'.

5

Tomados en conjunto los datos sugieren que los procesos transcripcionales requieren la formación de estructuras de orden superior, sin embargo las estructuras de orden superior, en el caso reportado que dependen de CTCF, existen sin la transcripción en curso. Estos hallazgos respaldan las observaciones de que las interacciones de ADN de largo alcance en el gen de β-globina se mantienen después de la inhibición de la transcripción (Palstra et al., 2008, PLoS ONE 3, e1661).

10

15

Los inventores proponen un modelo, en el que el establecimiento de la estructura de orden superior entre los extremos 5' y 3' de la pN-MycLuc de tipo silvestre es dependiente de CTCF (Fig. 11). En este modelo, la interacción entre dos moléculas de CTCF posicionadas en dos sitios distantes conduce a la formación del dímero CTCF y el establecimiento del bucle de ADN (Klenova et al., 2005, Cell Cycle, 4, 96-101). Los procesos transcripcionales pueden iniciarse después del establecimiento de esta configuración. Después de la inhibición de la transcripción y la remoción de Pol II, aún se puede detectar la estructura yuxtapuesta muy probablemente debido a su asociación con CTCF (Fig. 11).

Este ejemplo hace uso de un sistema de transcripción muy simple. En este sistema, la transcripción del constructo de luciferasa sin promotor fue conducida por CTCF interactuando con Pol II a través del sitio de enlazamiento de CTCF, N-Myc (Chernukhin et al., 2007, Mol Cell Biol, 27, 1631-1648). Se descubrió que el proceso de transcripción se basa en la yuxtaposición entre el 5' N-Myc y el extremo 3' del gen de la luciferasa, el segundo sitio de enlazamiento de CTCF se identificó dentro del extremo 3' yuxtapuesto. Se concluye que numerosas interacciones transitorias tienen lugar de forma continua entre las moléculas de CTCF enlazadas a ADN en cis y trans. La estabilización de tales asociaciones cromosómicas cuasi estables de orden superior puede ser un proceso regulado; se puede implicar el uso de poli ADP-ribosilación de CTCF en dicha regulación (Klenova et al., 2005, Cell Cycle, 4, 96-101; Yu et al., 2004, Nat Genet, 36, 1105-1110). Uno de los resultados de la formación de estructuras de orden superior puede ser el inicio de un proceso

transcripcional.
 El uso de un sistema de transcripción mínimo identificó que CTCF está involucrado en el establecimiento y mantenimiento de las estructuras de cromatina de orden superior, que a su vez son necesarias para la transcripción en curso por la ARN

Listado de secuencias

polimerasa II.

35 <110> Oxford BioDynamics Limited

<120> Método de Diagnóstico

<130> P037256WO

<150> GB 0810051.3 <151> 2008-06-02

- <160>73
- 45

40

<170> PatentIn versión 3.5

<210> 1 <211> 15

- 50 <212> ADN <213> Secuencia artificial
 - <220>

<223> Cebador

55

<400> 1 ctgtaaaaga gttgc 15

<210> 2

60 <211>14

<212> ADN

```
<213> Secuencia artificial
      <220>
      <223> Cebador
 5
      <400>2
                     14
      ctgcaacatc caat
      <210> 3
10
      <211>15
      <212> ADN
      <213> Secuencia artificial
      <220>
15
      <223> Cebador
      <400>3
      aaggaccact cttac
                      15
20
      <210>4
      <211>14
      <212> ADN
      <213> Secuencia artificial
25
      <220>
      <223> Cebador
      <400>4
      catgtgaaac caac
                     14
30
      <210>5
      <211> 15
      <212> ADN
      <213> Secuencia artificial
35
      <220>
      <223> Cebador
      <400> 5
40
      ctactttacc aaacg
                       15
      <210> 6
      <211> 15
      <212> ADN
45
      <213> Secuencia artificial
      <220>
      <223> Cebador
50
      <400>6
      caaggttaca caatc
                      15
      <210>7
      <211>14
55
      <212> ADN
      <213> Secuencia artificial
      <220>
      <223> Cebador
60
      <400>7
```

cttcaccagc agtc 14 <210> 8 <211>14 5 <212> ADN <213> Secuencia artificial <220> <223> Cebador 10 <400> 8 14 gcaagatagc aaac <210> 9 15 <211> 15 <212> ADN <213> Secuencia artificial <220> 20 <223> Cebador <400> 9 ttagctctac cacag 15 25 <210>10 <211>14 <212> ADN <213> Secuencia artificial 30 <220> <223> Cebador <400> 10 14 tggttatgaa gagg 35 <210> 11 <211>14 <212> ADN <213> Secuencia artificial 40 <220> <223> Cebador <400> 11 45 caaggctcat tgtc 14 <210> 12 <211> 14 <212> ADN 50 <213> Secuencia artificial <220> <223> Cebador 55 <400> 12 14 acggagtgac aata <210> 13 <211>14 60 <212> ADN <213> Secuencia artificial

<220> <223> Cebador <400>13 5 14 cttggttggt tttg <210> 14 <211>14 <212> ADN 10 <213> Secuencia artificial <220> <223> Cebador 15 <400>14 tggtgcttag aatc 14 <210> 15 <211> 15 20 <212> ADN <213> Secuencia artificial <220> <223> Cebador 25 <400> 15 15 ctacagattt tcctg <210> 16 30 <211> 15 <212> ADN <213> Secuencia artificial <220> 35 <223> Cebador <400> 16 15 aggtgatcta ttggt 40 <210> 17 <211> 14 <212> ADN <213> Secuencia artificial 45 <220> <223> Cebador <400>17 14 gcaagatttg tgac 50 <210> 18 <211>14 <212> ADN <213> Secuencia artificial 55 <220> <223> Cebador <400> 18 60 14 gttttggtta cagg

<210> 19 <211>14 <212> ADN <213> Secuencia artificial 5 <220> <223> Cebador <400> 19 10 14 gtgaagacga ggac <210> 20 <211>15 <212> ADN 15 <213> Secuencia artificial <220> <223> Cebador 20 <400> 20 15 caatcacttc ttctg <210> 21 <211> 19 25 <212> ADN <213> Secuencia artificial <220> <223> Cebador 30 <400>21 atagtcttcg gcgggcttc 19 <210> 22 35 <211>20 <212> ADN <213> Secuencia artificial <220> 40 <223> Cebador <400>22 cgatggaaat cctgcaccta 20 45 <210> 23 <211>26 <212> ADN <213> Secuencia artificial 50 <220> <223> Cebador <400> 23 cgtaataaac ttcaacagag cctaaa 55 <210> 24 <211>27 <212> ADN <213> Secuencia artificial 60 <220>

	<223> Cebador	
5	<400> 24 ttctagttat gtaagagtgg tcctttc	
5	<210> 25 <211> 20 <212> ADN <213> Secuencia artificial	
10	<220> <223> Cebador	
15	<400> 25 ttgtcactcc gttcaagtcg 20	
20	<210> 26 <211> 22 <212> ADN <213> Secuencia artificial	
	<220> <223> Cebador	
25	<400> 26 gcctcaaaga gatttaactt cg 22	
30	<210> 27 <211> 20 <212> ADN <213> Secuencia artificial	
25	<220> <223> Cebador	
55	<400> 27 accagtcgca ttcaaaggag 20	
40	<210> 28 <211> 22 <212> ADN <213> Secuencia artificial	
45	<220> <223> Cebador	
	<400> 28 tgaagtttcg caagaattga aa 22	
50	<210> 29 <211> 18 <212> ADN <213> Secuencia artificial	
55	<220> <223> Cebador	
60	<400> 29 gcgcttcgca atagttgt	18
00	<210> 30	

<211>23 <212> ADN <213> Secuencia artificial <220> 5 <223> Cebador <400> 30 ttgccagctt actatccttc ttg 23 10 <210> 31 <211>23 <212> ADN <213> Secuencia artificial 15 <220> <223> Cebador <400> 31 catcaatgta tctaccaggc tca 23 20 <210> 32 <211>20 <212> ADN <213> Secuencia artificial 25 <220> <223> Cebador <400> 32 30 aaattgactg ctggtgaagc 20 <210> 33 <211>19 <212> ADN 35 <213> Secuencia artificial <220> <223> Cebador 40 <400> 33 attttgaatg atgggtccc 19 <210> 34 <211>20 45 <212> ADN <213> Secuencia artificial <220> <223> Cebador 50 <400> 34 agatttcaag ccacgtttgc 20 <210> 35 55 <211>20 <212> ADN <213> Secuencia artificial <220> 60 <223> Cebador

	<400> 35 tggcgtattt cgtatgacca	20	
5	<210> 36 <211> 21 <212> ADN <213> Secuencia artificial		
10	<220> <223> Cebador		
	<400> 36 tgttgctgat aacctgtcga a	21	
15	<210> 37 <211> 20 <212> ADN <213> Secuencia artificial		
20	<220> <223> Cebador		
25	<400> 37 tggatggacg caaagaagtt	20	
23	<210> 38 <211> 18 <212> ADN <213> Secuencia artificial		
30	<220> <223> Cebador		
35	<400> 38 gctcggcggc ttctaatc	18	
40	<210> 39 <211> 24 <212> ADN <213> Secuencia artificial		
	<220> <223> Cebador		
45	<400> 39 cgaatcaaat taacaaccat ag	ga	24
50	<210> 40 <211> 25 <212> ADN <213> Secuencia artificial		
55	<220> <223> Cebador		
55	<400> 40 aatacaaact gaaaatgttg aa	agt	25
60	<210> 41 <211> 20 <212> ADN		

	<213> Secuencia artificial		
5	<220> <223> Cebador		
	<400> 41 aaaaccaaag actgcggaat	20	
10	<210> 42 <211> 22 <212> ADN <213> Secuencia artificial		
15	<220> <223> Cebador		
	<400> 42 tccgggttat agagttttgc tt	22	
20	<210> 43 <211> 24 <212> ADN <213> Secuencia artificial		
25	<220> <223> Cebador		
30	<400> 43 tgtcaataaa gtggaaatgt gtca		24
35	<210> 44 <211> 25 <212> ADN <213> Secuencia artificial		
00	<220> <223> Cebador		
40	<400> 44 gaattttagg aatacaatgc agctt		25
45	<210> 45 <211> 24 <212> ADN <213> Secuencia artificial		
	<220> <223> Cebador		
50	<400> 45 gaaaccaggc agttaataga aaaa		24
55	<210> 46 <211> 20 <212> ADN <213> Secuencia artificial		
60	<220> <223> Cebador		
50	<400> 46		

	gctgctgaaa aactaagaaa	20
5	<210> 47 <211> 19 <212> ADN <213> Secuencia artificial	
10	<220> <223> Cebador	
	<400> 47 ttaaaatcga ggcgaggtc	19
15	<210> 48 <211> 20 <212> ADN <213> Secuencia artificial	
20	<220> <223> Cebador	
	<400> 48 tgatttgttt gccgattacg	20
25	<210> 49 <211> 20 <212> ADN <213> Secuencia artificial	
30	<220> <223> Cebador	
35	<400> 49 gcggctcgtg ctatattctt	20
	<210> 50 <211> 23 <212> ADN <213> Secuencia artificial	
40	<220> <223> Cebador	
45	<400> 50 ttgctgtata acgaatttta tgc	23
50	<210> 51 <211> 20 <212> ADN <213> Secuencia artificial	
	<220> <223> Cebador	
55	<400> 51 ccagcagaca agaaatcacc	20
60	<210> 52 <211> 21 <212> ADN <213> Secuencia artificial	

5	<220> <223> Cebador <400> 52 ttgagggtac ggagattatg g	21
10	<210> 53 <211> 19 <212> ADN <213> Secuencia artificial	
	<220> <223> Cebador	
15	<400> 53 attaacgccg ttattaacg	19
20	<210> 54 <211> 21 <212> ADN <213> Secuencia artificial	
25	<220> <223> Cebador	
	<400> 54 ttcttggcta tgaaaatgag g	21
30	<210> 55 <211> 20 <212> ADN <213> Secuencia artificial	
35	<220> <223> Cebador	
	<400> 55 gggaatctcg tagcatcacc	20
40	<210> 56 <211> 20 <212> ADN <213> Secuencia artificial	
45	<220> <223> Cebador	
50	<400> 56 tgtgtgaccg aaaaggtctg	20
50	<210> 57 <211> 21 <212> ADN <213> Secuencia artificial	
55	<220> <223> Cebador	
60	<400> 57 tgttgtggaa atgtaaagag c	21

<210> 58 <211>22 <212> ADN <213> Secuencia artificial 5 <220> <223> Cebador <400> 58 10 22 gcaatgagca gttaagcgta tt <210> 59 <211>24 <212> ADN 15 <213> Secuencia artificial <220> <223> Cebador 20 <400> 59 tttttagcct tatttctggg gtaa 24 <210> 50 <211>21 25 <212> ADN <213> Secuencia artificial <220> <223> Cebador 30 <400> 60 aagtggttat gcagcttttc c 21 <210> 61 35 <211>20 <212> ADN <213> Secuencia artificial <220> 40 <223> Cebador <400> 61 gctcattgtc ggtgtcgtta 20 45 <210> 62 <211>23 <212> ADN <213> Secuencia artificial 50 <220> <223> Cebador <400> 62 tccgaagtta aatctctttg agg 23 55 <210>63 <211>20 <212> ADN <213> Secuencia artificial 60 <220>

	<223> Cebador		
5	<400> 63 ttgctttgcc tctccttttg		20
5	<210> 64 <211> 23 <212> ADN		
10	<213> Secuencia artificiai <220> <223> Cebador		
15	<400> 64 cgtttggtaa agtagagggg gta		23
20	<210> 65 <211> 20 <212> ADN <213> Secuencia artificial		
	<220> <223> Cebador		
25	<400> 65 cgcaccataa tctccgtacc	20	
30	<210> 66 <211> 19 <212> ADN <213> Secuencia artificial		
35	<220> <223> Cebador		
	<400> 66 cgcttcacca gcagtcaat	19	
40	<210> 67 <211> 21 <212> ADN <213> Secuencia artificial		
45	<220> <223> Cebador		
	<400> 67 gggcctacta atccgtatgg t	21	
50	<210> 68 <211> 21 <212> ADN <213> Secuencia artificial		
55	<220> <223> Cebador		
60	<400> 68 tcccagaaga atgtccctta g	21	
	<210> 69		

5	<211> 22 <212> ADN <213> Secuencia artificial <220> <223> Cebador	
	<400> 69 gaggaaaaat tggcagtaac ct	22
10	<210> 70 <211> 19 <212> ADN <213> Secuencia artificial	
15	<220> <223> Cebador	
20	<400> 70 gccccacaaa ccttcaaat	19
20	<210> 71 <211> 18 <212> ADN <213> Secuencia artificial	
25	<220> <223> Cebador	
30	<400> 71 atcaaagcca cgccaaac	18
35	<210> 72 <211> 23 <212> ADN <213> Secuencia artificial	
	<220> <223> Cebador	
40	<400> 72 cctacacgca aaggaactag aga	
45	<210> 73 <211> 20 <212> ADN <213> Secuencia artificial	
50	<220> <223> Cebador	
20	<400> 73 agcaaaccga acatcaaacc	20

Reivindicaciones

 Un método de control de los cambios epigenéticos en las interacciones cromosómicas condicionales de largo alcance en al menos un locus cromosómico donde las diferentes conformaciones de las interacciones de largo alcance están asociadas con una condición asociada con un cambio en el nivel de expresión de uno o más genes o un cambio en uno o más productos génicos en donde la condición se selecciona de entre cáncer, trastornos cardiovasculares, condiciones inflamatorias, incluyendo trastornos autoinmunes y respuestas inflamatorias a enfermedades infecciosas y trastornos genéticos heredados modulados por mecanismos epigenéticos, comprendiendo dicho método las etapas de:

- 10 (i) entrecruzamiento *in vitro* de dichas interacciones cromosómicas de largo alcance presentes en al menos un locus cromosómico;
 - (ii) aislamiento el ADN entrecruzado de dicho locus cromosómico;
 - (iii) sometimiento de dicho ADN entrecruzado a una digestión de restricción con una enzima que corta al menos una vez dentro de al menos un locus cromosómico;
- 15 (iv) ligar dichos extremos del ADN escindido entrecruzado para formar bucles de ADN;
 - (v) la identificación de la presencia de bucles de ADN;

en donde la presencia de bucles de ADN indica la presencia de una interacción cromosómica específica de largo alcance y en donde las interacciones cromosómicas de largo alcance no son interacciones de largo alcance entre genes y sus 20 elementos reguladores.

2. El método de acuerdo con la reivindicación 1, en donde la presencia de los bucles de ADN se identifican utilizando técnicas de PCR.

25 3. El método de acuerdo con la reivindicación 1 o la reivindicación 2, en el que la presencia de un bucle de ADN indica un estado de transcripción alterado indicativo de una condición fisiológica específica.







Figura 2



Figura 3



Figura 4



Figura 5

Figura 6











Figura 9

Figura 10



3

