

**UNIVERSIDAD DE LA FRONTERA
FACULTAD DE CIENCIAS AGROPECUARIAS Y FORESTALES**



**POSICIONAMIENTO NUCLEOSOMAL DEL GEN *Egr1* Y
MODIFICACIONES EPIGENÉTICAS QUE REGULAN SU
EXPRESIÓN**

Tesis presentada a la Facultad de Ciencias Agropecuarias y Forestales de la Universidad de La Frontera. Como parte de los requisitos para optar al título de Biotecnólogo.

ÁNGELA LETICIA RIFFO CAMPOS

TEMUCO - CHILE

2012

**UNIVERSIDAD DE LA FRONTERA
FACULTAD DE CIENCIAS AGROPECUARIAS Y FORESTALES**



**POSICIONAMIENTO NUCLEOSOMAL DEL GEN *Egr1* Y
MODIFICACIONES EPIGENÉTICAS QUE REGULAN SU
EXPRESIÓN**

Tesis presentada a la Facultad de Ciencias Agropecuarias y Forestales de la Universidad de La Frontera. Como parte de los requisitos para optar al título de Biotecnólogo.

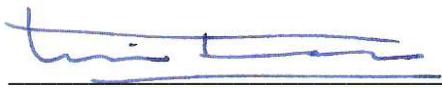
**ÁNGELA LETICIA RIFFO CAMPOS
PROFESOR GUIA: LUIS FRANCO VERA**

TEMUCO - CHILE

2012

**POSICIONAMIENTO NUCLEOSOMAL DEL GEN *Egr1* Y MODIFICACIONES
EPIGENÉTICAS QUE REGULAN SU EXPRESIÓN**

PROFESOR GUIA

:  _____

LUIS FRANCO VERA
DR. EN CIENCIAS QUÍMICAS
DEPARTAMENTO DE BIOQUÍMICA,
UNIVERSIDAD DE VALENCIA-ESPAÑA.

PROFESOR CO-GUIA

:  _____

ANDRES AVILA BARRERA
DR. EN INGENIERIA EN MATEMATICAS
DEPARTAMENTO DE INGENIERIA EN
MATEMATICAS, UNIVERSIDAD DE LA
FRONTERA-CHILE.

CALIFICACION PROMEDIO TESIS: **6,8**

A mi familia, porque sin ellos
ninguno de mis logros sería importante
y los esfuerzos no valdrían la pena.

AGRADECIMIENTOS

Sin duda alguna, esta tesis es producto de la ayuda tanto directa como indirecta de muchas personas. Agradezco y valoro el aporte de cada una de ellas, ya que contribuyeron en mi enseñanza y me apoyaron a lo largo de la carrera.

Continuaré por agradecer la ayuda incondicional del Dr. Luis Franco, quien me abrió las puertas de su laboratorio aún sin conocerme y me aconsejó amablemente durante del desarrollo del presente trabajo de investigación, compartiendo su conocimiento invaluable. Asimismo quiero agradecer al Dr. Andrés Ávila por el entusiasmo brindado en la codirección de esta tesis y por estar siempre dispuesto a apoyarme. Conjuntamente agradezco a la Dra. Maribel Rodrigo, quien me aconsejó en la realización y en las correcciones de este trabajo. A la Dra. Pepa Castillo quien me guió amablemente en el trabajo de laboratorio y en las correcciones de la parte experimental. A Paula Gonzales por su ayuda en la realización de parte de los experimentos. A la Dra. Alejandra Sandoval por haber solucionado mis problemas en más de una ocasión y hacer posible que haya presentando esta tesis aún desde tan lejos. A la Dra. Priscilla Brebi por acercarme al estudio de la epigenética. A la Sra. Mónica Pacheco por haberme aconsejado y animado a postular a las becas de movilidad estudiantil, y por guiarme en todo el proceso. A Daniel Medina por animarme a seguir adelante. Agradezco a mis padres, por todo su esfuerzo, perseverancia y apoyo, por dar siempre lo mejor de sí para que logre mis sueños.

Finalmente agradezco al Banco Santander Universidades por haberme concedido la beca que posibilitó económicamente la estancia en España. Al Laboratorio de Cromatina y a su director el Dr. Gerardo López por haber financiado esta tesis.

INDICE

Capítulo	Página
I RESUMEN.....	1
II SUMMARY.....	3
1 INTRODUCCION.....	5
2 REVISION BIBLIOGRAFICA.....	7
2.1 Mecanismo y regulación de la transcripción.....	7
2.2 Epigenética.....	11
2.3 Estructura y función de la cromatina.....	13
2.3.1 Complejos remodeladores de la cromatina.....	15
2.3.2 Mecanismo de remodelación de la cromatina.....	17
2.4 Código de histonas.....	19
2.5 Metilación y desmetilación.....	21
2.5.1 Enzimas metiltransferasas y desmetilasas de histonas.....	22
2.5.2 Residuos metilados en histonas.....	24
2.6 Acetilación y desacetilación.....	25
2.6.1 Enzimas acetiltransferasas y desacetilasas de histonas.....	26
2.6.2 Residuos acetilados en histonas.....	27
2.7 Gen <i>Egr1</i>	28
2.8 Posicionamiento y dinámica nucleosomal.....	34
2.8.1 Fundamentos bioinformáticos en la predicción nucleosomal.....	34
2.8.2 Programas bioinformaticos para la predicción de nucleosomas.....	38
2.8.3 Bases de datos de nucleosomas.....	43
3 HIPOTESIS Y OBJETIVOS.....	46
3.1 Hipótesis.....	46
3.2 Objetivo general.....	46
3.3 Objetivos específicos.....	46
4 MATERIALES Y METODOS.....	48
4.1 Línea celular MLP29.....	48

4.1.1	Reconstitución de la línea celular.....	48
4.1.2	Mantenimiento de la línea celular.....	48
4.1.3	Criopreservación de la línea celular.....	49
4.1.4	Tratamiento de la línea celular.....	49
4.2	Extracción de mononucleosomas con nucleasa micrococcal.....	50
4.3	Extracción de ADN genómico desde la línea MLP29.....	52
4.3.1	Sonicación de ADN genómico.....	52
4.4	Inmunoprecipitación de cromatina a partir de mononucleosomas.....	53
4.4.1	Incubación de los anticuerpos.....	53
4.4.2	Fragmentación de la cromatina.....	54
4.4.3	Prelavado de la cromatina.....	55
4.4.4	Incubación de la cromatina con los anticuerpos.....	55
4.4.5	Lavados y elución del ADN inmunoprecipitado.....	56
4.4.6	Purificación y cuantificación de las muestras.....	56
4.5	Reacción en cadena de la polimerasa en tiempo real (qPCR).....	57
4.5.1	Cuantificación del posicionamiento.....	58
4.5.2	Cuantificación de la inmunoprecipitación.....	59
4.6	Gel de agarosa.....	60
4.7	Diseño experimental.....	60
4.8	Análisis estadístico.....	64
4.9	Análisis bioinformático.....	64
5	PRESENTACION Y DISCUSION DE RESULTADOS.....	65
5.1	Obtención de mononucleosomas.....	65
5.2	Puesta a punto de la técnica de sonicación de ADN genómico.....	71
5.3	Verificación de los cebadores.....	72
5.4	Posicionamiento de nucleosomas en <i>Egr1</i>	76
5.5	Modificaciones epigenéticas.....	84
5.5.1	Remodeladores BRM y BRG1.....	85
5.5.2	Acetilación y trimetilación de H3K9 en <i>Egr1</i>	87
5.5.3	Enzimas acetiltransferasa y metiltransferasas de histonas.....	88
5.6	Predicción del posicionamiento nucleosomal.....	89
6.	CONCLUSIONES.....	97

7.	LITERATURA CITADA.....	99
8.	ANEXOS.....	108

I. RESUMEN

El proceso de transcripción que realiza la ARN polimerasa II comprende la unión, inicio, elongación y terminación de la transcripción. Para que la transcripción se lleve a cabo, en primer lugar la cromatina debe ser reestructurada por los remodeladores, los que mediante el uso de la energía del ATP pueden mover, desplazar, expulsar o reestructurar nucleosomas y de esta manera permitir el acceso a los elementos *cis*, a los cuales se unen los factores de transcripción. Asimismo, existen enzimas acetiltransferasas y metiltransferasas que regulan la transcripción de un gen mediante la adición de un grupo acetilo y metilo (respectivamente) a la cola de las histonas que conforman el octámero. Por lo general una marca de acetilación de histonas se relaciona con activación génica, mientras que una marca de metilación puede estar relacionada también con represión. Estas modificaciones que no alteran la secuencia de los genes, pero que afectan su expresión se denominan “epigenéticas”. Los procesos que ocurren a nivel de cromatina en nucleosomas individuales en el momento de la transcripción, por lo general, han sido poco estudiados, a pesar de que en primera instancia, es la estructura y posición de los nucleosomas lo que regula la transcripción génica.

A los pocos minutos de su activación, el gen *Egr1* se transcribe para producir la proteína EGR1, que participa en crecimiento, diferenciación y muerte celular. El funcionamiento anormal del gen causa un gran número de enfermedades, entre ellas algunos tipos de cáncer. Teniendo en cuenta que se desconocen los cambios que ocurren a nivel nucleosomal en el gen *Egr1* durante la activación transcripcional, en el presente trabajo se estudió la posición de los nucleosomas en el promotor proximal y al inicio de la zona transcribible del gen en el estado basal y luego de 30 minutos de inducción con el éster de forbol TPA. Para ello se utilizó la técnica de extracción de mononucleosomas y posterior PCR en tiempo real. A continuación, mediante la técnica de Nu-ChIP se estudió la presencia de los remodeladores BRM y BRG1, la modificación de acetilación y trimetilación del residuo H3K9 y la presencia de enzimas acetiltransferasa y metiltransferasa, en cada uno de los nucleosomas identificados. Además se comparó la posición experimental de

los nucleosomas, con la predicción realizada por los programas NXSensor, NuPop y nuScore, y con la posición indicada en la base de datos de nucleosomas del genoma de hígado de ratón.

Los resultados indican que la posición de los nucleosomas del promotor proximal y al inicio de la región transcribible queda definida entre los amplicones situados en -666 y -370 para el nucleosoma N-2, entre los amplicones -307 y -126 para el nucleosoma N-1 y entre el -69 y +305 para el nucleosoma N+1. En la región ocupada por los nucleosomas N-2, N-1 y N+1 disminuye la protección frente a MNasa luego de inducir la expresión de *Egr1* con TPA 30 min. El nucleosoma N-2, además, se desplaza aproximadamente 50 pb río abajo, dejando libre la región de acceso a EGR1 para su autorregulación. Este resultado demuestra la importancia que presenta el remodelamiento de la cromatina en la regulación de la transcripción. Los remodeladores BRM y BRG1 se encuentran presentes en el nucleosoma N+1 cuando *Egr1* se halla inactivo, sin embargo cuando el gen es activado con TPA 15 min, BRM abandona el nucleosoma y BRG1 aumenta. Los nucleosomas N-2 y N-1 se encuentran acetilados en H3K9, cuando el gen *Egr1* es activado con TPA 30 min, mientras que el nucleosoma N+1 se encuentra trimetilado.

Los programas bioinformáticos para predicción de la posición nucleosomal NXSensor, NuPop y NuScore, en este caso, no fueron asertivos a la hora de predecir la posición de los nucleosomas del promotor proximal y al inicio de la región transcribible de *Egr1* (N-2, N-1, N+1), en comparación a las posiciones identificadas experimentalmente. Por ello se hace patente la necesidad de diseñar nuevos programas bioinformáticos que permitan predecir con mayor exactitud la zona del ADN que forma parte de un nucleosoma. La posición experimental del nucleosoma +1 concuerda con la contenida en la base de datos GEO, la posición del nucleosomas N-2 concuerda parcialmente, mientras que la región del nucleosoma N-1 no concuerda.

II. SUMMARY

The transcription process carried out by the RNA polymerase II involves joining, initiation, elongation and termination of transcription. For transcription to take place, the chromatin must be first remodeled by remodelers, which by using the energy of ATP can move, remove or restructure nucleosomes and thereby allowing the access to the *cis* elements, which bind to transcription complexes. There are also acetyltransferases and methyltransferases that regulate the transcription of a gene by adding an acetyl or a methyl group (respectively) of the tail of octámero histones. Usually a mark of histone acetylation is associated with gene activation, whereas methylation mark may be related to either activation or repression. These modifications do not alter the sequence of the genes affecting their expression and are called "epigenetic". The processes that occur at the level of chromatin, that is to say in individual nucleosomes at the time of transcription, in general, are scarcely studied, although the structure and position of nucleosomes are the ultimate responsible for the regulation of gene transcription. The transcription of gene *Egr1* takes place within minutes of the activation signal.

Its product, the EGR1 protein is involved in growth, differentiation and cell death. Therefore, an abnormal functioning of the gene may cause a large number of diseases, including cancer. The changes occurring in individual nucleosomes of the *Egr1* gene upon activation are not unknown. In this context, the present work was first aimed at the study of the position of nucleosomes in the proximal promoter and the 3' end of the transcribed region. The study has been conducted by extracting mononucleosomes and quantifying their abundance at different positions by qPCR in control and TPA-induced cells. The presence of BRG1 and BRM remodelers as well as that of histone acetyltransferases and methyltransferases was studied by the Nu-ChIP technique. Finally, we studied the changes in the acetylation and trimethylation of H3K9 in each of the nucleosomes identified. In addition we compared the experimental positions of nucleosomes, with the predictions given by the programs NXSensor, NuPop and nuScore, and checked their position with the data shown in the database of nucleosomes in the genome of mouse liver.

The results indicated that nucleosome N-2 is delimited between amplicons centered at -666 and -370, nucleosome N-1 is localized, between -307 and -126 amplicons and N+1 nucleosome is found between -69 and +305. In the region occupied by nucleosomes N-2, N-1 and N+1 protection to MNase decreases after induction with TPA for 30 min. The nucleosome N-2, also slides 50 bp downstream, leaving free access to the binding site of EGR1 for self-regulation. These results demonstrate the importance of chromatin remodeling in transcriptional regulation. BRG1 and BRM remodelers are present in the nucleosome N +1 in untranscribed Egr1, but when the gene is activated with TPA for 15 min, BRM leaves the nucleosome while BRG1 is further recruited. Nucleosomes N-2 and N-1 are acetylated on H3K9, in the active gene, while the N +1 nucleosome is trimethylated.

The results of applying the bioinformatic programs NXSensor, NuScore and NuPop for prediction of nucleosome position do not agree with the experimentally identified positions of Egr1 N-2, N -1, N+1. The need to develop new bioinformatic programs to predict more accurately the sequences of DNA included in a nucleosome became evident from the present studies. The experimental position of nucleosome N+1 agrees with the data of the GEO database for mouse liver nucleosomes, but the region of nucleosomes N-1 not agrees and N-2 partially agrees.

1. INTRODUCCION

En las células eucariotas el ADN se encuentra empaquetado de forma compacta. El primer nivel de empaquetamiento corresponde a una partícula denominada nucleosoma, que se encuentra formada por un octámero de histonas (H2A, H2B, H3 y H4), enrolladas por 147 pares de bases de ADN genómico. Los nucleosomas se organizan entre sí hasta formar el cromosoma. Este sistema de almacenamiento de la información genética es de gran importancia en la regulación de la expresión de los genes. Hasta hace algunos años se creyó que la información se heredaba solo en la secuencia nucleotídica denominada en su conjunto “genoma” y que las alteraciones en la expresión sólo se debían a cambios en los nucleótidos dentro de la secuencia, denominadas “mutaciones”. Hoy se sabe que, además de estos cambios, existen factores que no alteran la secuencia de los genes, pero que afectan su expresión, pudiendo ser estables y heredados en la división celular e incluso de una generación a otra. El conjunto de estos factores constituyen los denominados “cambios epigenéticos”.

La Epigenética es una ciencia que estudia las modificaciones químicas que afectan la estructura de las proteínas, del ADN y en último término de la cromatina. Estas modificaciones ocurren de forma normal en la célula y permiten regular procesos como el desarrollo embrionario, la inactivación del cromosoma X, el desplazamiento nucleosomal, la expresión génica, entre otros. Entre los cambios epigenéticos más investigados se encuentran la metilación del ADN y las modificaciones de histonas. Estas últimas suelen ocurrir en la cola N-terminal de las histonas y son originadas por la incorporación de grupos químicos que no se encuentran formando parte de la estructura proteica nativa, pero que pueden llegar a afectar la interacción nucleosoma-nucleosoma o ADN-histonas, ya sea cambiando la carga eléctrica o por interacción con otros factores, originando una represión o activación de la transcripción. El funcionamiento anormal de la maquinaria epigenética de la célula causa graves daños y a distintos niveles en un organismo, como por ejemplo, el cáncer surge muchas veces como consecuencia de alteraciones

epigenéticas. Por esto, el conocimiento del “epigenoma” puede permitir comprender el origen de múltiples enfermedades y generar alternativas para combatirlas.

El gen *Egr1* del inglés “*early growth response 1*” pertenece a los genes de respuesta celular inmediata, y se activa rápidamente vía una cascada de señalización unos pocos minutos después de recibir un estímulo. El mecanismo de activación y la cascada de señalización por los cuales modifican su expresión varía de acuerdo al estímulo. El producto del gen *Egr1* es la proteína EGR1, la cual es responsable de la regulación de múltiples procesos celulares relacionados con la diferenciación, crecimiento celular y apoptosis, por lo que el gen se encuentra involucrado en el desarrollo de una gran variedad de enfermedades.

Los elementos *cis* que regulan a *Egr1* se encuentran descritos, sin embargo la mayor parte de estos estudios se realizan sin tener en cuenta la organización nucleosomal de la cromatina, que podría regular el acceso real a estos elementos en el promotor de *Egr1*. Los cambios que ocurren a nivel de nucleosomas en el gen *Egr1* aún no han sido descritos, por lo que investigar los cambios epigenéticos en nucleosomas particulares, involucrados en la regulación de la expresión del gen podría aportar un conocimiento básico imprescindible a la hora de fijar dianas para posibles tratamientos de enfermedades relacionadas con proliferación celular. Para realizar la presente investigación se ha escogido la línea progenitora de células hepatocíticas de ratón (MLP29).

2. REVISION BIBLIOGRAFICA

2.1 Mecanismo y regulación de la transcripción

Como es bien sabido, el material genético de los eucariotas está constituido por ADN. En último término, la información genética se emplea para dirigir la síntesis de proteínas, elementos estructurales y funcionales fundamentales en la dinámica celular. La hipótesis de Crick de que la información genética fluye desde el ADN hasta las proteínas a través del ARN, conocida como el “dogma central de la Biología Molecular”, abrió una serie fructífera de investigaciones, que condujeron a comprender el mecanismo de la transcripción, el proceso por el que se sintetizan los diversos ARNs, mensajero, ribosomal y de transferencia, así como el mecanismo de la traducción. En este último proceso, la información copiada en el ARN mensajero, dirige la biosíntesis de proteínas, que tiene lugar en los ribosomas, en la que los aminoácidos se adaptan a su lugar concreto gracias a la función de los ARN de transferencia.

Los genes son las unidades de información genética, aunque dar una definición más precisa constituye una tarea nada fácil. A la idea inicial de que un gen codificaba un carácter fenotípico, sucedió la concepción de Beadle y Tatum, según la cual un gen contenía la información necesaria para la síntesis de una enzima o, si se quiere, de una cadena polipeptídica. Esta concepción, esquematizada mediante el conocido aforismo “un gen, una enzima” se comenzó a desvanecer como consecuencia del descubrimiento de que algunos genes podían codificar más de una proteína a través de varios caminos, entre los que destacan los diferentes mecanismos de procesamiento del ARN mensajero. Pero la secuenciación de genomas eucarióticos, entre ellos el humano, arrinconó definitivamente la idea de Beadle y Tatum. El número de genes humanos ha resultado ser considerablemente menor que el de proteínas presentes en una célula, de modo que se ha comprendido que el hecho de que un gen codifique varias proteínas es algo relativamente común.

Un gen eucariótico típico se compone de una región codificante, que contiene la información para la síntesis de una o varias proteínas, una región situada en 5', que contiene el promotor, o lugar donde se inicia la transcripción (fig. 1). En el extremo 3' se encuentra otra región de terminación de la transcripción, a menudo dotada también de propiedades reguladoras. El promotor comprende además muchas secuencias reguladoras de la transcripción. Es también frecuente la existencia de regiones reguladoras distales, situadas, a veces, a varias kilobases de distancia de la región reguladora. Finalmente, hay que destacar que en la región codificante, las secuencias propiamente codificantes, los exones, se encuentran interrumpidas casi siempre por otras regiones, los intrones, que, aunque se transcriben para dar un precursor del ARN mensajero, se pierden de ordinario en el procesamiento del transcrito primario antes de dar lugar al ARN mensajero maduro. Precisamente, los diversos modos de procesamiento son la causa fundamental de que de un único gen se puedan obtener diversas proteínas.

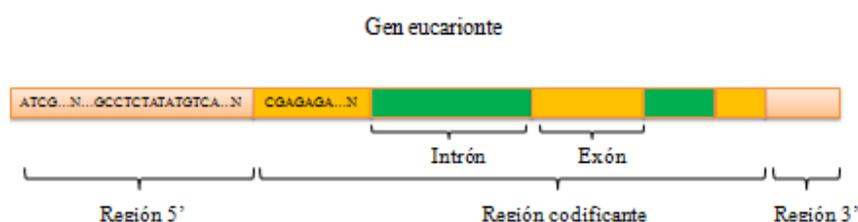


Figura 1. Organización de un gen eucariote típico.

La región promotora de un gen estructural controla en gran medida la transcripción del mismo, ya que en esta zona se encuentran los denominados elementos *cis*, secuencias específicas del ADN por donde se acopla e interactúa una gran variedad de factores, elementos *trans*, que representan un conjunto de señales de regulación. Entre los elementos *cis* se encuentra la región por donde se ensamblan la ARN polimerasa II y los factores generales de transcripción, TFIIA, TFIID, TFIIE, TFIIF y TFIIH, así como los coactivadores (por ej. el complejo mediador). En muchos genes, pero no en todos, esta región es la denominada caja TATA, por la presencia habitual de esa secuencia en el elemento. Por otra parte, los promotores contienen otras secuencias *cis*, específicas de cada gen o de un número restringido de genes, por las que se unen

otros factores específicos que actúan como activadores o inhibidores de la transcripción, dando lugar a un complejo entramado de interacciones que conduce a una transcripción altamente regulada.

La ARN polimerasa II es una enzima altamente conservada en estructura, función y mecanismo de acción. Posee 12 subunidades, entre ellas, la subunidad de mayor tamaño (RBP1) posee una cola carboxilo-terminal (CTD) de siete aminoácidos que se repite numerosas veces (hay 52 repeticiones en humanos) y se encuentra separada del cuerpo principal de la enzima. La fosforilación de esta cola por TFIIF, es la señal requerida para que la polimerasa abandone el promotor y comience la transcripción (Nelson y Cox, 2006).

El proceso de transcripción que realiza la ARN pol II comprende la unión, inicio, elongación y terminación de la transcripción. Los procesos de unión e inicio de la transcripción se llevan a cabo en la región promotora. De forma resumida, se puede decir que en los genes que poseen caja TATA se comienza a formar un complejo cuando la proteína de unión a TATA (TBP) se une a ésta. Posteriormente el factor de transcripción TFIIB se une a TBP y también al ADN (a ambos lados de TBP), luego se une TFIIA estabilizando el complejo, aunque su participación no es esencial. A continuación se une el complejo TFIIF y la ARN polimerasa II (TFIIF contribuye dirigiendo a la ARN polimerasa II hacia su promotor); finalmente se unen TFIIE y TFIIF y de esta manera se completa el complejo cerrado. TFIIF tiene actividad ADN helicasa que promueve el desenrollamiento del ADN cerca del sitio de inicio (Inr), creándose así el complejo abierto y finalizando el proceso de unión. El proceso de inicio de la transcripción se realiza por la actividad quinasa de TFIIF que junto al complejo del factor b positivo de elongación de la transcripción (pTEFb), del cual es parte CDK9, fosforilan la cola CTD de la subunidad RBP1 de la ARN pol II, lo que hace que la enzima abandone el promotor y comience la transcripción. (Nelson y Cox, 2006)

La región codificante de un gen comienza en el inicio de la transcripción a ARNm, en la posición +1, +2 o +3, dependiendo del marco de lectura abierto. Después de que la ARN pol II abandone el promotor del gen comienza el proceso de elongación. El TFIIF permanece asociado

a la ARN pol II durante toda la elongación. La acción de la ARN pol II se ve potenciada por los factores de elongación, que suprimen las pausas durante la transcripción. Para sintetizar una hebra de ARN, la enzima ARN pol II necesita un molde de ADN, los cuatro ribonucleósidos 5'-trifosfato (ATP, GTP, UTP y CTP) y el co-factor enzimático Mg^{2+} . Cataliza la adición de ribonucleótidos al extremo 3'-hidroxilo (que actúa como nucleófilo) de la cadena de ARN sintetizada en dirección 5'-3', de acuerdo a la lectura (siguiendo las interacciones de apareamiento de las bases nitrogenadas, en la que A se une a U, T, y estos a A, y C a G y G a C) determinada en la hebra molde de ADN, la cual es copiada en dirección 3'-5' durante la fase de elongación de la transcripción. Al principio de esta fase, el extremo en crecimiento de la nueva cadena de ARN se aparea temporalmente con el ADN molde formando una cadena híbrida de aprox. 8 pb, que durante el avance de la ARN pol II (que mantiene desenrollado el ADN unas pocas pares de bases mientras realiza la transcripción) se va despegando gradualmente para que el ADN ya transcrito se vuelva a hibridar con su hebra complementaria. Como se ha apuntado antes, la región codificante de un gen se encuentra compuesta de intrones y exones. Los intrones son segmentos intercalados de ADN que ordinariamente no codifican aminoácidos y se eliminan del ARN en el proceso de corte y empalme para la obtención del ARN maduro. Los exones son las secuencias del gen que codifican aminoácidos. En un gen eucariota típico existe mayor porcentaje de intrones que de exones (Nelson y Cox, 2006).

La región 3' corresponde a la zona en la cual se da término al gen. En esta región finaliza la transcripción, la polimerasa es liberada y el ARN mensajero se libera. El mecanismo involucrado en este proceso en eucariotas, aún no se encuentra completamente descrito, pero el fin de la síntesis del ARN está indicado por secuencias específicas, en las que se repite una secuencia de pares de bases en la hebra molde. Una vez sintetizado el ARN (en esta etapa se denomina transcrito primario) pasa por un proceso de maduración, en el que se somete a corte y empalme, en el cual los intrones son eliminados y los exones se unen formando una secuencia continua. También se modifica en ambos extremos. En el extremo 5' se añade un residuo modificado denominado casquete y en el extremo 3' se añaden entre 80 y 250 residuos de adenina, denominados "cola de poli(A)". El ARN mensajero final sale al citoplasma, se une a los ribosomas y en ellos se traduce a polipéptido. (Nelson y Cox, 2006).

Como se ha descrito anteriormente la estructura del ADN determina en gran medida la obtención de un polipéptido. Por ello, los mecanismos involucrados en la transcripción y traducción son de suma importancia, siendo en última instancia los que determinan, no solo que la obtención del polipéptido se lleve a cabo correctamente, sino también el tipo de polipéptido y la cantidad del mismo. Pero, como se mencionará más adelante, para que la transcripción se lleve a cabo, en primer lugar la cromatina debe ser remodelada, proceso en el cual se desplazan algunos nucleosomas. Además la estructura del nucleosoma y los cambios epigenéticos en las histonas también cumplen un papel importante en la regulación de la expresión.

2.2 Epigenética

Epigenética es un término acuñado hace alrededor de 70 años por Conrad Waddington, que la definió como “la rama de la Biología que estudia las interacciones causales entre los genes y su entorno que conducen a la existencia del fenotipo”. Como esas interacciones modifican la estructura del material genético sin producir cambios en la secuencia del ADN, alternativamente se suele definir la Epigenética como “el estudio de los cambios, hereditarios o al menos potencialmente heredables, experimentados por el material genético y no debidos a alteraciones en la secuencia del ADN” (Jiang et al., 2004). En los años 70 ya se sabía que los mecanismos de regulación de la actividad génica están relacionados con los cambios estructurales de la cromatina (proceso epigenético) y se comenzaba a discutir sobre el papel regulador de las histonas en la expresión génica (Khesin y Leibovich, 1976). En 1981 Braun postuló un modelo epigenético para el origen del cáncer, poniendo de manifiesto que la epigenética no solo afectaba a invertebrados y plantas, sino también a mamíferos, mencionando además su carácter heredable. Posteriormente en 1987 Edelman plantea la existencia de “normas” epigenéticas que se encuentran regulando la expresión de moléculas de adhesión celular y señala a la epigenética como un mecanismo evolutivo que regula procesos como la diferenciación, el movimiento y la muerte celular. Ese mismo año Holliday postula que los carcinógenos pueden actuar mediante la alteración de los patrones epigenéticos normales alterando la actividad génica en células especializadas y por lo

tanto pueden producir fenotipos aberrantes hereditarios (Holliday 1987). A partir de los años 90, la epigenética resurgió como una ciencia novedosa e interesante para el estudio de todos los procesos celulares y en todos los organismos, siendo *Saccharomyces cerevisiae*, *Drosophila*, *Mus musculus* y *Arabidopsis* los modelos de estudio más empleados (Inglis 2007; Tollefsbol, 2011).

Uno de los factores que ha contribuido a la expansión de la Epigenética ha sido la secuenciación de numerosos genomas. Como consecuencia de estos estudios, se comprobó que en organismos superiores el número de genes es menor de lo que se estimaba a partir de la variedad de proteínas y de funciones celulares, por lo que hubo que buscar mecanismos causantes de dicha variedad, compatibles con una información genética más simple de lo esperado. Lo que encuadra a una determinada célula dentro de un tipo celular es su proteoma, el conjunto de sus proteínas constituyentes, que, a su vez, viene determinado por su transcriptoma, es decir, por el conjunto de los ARN mensajeros característicos de esa célula. Y son, en definitiva, los factores epigenéticos los que hacen posible que, teniendo el mismo genoma todas las células de un organismo humano —con la evidente excepción de los linfocitos B y T—, tengan un transcriptoma y un proteoma específicos del tipo a que pertenecen.

Los factores epigenéticos son numerosos, siendo los más estudiados la metilación del ADN y las modificaciones de histonas. En los últimos años han cobrado especial interés los micro ARN y el ARN de interferencia (Rivera y Bennett, 2010). Si se tiene en cuenta que el ADN genómico se encuentra en un entorno definido por la estructura de la cromatina, el genotipo puede dar lugar al fenotipo solamente a través del epigenotipo. En este sentido, los estudios epigenéticos son también susceptibles de aproximaciones “ómicas”, lo que ha dado origen al neologismo “epigenómica”, y se propone, además, la creación de un proyecto denominado epigenoma humano, ya que la epigenética podría proveer la información necesaria para el esclarecimiento, diagnóstico y tratamiento contra numerosas enfermedades y patologías humanas. Ya existen bases de datos para almacenar la información al respecto; por ejemplo la NCBI (National Center for Biotechnology Information) posee una base de datos denominada “Epigenomics”, en la cual se recopila la información sobre mecanismos de reprogramación nuclear, control de la expresión génica y enfermedades ligadas a los cambios epigenéticos.

2.3 Estructura y función de la cromatina

En los procariontes, el ADN no se encuentra completamente desnudo y el problema que surgiría en la célula por las dimensiones que alcanza extendido, se soluciona mediante el empaquetamiento, por asociaciones dinámicas e inestables de proteínas estructurales y reguladoras (HU, IHF, H1 y P) que interactúan con el ADN formando el nucleoide. En células eucariotas el ADN interacciona con una gran variedad de proteínas enrollándose y condensándose hasta formar el cromosoma. Los niveles de empaquetamiento hasta formar el cromosoma, comienzan cuando la doble hélice de ADN desnudo de 2 nm se enrolla en casi dos vueltas alrededor del octámero (ocho moléculas) de histonas. La estructura del octámero se conoce gracias al trabajo del grupo de Moudrianakis (Arents et al., 1991), conteniendo dos subunidades de cada una de las histonas H2A, H2B, H3 y H4. Esta estructura de ADN más histonas, se denomina nucleosoma (fig. 2) y posee un diámetro de 11 nm. El modelo nucleosomal para la estructura de la cromatina fue propuesto por Kornberg en 1974.

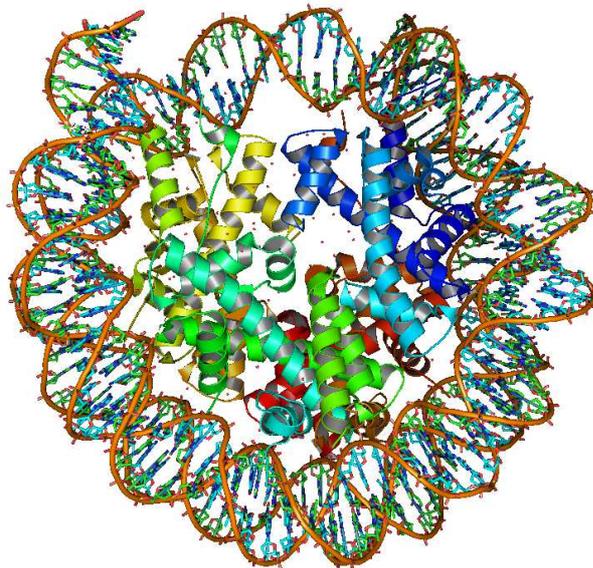


Figura 2. Estructura del nucleosoma. Un octámero de histonas (H2A, H2B, H3 y H4) rodeado por casi 2 vueltas de ADN. Cristalografía 3AV1 visualizada con el programa PyMOL y extraída desde “Protein Data Bank” (<http://www.rcsb.org/pdb/home/home.do>).

La longitud de ADN ocupada por un nucleosoma es aproximadamente 200 pares de bases (pb), de los cuales 147 están fuertemente unidos al núcleo de histonas y el resto sirve de unión entre los nucleosomas. Posteriormente los nucleosomas, que contienen adherida la histona H1 se empaquetan de forma compacta y ordenadas formando una estructura de 6 a 7 nucleosomas por vuelta denominada fibra de cromatina que poseen un diámetro de 30 nm. La fibra de cromatina se enrolla para compactar aun más el ADN, pero los siguientes niveles de empaquetamiento no se encuentran completamente definidos y podrían cambiar de una célula a otra y de un tejido u organismo a otro. Sin embargo un esquema general sugiere que la fibra de cromatina se enrolla en sí misma y forma un bucle de entre 20000 pb a 100000 pb y que este continúa enrollándose hasta 6 veces formando una roseta que en su interior contiene una estructura denominada armazón nuclear, las rosetas se envuelven para finalmente formar el cromosoma metafásico de 1.400 nm (fig. 3) (Nelson y Cox, 2006).

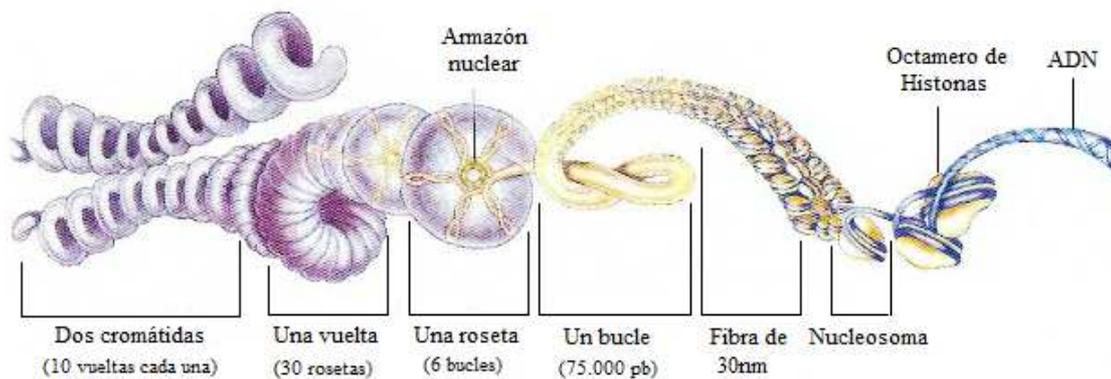


Figura 3. Niveles de empaquetamiento del ADN en células eucariontes. Modificado desde Nelson y Cox (2006).

La estructura típica de cada cromosoma se hace patente en la célula solo durante la mitosis, por lo que en la mayor parte el ciclo vital de las células eucarióticas (interfase y células en G0) no se distinguen los cromosomas, por esto la cromatina debe desenrollarse, adoptando una estructura amorfa y dispersándose en diversas regiones del núcleo, para permitir el acceso de los mecanismos celulares al ADN (Li *et al*, 2004; Nelson y Cox, 2006).

La estructura de la cromatina cumple funciones esenciales, no solo por el empaquetamiento y protección del ADN, sino también en la preservación de la información genética y el control de la expresión de los genes, formando el andamio de los procesos nucleares fundamentales. Las células han desarrollado mecanismos que alteran la estructura de la cromatina, entre los que se incluyen la movilización nucleosomal ATP dependiente (remodelación de la cromatina) y la modificación post-translacional de las histonas que pueden ocasionar modificación en la cromatina. En efecto, las modificaciones de histonas, en particular la acetilación, pueden alterar las interacciones entre nucleosomas o entre el octámero de histonas y el ADN. Por otro lado, las marcas epigenéticas pueden actuar como señales para reclutar otros factores que alteren la estructura de la cromatina, conduciendo a estados represivos o permisivos de la transcripción (Li et al., 2004, Young y Kirchmaier, 2011).

En la cromatina se pueden distinguir dos niveles estructurales distintos, una forma condensada que por lo general reprime todas las funciones del ADN, denominada heterocromatina y una forma más descondensada, que proporciona el entorno de los procesos funcionales del ADN, llamada eucromatina. Las regiones heterocromáticas constitutivas, que contienen fundamentalmente ADN repetitivo, incluyendo secuencias satélite repetidas en tándem y secuencias relacionadas con transposones y retrovirus, se encuentran casi desprovistas de genes. La heterocromatina facultativa es aquella en la que se encuentran todos aquellos genes que no se expresan en el tipo celular considerado. Esta organización diferencial de la cromatina se debe por un lado a la composición del ADN y por otro a las modificaciones que suelen poseer las histonas en su secuencia aminoacídica que ocasionan la inactivación de genes en zonas eucromáticas, un ejemplo es la inactivación del cromosoma X en mamíferos, que en su estado condensado da origen al corpúsculo de Barr (Dillon, 2004).

2.3.1 Complejos remodeladores de la cromatina. Los complejos remodeladores son aquellos que utilizan la energía de la hidrólisis del ATP para mover, desplazar, expulsar o reestructurar nucleosomas dentro de la cromatina, por lo que son necesarios para empaquetar el genoma completo tras la replicación y para desenrollar correctamente los nucleosomas antes de la

siguiente replicación. Además los remodeladores también proporcionan accesibilidad regulada a las regiones del ADN empaquetado y de esta manera la maquinaria transcripcional puede acceder a los genes que se han de transcribir (Clapier y Cairns, 2009; Piatti *et al*, 2011). Actualmente se conocen cuatro familias de complejos remodeladores de la cromatina, cuyos miembros comparten la afinidad por el nucleosoma, dominios que reconocen las modificaciones covalente de las histonas y un dominio ATPasa similar, que utiliza la energía de la hidrólisis del ATP proporsionando la energía necesaria para alterar los contactos entre las histonas y el ADN. También comparten dominios de interacción con otras proteínas, pero difieren en los dominios de las regiones flanqueantes, permitiendo su especialización en distintos contextos biológicos y diferenciando a las cuatro familias (cuadro 1): SWI/SNF, ISWI, CHD, INO80 (Bao y Shen, 2007; Clapier y Cairns, 2009).

Cuadro 1. Familias de remodeladores y sus miembros principales en la especie humana. Adaptado de Bao y Shen (2007) y de Clapier y Cairns, (2009).

Familia y composición		Humano		
SWI/SNF	Complejo	BAF	PBAF	
	ATPasa	hBRM or BRG1	BRG1	
	Subunidades homologas no catalíticas	BAF250/hOSA1		
			BAF180 BAF200	
		BAF155, BAF170		
		BAF60a,b o c		
		hSNF5/BAF47/INI1		
		BAF57		
		BAF53a o b		
		β-actina		
ISWI	Complejo	NURF	CHRAC	ACF
	ATPasa	SNF2L	SNF2H	
	Subunidades homologas no catalíticas	BPTF	hACF1/WCRF180	
			hCHRAC17	
		RbAp46 o 48	hCHRAC15	
CHD	Complejo	CHD1	NuRD	
	ATPasa	CHD1	Mi-2α/CHD3, Mi-2β/CHD4	
	Subunidades homologas no catalíticas		MBD3	
			MTA1,2,3	
			HDAC1,2	
			RbAp46 o 48 p66α,β	

Familia y composición		Humano			
INO80	Complejo	INO80	SRCAP	TRRAP/Tip60	
	ATPasa	hIno80	SRCAP	p400	
	Subunidades homologas no catalíticas	RUVBL1,2/Tip49a,b			
		BAF53a			
		Arp5,8	Arp6	Actina	
		GAS41			
		hles2,6			
		DMAP1			
		YL-1			
		Brd8/TRC/p120			
	H2AZ,H2B				
	ZnF-HIT1				
		TRRAP	Tip60		
		MRG15	MRGX		
		FLJ11730	MRGBP		
		EPC1,	EPC-like		
		ING3			

Los remodeladores BRM y BRG1, pertenecientes a la familia SWI/SNF, pueden regular la transcripción ya sea positiva o negativamente dependiendo del promotor del gen sobre el cual actúen. Por ejemplo en el caso del gen *Egr1*, en el que se centra este trabajo, BRM regula negativamente la expresión del gen, mientras que BRG1 parece estar presente cuando *Egr1* se expresa (Sakurai et al, 2011).

2.3.2 Mecanismo de remodelación de la cromatina. El mecanismo de remodelación varía según la forma en la que la hidrólisis de ATP altera la interacción de las histonas con el ADN y según este proceso se utiliza apropiadamente para modificar la estructura y posición del nucleosoma. El modelo más reciente explica que el proceso comienza cuando el remodelador se ancla al nucleosoma y el dominio ATPasa responsable del desplazamiento se une al ADN a dos vueltas de su eje diádico (eje de simetría pseudobinaria).

En la figura 4 se describe éste modelo, en el estado 1 el remodelador se encuentra anclado al nucleosoma con su dominio ATPasa permaneciendo en el interior del octámero de histonas y desde allí lleva a cabo el desplazamiento direccional del ADN. En el estado 2, mediante la alteración de interacciones que cambian la afinidad ADN-Histonas y de esta manera transportan el ADN hacia la díada ocasionando la formación de un bucle transitorio de ADN, luego el dominio Tr desplaza el ADN y un dominio DBD empuja al ADN y propaga el bucle en dirección a la díada, como se muestra en el estado 3. En el estado 4 el dominio DBD del remodelador libera una región del ADN y vuelve a enlazar otra, para pasar nuevamente al estado 1. (Clapier y Cairns 2009).

Modelo de movimiento del ADN en la remodelación

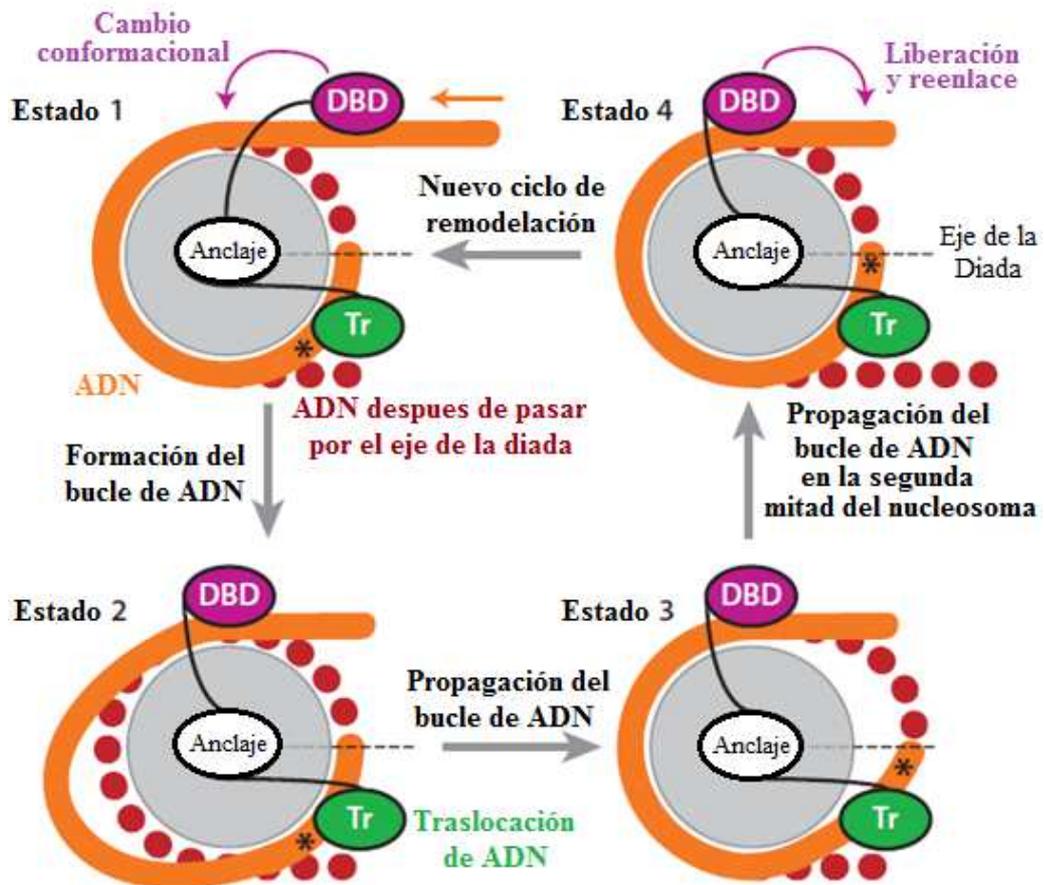


Figura 4. Mecanismos de remodelación de la cromatina, a través del desplazamiento de ADN en el nucleosoma, separado en estados 1, 2, 3 y 4. Figura modificada de Clapier y Cairns (2009).

Este modelo requiere la interrupción del contacto del ADN con una o dos histonas a la vez y se desconoce el tamaño de los bucles que se forman para lograr el desplazamiento, pudiendo ser de 1pb a 100pb. También existen pruebas de que la direccionalidad del desplazamiento puede cambiar, pero no se conocen los mecanismos implicados. Un aspecto importante en el estudio de los mecanismos de remodelación es el modo en que se regula la formación del bucle de ADN, puesto que de ello dependen los diferentes resultados que se pueden obtener en la remodelación de la cromatina (Clapier y Cairns 2009).

2.4 Código de histonas

Las histonas H2A, H2B, H3 y H4 se encuentran en el interior del nucleosoma formando un octámero, se sabe que estas proteínas pueden sufrir modificaciones, se han identificado por lo menos 60 residuos diferentes donde se han detectado modificaciones de histonas y es probable que este número este subestimado, el desarrollo de nuevas técnicas sin duda ayudan a identificar nuevas modificaciones en los residuos. Hay hasta la fecha, por lo menos ocho tipos diferentes de modificación de histonas, los cuales son acetilación, metilación, fosforilación, ubiquitinación, sumoilación, ADP ribosilación, deiminación y prolina isomerización (cuadro 2). Se cree que estas modificaciones tienen como función regular dos mecanismos, que son la interacción núcleo-núcleo y la interacción ADN-histonas, los cuales tienen lugar mediante la adición de las entidades físicas que no cambia la carga de las histonas, como por ejemplo la de un grupo metilo, o cambiando la carga de las histonas, como en el caso de la adición de un grupo acetilo (Kouzarides, 2007a).

Cuadro 2. Modificaciones identificadas en histonas, localización y función. Extraído desde Kouzarides (2007a).

Modificación	Residuo	Función
Acetilación	K-ac	Transcripción, reparación, replicación, condensación
Metilación (lisina)	K-me1, me2, me3	Transcripción, reparación
Metilación (arginina)	R-me1, me2a, me2s	Transcripción
Fosforilación	S-ph, T-ph	Transcripción, reparación, condensación
Ubiquitinación	K-ub	Transcripción, reparación
Sumoilación	K-su	Transcripción
ADP ribosilación	E-ar	Transcripción
Deiminación	R>Cit	Transcripción
Prolina isomerización	P-cis > P-trans	Transcripción

Diferentes marcas podrían presentar un sitio de acoplamiento para reclutar proteínas específicas que podrían generar resultados celulares distintos. Además numerosos informes

plantean la posibilidad de que todas estas modificaciones son combinatorias e interdependientes, formando el llamado “código de histonas”. Strahl y Allis propusieron el año 2000 la hipótesis del “código de histonas”, planteando que las múltiples marcas observadas en las colas de las histonas podrían ampliar la lectura de las vías de señalización río arriba, provocando cambios en densidad de cargas que conducen a grandes cambios en la estructura de la cromatina de los genes diana, en especial los de respuesta inmediata. Según esto las modificaciones en las colas de las histonas no son eventos azarosos, tienen un propósito determinado y existen enzimas y otros factores específicos que actúan en cada modificación (Allis *et al*, 2007). En la figura 5 se indican las modificaciones conocidas que sufren las colas de la histona H3 y las enzimas que pueden estar implicadas en estas modificaciones, señaladas con la nueva nomenclatura.

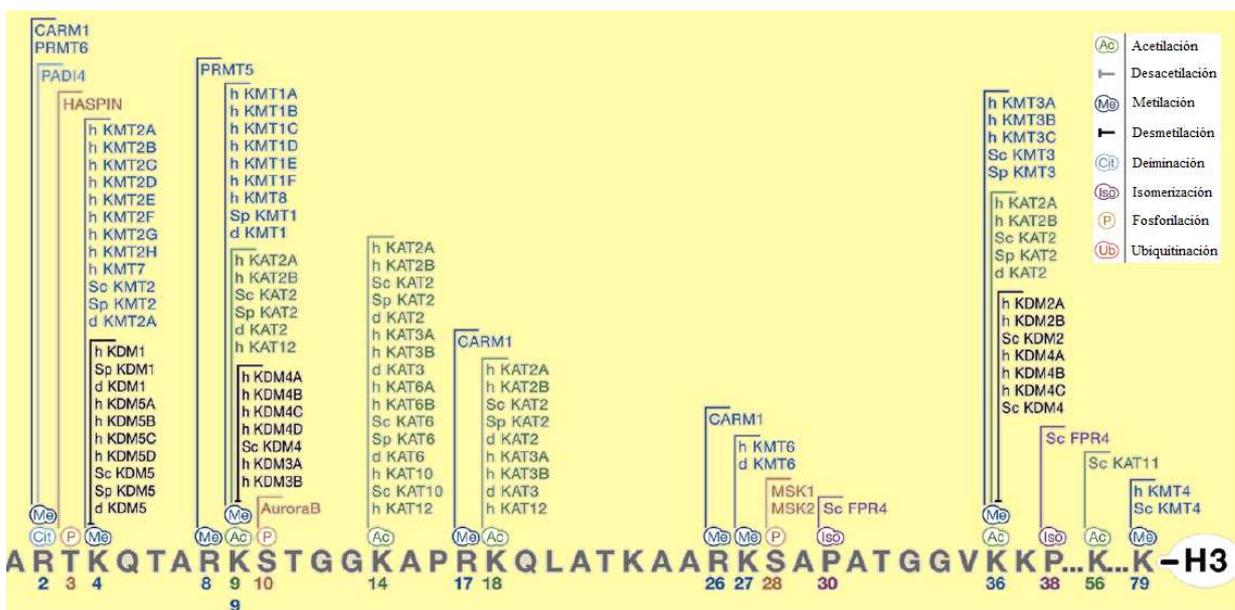
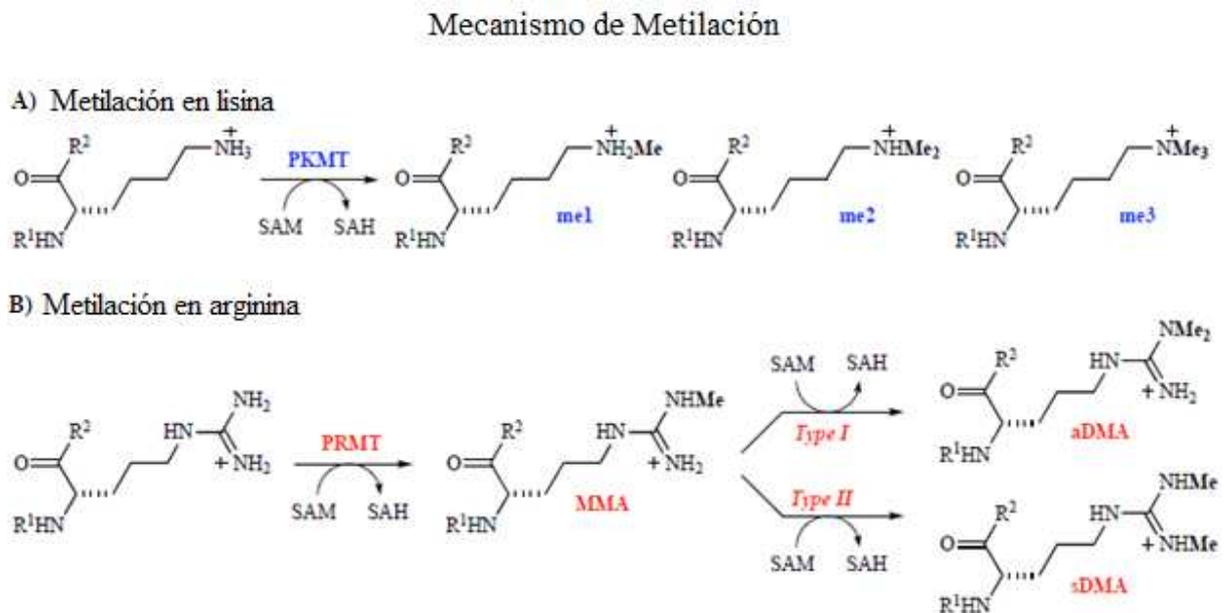


Figura 5. Modificaciones más comunes en la histona H3 y enzimas que las generan. Modificada desde Kouzarides (2007b). Ver nomenclatura en anexo 1.

2.5 Metilación y Desmetilación

El término metilación se emplea para describir el fenómeno en el que S- adenosilmetionina (SAM) dona un grupo metilo al carboxilo de los residuos glutamato, cisteína y leucina, o en los átomos de nitrógeno de lisina, arginina e histidina, dejando un grupo S-adenosil-lHomocisteína (SAH) como resultado. Sin embargo la metilación en histonas se produce únicamente en el grupo amino de los aminoácidos arginina y lisina (figura 6). La arginina puede estar mono o di-metilada, mientras que la lisina puede estar mono, di o tri-metilada, en el caso de la lisina este proceso lo realizan las enzimas lisina metiltransferasas (PKMTs), mientras que en arginina lo realizan las arginina metiltransferasas (PRMTs), recientemente se ha descubierto que también son capaces de metilar otras proteínas importantes como P53 (Yost *et al*, 2011). El ADN también se puede encontrar metilado, pero esa modificación no será revisada en la presente investigación.



El proceso por el cual se revierte la marca de metilación se conoce como desmetilación y se lleva a cabo por las enzimas Histona desmetiltransferasas. Shi *et al.* (2004) identificaron por primera vez una enzima de este tipo y la llamaron Lisina desmetilasa específica 1 (LSD1). Hasta ahora se conocen dos tipos de histonas desmetilasas, una familia presenta el dominio LSD1 y la otra familia o tipo presenta el dominio JmjC, ambas familias realizan reacciones catalíticas diferentes (Kouzarides, 2007a). La metilación no altera la carga de las proteínas y puede estar relacionada tanto con activación como con represión génica.

2.5.1 Enzimas metiltransferasas y desmetilasas de histonas. Las enzimas que catalizan la metilación en histonas son la familia lisina metiltransferasas (PKMTs) y la familia arginina metiltransferasas (PRMTs). Las enzimas que pertenecen a la familia lisina metiltransferasas pueden añadir uno (mono), dos (di) o tres (tri) grupos metilo a los residuos lisina de una gran gama de proteínas, incluyendo las Histonas. Desde el año 2000 se han identificado más de 50 proteínas de esta familia en humanos. PKMTs se caracteriza por presentar el dominio SET (Supresor de variegación, Enlazador de zeste y Tritorax), este dominio divide a la familia en 5 subfamilias nombradas por sus miembros principales: RIZ, SET1, SET2, SMYD3 y SUV39 (Yost *et al.*, 2011). La modificación del residuo determina si se acoplan los complejos activadores o represores de la transcripción, esto depende del dominio de reconocimiento que posee cada proteína, por ejemplo SUV39 posee un cromodominio y se sabe que actúa metilando Lisina 9 en la histona H3 (Selvi *et al.*, 2010).

La familia de enzimas arginina metiltransferasas, introducen marcas de mono o di-metilación al residuo arginina de las proteínas y dependiendo del tipo de PRMT que di-metila el residuo puede ser simétrica (un grupo metilo en cada uno de los dos átomos de N-terminal) o asimétrica (los dos grupos metilo en el mismo átomo N-terminal). Hasta el momento se han identificado nueve tipos de PRMTs en humano y se clasifican principalmente como tipo I y tipo II. Ambos tipos catalizan la reacción de arginina mono-metilación (MMA), las enzimas de tipo I (PRMT1, 2, 3, 4, 6 y 8) catalizan la arginina di-metilación asimétrica (ADMA), mientras que las de tipo II (PRMT5, 7 y 9/FBXO11) originan la di-metilación simétrica (SDMA) (Yost *et al.*, 2011).

Las proteínas desmetilasas también se clasifican en dos tipos. Las lisina desmetilasas específicas 1 (LSD1), que requieren nitrógeno protonado y por lo tanto solo pueden desmetilar lisina mono y di-metiladas, catalizando la eliminación de grupos metilo en H3K4 mono y di metilado, pero no cuando se encuentra en el nucleosoma, solo cuando forma un complejo con el represor Co-REST puede desmetilar histonas nucleosomales. Cuando LSD1 forma un complejo con el represor andrógeno puede desmetilar H3K9, cambiando la función de LSD1 de represor a co-activador (Bannister y Kouzarides, 2011).

El segundo tipo de desmetilasas se comenzaron a estudiarse el 2006, al identificar la proteína desmetilasa tri-metil-lisina (JMJD2), esta proteína presenta un dominio Jumonji (JmjC) que tiene la capacidad de desmetilar residuos trimetilados como H3K9 me3 y H3K56 me3, empleando un mecanismo de catalizador similar al utilizado por LSD1, con Fe (II) y a-cetoglutarato como co-factor y un mecanismo de ataque a radicales libres. Ahora se conocen muchas proteínas desmetilasas (Cuadro 3) y a excepción de LSD1 todas poseen el dominio Jumonji (Bannister y Kouzarides, 2011).

Cuadro 3. Enzimas lisina desmetilasas y residuos sobre los que actúan. Modificado desde Bannister y Kouzarides (2011).

Desmetilasas de lisina	síto de la modificación
LSD1/BHC110	H3K4
JHDM1a y b	H3K36
JHDM2a y b	H3K9
JMJD2A/JMDM3A	H3K9, H3K36
JMJD2B	H3K9
JMJD2C/GASC1	H3K9, H3K36
JMJD2D	H3K9

2.5.2 Residuos metilados en histonas. Las marcas de metilación en Histonas (Fig. 7) pueden estar relacionadas con la activación, elongación o represión de la expresión génica, dependiendo del tipo de aminoácido e histona involucrados en la modificación.

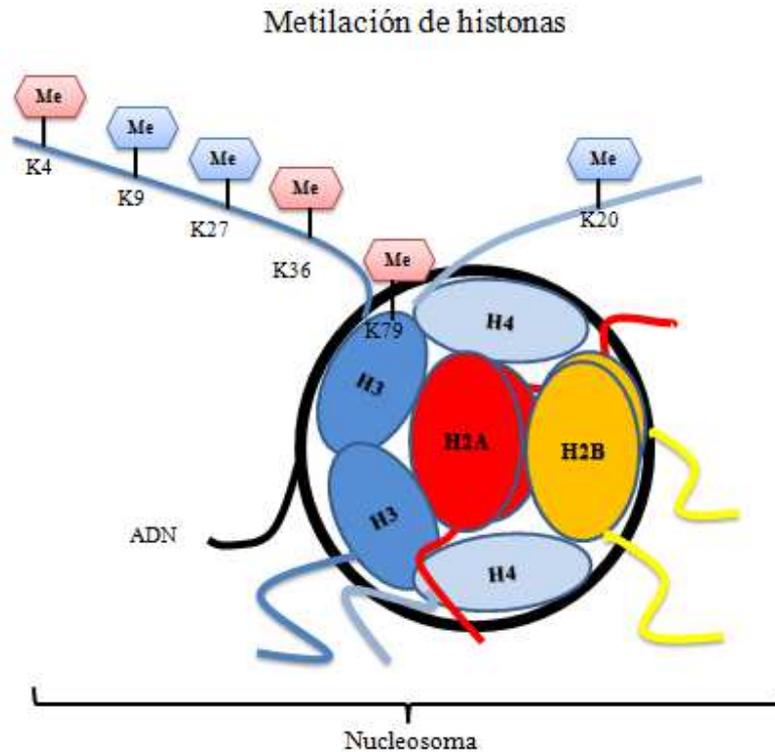


Figura 7. Marcas de metilación en los residuos de las histonas H3 y H4 en nucleosomas. En color rosa se representan las marcas de metilación que indican una activación transcripcional y en color azul las marcas de represión. Modificado desde Inglis *et al.*, 2007, p195.

Se ha visto que la metilación en lisina 4 de la histona 3 (H3K4m) se relaciona con la activación transcripcional, al igual que H3K36m y H3K79m. Mientras que las marcas H3K9m, H3K27m y H4K20m se asocian con la represión de la transcripción. Además se ha visto que la tri-metilación de H3K9, es catalizada principalmente por SUV39 (cuadro 5) (Inglis *et al.*, 2007; Selvi 2010).

Cuadro 5. Metilación en histonas, enzimas responsables y función de la modificación. Modificado desde Selvi *et al* (2010).

Modificación	Enzima	Función
Metilación de Lisina		
H3K4	MLL1-6 SET7-9	Activación
H3K9	G9a, EuHMTasa1, ESET SUV39H1 Ash1 RIZ1	Represión de la transcripción Silenciamiento mediado por Rb Activación Represión de tumor
H3K27	EZH2	Desarrollo temprano de células B Inactivación del cromosoma X
H3K36 H3K79 H3K20	SET2, NSD1, SYMD2 DOT1 Pr-SET 7/8	Activación de la transcripción Activación de la transcripción Silenciamiento de la transcripción Condensación mitótica
Metilación de Arginina		
H3R2 H3R8 H3R17 H3R26 H4R3 H4R3	CARM1, PRMT2 PRMT5 CARM1 CARM1 PRMT1 PRMT5	Expresión génica Represión de la transcripción Activación de la transcripción Activación de la transcripción Represión de la transcripción

2.6 Acetilación y desacetilación

El término acetilación describe una reacción de transferencia de un grupo acetilo funcional dentro de un compuesto orgánico. Tanto las proteínas Histonas, como las no histonas pueden estar acetiladas. La acetilación en histonas (H2A, H2B, H3 y H4) es llevada a cabo por las enzimas histona acetiltransferasas (HATs) y consiste en la transferencia de un grupo acetilo del donante acetyl-CoA al grupo amino de los residuos de lisina en la cola N-terminal. La acetilación se correlaciona en general con la activación transcripcional. El proceso contrario por el cual la marca de acetilación se elimina del residuo se denomina desacetilación y es llevada a cabo por las

enzimas histonas desacetilasas (HDACs) y se correlaciona con la represión transcripcional. Esta modificación altera la carga de las proteínas (Selvi *et al.*, 2010; Bannister y Kouzarides, 2011).

2.6.1 Enzimas acetiltransferasas y desacetilasas de histonas. La reacción de acetilación es catalizada por las enzimas histona acetiltransferasas (HATs), que se pueden clasificar según su estructura en tres familias diferentes, la familia GNAT, la familia MYST y la familia CBP/p300. Si bien las proteínas HAT pueden acetilar residuos de lisina en las cuatro histonas, las diferentes familias de enzimas poseen distinta especificidad de sustrato. Al parecer las proteínas HAT miembros de GNATs tienen una mayor especificidad para acetilar residuos de lisina de H3 como sustrato principal. Mientras que la superfamilia MYST prefieren como sustrato principal los residuos de lisina de H4. Finalmente la familia p300/CBP se dirige tanto a los residuos de lisina de H3, como de H4 (Sterner y Berger, 2000; Kouzarides, 2007a).

La superfamilia GNATs es muy numerosa y sus miembros se encuentran distribuidos en los distintos tipos de organismos. Esta familia posee cuatro motivos compartidos (A, B, C y D) con diferente grado de conservación y que también incluye acetiltransferasas no histonas (AT), el motivo más conservado es el de reconocimiento y unión a acetil-CoA en las HAT, siendo GCN5, PCAF, Hat1, Elp3 y Hpa2 los miembros más relevantes de HATs (Bannister y Kouzarides, 2011).

La superfamilia MYST debe su nombre a sus miembros fundadores MOZ, Ybf2/Sas3, Sas2 y Tip60 y recientemente se han unido MOF, Esa1, HBO1 y MORF. Estas enzimas se agrupan por su similitud de secuencia y por que poseen una región homóloga acetiltransferasa particular (parte del motivo A de la superfamilia GNAT) que se une a acetil-CoA, además de un dominio de dedos de zinc (C2HC) y un motivo “E-R” (Esa1-Rpd3), que son necesarios para el reconocimiento del sustrato y actividad enzimática. A pesar de su semejanza estructural los miembros de esta familia tienen varias funciones en diversos organismos (Selvi *et al.*, 2010).

Las proteínas p300 y CBP también están consideradas en la acetilación de las histonas y son el tercer tipo de HATs, ya sea actuando juntas o separadas, algunos estudios indican que son co-activadores de la acetilación preferentemente de K12 y K15 en H2B, de K8 en H4 y de K5 (y también K14, K18 y K56) en H3 (Schiltz, 1999; Bannister y Kouzarides, 2011).

Las histonas desacetilasas se agrupan en tres familias llamadas clase I, clase II y clase III, en general estas familias no parecen mostrar especificidad por desacetilar algún residuo determinado, aunque se sabe que en levadura la enzima Hda1 desacetila H3 y H2B, mientras que Hos2 actúa en H3 y H4 (Kouzarides, 2007a).

2.6.2 Residuos acetilados en histonas. En general la acetilación en Histonas se asocia con la activación de la transcripción, aunque también se investiga su papel en los procesos de replicación y reparación del ADN. Las marcas de acetilación que se conocen se muestran en la figura 8 y en el cuadro 6 se menciona las acetiltransferasas involucradas en la acetilación de estos residuos y la función asociada a la modificación (Inglis *et al.*, 2007).

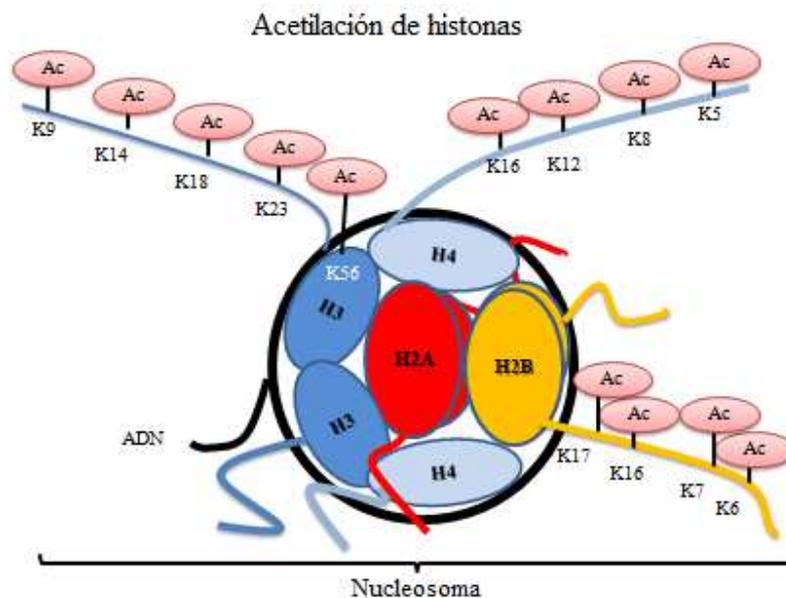


Figura 8. Marcas de acetilación en residuos de las histonas H4, H3 y H2B en un nucleosoma. Modificado desde Inglis *et al.*, 2007, p195.

Cuadro 6. Acetilación en histonas, enzimas responsables y función de la modificación. Modificado desde Selvi *et al* (2010).

Modificación	Enzima	Función
H2AK5 H3K9	TIP60, p300/CBP, Hat1 SRC1	Activación de la transcripción Receptor de señales
H3K14	PCAF, p300/CBP	Activación de la transcripción
H3K18	TAFII250, p300/CBP, PCAF, TIP60	Activación de la transcripción
H3K23	p300/CBP, PCAF	Activación de la transcripción
H3K56	CBP	Activación de la transcripción
H4K5	p300/CBP	Activación de la transcripción
H4K8	Hat1, HBO1, p300/CBP, ATF2, TIP60	Deposición de histonas Activación de la transcripción
H4K12	HBO1, ATF2, TIP60, PCAF	Activación de la transcripción
H4K16	p300/CBP, Hat1, HBO1	Activación de la transcripción
	p300/CBP, PCAF, TIP60, Mof, ATF2	Activación de la transcripción y estructura de la cromatina

2.7 Gen *Egr1*

El gen *Egr1* pertenece a la familia de los genes de respuesta temprana de crecimiento (*Egr*) y fueron descritos por primera vez en 1987 por Sukhatme *et al*, quienes observan que esta familia de genes, hasta entonces desconocidos a la cual llamaron “*egr*” (del inglés “*Early Growth Response*”) es altamente inducible por diversos mitógenos en una amplia gama de tipos celulares (como por ejemplo en hepatoma de rata), se activan con rapidez y de forma transitoria, son conservados evolutivamente, por lo que dedujeron que deberían jugar un papel clave en el crecimiento celular, en la regulación de la proliferación celular y tal vez en la oncogénesis, ya que compartía las propiedades de *c-myc* y *c-fos* (únicos genes conocidos hasta ese momento con las características antes mencionadas). Hoy se sabe que El gen *Egr1*, pertenece al tipo de genes que se activan con rapidez y de forma transitoria denominados IEGs (genes inmediato tempranos), por lo que su transcripción ocurre a los pocos minutos de su activación (Poirier *et al*, 2008).

La familia *egr* codifica factores transcripcionales con motivos de unión al ADN de dedos de zinc y está integrada por *Egr1*, *Egr2* (o *Krox-20*), *Egr3* (o *Pilot*) y *Egr4* (o *NGFI-C*). Entre ellos presentan un 90% de homología en el dominio dedos de zinc Cys2-Hys2. El hecho de que los cuatro miembros de la familia compartan el sitio de unión al ADN sugiere que pueden unirse a regiones de regulación en *cis* de al menos un subconjunto de los mismos genes diana. Los miembros de la familia difieren en el resto de su secuencia, rescatándose pocos dominios conservados fuera del dominio dedos de zinc (como por ejemplo el dominio de unión al represor Nab), lo que indica que pueden ser regulados por diferentes vías y por lo tanto cumplen funciones diferentes. (Sukhatme, 1990; Bozon *et al.*, 2002; Poirier *et al.*, 2008).

Muthukkumar *et al.* (1995) demostraron que *Egr1* se encuentra involucrado en el control de crecimiento celular y en apoptosis, posteriormente Nair *et al.* (1997) presentaron evidencia sobre el papel de *Egr1* como transactivador del promotor del gen p53 regulando su transcripción. Hasta ahora se ha demostrado la importancia del papel que juega *Egr1* en apoptosis en distintas líneas celulares, una investigación reciente realizada por Xie *et al.* (2011) demuestra que *Egr1* también posee un papel importante en la apoptosis en neuronas. Yu *et al.* (2004) demostraron que *Egr1* además de apoptosis y crecimiento se encuentra involucrado en la supervivencia celular. También es importante como modulador de la respuesta inmune (Gómez *et al.*, 2010), en plasticidad sináptica y memoria (Poirier *et al.*, 2008) y otros procesos que actualmente se investigan.

Egr1 también se conoce como *A530045N19Rik*, *Egr-1*, *ETR103* (Shimizu *et al.* 1992), *Krox-1*, *Krox-24*, *Krox24*, *NGFI-A*, *NGFI-A*, *NGFIA* (DeFranco *et al.* 1993), *TIS8* (Sakamoto *et al.*, 1991), *Zenk* (Mello y Clayton, 1994), *Zfp-6* (Darland *et al.*, 1991), *Zif268* (Christy y Nathans, 1989). Se localiza en el cromosoma 18 de *Mus musculus* y se encuentra flanqueado por los genes *D18Mit70* y *Gm17507*. En humanos *Egr1* se encuentra en el cromosoma 5 (q31.1) en el que posee los sinónimos *AT225*, *GOS30*, *KROX-24*, *NGFI-A*, *TIS8*, *ZIF-268*, *ZNF225*. En *mus musculus* *Egr1* posee 3778 pares de bases y su ARN mensajero 3100 pb, mientras que la proteína codificada es de 533 aminoácidos (base de datos MGI, N° de identificación MGI: 95295)¹.

1. http://www.informatics.jax.org/searches/accession_report.cgi?id=MGI:95295

La proteína EGR1 de ratón posee una alta homología (fig. 9) con la de humanos, por lo que hace del ratón un modelo de estudio adecuado para extrapolar los resultados obtenidos (Deguchi *et al*, 2009).

Comparación entre la secuencia de EGR1 de humanos y ratón

<i>Homo sapiens</i> EGR1	1 : MAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMLLSNGAPQFLGAAGAPEGSGS	60
<i>Mus musculus</i> EGR1	1 : MAAKAEMQLMSPLQISDPFGSFPHSPTMDNYPKLEEMLLSNGAPQFLGAAGTPEGSGG	60
<i>Homo sapiens</i> EGR1	61 : NSSSSSSGGGGGGSSNSSSSSTFNPOADTGEQPYEHLTAESFPD SLNNEKVLVETS	120
<i>Mus musculus</i> EGR1	61 : NSSSSTSSGGGGGG—SNGSSAFNPQGEPEQPYEHLTTESFSD ALNNEKAMVETS	117
<i>Homo sapiens</i> EGR1	121 : YPSQTTRLPP TYTGRFSLEPAPNSGNTLWPEPLFSLVSGLVSMNPPASSSSAPSPAAS	180
<i>Mus musculus</i> EGR1	118 : YPSQTTRLPP TYTGRFSLEPAPNSGNTLWPEPLFSLVSGLVSMNPPSSSSAPSPAAS	177
<i>Homo sapiens</i> EGR1	181 : SAS-ASQSPP—————LSCAVPSNDSSP YSAAPTFTPTNTD FPEPQSQAF	225
<i>Mus musculus</i> EGR1	178 : SSSASQSPP—————LSCAVPSNDSSP YSAAPTFTPTNTD FPEPQSQAF	223
<i>Homo sapiens</i> EGR1	226 : PGSAGTALQYPPPAYP—AAKGGFQVPM PDYLFQQQGDLSLGTDPQKPFQGLERSTQ	282
<i>Mus musculus</i> EGR1	224 : PGSAGTALQYPPPAYP—ATKGGFQVPM PDYLFQQQGDLSLGTDPQKPFQGLERSTQ	280
<i>Homo sapiens</i> EGR1	283 : QPSLTPLST KAFATQSGSQDLKALNTSYQSQL KPSRMRKYPNRPSKTPPHERPYACPV	342
<i>Mus musculus</i> EGR1	281 : QPSLTPLST KAFATQSGSQDLKALNTTYQSQL KPSRMRKYPNRPSKTPPHERPYACPV	340
<i>Homo sapiens</i> EGR1	343 : ESCDRRF SRSDLTRH R HTGQKPFQCR CMRNF SRSDHLTTH RTHTGEKPFACD CG	402
<i>Mus musculus</i> EGR1	341 : ESCDRRF SRSDLTRH R HTGQKPFQCR CMRNF SRSDHLTTH RTHTGEKPFACD CG	400
<i>Homo sapiens</i> EGR1	403 : RKFARSDERKRHTK HLRQKDKKADKSVVASSATSSLS—SYSPVATSYSPVTTSTYP	459
<i>Mus musculus</i> EGR1	401 : RKFARSDERKRHTK HLRQKDKKADKSVVASPAASSLS—SYSPVATSYSPS—	450
<i>Homo sapiens</i> EGR1	460 : SPATTSYSPVPTSFSSPGSSTYSPVHSGFSPSVATTYSSVPP—AFPAQVS—SFPSS	516
<i>Mus musculus</i> EGR1	451 : -PATTSYSPVPTSYSSPGSSTYSPAHSFGFSPSVATTAFASVPP—AFPTQVS—SFPSS	506
<i>Homo sapiens</i> EGR1	517 : AVTNSFS—ASTGLSDMTATFSPT E C	543
<i>Mus musculus</i> EGR1	507 : GVSSFS—TSTGLSDMTATFSPT E C	533

Figura 9. Comparación entre la proteína EGR1 de *Homo Sapiens* y *Mus musculus*. Modificado desde Deguchi y *et al.* (2009).

La proteína EGR1, consta de dominios funcionalmente independientes (Fig. 10), la región N-terminal es rica en serina y treonina, la región C-terminal es además rica en prolinas. El dominio central de unión al ADN (DBD) se encuentra desde el aminoácido 331 al 419 y consiste en tres motivos repetidos de dedos de zinc que se caracterizan por unirse a secuencia de promotores ricas en CG, por lo que se nombran como secuencia de unión Egr (EBS). El dominio dedos de zinc presenta una estructura antiparalela de hoja beta plegada y alfa hélice. La proteína además posee una secuencia bipartita de localización nuclear (NLS), que consta de 15 aminoácidos en de la región flaqueante N-terminal dentro de DBD. La proteína presenta además un dominio activador

y uno represor, un co-represor nuclear llamado proteína de unión NGFI-A (NABs) se adhiere al dominio represor (ubicado desde el aminoácido 281 al 314) regulando negativamente la actividad transcripcional de *Egr1* (Pagel y Deindl, 2011).

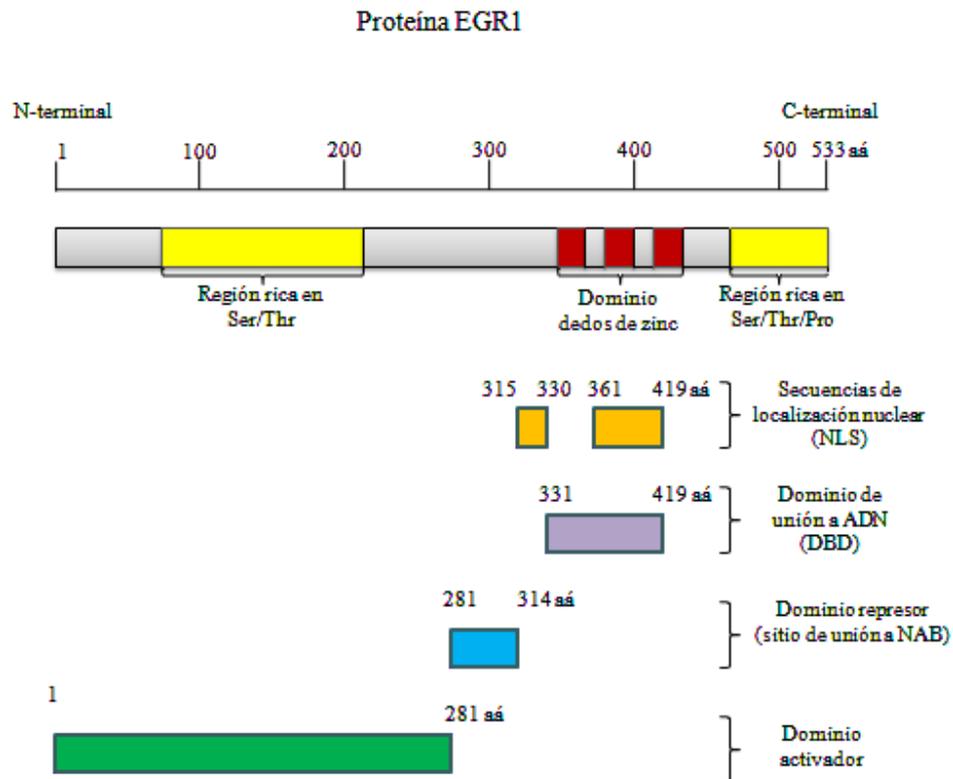


Figura 10. Dominios funcionales de la proteína EGR1 y su localización a lo largo de la cadena de aminoácidos (aa). Modificado desde Pagel y Deindl, (2011).

En el promotor del gen *Egr1* (fig. 11) se han identificado y caracterizado varios elementos funcionales, al lado de la tentativa caja TATA se localizan cinco elementos de respuesta a suero (SRE), adyacente a los SER se ubican cinco miembros de la familia de factores transcripcionales Ets, ambos se organizan en dos grupos. Además hay dos elementos de respuesta a adenosina

monofosfato cíclico (CREs), un sitio de unión a APETALA1 (AP1), dos sitios de unión a Sp1 (proteína específica 1) y el sitio de unión donde se une EGR1 (EBS) (Pagel y Deindl, 2011).

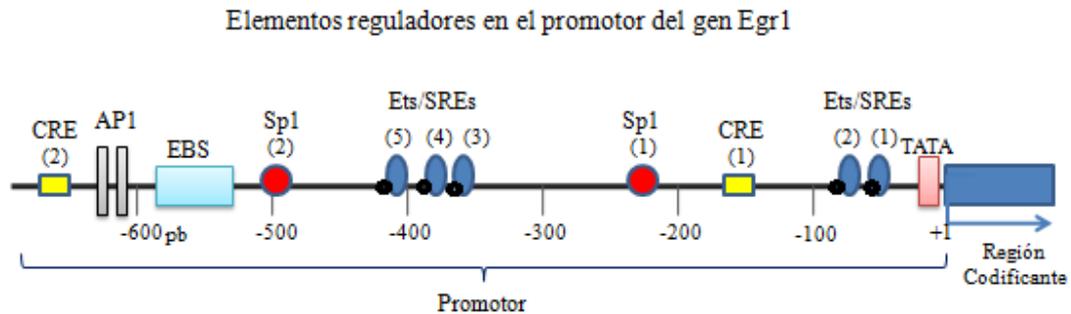


Figura 11. Elementos en el promotor de *Egr1*. En óvalos azules se muestran los cinco elementos SRE y en color negro los 5 elementos Ets adyacentes. En amarillo se indica la ubicación de los dos elementos CRE y en rojo los dos Sp1, además de la ubicación de AP1 y EBS. Modificado desde Pagel y Deindl (2011).

El mecanismo de activación en *Egr1* es mediado con mayor frecuencia por la vía de señalización MAP-K (Fig. 12). Comúnmente las familias ERK1/2, JNK y p38maP-K se encuentran involucradas en esta vía. La familia ERK1/2 media la expresión de *Egr1* en respuesta a factores de crecimiento, mientras que una combinación entre las tres familias induce la expresión de este gen en respuesta al estrés. La activación de *Egr1* comienza cuando los factores activados de la vía MAP-K se traslocan al núcleo, fosforilando y activando la transcripción de los factores de las familias Elk-1/SAP-1/2. El factor Elk-1 forma un complejo ternario con la proteína de unión a CREB (CBP) y el factor de respuesta al suero (SRF), el cual se une y transactiva al promotor del gen *Egr1*. Una vía alternativa es la activación por NFκB en respuesta a radiación UV. Otro estudio indica que es p53 el que se une al promotor de *Egr1* en respuesta al daño del ADN y conduce a apoptosis (Boone *et al.*, 2011). Por último EGR1 se une a su propio promotor y reprime la transcripción (Gitenay y Baron, 2009).

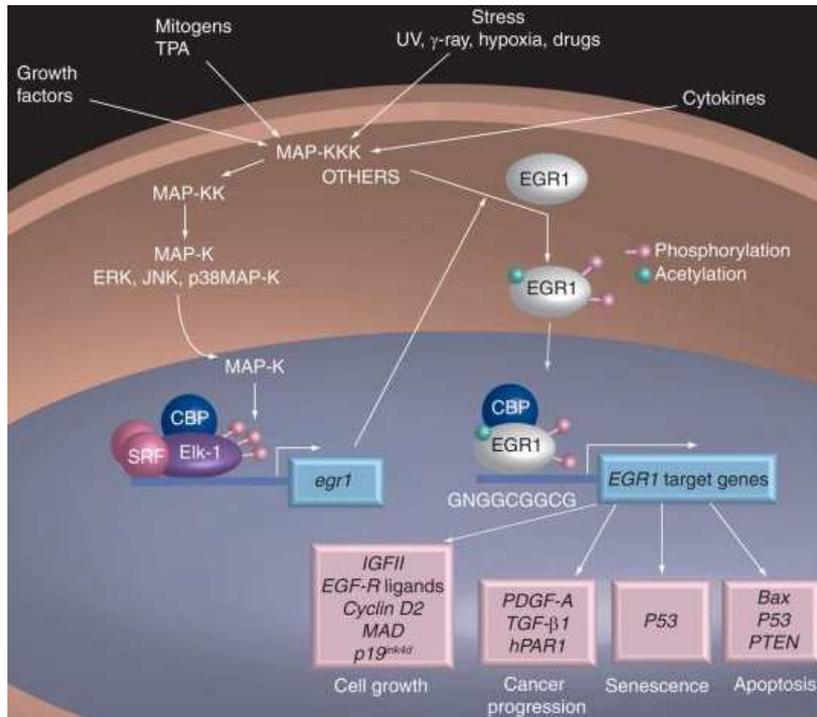


Figura 12. Vía de señalización MAP-K en la activación del gen *Egr1*. Extraído desde Gitenay y Baron (2009).

Debido a la importancia de *Egr1* en las vías de señalización intracelulares, está implicado en un gran número de enfermedades, entre la cuales se encuentra la esclerosis (Bhattacharyya *et al*, 2011a), la enfermedad de Parkinson (Ebihara *et al* 2011), cáncer epitelial (Kataoka *et al*, 2011), cáncer de mama (Einbond *et al*, 2010), aterosclerosis (Hamada *et al*, 2011), la enfermedad proliferativa vascular (Fukuda, 2010), entre muchas otras.

Este gen es de gran importancia, no solo para caracterizar procesos como desdiferenciación celular, apoptosis, respuesta inmune, entre otros, sino también por su posible uso a nivel terapéutico en enfermedades como fibrosis (Bhattacharyya *et al*, 2011b), terapia contra el cáncer (Gitenay y Baron, 2009; Yang *et al*, 2010) entre otras, pero el mecanismo involucrado en su transcripción a nivel de cromatina no se encuentra caracterizado.

2.8 Posicionamiento y dinámica nucleosomal

Los cambios o modificaciones que sufren las histonas y el ADN en ocasiones traen como consecuencia una alteración en la estructura del nucleosoma, pudiendo afectar la transcripción, al impedir o favorecer la entrada de los factores involucrados, o bien al no permitir que la elongación continúe en un cierto punto dentro del gen (Herman *et al*, 2006). Pero en otros casos, las modificaciones epigenéticas sirven para reclutar remodeladores de la cromatina que pueden provocar un desplazamiento del nucleosoma que modifica la accesibilidad de determinadas secuencias del ADN (Längst y Becker, 2001). De este modo, la posición que ocupa un nucleosoma constituye un elemento fundamental en la regulación de la expresión génica (Cui *et al*, 2010).

El posicionamiento del nucleosoma no es aleatorio, lo que indica un grado de especificidad en el modo en que la secuencia de ADN se enrolla en el octámero de histonas. Teniendo en cuenta las preferencias de secuencia que experimentalmente se observan en el ADN nucleosomal, así como los parámetros que intervienen en la estructura del nucleosoma, se han elaborado programas bioinformáticos con el fin de determinar la posición dentro de la secuencia de ADN que tiene una mayor probabilidad de unirse al octámero de histonas. Puesto que el nucleosoma tiene importancia no solo en la organización estructural de la cromatina, sino también en la regulación de la accesibilidad de los complejos proteicos implicados en la transcripción del ADN, la elaboración de nuevos programas bioinformáticos más precisos contribuirá a la investigación en epigenética, al proporcionar una importante herramienta predictiva.

2.8.1 Fundamento bioinformático en la predicción nucleosomal. La estructura del ADN depende sutilmente de la secuencia. Las clásicas estructuras canónicas A y B representan sólo situaciones extremas, que se encuentran en algunas regiones concretas del ADN. Pero, de ordinario, la estructura del ADN presenta una gran microheterogeneidad, en función de la secuencia. Esto se debe a que las variables conformacionales que definen la posición relativa de

dos pares de bases consecutivos dependen de la naturaleza de las bases. En algunos casos, como por ejemplo en las secuencias AA/TT, las bases tienden a situarse perfectamente perpendiculares al eje de la doble hélice, con los parámetros característicos de la estructura B de Watson y Crick. En otros casos, la secuencia de dos pares de bases consecutivos favorece que se sitúen en cuña, abiertos hacia el surco ancho. Hay también secuencias que prefieren la cuña abierta hacia el surco estrecho y otras que pueden adoptar diversas conformaciones, confiriendo, por tanto, una cierta flexibilidad a la estructura.

Aunque los nucleosomas pueden situarse al azar sobre el ADN, en gran parte de los casos se encuentran posicionados, es decir, el octámero de histonas interacciona con la misma secuencia de 147 pb en todas las moléculas de ADN (de las diferentes células que existan en la muestra) o en todas las regiones de idéntica secuencia dentro de la misma molécula de ADN. Hay dos causas que determinan el posicionamiento de nucleosomas. La primera es la existencia de regiones excluyentes de nucleosomas. A veces se trata de secuencias a las que se une una proteína tan fuertemente que impide, por competencia, la unión de un octámero para ensamblar el nucleosoma. Otras veces, el motivo excluyente es una secuencia relativamente larga de ADN que, por su rigidez o por otras características, impide la adquisición de la curvatura requerida para enrollarse alrededor del octámero de histonas.

La segunda causa de posicionamiento es la debida a la secuencia del ADN en una región, que puede implicar, o al menos permitir, la curvatura necesaria para enrollarse alrededor de un octámero. Hay que tener en cuenta que deformar el ADN para curvarlo alrededor del octámero consume energía, que se compensa por la liberada al establecerse las interacciones histona-ADN en el nucleosoma. Pero si el ADN está intrínsecamente curvado con el diámetro adecuado para formar un nucleosoma, o al menos puede hacerlo, su ensamblaje será claramente favorable, ya que desaparecería el componente energético endergónico.

Los motivos excluyentes de nucleosomas se suelen denominar determinantes translacionales de posicionamiento. La razón es que imponen una barrera a la presencia de nucleosomas, de modo que si el borde de esa barrera está en la posición i , el borde del nucleosoma puede

encontrarse en i , en $i + 1$, en $i + 2$, etc. pero nunca en $i - 1$, $i - 2$, etc. Dicho de otro modo, el nucleosoma podría idealmente deslizarse en el sentido opuesto a la barrera, pero nunca podría atravesarla. La situación es distinta cuando el nucleosoma está posicionado por la curvatura inducida por la secuencia del ADN. Un segmento de ADN que se curva hacia la derecha no puede hacerlo hacia la izquierda. Por eso, si un nucleosoma se ensambla con su eje diádico en la posición i , podría hacerlo también en la posición $i + 10$ o $i - 10$, pero no en las posiciones $i \pm 5$, $i \pm 4$, etc. Por este motivo, se suele hablar de los motivos de secuencia como determinantes rotacionales.

Experimentalmente, al analizar las secuencias de ADN procedentes de nucleosomas, se ha comprobado que las secuencias AA/TT se organizan de forma periódica cada 10pb, con el surco menor del ADN orientado hacia la superficie de las histonas, mientras que las secuencias GC/GC tienden a situarse en la cara externa del ADN y no interactúan con las histonas (Battistini et al, 2010). Además, las regiones que confieren flexibilidad al ADN se encuentran en abundancia en la zona del eje diádico, donde la curvatura del ADN en el nucleosoma es máxima y llega a alcanzar los 50° .

El análisis estadístico de la frecuencia de esos motivos de secuencia en las diversas localizaciones de los nucleosomas es la base de algunos programas de predicción de posicionamiento de nucleosomas. En ellos, se elabora un algoritmo que permite analizar una secuencia dada en busca de esos motivos característicos, con lo que se puede llegar a predecir la localización más probable de un nucleosoma basada en determinantes rotacionales. También se pueden analizar las secuencias en búsqueda de determinantes traslacionales, como puede ser la existencia de secuencias poli(dA/dT), que, por su rigidez, excluyen nucleosomas.

Como una variante de estos procedimientos de predicción, Teif y Rippe (2009) desarrollaron un modelo en *Drosophila* en el que, además de las características de la secuencia, también toma en cuenta la afinidad con los remodeladores (a los cuales clasifican en amplificador, removedor y espaciador).

Otro parámetro que se puede considerar en la predicción de la posición que ocupa el nucleosoma en el ADN es el costo energético que conlleva flexionar la molécula de ADN. Como se ha comentado antes, ese componente energético depende de la secuencia del ADN. Como es evidente que un nucleosoma tenderá a colocarse en la posición de mínima energía, el cálculo teórico de la energía necesaria para curvar el ADN con los parámetros requeridos para que cada segmento de 147 pb pueda envolver un octámero de histonas, permite encontrar la posición más favorable desde un punto de vista teórico.

Se sabe que regiones con repeticiones cortas de A/T en la ranura menor y G/C en la ranura mayor, permiten que el ADN pueda doblarse con mayor facilidad, cuando la región rica en A+T se organiza de forma periódica cada 10pb (en donde se observa una curvatura de 30°), la ranura menor del ADN se encuentra orientada a la superficie de las histonas, mientras que la región rica en G+C no interactúa con las histonas (Battistini *et al.*, 2010). Además, estudios en eritrocitos de pollo y en levadura encontraron dinucleótidos AA/TT cada 10 pb y GC cada 5 pb en la zona del nucleosoma. Estas regiones que le confieren flexibilidad al ADN se enriquecen en la zona de la diada (zona de mayor curvatura del ADN en el nucleosoma alcanzando los 50°). Siendo el dinucleótido TA cada 67 pb el que indica con mayor frecuencia la presencia de un nucleosoma y ubicándose a 15pb del centro del nucleosoma, luego le siguen (a distancia de repetición diferente) los dinucleótidos TG...N...TT con mayor frecuencia que TC, mientras que este último con la misma frecuencia que GG, GA y AT (Fernández y Anderson, 2007).

Otra característica importante a tener en cuenta es la composición de bases, ya que el contenido en C+G es diferente entre el nucleosoma y la secuencia enlazadora, existiendo un mayor contenido de A+T en la secuencia enlazadora y un mayor contenido de C+G en el nucleosoma. Las secuencias polidA-dT impiden la unión de octámeros de histonas, porque adoptan una estructura rígida, incapaz de curvarse. Al igual que en la caja TATA del promotor de un gen. También se ha descrito el motivo GRAAATTTYC que se repite en los nucleosomas de eucariotas, en base a lo anterior Xing *et al.* (2011) propusieron un parámetro $M_k(i)$ para analizar las frecuencias de las zonas repetidas en nucleosomas y en la secuencia enlazadora en el genoma de *S. cerevisiae*. Sus resultados indican que los ocho dinucleótidos más frecuentes en los

nucleosomas son AA, TT, AT, TG, CA, CT, GA y AG, mientras que en la secuencia entre nucleosomas son TT, AA, AT, TA, TG, CA, GA y TC, sin diferencias significativas entre ambas. Los tetranucleótidos más frecuentes son AAGA, TCTT, TTCT, AGAA, TGAA, TTCA, TCAA y TTGA en el nucleosoma y AAAA, TTTT, ATTT, AAAT, ATAT, AATA, TATT y TATA en la secuencia entre nucleosomas, siendo más frecuentes A y T en la zona de enlace y en el nucleosoma G y C.

Otro parámetro que se debe considerar en la determinación de la posición que ocupa el nucleosoma en el ADN es el costo energético que conlleva flexionar la molécula de ADN. Se sabe que el ADN se enrolla alrededor del octámero de histonas en una super-hélice plana, además un nucleosoma puede formarse en casi cualquier secuencia de ADN, pero el costo energético es dependiente de la afinidad del ADN con el octámero de histonas, por lo que adoptaría la forma de mínima energía. Por lo observado en la composición de bases que forman al nucleosoma se cree que las zonas ricas en GC pueden flexionarse formando el nucleosoma con un menor costo energético. En base a esto Chung y Vingron (2009), observaron que el enriquecimiento periódico (10pb) de AT (surco menor) y GC (surco mayor) es una característica reproducible en todos los nucleosomas del ADN, además determinaron que el costo energético de la traslación (78% de la energía total) en la formación del nucleosoma es menor cuando la secuencia es rica en GC, mientras que el costo de rotación es menos dependiente de la composición de la secuencia.

Teif y Rippe (2009) desarrollaron un modelo de predicción de nucleosomas que además de las características de la secuencia también toma en cuenta la afinidad con los remodeladores (a los cuales clasifican en amplificador, removedor y espaciador) en *Drosophila*.

2.8.2 Programas bioinformáticos para la predicción de nucleosomas En base a todos los patrones descritos anteriormente es posible diseñar programas que predicen la posición del nucleosoma dentro del genoma y crear bases de datos de nucleosomas. Un ejemplo es el programa “**nuScore**” desarrollado por Tolstorukov *et al.* (2008), el cual se basa en las en la estimación del costo energético que supone curvar al ADN alrededor del octámero de histonas

(E) y de este modo predice el posicionamiento (P). El costo energético depende de la composición de la secuencia y su interacción directa con las histonas.

La interfaz (fig. 13) de este programa, dispone de dos formatos para introducir la secuencia de nucleótidos, además se debe especificar el tipo de nucleosoma (145, 146 y 147 pb) y el organismo al cual pertenece la secuencia de inicio, además de la orientación de la misma.

nuScore
Deformation energy and nucleosome-positioning score calculator

Sequence format Fasta Plain text [i](#) ← Formato de la secuencia

Copy and paste sequences for analysis in the text box below
Sequences should be no shorter than the size of the window for nucleosome-positioning score calculations (see below)

OR

Submit a file containing sequences for analysis
This is a preferred option if the number of sequences is large.

Examinar...

OR

Submit an example sequence (sea urchin 5S rDNA)

Calculate average and standard deviation [i](#)

Random sequences: [i](#)
Adding random sequences will increase the computation time

Template: 1kx5 (ncp147, xenopus) - best-resolved Complete [i](#) ← Tipo de nucleosoma de referencia

Symmetrical

Direct orientation Reverse orientation Average Best of two [i](#) ← orientación de la secuencia de entrada

Window size for nucleosome-positioning score calculations (in base pairs): 165 [i](#)
No less than the selected template size. Recommended window size: Template size + 18 bp
ODD number for ncp147 and ncp145 templates; EVEN number for ncp146 templates

Submit Data Reset

Figura 13. Vista general de la interfaz de nuScore. En donde se indican las opciones que posee el programa.

Una vez que los parámetros de inicio se definen, el programa entrega el análisis mediante los dos parámetros antes mencionados energía (fig. 14a) y posicionamiento (fig 14b), mostrando en los archivos de salida la desviación media de los datos para cada parámetro, adicionalmente calcula el porcentaje CG de la secuencia de inicio.

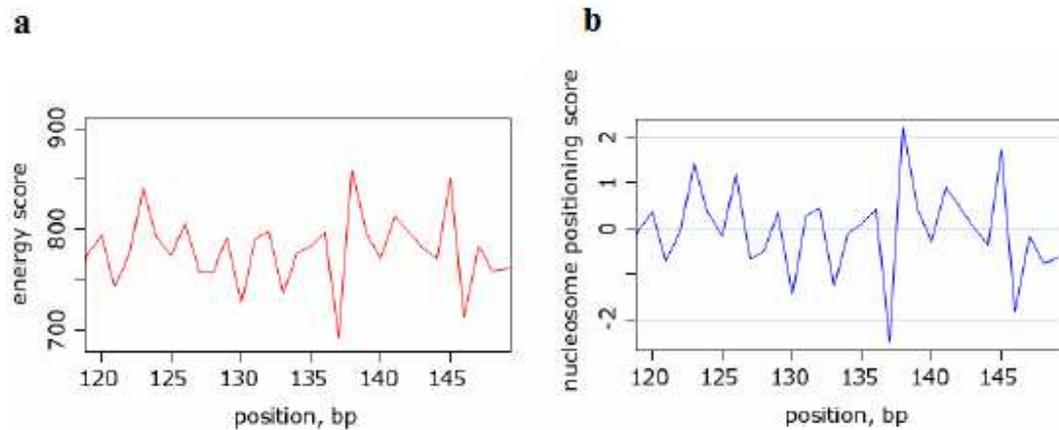


Figura 14. Gráficos de salida de nuScore, a) puntuación de energía y b) puntuación de posición del nucleosoma.

El programa NuPoP (del inglés “Nucleosome Positioning Prediction Engine”), fue desarrollado por Xi *et al.* (2010) y utiliza el modelo de cadenas de Markov (HMM) de primer y cuarto orden, teniendo en cuenta dos estados, el nucleosoma (N) y el enlazador (L), el nucleosoma tiene una longitud fija de 147 pb y el enlazador una longitud variable. El programa se encuentra disponible para Windows XP, Linux y Mac OSX, además se ofrece una interfaz web (fig. 15), que permite secuencias de hasta 500 Kb, se debe especificar la especie a la cual pertenece la secuencia, ya que si bien este programa se basa en la composición de bases en levadura, se reestructura para otras especies. Los resultados luego del análisis de la secuencia se presentan de dos formas, en un gráfico (fig. 16a) que en azul indica la probabilidad de que el inicio de un nucleosoma se encuentre en ese nucleótido, y en rojo se muestra el puntaje de ocupación de un nucleosoma. En una tabla (fig. 16b) que muestra por cada nucleosoma los parámetros utilizados (N/P, afinidad, puntaje de inicio y probabilidad de ocupación) para determinar la presencia del nucleosoma.

NuPoP: Nucleosome Positioning Prediction Engine

Species: S. cerevisiae DHMM: 4th order Markov Chain

Please paste in or upload one, and only one, query sequence of length 148-50,000 bp to predict the nucleosome positioning.
 For longer sequences, please use the stand-alone [Fortran program](#) or [NuPoP R package](#).

*FASTA format requires the first line should be ">sequenceName", each line is of the same length (the last line can be shorter).

Figura 15. Interfaz de NuPoP.

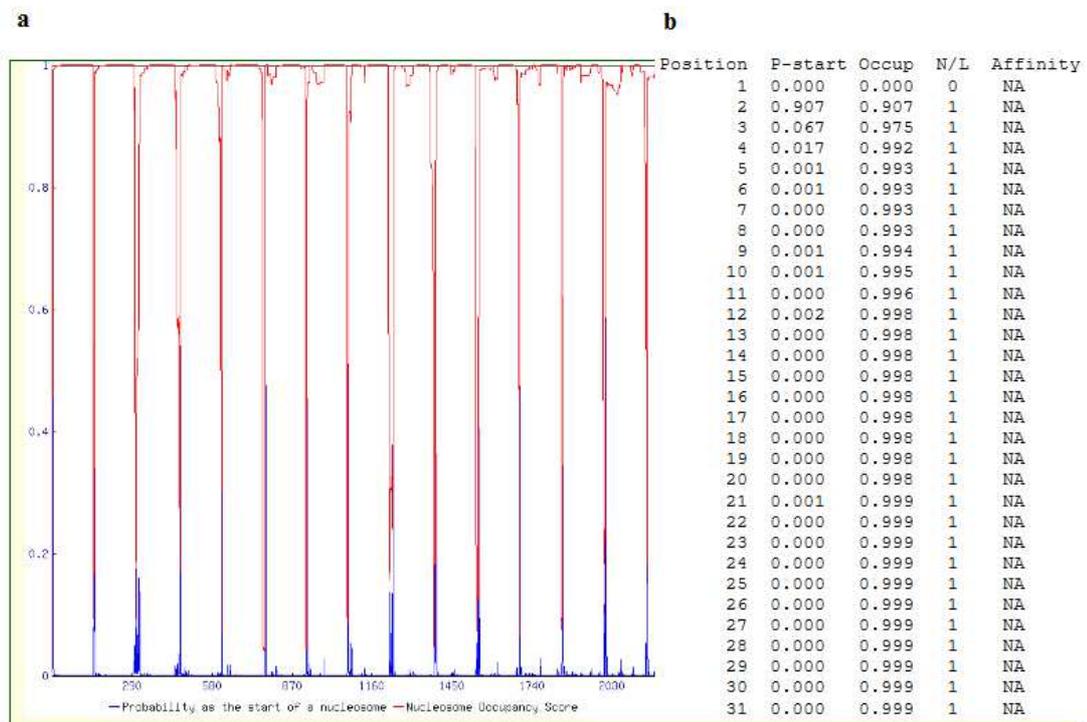


Figura 16. Resultados de NuPoP. a) Grafico de salida, b) Resultados por cada nucleosoma.

El último programa que se describe es NXsensor (del inglés (Nucleosome eXclusion Sensor)), creado por Luykx *et al.* (2006), es una herramienta para encontrar las regiones en la secuencia de ADN en la que es probable que no se formen nucleosomas. Este programa se basa en que la secuencia de ADN debe poseer una cierta flexible para formar el nucleosoma, la curvatura intrínseca de un fragmento del ADN depende de la composición de su secuencia, bajo este principio el programa busca las zonas del ADN que son demasiado rígidas para formar parte del nucleosoma, calculando el porcentajes de A/T y C/G para discriminar.

La interfaz de NXSensor (fig. 17) permite la entrada de secuencias en formato FASTA, además posee la opción de elegir que los resultados muestren la secuencia enlazadora, arrojando los resultados en rojo (Fig. 18a) o la secuencia que posee una mayor posibilidad de formar un nucleosoma, resultados en gris (Fig. 18b) los repetidos en el nucleosoma y en la secuencia enlazadora.

NXSensor
version 1.3.1

- [Home](#)
- [What it is](#)
- [How it works](#)
- [Output help](#)
- [Who we are](#)
- [Thanks](#)

Input DNA sequence(s) in [FASTA](#) format in the window below. Please ensure that:

- Each DNA sequence is preceded by a single header line starting with ">" (without quotes).
- Each header line is no longer than 250 characters, and ends with a new-line character.

Click [here](#) for an example of a valid input sequence. Please note: on some systems, very long sequences may results will be returned anyway.

Please enter your sequence here

Window length:

Minimum number of exclusion sequences considered significant:

Minimum length of open segments:

Output: Exclusion sequences
 Nucleosome segments

Figura 17. Interfaz de NXSensor.

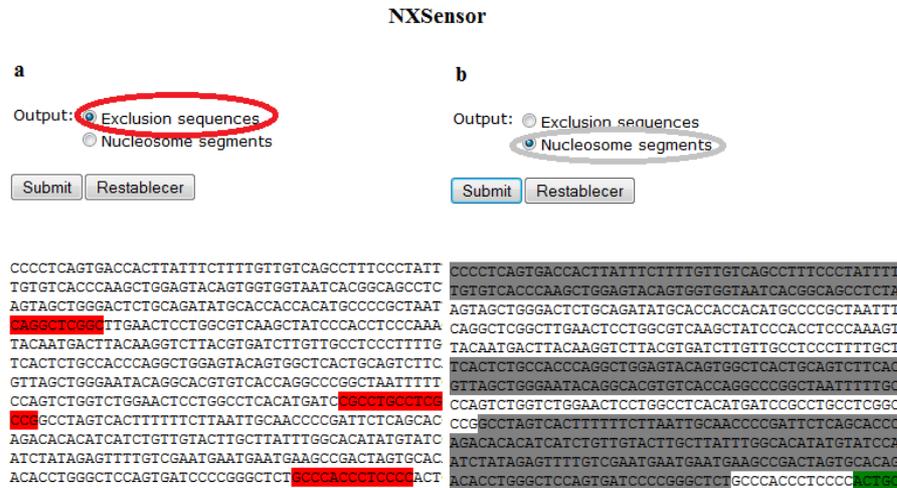


Figura 18. Resultados de NXSensor. a) En rojo muestra las zonas con probabilidad de ser parte de la secuencia de enlace. b) Zonas con mayor probabilidad de formar un nucleosoma.

2.8.3 Bases de datos de nucleosomas. La información sobre la ubicación de los nucleosomas en los distintos genomas se almacena en bases de datos. La base de datos de las regiones de posicionamiento nucleosomal (NPRD) fue desarrollada para almacenar la información sobre los sitios de formación de los nucleosomas y su correlación con la actividad del gen en el que se encuentran y en distintos organismo (fig. 19), para acceder a la base de datos hay que dirigirse al sitio web <http://srs6.bionet.nsc.ru/srs6bin/cgi-bin/wgetz?&-page+LibInfo+-id+ZGe1f5vNU+-lib+NUCLEOSOME> (Levitsky *et al*, 2005).

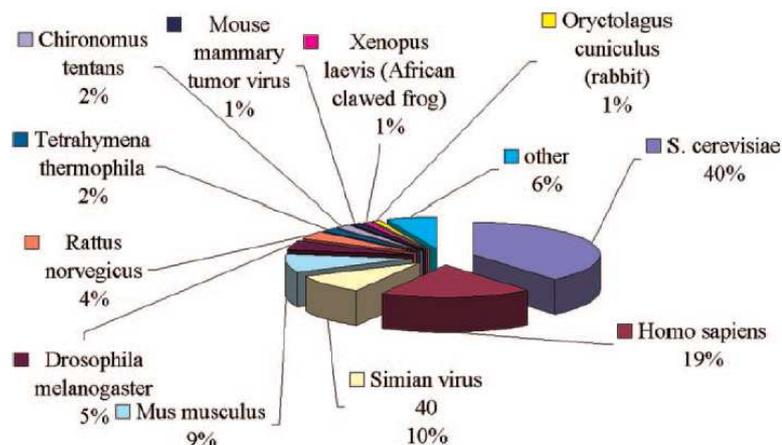


Figura 19. Representación de las especies en la base de datos NPRD. Extraído desde Levitsky *et al.*, 2005.

Otra base de datos de nucleosomas, más específica, es la que se encuentra disponible en GEO (del inglés “Gene Expression Omnibus”) de la NCBI (fig. 20), es esta base de datos se almacenó la información de la posición de los nucleótidos en el ADN genómico de células de hígado de *Mus musculus*. Fue realizada mediante la técnica de digestión con MNasa. En donde la cromatina nativa sin entrecruzar es sometida a 0,1 N de CaCl₂ y luego fue digerida parcialmente con MNasa por 15 min y digerida completamente con MNasa por otros 30 min. Obteniendo de esta manera mononucleosomas del genoma de cuatro hígados de ratón. Posteriormente fueron purificados y tratados con proteinasa K, el ADN fue modificado y sometido a ChIP-seq, después de la unión la cromatina fue sonicada para reducir su tamaño a fragmentos de entre 100 y 1000pb, modificándolo nuevamente para la secuenciación final. Los datos del mapa de nucleosomas se pueden descargar libremente, el archivo de descarga es un documento de texto que contiene la posición (en pares de bases) de los nucleosomas a lo largo del genoma de ratón (fig. 20b). Esta base de datos fue publicada el 30 de abril del 2011 y se creó en el Departamento de Medicina de la Universidad de Pennsylvania (Li *et al.*, 2011).

a)

NCBI > GEO > **Accession Display** [?](#) Not logged in | [Login](#) [?](#)

Scope: Format: Amount: GEO accession:

Sample GSM717558 [Query DataSets for GSM717558](#)

Status Public on Apr 30, 2011
 Title Nucleosomal DNA_Mnase-Seq
 Sample type SRA

Source name Control liver
 Organism [Mus musculus](#)
 Characteristics tissue: liver
 Extracted molecule genomic DNA

b)

Supplementary file	Size	Download	File type/resource
SRX/SRX059/SRX059741		(ftp)	SRA Experiment
GSM717558_nucleosomes.bed.gz	55.2 Mb	(ftp)(http)	BED
GSM717558_ucsc_bed.tar.gz	1.1 Gb	(ftp)(http)	TAR

Figura 20. Base de datos de nucleosomas de hígado de ratón (a). En b se muestra los archivos de descarga de la base de datos.

A pesar de la gran cantidad de datos que se encuentran almacenados en las bases de datos sobre el lugar en donde es más probable que se forme un nucleosoma, las predicciones realizadas por programas bioinformáticos no son confirmadas por los datos experimentales, esto indica que la posición no solo depende de la secuencia, sino que hay otros factores que intervienen en la formación del nucleosoma, los algoritmos bioinformáticos intentan integrar estos factores para lograr una predicción exacta de la ubicación del nucleosoma.

3. HIPOTESIS Y OBJETIVOS

3.1 Hipótesis

La posición de los nucleosomas del promotor proximal y al inicio de la región transcribible y las modificaciones epigenéticas, regulan la transcripción de *Egr1* cuando este es activado con 12-*O*-Tetradecanoilforbol-13-acetato (TPA).

3.2 Objetivo general

Estudiar los cambios epigenéticos y las modificaciones en los nucleosomas particulares de la región promotora proximal y al inicio de la zona transcribible, cuando el gen *Egr1* se expresa.

3.3 Objetivos específicos

- Estandarizar un protocolo de extracción de mononucleosomas a partir de la línea celular MLP29.
- Establecer el posicionamiento nucleosomal en el promotor proximal y al inicio de la región transcribible, utilizando la técnica de PCR en tiempo real.

- Determinar si se producen cambios en la posición o en la cantidad de nucleosomas cuando el gen es activado por éster de forbol 12-*O*-Tetradecanoilforbol-13-acetato (TPA) durante 30 min.
- Estudiar los complejos remodeladores de cromatina que podrían estar implicados si se observase movimiento nucleosomal, mediante la técnica de inmunoprecipitación de mononucleosomas (Nu-ChIP) y posterior cuantificación en PCR a tiempo real.
- Estudiar las modificaciones de acetilación y tri-metilación en Lys9 en la histina H3. Así como la implicancia de alguna histona acetiltransferasa o metiltransferasa, mediante la inmunoprecipitación de mononucleosomas (Nu-ChIP) y posterior cuantificación en PCR a tiempo real.
- Finalmente y a partir del posicionamiento nucleosomal de *Egr1* y de la base de datos de nucleosomas, se plantea la necesidad de crear un nuevo programa para la predicción de la posición de los nucleosomas.

4. MATERIALES Y METODOS

4.1 Línea celular MLP29

La línea de células progenitoras de hepatocitos de ratón MLP29, se cultiva en medio DMEM del inglés “*Dulbecco's Modified Eagle Medium*” (Sigma) suplementado con un 10% de suero fetal bobino 1 (Gibco), 1% de penicilina/estreptomicina (Sigma) y un 1% de fungizona (Gibco), a una temperatura de 37°C y en una atmósfera húmeda que contiene un 5% de CO₂.

4.1.1 Reconstitución de la línea celular. Para descongelar y comenzar el cultivo, el criotubo que contiene las células, se extrae del tanque de nitrógeno, se descongela rápidamente y las células se transfieren a un tubo que contiene medio de cultivo DMEM suplementado. De esta manera se diluye el DMSO que contiene el medio de criopreservación, evitando así que las células se dañen. Las células se centrifugan 5 min a 400 g, se resuspenden en medio DMEM suplementado y se depositan en la botella o placa de cultivo. Al día siguiente se cambia el medio a las células.

4.1.2 Mantenimiento de la línea celular. Las células cultivadas en una botella de 75 cm² (Corning), se lavan con tampón fosfato (PBS de Gibco), posteriormente las células se despegan con 1,5 ml de tripsina-EDTA 0,25% (Invitrogen) tratando 5 min a 37°C, finalmente las células se diluyen como máximo 1/6 en DMEM suplementado. Se cultivan a 37°C y un 5% de CO₂. En el dorso de la botella es anotado el número de pase, la fecha y la línea celular. Este procedimiento se realiza habitualmente cada tres días. Todas las soluciones empleadas deben ser previamente temperadas a 37°C en el baño termo regulado.

4.1.3 Criopreservación de la línea celular. Para la criopreservación las células deben despegarse de la botella de cultivo con tripsina-EDTA, se recuperan en 10 mL de DMEM suplementado y se centrifugan a 400 g 5 min. Se elimina el sobrenadante, se resuspenden en suero fetal bovino con un 10% de DMSO (preparado y enfriado previamente) y se transfiere a criotubos, donde se congelan en un contenedor con 2-propanol a -80°C durante una semana antes de almacenarlas en el tanque de N_2 líquido.

4.1.4 Tratamiento de la línea celular. Las células se despegan tal y como se indica en el apartado 4.1.2, se recuperan en 6 mL de DMEM y se cuenta el número de células. Para el recuento celular se extraen $10\ \mu\text{L}$ del cultivo y se deposita en $90\ \mu\text{L}$ de PBS (factor de dilución 10), extrayendo la cantidad necesaria para cubrir la cámara de Neubauer (fig. 21) y realizar el recuento de células en un microscopio invertido.

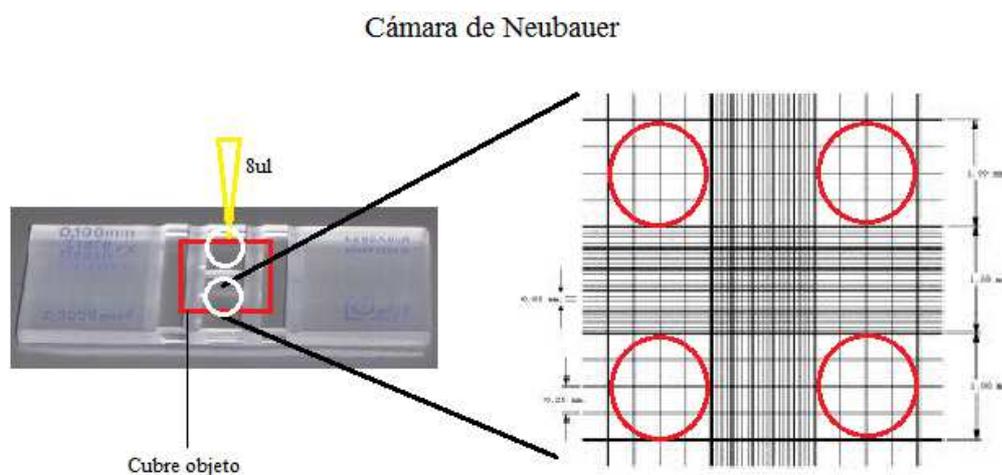


Figura 21. Cámara de Neubauer para el conteo de células. Como se muestra en la figura entre la cámara y el cubre objetos se depositan aprox. $8\ \mu\text{L}$ del preparado celular a cada lado de la división, la zona que se emplea en el recuento está señalada en círculos rojos.

Para calcular el número de células presentes en la muestra (N°) se cuenta el número de células de los cuatro cuadrantes (X), después se divide por 4 para calcular la media, el resultado se

multiplica por el factor de dilución que es 10 y luego por el factor de la cámara que es 10^4 , lo que indica el total de células por mililitro que hay en la muestra.

$$4.1.2 \quad (1) \quad N^{\circ}\text{cel/ml} = ((X/4)*10*10^4)$$

Se depositan $1,5 \times 10^6$ células por cada botella de 75 cm^2 y se llevan a un volumen final de 12 mL con medio DMEM suplementado. Las células son cultivadas a 37°C con un 5% de CO_2 , hasta que alcanzan una confluencia del 70% al 80%, necesaria para privarlas de suero. Para este procedimiento se realiza un lavado con 8mL de PBS y se agregan 12 mL de DMEM suplementado con un 0,5% de suero fetal bovino y se cultivan a 37°C con un 5% de CO_2 durante 20 a 22 horas. Transcurrido este tiempo las células se tratan con 50 nM del compuesto 12-O-tetradecanoil-forbol-13-acetato (TPA) y a las células control se le agrega la misma concentración del vehículo o disolvente DMSO, durante 30 min.

Las condiciones de uso de TPA en la línea celular MLP29 fueron establecidas por Tur (2007), indicando que es un potente inductor de la expresión de *Egr1* al usarlo en la concentración señalada. Esta concentración permite una inducción de la expresión de este gene en un corto periodo de tiempo y sin afectar la morfología de las células.

4.2 Extracción de mononucleosomas con nucleasa micrococcal

Este protocolo no se encuentra estandarizado por lo que se debió poner a punto a partir de los protocolos ya existentes de extracción de nucleosomas, con los cuales se obtienen mono, di, tri y polinucleosomas. Para poner a punto la extracción de mononucleosomas se tomó como base el protocolo de extracción de nucleosomas que se emplea en el laboratorio de Cromatina del Departamento de Bioquímica de la Universidad de Valencia y parte de otros protocolos obtenidos de la bibliografía. El razonamiento de la puesta a punto se describe con detalle en los resultados, el protocolo que se obtuvo se describe a continuación.

Tras el tratamiento con TPA, las células son lavadas con PBS y se fijan durante 5 min con formaldehído 1% en PBS y en agitación a temperatura ambiente. La fijación se detiene con glicina 0,125 M y se incuban en agitación a temperatura ambiente durante 5 min. Posteriormente se lavan dos veces con 10 ml de PBS frío y se añaden otros 10 ml de PBS frío para rascar y recuperar las células en un tubo falcon. Se vuelven a añadir 5 ml de PBS frío para rescatar las células que pudieran haber quedado en la botella de cultivo. Se centrifuga a 400 g 5 min a 4°C y se elimina el sobrenadante. Las células se resuspenden en 5 ml de tampón de lisis celular (Tris-HCl 0,01 M (pH 7.5), NaCl 0,1 M, MgCl₂ 0,003 M, sacarosa 0,03 M, EDTA 10 mM, NP-40 0,5%) con 2µl/ml de inhibidores de proteasas (Sigma) por cada botella de cultivo de 75 cm² inicial, a continuación se lisan las células en un homogeneizador de Potter y se incuban en hielo durante 15 min. Se centrifuga a 500 g por 5 min a 4°C y se elimina el sobrenadante. Los núcleos se resuspenden en 500 µl de tampón de lavado (Tris-HCl 15 mM (pH 7.5), NaCl 15 mM, MgCl₂ 3 mM, glicerol 20%, KCl 60 mM) por cada botella de cultivo de 75 cm² inicial. A continuación se extraen 10 µl de la suspensión de núcleos y se mezclan en 990 µl de SDS 1% para medir la absorbancia a 260 nm. Cada unidad de absorbancia de la cromatina se trata con 50 U de MNasa y se lleva a un volumen final de 500 U de MNasa (Roche) por mililitro en tampón de digestión (Tris-HCl 15 mM (pH 7.5), NaCl 15 mM, MgCl₂ 3 mM, glicerol 20%, KCl 60 mM, CaCl₂ 3 mM). La muestra es tratada durante 15 min a 37°C y la reacción se detiene con EDTA 10 mM.

Para realizar posicionamiento la muestra es tratada con ARNasa 5 µg/mL (Sigma) durante 60 min a 65°C. Posteriormente es tratada con proteinasa K 100 µg/ml (Roche) durante 90 min a 65°C. Finalmente la muestra se purifica mediante “Kit Gene Jet Purification” (Fermentas). Para realizar NuChIP seguir en el apartado 4.4.2. Para verificar la extracción exclusivamente de mononucleosomas, se extrae una alícuota de la muestra y se verifica en gel de agarosa tal como se indica en el apartado 4.6.

4.3 Extracción de ADN genómico desde la línea MLP29

El protocolo de extracción de ADN genómico desde un cultivo celular se llevo a cabo de acuerdo con el protocolo descrito en Current Protocols (1998), como se explica a continuación.

Se elimina el medio de cultivo DMEM del Corning que contiene la línea celular MLP29 y se lava con 8 mL de PBS, luego se despegan las células con 1,5 mL de tripsina, se recuperan en 6 mL de medio DMEM suplementado y se extraen 10 μ L de la suspensión celular para realizar recuento de células iniciales. El resto de las células se centrifugan a 500g 5 min. a 4°C y se elimina el sobrenadante. Las células se resuspenden en 1 volumen de tampón de digestión de manera que para un número menor de 3x10⁷ células se utilizan 300 μ L de tampón de digestión, para un número mayor de células se utiliza 1mL de tampón de digestión por cada 1x10⁸ células. Se traspasan las células contenidas en tampón de digestión a un tubo eppendorf y se incuban a 50°C en agitación durante 12 a 18 horas. A continuación se agrega 1 volumen de fenol-cloroformo-isoamílico en proporción (25:24:1) y se centrifugan a 1700 g 10 min a 4°C. La fase acuosa se transfiere a un tubo nuevo y se agrega ½ volumen de acetato de amonio 7,5 M y 2 volúmenes de etanol al 100%. Se centrifuga a 1700g 2 min a 4°C, se elimina el sobrenadante y se lavan con etanol al 70%. El pellet se seca a temperatura ambiente y finalmente se resuspende en agua desionizada estéril.

4.3.1 Sonicación de ADN genómico. En la sonicación del ADN genómico se ha puesto a punto la amplitud, el número de pulsos y el tiempo de duración por pulso en el sonicador Vibra cell (SONIC), con el fin de obtener fragmentos entre 200pb y 500pb. El resultado de la sonicación se verifica en un gel al 1,7% de agarosa en tampón TAE como se describe en el apartado 4.6. La estandarización de este protocolo se desarrolla en el apartado de resultados y discusión, pero aquí se resumen las condiciones óptimas de sonicación: amplitud del 35% y con 8 pulsos totales de 10 seg cada uno.

4.4 Inmunoprecipitación de cromatina a partir de mononucleosomas (Nu-ChIP)

El procedimiento de inmunoprecipitación de cromatina se extrajo de Tur *et al.* (2010) (ChIP) y se modificó para los propósitos de la presente investigación. A continuación se describe el protocolo de inmunoprecipitación de la cromatina, tomando como material de partida cromatina digerida con MNasa hasta obtener tamaño mononucleosomal (Nu-ChIP).

4.4.1 Incubación de los anticuerpos. Del stock de bolitas magnéticas con proteína G (Invitrogen), se extraen 40 μ L por cada muestra y cada anticuerpo (fig. 22). Se centrifuga a 1600 g 1 min. a 4°C y se colocan los tubos en el imán (Invitrogen) para eliminar el sobrenadante. Las bolitas se lavan tres veces con 1mL de PBS/BSA frío (5 mg/mL), se mezclan invirtiendo sin el imán y se elimina el PBS/BSA con el imán puesto. En el último lavado se centrifugan a 1600 g 1min. a 4°C y se elimina el PBS/BSA.

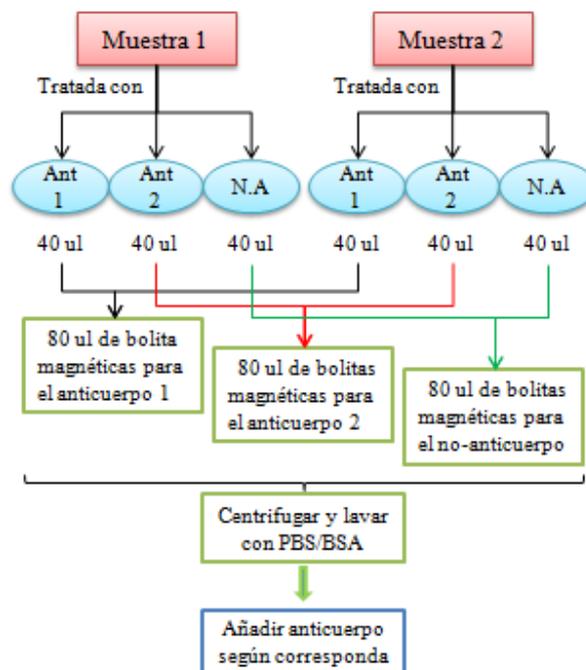


Figura 22. Distribución de las bolitas magnéticas con proteína G para la incubación con los anticuerpos.

A estas bolitas se les añade la cantidad de anticuerpo necesaria para inmunoprecipitar 3,5 unidades de cromatina por muestra, 2 μ L de tRNA (10 mg/mL) por muestra y finalmente se lleva a un volumen de 100 μ L con PBS/BSA. Se incuba en agitación a 4°C toda la noche.

La cantidad de anticuerpo se determinó de manera experimental durante el desarrollo de este trabajo de investigación y se detalla en el cuadro 7.

Cuadro 7. Anticuerpos utilizados y concentración empleada por cada 3,5 U de cromatina.

Anticuerpo	Código y Casa comercial	ug/3,5 U de cromatina
Anti-H3K9 acetilado	ab-4441 (Abcam)	8 ug
Anti-H3K9 tri-metilado	ab-8898 (Abcam)	4 ug
Anti-GCN5	sc-6303 (Santa Cruz)	12 ug
Anti-SUV39	ab-12405 (Abcam)	16 ug
Anti-BRG1	sc-10768 (Santa Cruz)	12 ug
Anti-BRM	sc-28710 (Santa Cruz)	12 ug

4.4.2 Fragmentación de la cromatina. Una vez estandarizado el protocolo para la extracción de mononucleosomas mediante la digestión con MNasa en la línea celular MLP29, se puede realizar la inmunoprecipitación de mononucleosomas (Nu-ChIP). Después de la digestión con MNasa y se detiene la reacción con EDTA 10 mM y la muestra debe ser tratada agregando 1/5 del volumen de tampón de suplementación 5 X (Tris-HCl 25 mM pH 8.6, NaCl 0.675 M, SDS 0,5%, triton X-100 5%). A continuación se incuba 10 min en hielo y se centrifuga a 26000 g 10 min a 4°C. Se recupera el sobrenadante, se deja en hielo y se extrae una alícuota de cada muestra para medir la absorbancia en el espectrofotómetro a 260 nm. Posteriormente la muestra se diluye en tampón IP (Tris-HCl pH 8.1 20 mM, EDTA 1 mM, 0,1% SDS, 1% de Tritón X-100 y NaCl 150 mM) más 2 μ l/ml de inhibidor de proteasas, de modo que la cromatina quede a una

concentración de 7 U/ml. Para verificar que solo se obtuvieron mononucleosomas en la digestión con MNasa se extrae una alícuota de cada muestra antes de añadir el tampón IP y se trata con RNasa y proteinasa K tal y como se indica al final del apartado 4.2. Las muestras se corren en un gel al 2% de agarosa como se describe en el apartado 4.6.

4.4.3 Prelavado de la cromatina. Para realizar el prelavado de la cromatina se extraen 20 μ l de bolitas magnéticas con proteína G por cada mililitro de cromatina que se quiera inmunoprecipitar. Las bolitas se lavan 3 veces con PBS/BSA y se resuspenden en 50 μ l de PBS/BSA por muestra. A las muestras de cromatina diluidas anteriormente en buffer IP (4.4.2) se les añaden estos 50 μ l de bolitas magnéticas y se incuban a 4°C en agitación durante 60 min.

4.4.4 Incubación de la cromatina con los anticuerpos. 20 minutos antes de que termine la incubación del prelavado, se extraen los anticuerpos incubados desde el día anterior con bolitas magnéticas unidas a proteínas G. Se centrifugan a 1600 g durante 1 min a 4°C, se colocan en el imán, se elimina el sobrenadante y se lavan 3 veces con 1mL de PBS/BSA. Las bolitas con los anticuerpos se resuspenden en 30 μ L de PBS/BSA por muestra.

Por otro lado, tras los 60 min del prelavado de la cromatina, las muestras se colocan en el imán y se distribuyen en tubos siliconizados, colocando 500 μ L (que contienen 3,5 U de cromatina) por cada 30 μ L de anticuerpo y no anticuerpo según corresponda. Posteriormente se incuban las muestras a 4°C en agitación durante 2 horas. Tras la incubación con los anticuerpos se les da un spin y se colocan en el imán. El sobrenadante de todos los no anticuerpos se deposita en tubos eppendorf no siliconizados, estos corresponden al INPUT de cada muestra. A los INPUT se les añade 5 μ g/mL de RNasa y se incuban a 65°C durante 2 horas. El resto de sobrenadantes se elimina y a continuación se realizan los lavados.

4.4.5 Lavado y elución del ADN inmunoprecipitado. Primero se lava con 1mL de tampón de baja sal (50 mM de Tris-HCl a pH 8, 15 mM de NaCl, 0,1% de SDS, 1% de Triton X-100 y 2 mM de EDTA), quitando el imán para resuspender las bolitas y colocando el imán para eliminar el sobrenadante. Luego se resuspende nuevamente en 1mL del tampón de baja sal y los tubos se agitan a temperatura ambiente durante 4 minutos, se elimina el sobrenadante. A continuación se repite el proceso anterior con el tampón de alta sal (50 mM de Tris-HCl a pH 8, 0,5 mM de NaCl, 0,1% de SDS, 1% de Triton X-100 y EDTA 2 mM). Posteriormente se lavan los tubos con 1mL del tampón LiCl (25 mM de Tris-HCl a pH 8, 0,25 mM de LiCl, 1% de NP40 y 1 mM de EDTA), se quita el imán para resuspender las bolitas y se coloca para eliminar el sobrenadante. Se hace un segundo lavado con 1mL del tampón LiCl y se centrifuga a 1600 g 1 min a 4°C. Los tubos se colocan en el imán y se elimina el sobrenadante. Finalmente se realiza un último lavado con 1mL de tampón TE (10 mM de Tris-HCl a pH 8 y 1mM de EDTA) y se traspasa el contenido a tubos eppendorf no siliconizados, se colocan en el imán y se descarta el sobrenadante.

Para la elución del ADN inmunoprecipitado se añaden 50µL de tampón de elución (50 mM de NaHCO₃ y 1% de SDS) y se da un vortex al máximo. Se incuba 10 min a 65°C y se centrifuga a 1600 g 1 min a 4°C. A continuación las muestras se colocan en el imán y se rescata el sobrenadante a un tubo nuevo bien rotulado. Se repite el paso anterior rescatando la muestra en el mismo tubo anteriormente rotulado. Tanto las muestras como el INPUT se tratan con proteinasa K 200 µg/mL incubando a 65°C toda la noche.

4.4.6 Purificación y cuantificación de las muestras. Tras la incubación a 65°C, se da un pulso de centrifuga a las muestras y se extrae 1/5 de cada INPUT para la purificación y el resto se almacena a 4°C. Todas las muestras (incluyendo 1/5 de los input) son purificadas mediante el Kit Gene Jet Purification (Fermentas) y se eluyen en 55 µL de tampón de elución. Los INPUT se cuantifican en el nanodrop y se almacenan junto a las muestras a 4°C.

4.5 Reacción en cadena de la polimerasa en tiempo real (qPCR)

Para la reacción en cadena de la polimerasa en tiempo real se utilizó SYBR Green Platinum (Invitrogen, N° Cat. 11744-100). La enzima ADN Taq polimerasa Platinum amplifica desde 10 copias de una secuencia, es decir 1 pg de ADN o ARN.

Por cada reacción de PCR en tiempo real se utilizan 10 µL de la mezcla de SYBR Green Platinum, 0,4 µL del cebador sentido (15 µM) y 0,4 µL del cebador antisentido (15 µM), 7,2 µL de agua desionizada estéril y 2µL de ADN molde.

Las condiciones de amplificación se programaron como se indica a continuación:

- Paso 1: 2 min. a 50°C
- Paso 2: 2 min. a 95°C
- Paso 3: 30 seg a 95°C
- Paso 4: 30 seg. a 60°C (lectura)
- Paso 5: 5 seg a 80°C (lectura)
- Paso 6: Repetir 37 ciclos desde el paso 3.
- Paso 7: Curva de Fusión de 65°C a 95°C incrementando 0.5°C cada 5 seg. (lectura).
- Fin.

Realizando lectura de datos (ct) a 60°C y a 80°C.

Para analizar los resultados que se obtuvieron en el qPCR se utilizó la siguiente fórmula:

$$4.5 \quad (1) \quad 2^{-(Ct \text{ genómico} - Ct \text{ muestra})}$$

Esta fórmula se basa en el supuesto de que los cebadores tienen un 100% de eficiencia y fue empleada para el análisis de posicionamiento.

4.5.1 Cuantificación del posicionamiento. Para estimar la posición de los nucleosomas se obtuvo la secuencia del gen *Erg1* en la base de datos de genes de NCBI, el código de identificación del gen es 13653. Y se diseñaron cebadores para la región promotora y al inicio de la región transcribible (fig. 23).

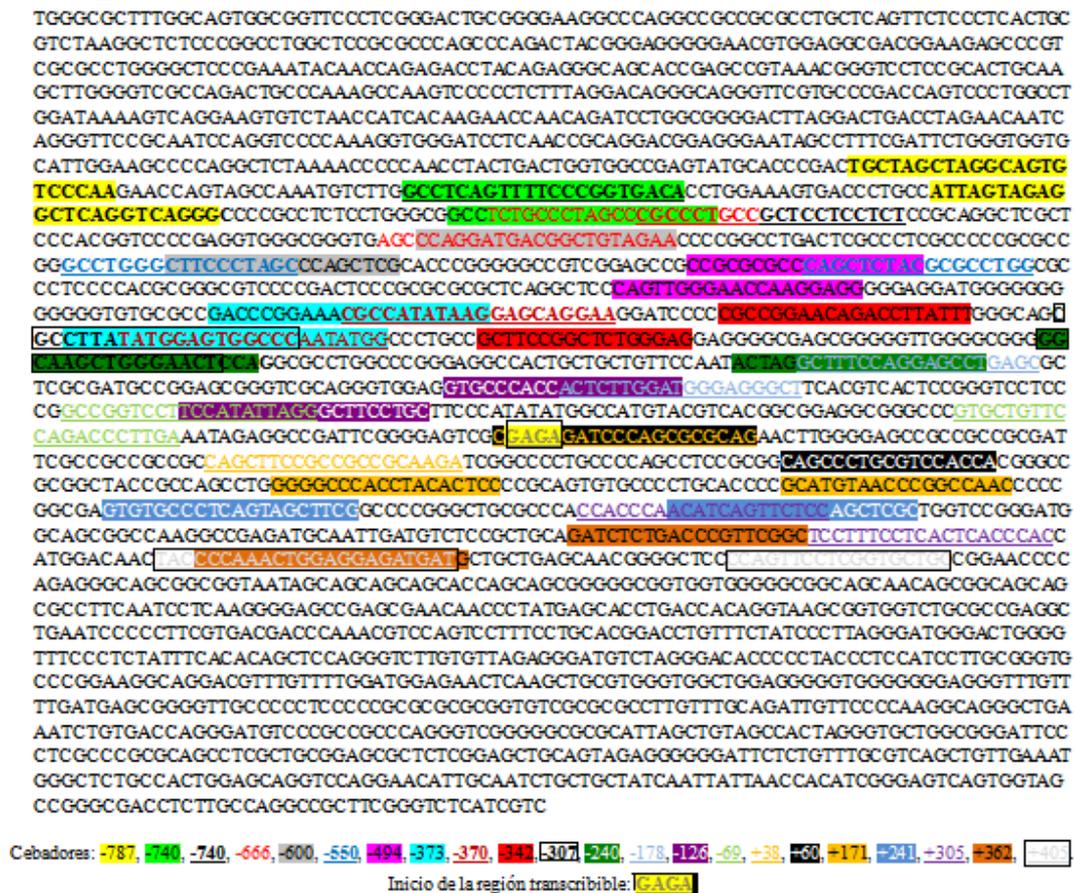


Figura 23. Representación de la posición de los cebadores en la región promotora y al inicio de la región transcribible (recuadro con fondo amarillo) del gen *Erg1*.

La metodología empleada experimentalmente consiste en diseñar cebadores (fig. 23 y cuadro 8) a lo largo de la secuencia del gen, que abarquen fragmentos de aproximadamente 150 pb solapantes. Los núcleos se digieren con MNasa, de manera que la región del ADN en donde hay un nucleosoma se amplificará normalmente, mientras que aquellas zonas que no están protegidas

se amplificaran deficientemente o no se amplificaran (Steward y Sano, 2004). Como normalizador de los datos se utiliza ADN genómico sonicado a una concentración de 5 ng/μL al igual que las muestras. El orden de las muestras en la placa de PCR se muestra en el diseño experimental (fig. 24).

Cuadro 8: Cebadores utilizados en la presente investigación. Los cebadores-69, -126, -178, -342, -370, -494, -550 y -600 fueron descritos por Tur (2007). El diseño de los demás forma parte del desarrollo de este trabajo.

Cebador	Forwad	Reverso	Tamaño amplicon
-787	TGCTAGCTAGGCAGTGTCCCAA	GCCCTGACCTGAGCCTCTACTAAT	107 pb
-746	GCCTCAGTTTTCCCGGTGACA	AGGGCGGGCTAGGGCAGAGGC	102 pb
-740	GCCTCAGTTTTCCCGGTGACA	AGAGGAGGAGCGGCAGGGCG	116 pb
-666	TCTGCCCTAGCCCGCCTGCC	TTCTACAGCCGTCATCCTGGGCT	94 pb
-600	CCAGGATGACGGCTGTAGAA	CGAGCTGGGCTAGGGAAG	78 pb
-550	GCCTGGGCTTCCTAGC	CCAGGCGCGTAGAGCTG	73 pb
-494	CCGCGCGCCAGCTCTAC	CCTCCTTGGTTCCCAACTG	93 pb
-373	GACCCGAAACGCCATATAAG	CCATATTGGGCCACTCCATATAAG	91 pb
-370	CGCCATATAAGGAGCAGGAA	CCATATTGGGCCACTCCATA	79 pb
-342	CGCCGGAACAGACCTTATT	CTCCAGAGCCGGAAGC	77 pb
-307	CGCCTTATATGGAGTGGCC	CTGGAGTTCCAGCTTGCC	96 pb
-240	GGCAAGCTGGGAACTCCA	AGGCTCCTGGAAAGCCTAGT	75 pb
-178	GCTTTCCAGGAGCCTGAGC	AGCCCTCCCATCCAAGAGT	81 pb
-126	GTGCCCAACTCTTGGAT	GCAGGAAGCCCTAATATGGA	80 pb
-69	GCCGGTCTTCCATATTAGG	TCAAGGGTCTGGAACAGCAC	90 pb
+38	CGAGAGATCCCAGCGCGCAG	TCTTGCGGCGGAAGCTG	77 pb
+60	CGAGAGATCCCAGCGCGCAG	TGGTGGACGCAGGGCTG	119 pb
+171	GGGGCCACCTACACTCC	GTTGGCCGGGTTACATGC	58 pb
+241	GTGTGCCCTCAGTAGCTTCG	GCGAGCTGGAGAACTGATGT	64 pb
+305	CCACCCAACATCAGTTCTCC	GTGGGTGAGTGAGGAAAGGA	119 pb
+362	GATCTTGACCCGTTCCGGC	CATCATCTCTCCAGTTTGGG	73 pb
+405	TACCCCAAACCTGGAGGAGATGAT	GCAGCACCGAGGAACTGG	62 pb

4.5.2 Cuantificación de la inmunoprecipitación. Para cuantificar las muestras obtenidas de la inmunoprecipitación de cromatina se mantuvieron las condiciones señaladas para posicionamiento, con excepción del número de cebadores empleados, para este experimento se utilizan los cebadores -494, -550, -178, -240, +171 y +241, ya que mediante posicionamiento se

comprobó que en la zona de amplificación de estos cebadores se encuentran los nucleosomas. Se mantienen los 2µL de muestra por pocillo, pero estas corresponden al ADN inmunoprecipitado tras la incubación con los anticuerpos. Como control negativo se utiliza el no anticuerpo y como control positivo el INPUT, que corresponde a la cromatina tratada sin inmunoprecipitar.

La fórmula para el análisis de los datos es la siguiente:

$$4.5.2 \quad (2) \quad 2^{(Ct \text{ input} - Ct \text{ muestra})}$$

4.6 Gel de agarosa. Para verificar la obtención de mononucleosomas tras la digestión con MNasa y para verificar fragmentos tras la PCR en tiempo real se utiliza un gel al 2% de agarosa en tampón TAE (tris-acetato-EDTA), con 0,2 µg/mL de bromuro de etidio. Para cargar la muestra en el gel se utiliza un tampón de carga 10X (0,1% de Xilencianol, 30% de glicerol, 1% de SDS y 0,1 M de EDTA). Para analizar el tamaño de los fragmentos se utilizan el marcador de peso molecular 100pb (Fermentas). La muestra migra en el gel a 100 volt durante 25 o 30 min. Para verificar el tamaño de los fragmentos tras la sonicación, se utiliza un gel al 1,7% de agarosa en TAE.

4.7 Diseño experimental

El diseño experimental corresponde a un diseño unifactorial completamente aleatorizado. La población de estudio corresponden a células de hígado de ratón (línea MLP29), de la cual se extraen dos muestras, la primera corresponde a dos botellas de 75 cm² con aproximadamente un 90% de confluencia, estas se tratan con TPA durante 30 minutos. La segunda muestra también corresponde a dos botellas de 75 cm² con la misma confluencia pero sin tratar (control), ambas condiciones son sometidas a la digestión con MNasa para posicionamiento nucleosomal o para Nu-ChIP y posteriormente las muestras son amplificadas por triplicado mediante qPCR con los cebadores correspondientes para el gen *Egr1*. Los resultados que se obtienen son analizados con

la fórmula correspondiente (4.5.1 y 4.5.2 respectivamente). El experimento se repite por completo tres veces. La variable de respuesta para posicionamiento y Nu-ChIP corresponde a la cantidad de ADN que resulta de la amplificación de la zona del gen *Egr1* determinada por los cebadores. La variable fija o independiente corresponde a los cebadores del gen *Egr1* de la línea celular MLP29.

En la figura 24 se muestra un esquema del diseño experimental para posicionamiento de los nucleosomas de la región promotora proximal y al inicio de la región transcribible. En la figura 25 se señala el orden de las muestras en la placa de qPCR.

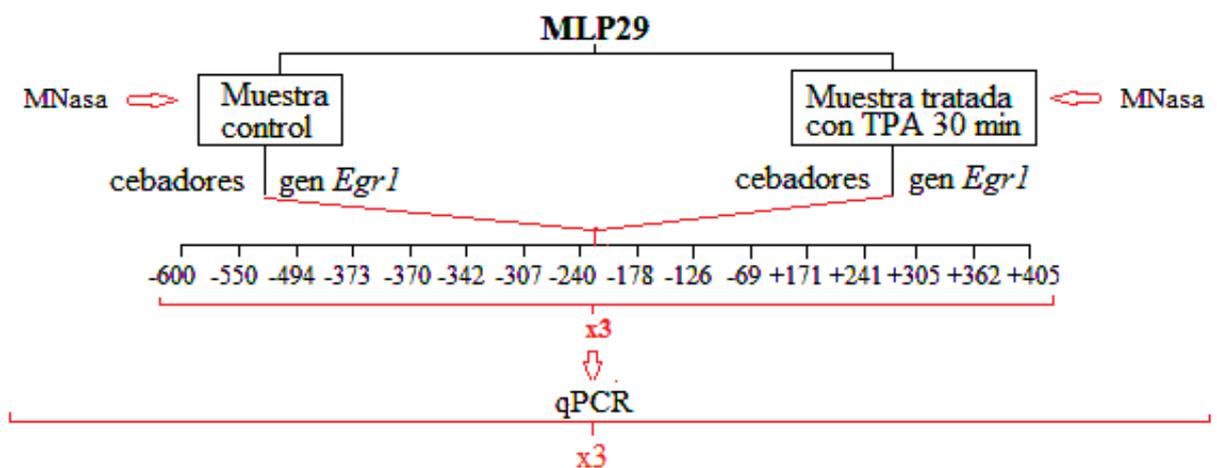


Figura 24. Diseño del experimento de posicionamiento.

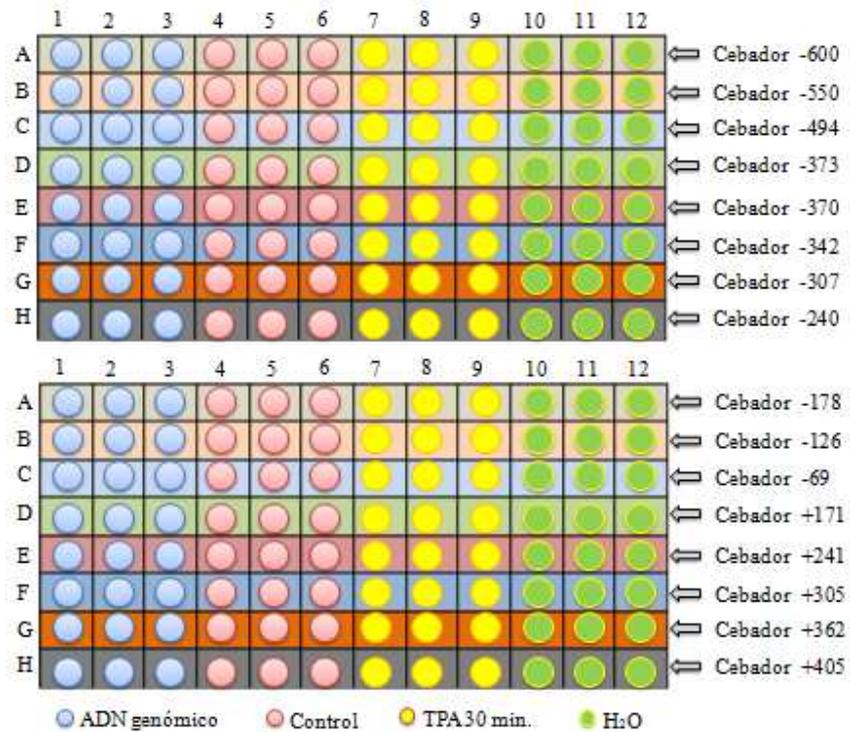


Figura 25. Orden de las muestras y cebadores para posicionamiento en la placa de PCR.

En los ensayos de Nu-ChIP el diseño experimental (fig. 26) es el mismo que para posicionamiento hasta el paso de obtención de mononucleosomas, a partir de aquí se inmunoprecipita con los anticuerpos y posteriormente las muestras son amplificadas por triplicado con los cebadores de cada nucleosoma (-2, -1 y +1) por la técnica de qPCR, los datos son analizados con la fórmula 4.3.2. Cada experimento se realiza por triplicado (fig. 27).

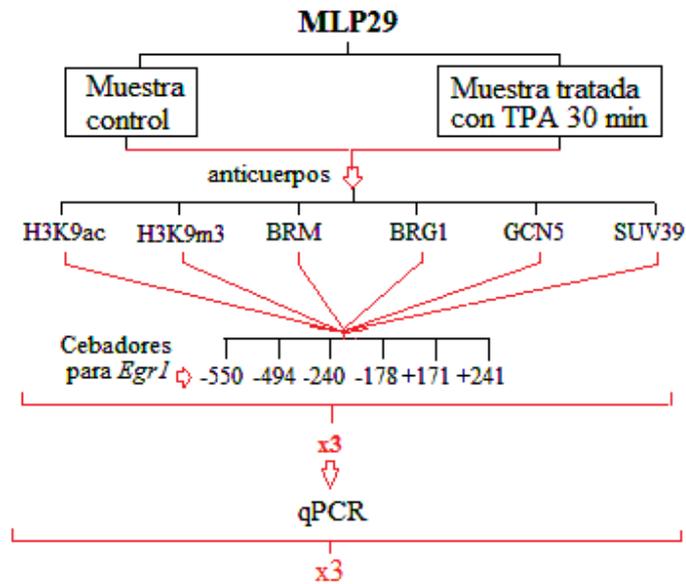


Figura 26. Diseño del experimental para los ensayos de Nu-ChIP.

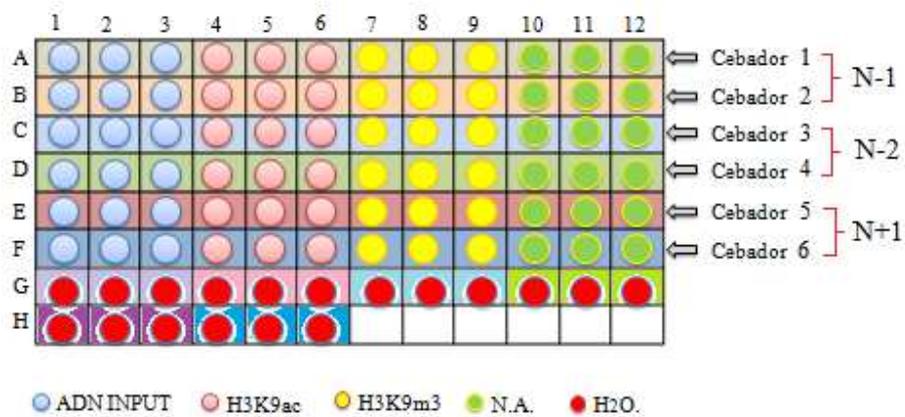


Figura 27. Orden de las muestras y cebadores en la placa de qPCR para Nu-ChIP.

4.8 Análisis estadístico

En el caso del análisis de posicionamiento nucleosomal que consideramos un diseño experimental unifactorial completamente aleatorizado, en el que la variable de respuesta (cantidad de ADN obtenida en la amplificación con cebadores específicos del gen *Egr1*) depende de un único factor o tratamiento, en una muestra completamente aleatorizada (línea celular MLP29 en condiciones de laboratorio), de forma que el resto de las causas de variación se engloban en el error experimental, los análisis estadísticos que se realizan son el cálculo del promedio de cada repetición muestral con su correspondiente desviación o error estándar. Por lo que se realiza un análisis estadístico univariado simple. Lo mismo para los datos del experimento de inmunoprecipitación de cromatina (Nu-ChIP).

4.9 Análisis bioinformático

Para comparar las posiciones de los nucleosomas N-2, N-1 y N+1 del gen *Egr1*, obtenidos experimentalmente, con la predicción realizada por programas bioinformáticos, se recurre a la versión 1.3.1 de “NXSensor”, disponible en la página web <http://www.sfu.ca/~ibajic/NXSensor/>. Igualmente se utiliza el programa “nuScore” disponible en la página web <http://compbio.med.harvard.edu/nuScore/>. Además se emplea el programa NuPop disponible en la página web <http://nucleosome.stats.northwestern.edu/>.

Se recurre a la base de datos de nucleosomas de hígado de *Mus musculus*, para comparar los resultados de la posición de los nucleosomas N-2, N-1 y N+1 del gen *Egr1*, obtenidos experimentalmente, con la información de la posición de los nucleosomas para este gen almacenada en la base de datos. La base de datos se encuentra disponible en la dirección <http://www.ncbi.nlm.nih.gov/projects/geo/query/acc.cgi?acc=GSM717558>.

5. PRESENTACION Y DISCUSION DE RESULTADOS

5.1 Obtención de mononucleosomas

El estudio de la regulación de la transcripción, habitualmente, se realiza mediante la investigación de los elementos *cis* y *trans*, sin tener en cuenta la organización a nivel de cromatina, cuándo en primera instancia, es la posición de los nucleosomas individuales (mononucleosomas), lo que regula el acceso real a los elementos *cis* en el promotor de los genes.

En el presente trabajo, para investigar el posicionamiento de los nucleosomas en el gen *Egr1*, se siguió el método de protección a la digestión con MNasa. Este método se basa en que el ADN ocupado por un nucleosoma es resistente a la digestión. Para aplicar el método es necesario aislar mononucleosomas y determinar qué secuencias de ADN protegen. Una posibilidad consiste en realizar una digestión convencional con MNasa, que da lugar a mono-, di-, tri- y oligonucleosomas, para después resolver los productos de digestión mediante electroforesis en gel de agarosa y recuperar la banda correspondiente a los mononucleosomas (Lefevre y Bonifer, 2006). Pero este método, aunque proporciona mononucleosomas de gran pureza, tiene un rendimiento muy bajo, lo que dificulta la utilización de esos nucleosomas para estudiar las modificaciones epigenéticas por inmunoprecipitación. Por esos motivos, el primer objetivo fue poner a punto el protocolo de obtención de mononucleosomas, mediante digestión con MNasa, sin la necesidad de aislarlos desde un gel de agarosa. Esto permite no solo el estudio de la posición nucleosomal de manera directa (sin la necesidad de realizar la extracción desde un gel de agarosa), sino que también permite estudiar las modificaciones epigenéticas en los nucleosomas individuales.

En la bibliografía es posible encontrar numerosos protocolos para la obtención de mononucleosomas por digestión con MNasa. En el laboratorio de Cromatina del Departamento de Bioquímica de la Universidad de Valencia se utiliza un protocolo para extracción de nucleosomas a partir de cultivos celulares (el cual se describe en anexo 2), y recientemente se ha

empleado con algunas variantes en las tesis de Tur (2007) y Sacilotto (2010) (ver anexo 3). No obstante, el protocolo debe adaptarse a cada situación concreta; en particular, el tipo de tejido o de células empleadas puede afectar considerablemente a determinadas variables experimentales. Por ese motivo, el protocolo que se describe en el anexo 2 fue adaptado y puesto a punto para las necesidades de la investigación presente.

Para optimizar el protocolo se modificó el tiempo de entrecruzamiento con formaldehído, el tiempo de incubación con MNasa y las unidades totales de MNasa en la muestra, tal como se observa en la figura 28 (A). Estas modificaciones se realizaron debido a que un menor tiempo de entrecruzamiento podría facilitar la digestión enzimática, al igual que aumentar la concentración de MNasa y el tiempo de incubación con la misma. Como se observa en la figura 28 (B), en los carriles C4, C5 y C6, fijados 5 min con formaldehído al 1%, se obtuvo una mayor cantidad de muestra, además, al tratar con 150 U de MNasa por 60 min (C4) los resultados fueron similares que al tratar con 300 U por 15 min (C6). Estos resultados sugieren que al tratar la muestra se digiere mejor con una mayor concentración de MNasa, que al emplear una concentración baja por más tiempo. Sin embargo, en este primer experimento no se logró obtener exclusivamente mononucleosomas, posiblemente porque algún componente del tampón de digestión podría estar afectando la actividad de la MNasa.

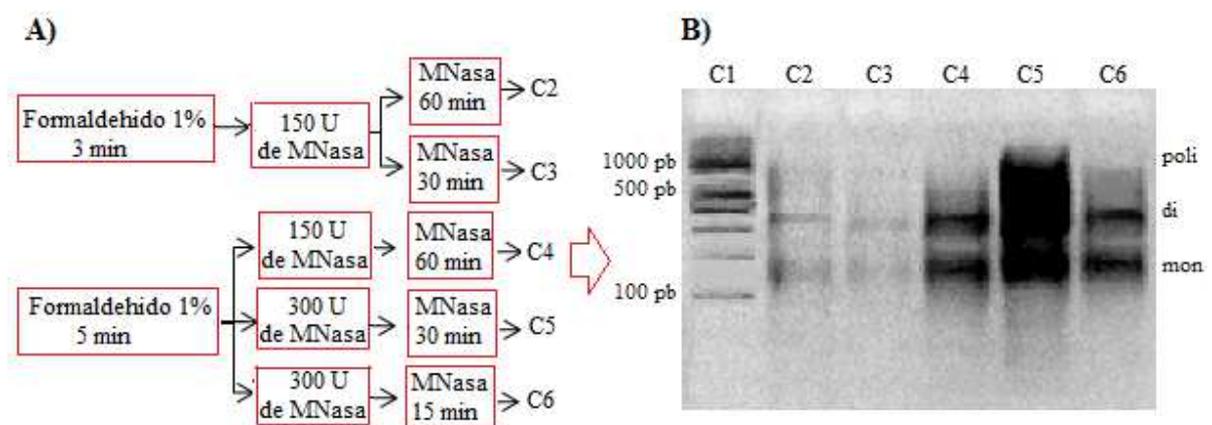


Figura 28. Modificaciones en el protocolo de extracción de nucleosomas. A) Esquema de las modificaciones en el tiempo de entrecruzamiento con formaldehído, unidades y tiempo de incubación con MNasa, y los carriles (C2 a C6) correspondientes en cada caso. B) Electroforesis en gel de agarosa al 2% en TAE. En C1 se encuentra el marcador de peso molecular de 100pb, de C2 a C6 se corresponde con las modificaciones descritas en el panel A.

Para comprobar el planteamiento anterior se realizó un ensayo modificando la composición del tampón RSB de digestión, se mantuvo el tiempo de tratamiento con formaldehído 1% por 5 min y se digirió con 100 U de MNasa. Las variantes del tampón de digestión (A, B, C y D) se indican en la figura 29A, donde A es el tampón de digestión RSB sin modificar, B corresponde al RSB sin espermidina y espermina, C es el RSB sin β -Mercaptoetanol y D no contiene protectores (espermina y espermidina), ni tampoco β -Mercaptoetanol. Como se muestra en la figura 29B, si bien no se obtuvieron exclusivamente mononucleosomas, se observó una variación en la cantidad de mono, di y trinucleosomas obtenidos. En el carril 5, se observa que en el tampón D se obtuvo una mayor cantidad de mono y dinucleosomas, obteniendo además una disminución casi completa de los trinucleosomas. Esto se reafirma al observar los carriles 3 y 4 y compararlos con el carril 2 (RSB sin modificar) que contiene mayor cantidad de dinucleosomas y presenta de forma notoria trinucleosomas.

De los resultados obtenidos se puede concluir que la MNasa digiere mejor la cromatina en un medio que no posee los protectores y tampoco β -Mercaptoetanol. Pero las variantes anteriormente estudiadas, si bien permiten una mejor digestión de la muestra, no son los únicos factores que intervienen en obtener exclusivamente de mononucleosomas en el protocolo. Por lo que se planteó la posibilidad de que la enzima pudiera no estar actuando en toda su eficiencia y que otros factores endógenos pudieran estar contribuyendo a la fragmentación de la cromatina.

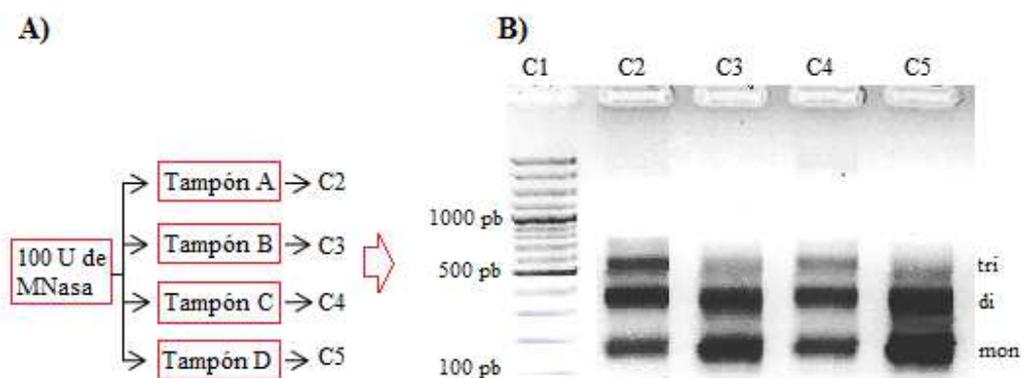


Figura 29. Prueba de modificación en la composición del tampón de digestión. A) Tampones de digestión A, B, C y D y carriles correspondientes. B) Electroforesis en gel de agarosa al 2% en TAE. En C1 se encuentra el marcador de peso molecular de 100pb.

Para estudiar la capacidad de digestión de la MNasa, se realizó una prueba con el tampón A sin MNasa, con 50 U de MNasa, 200 U de MNasa y 300 U de MNasa (fig. 30A). Como se indica en el carril 2 de la figura 30B en ausencia de MNasa la muestra no es digerida, por lo que se descartó la presencia de nucleasas endógenas, al tiempo que se demostraba que la enzima utilizada mantiene su actividad catalítica. No obstante se utilizó un lote nuevo de enzima para las pruebas posteriores.

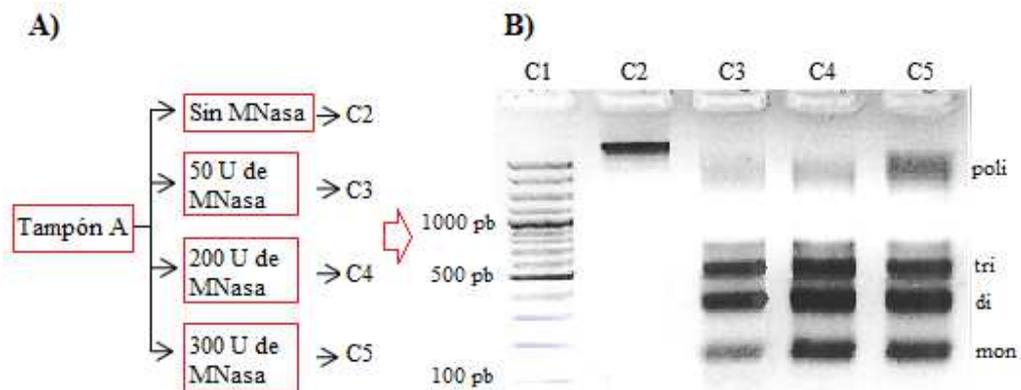


Figura 30. Modificación en la concentración de MNasa. A) Pruebas sin MNasa y con 50 U, 200 U y 300 U de MNasa y el orden en el que se cargaron en los carriles. B) Electroforesis en gel de agarosa al 2% en TAE. En C1 se encuentra el marcador de peso molecular de 100pb.

Dado que al modificar la composición del tampón de digestión se observa un aumento en la obtención de mononucleosoma, se plantea continuar las pruebas desde este punto. Para lo cual se utilizaron dos tampones de digestión, el primero modificado desde Tur (2007), al que se le denominó con las siglas BD y el segundo tampón fue modificado desde Sacilotto (2010) al que se denomina tampón denominado BG. Se realizó una prueba con muestra fijada y sin fijar, las muestras fueron digeridas a distintas concentraciones de MNasa por 30 min a 37°C. A las muestras fijadas que posteriormente se trataron con los tampones, se les antepuso la letra F siendo denominadas FBD y FBG.

Otro componente importante en la optimización de la digestión con MNasa es el co-factor CaCl_2 , que por su función de permitir la catálisis enzimática, la concentración en la que se

emplea podría ser fundamental para la obtención exclusiva de mononucleosomas, por esto se consideró aumentar su concentración de 1 mM a 3 mM en el tampón de digestión.

En la figura 31A se indica la modificación realizada y la correspondiente denominación en los carriles en la figura 31B. Al emplear el tampón BD fijado (FBD, carriles C2, C3 y C4) y sin fijar (C8 y C9) se obtuvieron exclusivamente mononucleosomas en todas las concentraciones de MNasa, pero la obtención de los mismos es baja. En el caso de FBG, solo a una concentración de 200 U/ml de MNasa se obtuvieron exclusivamente mononucleosomas (carril 7), permitiendo obtener una mayor cantidad de muestra.

La modificación en la concentración de CaCl_2 , mejoró la eficiencia de la enzima MNasa y al combinar esta modificación con el tampón FBD, se obtuvieron mononucleosomas en todos los casos e incluso a bajas concentraciones de MNasa, por lo que para el protocolo final se mantiene a una concentración de 3 mM.

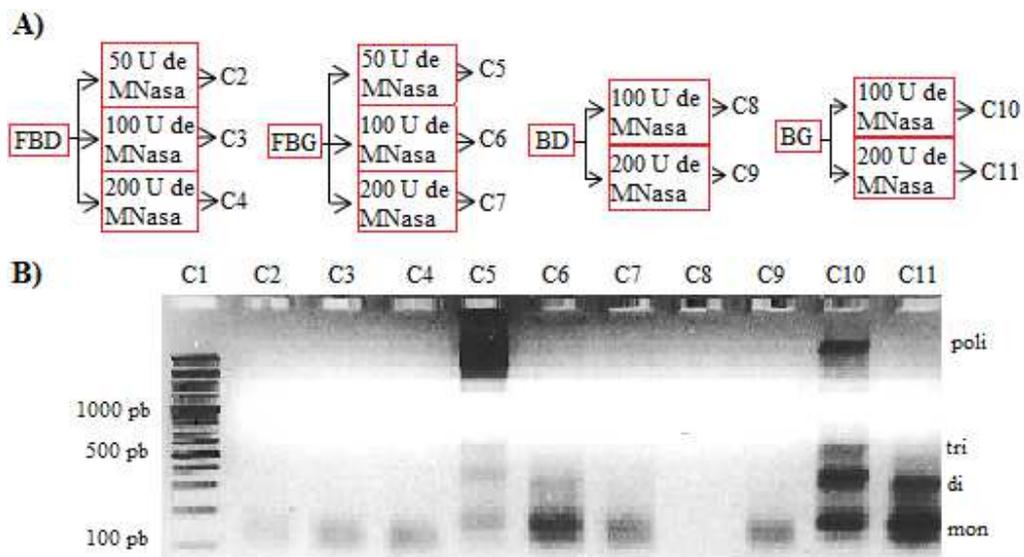


Figura 31. Prueba de nuevos tampones de digestión con CaCl_2 3 mM . A) Tampones y unidades de MNasa empleados en cada carriles. B) Electroforesis en gel de agarosa al 2% en TAE, en C1 se encuentra el marcador de peso molecular de 100pb.

Con la finalidad de obtener un mayor rendimiento en la extracción, se probaron distintas concentraciones de MNasa y tiempo de incubación en los tampones de digestión BD y BG (fig. 32 A). Las muestras se fijaron 5 min previamente a la digestión con MNasa, se emplea una concentración de 3 mM de CaCl₂. Bajo estas condiciones se obtuvieron mononucleosomas en todos los casos (fig. 32 B), determinándose que a una concentración de 500 U/ml de MNasa en tampón BG, incubando 20 min a 37°C (carril 14) se obtiene un mayor rendimiento en la extracción de mononucleosomas exclusivamente.

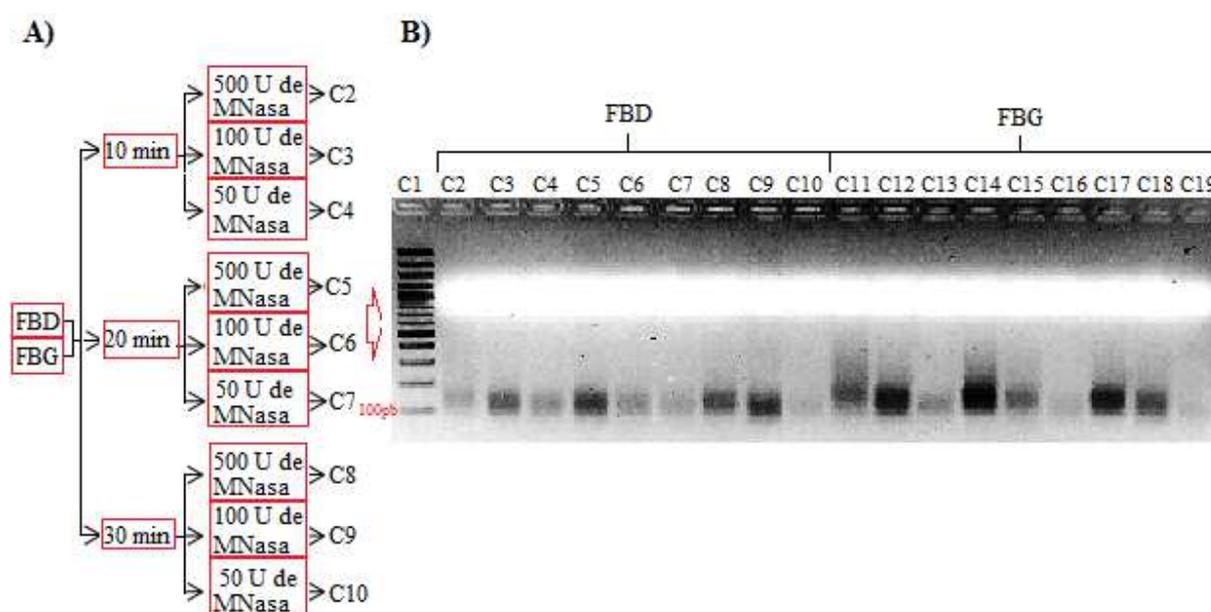


Figura 32. Prueba del tampón FBD y FBG a distinto tiempo de incubación y concentración de MNasa. A) carriles correspondientes en cada caso. B) Electroforesis en gel de agarosa al 2% en TAE, en C1 se encuentra el marcador de peso molecular de 100pb.

El protocolo estandarizado de extracción de mononucleosomas se expone en materiales y métodos en el apartado 3.2. Este protocolo permite la obtención de mononucleosomas mediante la digestión con MNasa, sin la necesidad de purificar el fragmento de 147pb a partir de un gel de agarosa, tal como se indica en la figura 33.

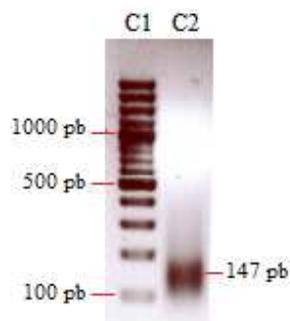


Figura 33. Electroforesis en gel de agarosa al 2% en TAE. En C1 se encuentra el marcador de peso molecular de 100pb y en C2 el resultado del protocolo de extracción de mononucleosomas.

5.2 Puesta a punto de la técnica de sonicación de ADN genómico

La sonicación del ADN genómico se realiza debido a que este se utiliza con finalidades de normalización en el posicionamiento nucleosomal. Para la puesta a punto de la sonicación del ADN genómico se sónico en todos los casos a un 35% de amplitud y cada pulso de sonicación fue de 10 seg. Para la primera prueba de sonicación se dieron 3 pulsos, extrayendo 10 μ l de la muestra para comprobar el tamaño de los fragmentos de ADN. Como se observa en la figura 34A se obtuvieron fragmentos desde 200 pb hasta por encima del marcador de tamaño. La muestra se sónica nuevamente, aplicando otros 3 pulsos, se extrae 10 μ l de la muestra y se vuelve a verificar el tamaño de los fragmentos, obteniendo esta vez fragmentos desde aproximadamente 180 pb hasta 800 pb (fig. 34 B). Finalmente se le aplican dos pulsos más de sonicación, obteniendo fragmentos de entre 150 pb y 600 pb (fig 34 C).

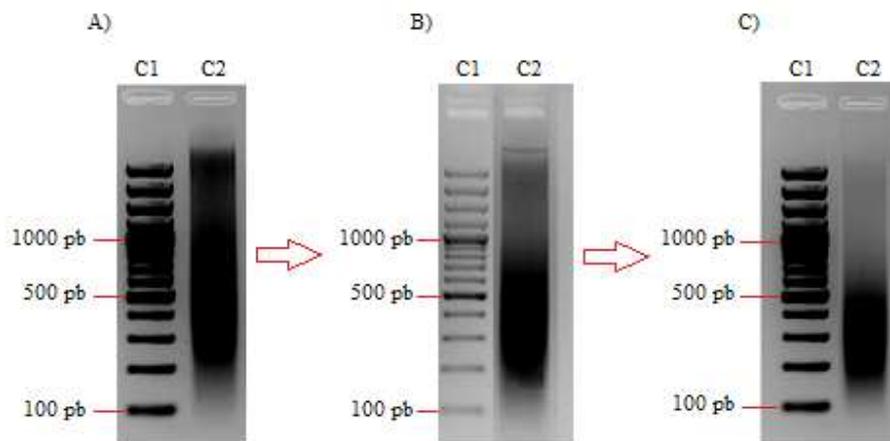


Figura 34. Gel de agarosa al 1,7% en TAE. A) 3 pulsos de sonicación de 10 seg del ADN genómico. B) 6 pulsos de sonicación de 10 seg cada uno. C) 8 pulsos de sonicación. En el carril 1 se indica el marcador de peso molecular 100 pb y en el carril 2 la fragmentación de la cromatina.

5.3 Verificación de los cebadores

Para verificar que los cebadores amplificaran correctamente, se utilizaron 5 ng/ μ l de ADN genómico sonicado, 5 ng/ μ l de una muestra tratada 30 min con TPA y 5 ng/ μ l de una muestra control. Ambas muestras corresponden a cromatina fijada y posteriormente sonicada. El análisis se realizó mediante PCR en tiempo real.

En la figura 35 se muestran los resultados de la curva de fusión del ADN amplificado con cada pareja de cebadores. Como se observa, la mayoría de los cebadores no forman dímeros y poseen un único pico de amplificación. En esta situación se encuentran las parejas de cebadores -69, -126, -178, -342, -370 y -180. Las parejas de cebadores -494, -550 y el -600 presentan una tendencia a formar dímeros, aunque solo amplifican una especie definida de ADN. El par de cebadores -425, al igual que el -269, solo proporcionan productos a ciclos altos de amplificación y la altura de los picos varía dentro de la misma muestra. Todo esto sugiere que están produciendo una amplificación inespecífica por lo que fueron descartados.

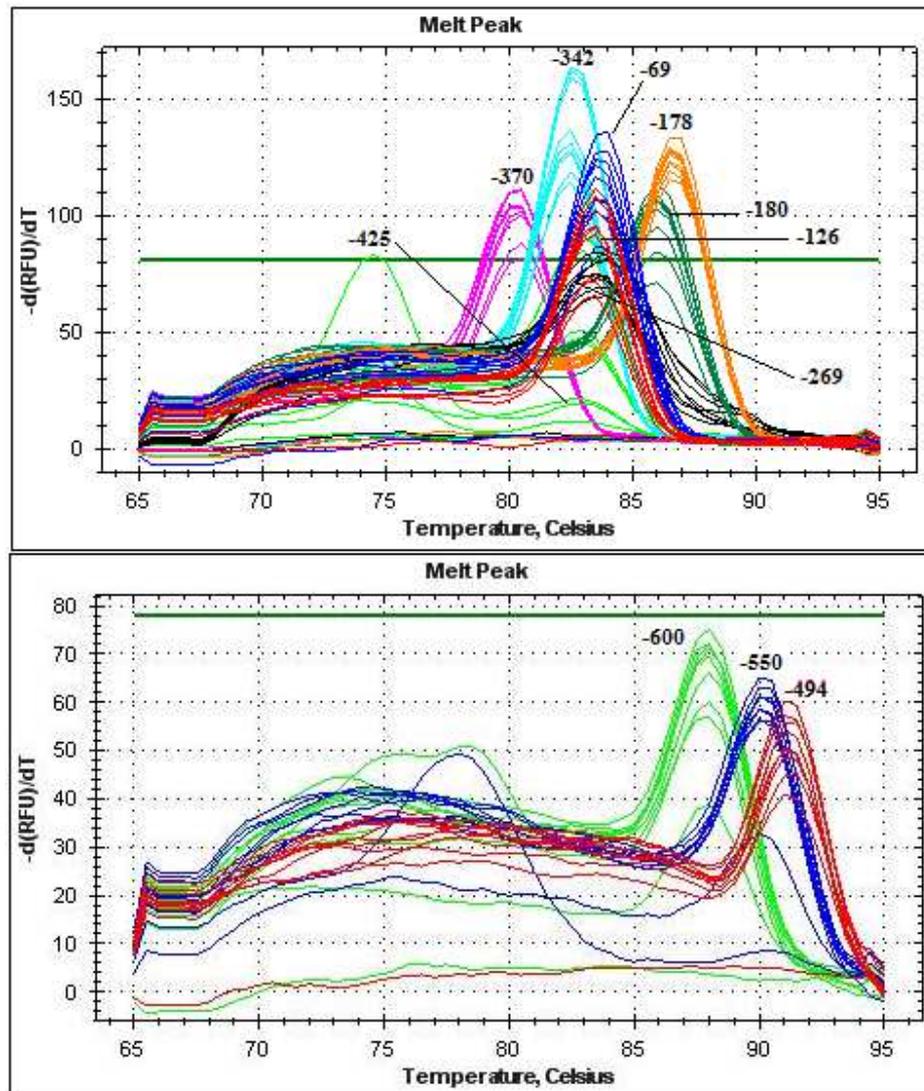


Figura 35. Curva de fusión del qPCR. Prueba de los cebadores extraídos desde la tesis doctoral de Gema Tur (2007).

La amplificación mediante qPCR se verificó en un gel de agarosa al 2% en TAE, tal como se observa en la figura 36. En los carriles 14, 15 y 16 se encuentran los fragmentos amplificados por el par de cebadores -269, lo que permite confirmar que no funciona correctamente, por lo que se suprime para futuros ensayos. En los carriles 21 y 22 no se observan los fragmentos amplificados por el par de cebadores -425, por lo que se elimina este par de cebadores para investigaciones posteriores. El resto de cebadores amplifican correctamente, puesto que amplifican solo el

fragmento correspondiente, que por lo general es de 100 pb. Debido a que los cebadores -180 y -178 amplifican los mismos fragmentos, ya que sus centros respectivos se encuentran separados solo por 2 pb, se selecciona para estudios posteriores el par de cebadores -178, por presentar una mejor curva de fusión (fig. 35).

En resumen, de los cebadores se utilizaran los siguientes: -69, -126, -178, -342, -370, -494, -550 y el -600. Para reemplazar los cebadores que no amplificaron correctamente y para amplificar el ADN al inicio de la región transcribible (hasta 450 pb río abajo) se diseñaron nuevos cebadores solapantes entre sí. Se verificó la amplificación mediante qPCR en las mismas condiciones señaladas anteriormente.

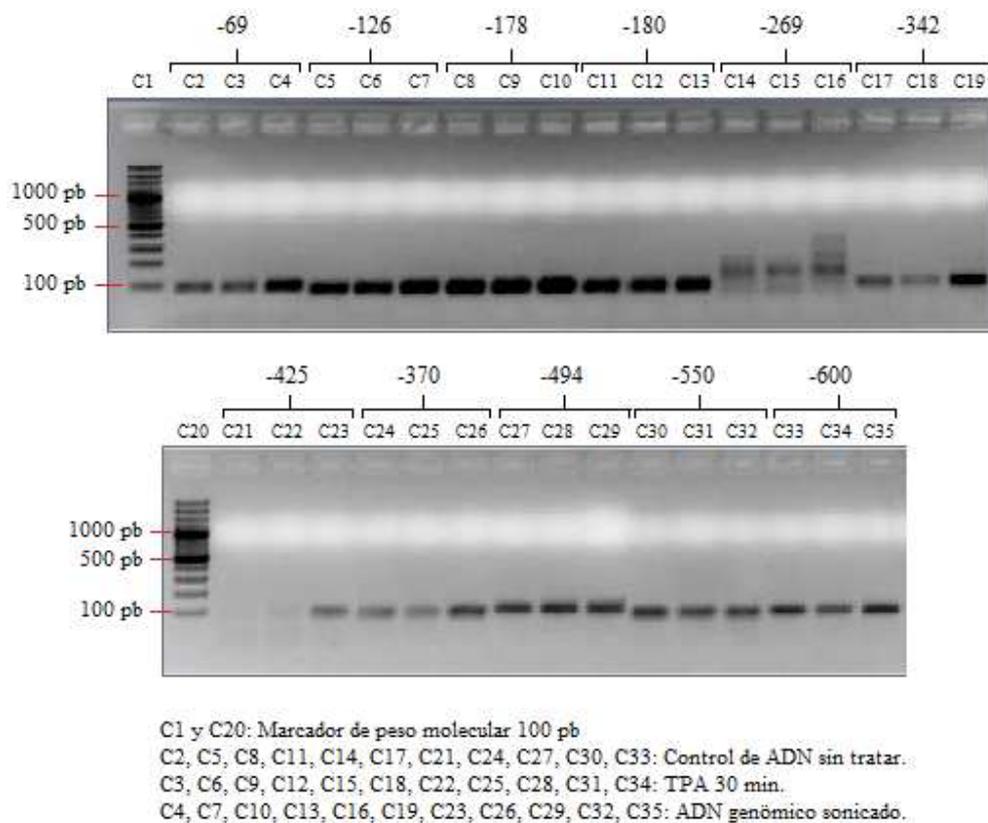


Figura 36. Gel de agarosa al 2% en TAE. Se muestra el resultado de la amplificación por qPCR de ADN genómico sonicado, muestra sonicada tratada con TPA 30 min y una muestra control sin tratar. Los cebadores fueron extraídos desde la tesis doctoral de Gema Tur (2007).

En la figura 37 se observan las curvas de fusión del ADN obtenido con los nuevos cebadores. Como se muestra en la curva la mayoría de los cebadores dan lugar a un único pico, pero en algunos casos aparece un doble pico, que podría deberse a la formación de dímeros.

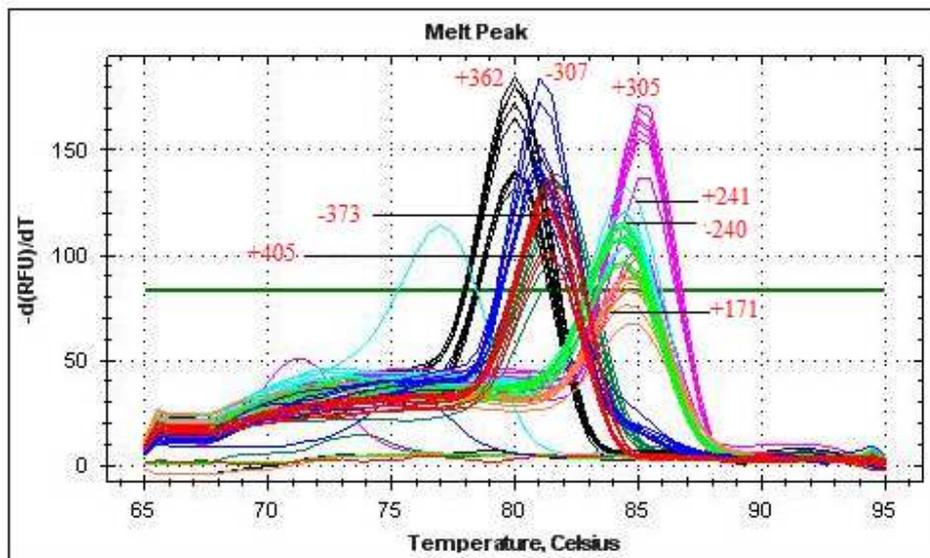


Figura 37. Curva de fusión del qPCR. Prueba de los cebadores nuevos.

Para comprobar la amplificación de los cebadores, los productos de qPCR se corrieron en un gel de agarosa al 2%, tal como se muestra en la figura 38. En los carriles 5, 6 y 7 se encuentra el resultado de la amplificación del par de cebadores -307, como se observa en el gel parece amplificar fragmentos inespecíficos de ADN (C5 y C6), aunque también podría deberse a la formación de dímeros. A pesar de la posible amplificación de fragmentos inespecíficos, el par de cebadores -307 no será descartado, debido a que es el único que amplifica de forma solapante la zona del ADN comprendida entre el nucleótido -355 y el -260. Los demás cebadores amplifican correctamente, por lo que todos son empleados posteriormente en el estudio del posicionamiento.

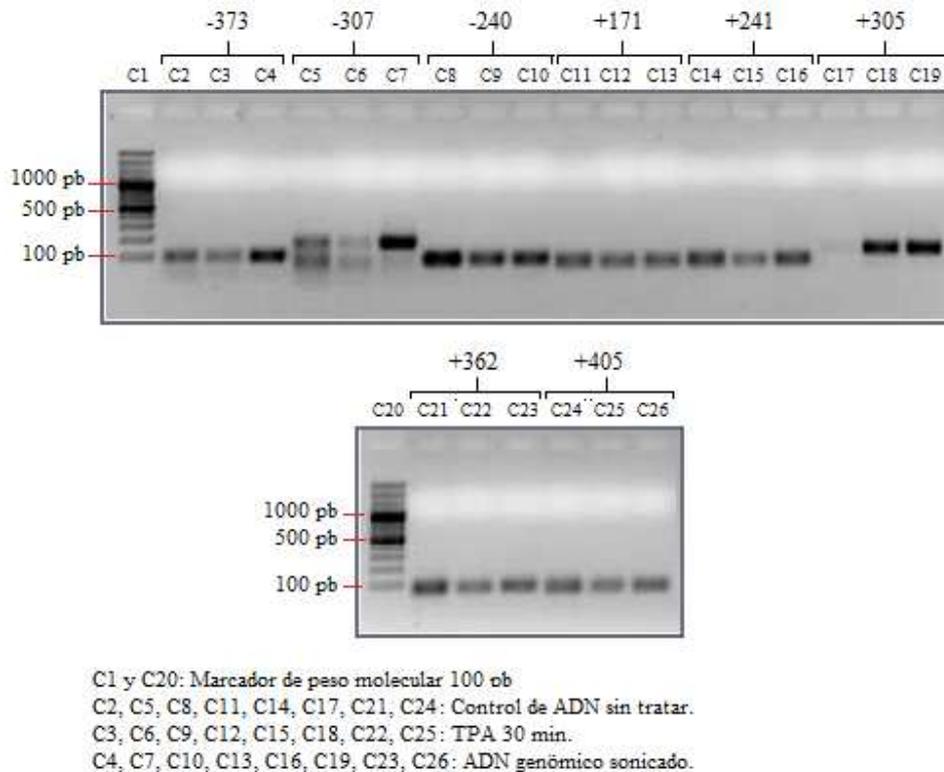


Figura 38. Gel de agarosa al 2% en TAE. Se muestra el resultado de la amplificación por qPCR de ADN genómico sonicado, muestra sonicada tratada con TPA 30 min y una muestra sonicada control sin tratar. Los cebadores fueron diseñados en el programa AmplifX.

5.4 Posicionamiento de nucleosomas en *Egr1*

Un nucleosoma se compone por ADN e histonas (H2A, H2B, H3 y H4). La longitud de ADN ocupada por un nucleosoma es aproximadamente 200 pb, de los cuales 147 están fuertemente unidos al núcleo de histonas (mononucleosoma) y el resto sirve de unión entre los nucleosomas. El nucleosoma es la partícula fundamental de la cromatina, según el modelo propuesto por Kornberg en 1974. La posición que ocupa un nucleosoma constituye un elemento fundamental en la regulación de la expresión génica (Cui *et al*, 2010). Esto se debe a que puede, bajo condiciones determinadas, impedir o favorecer la entrada de los factores involucrados en la transcripción, o bien al no permitir que la elongación continúe en un cierto punto dentro del gen (Herman *et al*,

2006). Por lo que el segundo objetivo de este trabajo es establecer el posicionamiento nucleosomal en el promotor proximal y al inicio de la región transcribible del gen *Egr1*. Para cumplir este objetivo, se utilizó el ADN mononucleosomal obtenido por digestión con MNasa (apartado 3.2) y mediante qPCR se amplificó con las parejas de cebadores seleccionados (Cuadro 8). Los datos obtenidos fueron procesados según la fórmula 3.5.

Tras la fijación de la cromatina y la digestión con MNasa, el ADN internucleosomal se pierde en el protocolo de extracción de mononucleosomas, y solo se obtiene el ADN adherido al octámero de histonas. Los cebadores se solapan por lo que amplifican fragmentos de toda la región estudiada del gen *Egr1*. Aquellas regiones que se perdieron en el proceso de extracción no serán amplificadas, mientras que las regiones de los nucleosomas serán amplificadas por los cebadores. Cuanto mayor es la cantidad de ADN amplificado, más representada está su secuencia en el molde y, por tanto, más protegida está esa región a la digestión con MNasa. En la figura 40 se exponen los resultados obtenidos en el ensayo realizado para determinar la posición de los nucleosomas (este ensayo fue realizado por triplicado). Los resultados del análisis se grafican de acuerdo a la protección relativa a MNasa (eje Y), frente al valor medio del amplicon (fragmento amplificado por cada pareja de cebadores) (eje X).

De la figura 39 se puede deducir la posición de tres nucleosomas, denominados a partir de ahora como N-2 y N-1, los situados río arriba del inicio de la transcripción (TSS) y el N+1, localizado río abajo del TSS. Los amplicones -600, -550 y -494 presentan una mayor protección frente a la digestión con MNasa por lo que en esta zona se ubica el nucleosoma N-2. Los amplicones -240 y -178 indican la posición del nucleosoma N-1. El nucleosoma +1 parece estar ubicado en la zona amplificada por +171 y +241. El resto de los cebadores parecen amplificar zonas internucleosomales, por lo que no se encuentran protegidos frente a la digestión con MNasa y por ende, no amplifican.

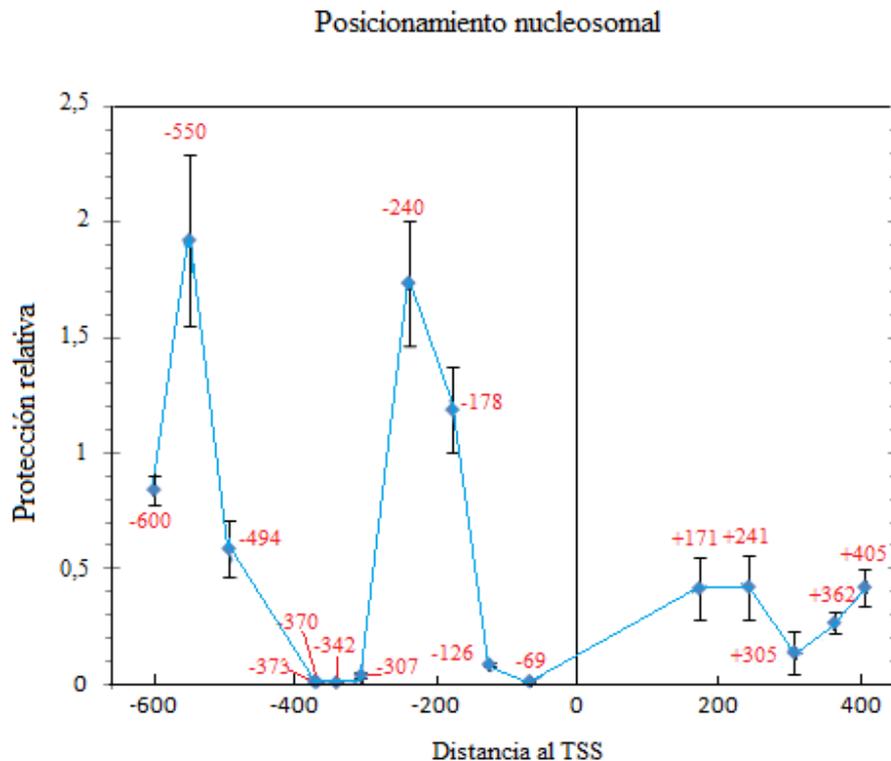


Figura 39. Posicionamiento de los nucleosomas de la región promotora y al inicio de la región transcribible. Los puntos representan la media de tres determinaciones de qPCR y las barras dan el correspondiente error estándar

Los resultados sugieren que es necesario diseñar nuevos cebadores en la región comprendida entre -100 y el +150 para delimitar la ubicación del nucleosoma +1 y en la región entre el nucleótido -600 y el -850 para definir la ubicación del nucleosoma -2. Para esto se empleó la versión de prueba del programa AmplifX, con el cual se diseñó el amplicon -787 y los cebadores en 5' de -746, -740, +38 y +60. Los cebadores antisentido de estos amplicones y del amplicon -666 fueron diseñados a mano, debido a que los programas bioinformáticos no permitían hacerlo por ser una zona con alto contenido en CG. Para el diseño manual se calcula el % de CG y de AT en la secuencia y se relaciona con la temperatura de fusión esperada, además de considerar el largo del fragmento a amplificar y la longitud de cada cebador.

En la figura 40 se pueden observar los resultados de la amplificación con los amplicones +38 y +60. A pesar de que estos cebadores forman dímeros (fig. 40 A y B), dan lugar a un solo pico definido de amplificación (fig. 40 A), por lo que esta circunstancia no interfiere en los datos. Además como se encuentran en una zona en la que los programas bioinformáticos no permiten el diseño de cebadores, son los únicos que hacen posible la amplificación de esta zona.

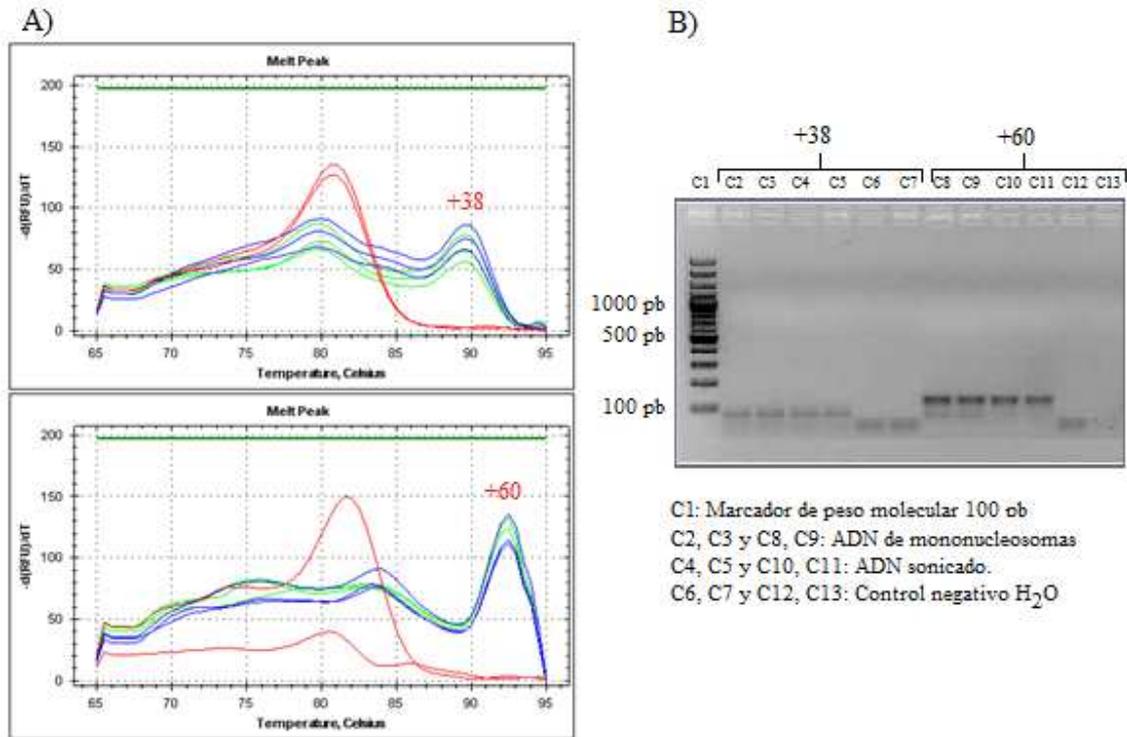


Figura 40. Verificación de la amplificación por qPCR de los cebadores +38 y +60. A) Curva de temperatura de fusión. Gel de agarosa al 2% en TAE. El ADN molde corresponde a genómico sonificado (5 ng/μl) y mononucleosomal (5 ng/μl).

Los cebadores diseñados para definir la posición del nucleosoma -2 fueron sometidos a un gradiente de temperatura, tras su primera prueba, debido a que no amplificaron correctamente. En la figura 41 se presentan los resultados del gradiente de temperatura de fusión. A 67°C (fig. 41A) y a 64,5°C (fig. 41B), en el paso 4 del protocolo de amplificación para qPCR anteriormente descrito (apartado 3.5.1), todos los cebadores dan lugar a un único pico de amplificación. Por el contrario a 62°C (fig. 41C) el par de cebadores -740 ya no amplifica adecuadamente y a 59,5°C (fig. 41D) solo -787 amplifica adecuadamente. Según lo anterior y puesto que a una temperatura

muy elevada la especificidad de los cebadores no es la adecuada, se concluyó que la temperatura de 64,5°C es la óptima para que los cebadores diseñados amplifiquen correctamente el fragmento de interés.

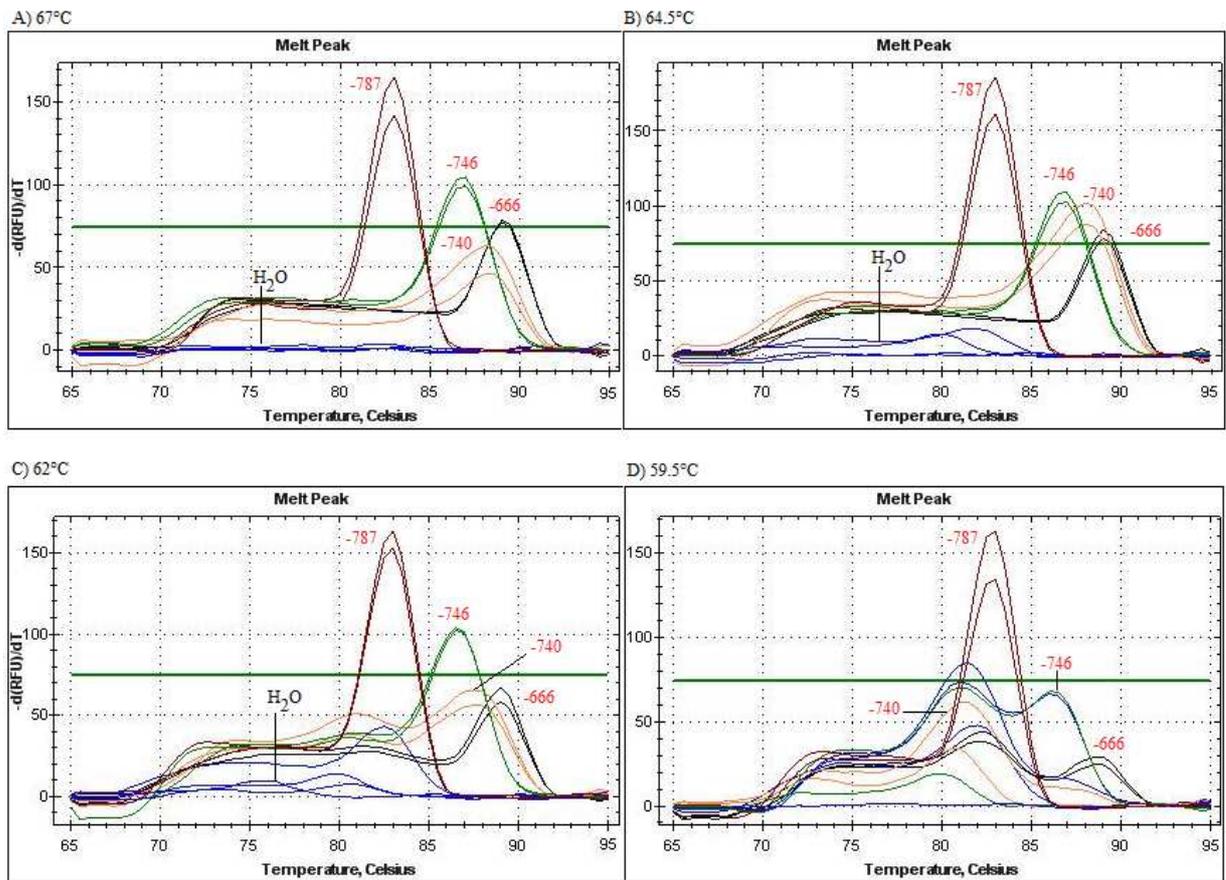


Figura 41. Gradiente de temperatura y curva de merlting en qPCR, para los cebadores -787, -146, -740 y -666. A) Amplificación a 67°C. B) Amplificación a 64,5°C. C) Amplificación a 62°C. D) Amplificación a 59,5°C.

Los resultados para la temperatura de 64.5°C se comprobaron posteriormente en un gel de agarosa al 2% en TAE (fig. 42), en la que se demuestra la amplifcan de un único fragmento en todos los casos.

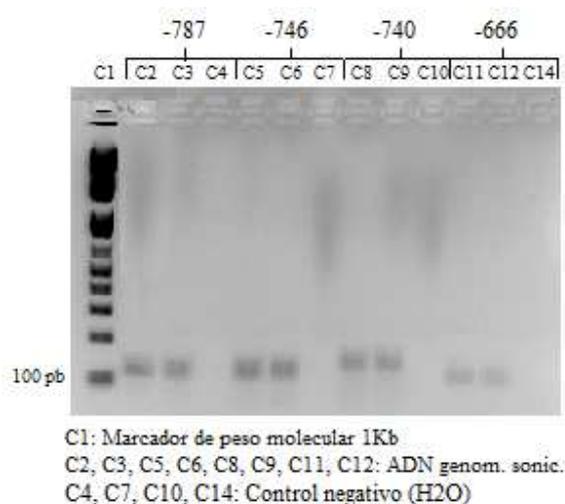


Figura 42. Gel de gel de agarosa al 2% en TAE. Resultado de la prueba de amplificación con una temperatura de fusión de 64.5 en qPCR.

Una vez verificada la buena calidad de los cebadores, se realizaron los ensayos para perfilar la posición de los nucleosomas -2 y +1 de las regiones estudiadas del gen *Egr1* (promotor proximal e inicio de la zona transcribible). En la figura 43 se muestra el grafico de protección relativa frente a la digestión con MNasa y la posición de los nucleosomas. Según estos resultados, la posición del nucleosoma -2 se encuentra definida en los amplicones -666, -600, -550 y -494. La posición del nucleosoma -1 se encuentra definida por los amplicones -240 y -178. La posición del nucleosoma +1 no queda totalmente delimitada. El máximo de protección a MNasa está centrado sobre el amplicon +241, pero se observa una caída del pico con una pequeña meseta centrada en +38 y +171 que puede indicar que existe una cierta movilidad del nucleosoma, de modo que, aunque ocupe preferentemente la posición centrada en +241, en algunas células se encuentra más desplazado en el sentido 5'. Alternativamente, puede ocurrir que la protección a MNasa en la región de la meseta entre +38 y +171 corresponda a otra causa distinta de la presencia del nucleosoma +1. Una explicación plausible sería la presencia del complejo de la ARN polimerasa II, pausado precisamente por la situación del nucleosoma +1. De hecho, se ha demostrado que la ARN polimerasa II está ensamblada en el promotor proximal del gen *Egr1* aún en ausencia de transcripción (Tur *et al.* 2010). Los experimentos presentes no permiten decidir entre esas posibles causas.

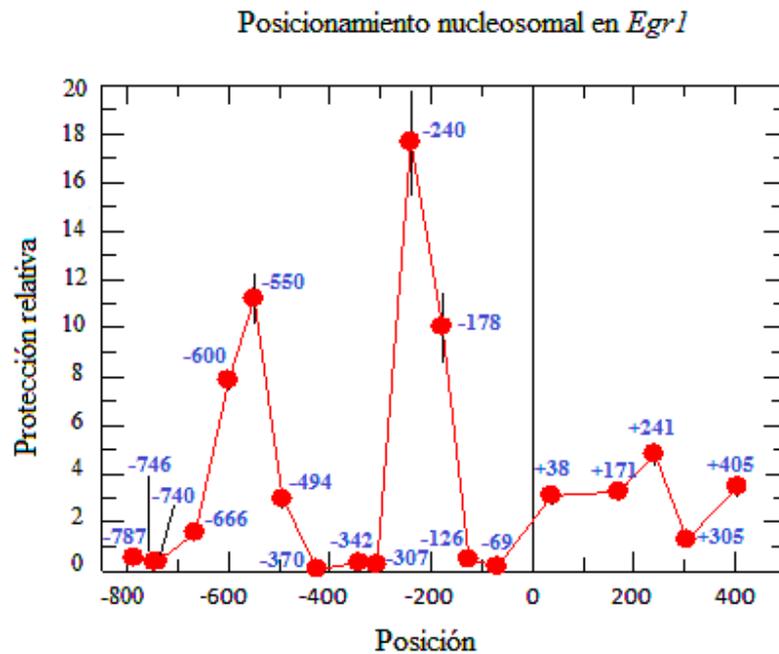


Figura 43. Definición de la posición de los nucleosomas -2 y +1. Grafico de puntos obtenido del análisis de la cuantificación por qPCR desde ADN mononucleosomal y normalizado con ADN sonificado. Los puntos representan la media de tres determinaciones de qPCR y las barras dan el correspondiente error estándar.

En la figura 44 se esquematiza el resultado del posicionamiento. En rojo se resaltan los amplicones en los que se consigue amplificación y por ende son los que se encuentran en una región protegida frente a la digestión con MNasa, indicando la posición de los nucleosomas. Estas regiones se representa con un ovalo azul a lo largo del promotor proximal y del inicio de la transcripción del gen *Egr1*.

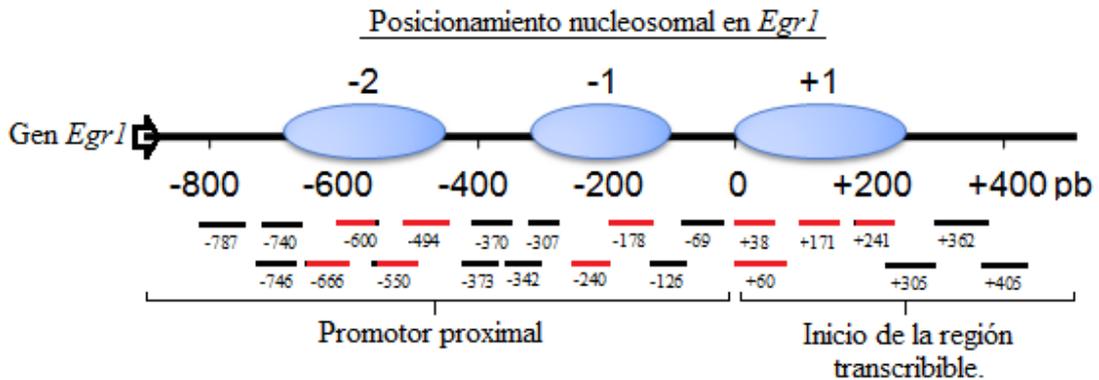


Figura 44. Esquema de la localización de los nucleosomas en la región del promotor proximal (-2 y -1) y al inicio de la región transcribible (+1) del gen *Egr1* en la línea celular MLP29.

Considerando los resultados obtenidos por Tur *et al* (2010), que indican que el gen *Egr1* posee una máxima transcripción a los 30 min de inducción con TPA (empleando cuantificación del ARNm por RT-PCR), resulta interesante estudiar si la posición y cantidad de protección de los nucleosomas identificados (N-2, N-1 y N+1) varía cuando el gen se encuentra en nivel máximo de expresión. Por lo que el tercer objetivo es determinar si se producen cambios en la posición o en la cantidad de nucleosomas cuando el gen es activado por éster de forbol 12-*O*-Tetradecanoilforbol-13-acetato (TPA) 30min.

Los resultados obtenidos, los que se exponen en la figura 45, indican que el N-2 con un pico de protección a -550, cuando se induce con TPA durante 30 minutos, disminuye su protección a la vez que aumenta la protección en -494. Esto sugiere un desplazamiento del N-2 unos 50 pb río abajo, permitiendo la accesibilidad al sitio EBS que se encuentra en cerca de -600. Gitenay y Baron, 2009 y Tur *et al.*, (2010) han descrito que EGR1 se une a este elemento *cis* en su propio promotor y reprime la transcripción de *Egr1*, en este sentido los resultados demuestran la importancia de la estructura de la cromatina como modulador de los procesos de regulación de la transcripción.

La disminución de la protección en la zona ocupada por los nucleosomas -1 y +1 seguramente es señal de que esos nucleosomas se pierden parcialmente, para facilitar el ensamblaje de la maquinaria de transcripción y la posterior elongación.

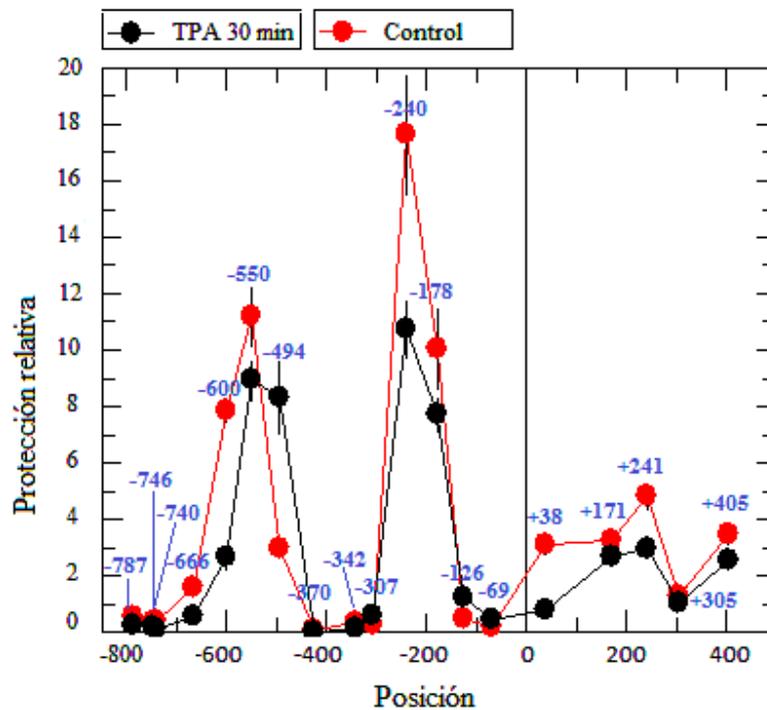


Figura 45. Posición de los nucleosomas del gen *Egr1* en las células MLP29 control y a los 30 min de la inducción con TPA. Cuantificación por qPCR desde ADN mononucleosomal y normalizando con ADN genómico sonicado. Los puntos representan la media de tres determinaciones de qPCR y las barras dan el correspondiente error estándar.

5.5 Modificaciones epigenéticas

El principal método empleado para estudiar las modificaciones epigenéticas es la inmunoprecipitación de la cromatina fragmentándola por sonicación (ChIP), pero con este método se obtienen fragmentos de distinta longitud, por lo que no permite el estudio directo exclusivamente de mononucleosomas. Además, puede ocurrir que la inmunoprecipitación de una

modificación que se encuentre en un nucleosoma en concreto, arrastre a los nucleótidos vecinos, por el hecho de estar en el mismo complejo di o tri nucleosomal. En cambio, utilizar la inmunoprecipitación directa de mononucleosomas (Nu-ChIP), permite estudiar las modificaciones epigenéticas a nivel de un único nucleosoma.

Para cumplir el cuarto y quinto objetivo, primero se estudian los remodeladores de la cromatina BRG1 y BRM por su posible participación en el desplazamiento y/o disminución de la protección en los nucleosomas identificados en el gen *Egr1*. Posteriormente se estudia la trimetilación y la acetilación del residuo lisina 9 de H3, por su posible participación en la activación y represión de la transcripción en cada uno de los nucleosomas identificados. Finalmente se investiga si la acetiltransferasa GCN5 y la metiltransferasa SUV39 son las enzimas responsables de las modificaciones de estos residuos a nivel nucleosomal.

5.5.1 Remodeladores BRM y BRG1. Los remodeladores utilizan la energía de la hidrólisis del ATP para mover, desplazar, expulsar o reestructurar nucleosomas dentro de la cromatina (Clapier y Cairns, 2009; Piatti *et al*, 2011). En el estudio anterior se comprobó que los nucleosomas identificados en *Egr1* sufren cambios en su posición y cantidad (fig. 46), al inducir su expresión con TPA 30 min. Por lo que, los remodeladores debieran estar involucrados en estos cambios. Se ha visto que BRM regula negativamente la expresión del gen, mientras que BRG1 parece estar presente cuando *Egr1* se expresa (Sakurai *et al*, 2011). También se ha visto que estos remodeladores se encuentran presentes cuando *Egr1* no se está expresando, pero cuando el gen es activado estos remodeladores abandonan el promotor a los 15 min de inducción con TPA (Tur, 2007).

En el presente trabajo, se estudió empleando la técnica de Nu-ChIP, la participación de BRM y BRG1 en los cambios vistos en los nucleosomas N-2, N-1 y N+1 del gen *Egr1*. Se tomó como referencia la inducción con TPA durante 15 min. En la figura 46 A se observa que BRM se encuentra posicionado en el nucleosoma N+1, específicamente en el amplicon +241, cuando el gen no se expresa, en cambio, cuando el gen es activado (mediante la inducción con TPA 15 min) BRM abandona el nucleosoma N+1.

Los resultados obtenidos para BRG1 (fig. 46 B), indican que este remodelados se encuentra presente en N-2 y N-1, pero se encuentra mayormente representado en N+1, aumentando su representación en este nucleosoma luego de la activación con TPA 15 min.

Estos resultados sugieren que BRG1 podría estar alterando la interacción de las histonas con el ADN para modificar la estructura del nucleosoma N+1 y de esta manera permitir el acceso de la RNA pol II para que continúe la transcripción. Por esta razón N+1 disminuye su protección frente a la digestión con MNasa, y el remodelados se encuentra más representado, cuando el gen se encuentra activo. En el caso de BRM, los resultados sugieren que este remodelados podría estar implicado en mantener la estructura del nucleosoma N+1, por lo que, en el momento en que *Egr1* es activado, el remodelados abandona el nucleosoma permitiendo que BRG1 ingrese. Tanto BRM como BRG1 no parecen estar involucrados en la remodelación de los nucleosomas N-1 y N-2.

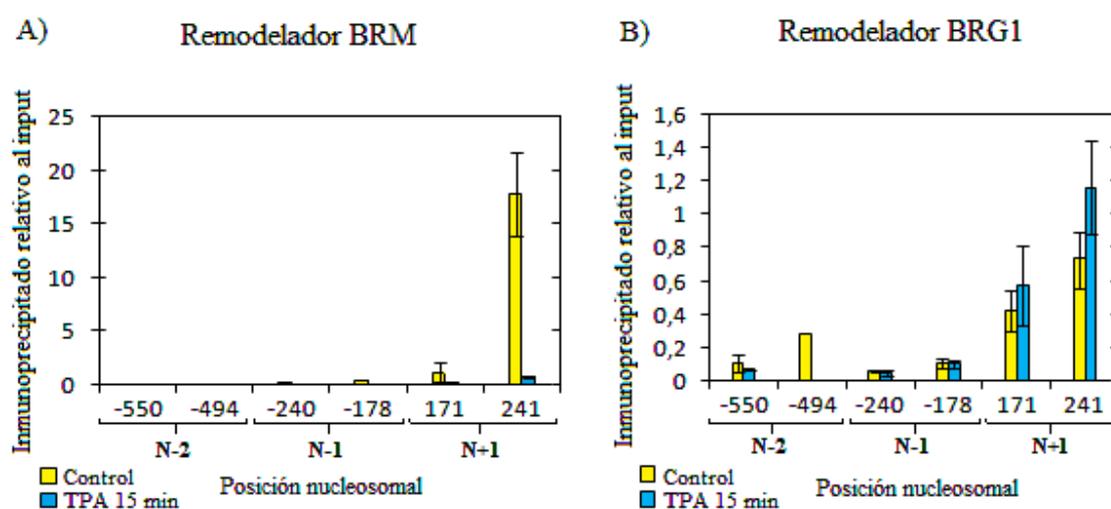


Figura 46. Cuantificación de Nu-ChIP de los remodeladores BRM (en A) y BRG1 (en B), en los nucleosomas N-2, N-1 y N+1 del gen *Egr1*, en el control y a los 15 min de inducción con TPA. Los histogramas representan la media de tres determinaciones de qPCR y las barras dan el correspondiente error estándar.

5.5.2 Acetilación y tri-metilación de H3K9 en *Egr1*. Considerando la teoría del código de histonas, las modificaciones epigenéticas puntuales cumplen determinadas funciones en la regulación genómica (Strahl y Allis, 2000). Por lo general la marca de acetilación en histonas se encuentra relacionada con la transcripción génica, mientras que la marca de metilación con represión o activación dependiendo del residuo modificado, siendo el residuo de lisina 9 en la histona H3 una de las marcas epigenéticas más estudiadas (Kouzarides, 2007a).

En la figura 47A se muestran los resultados sobre la acetilación en H3K9, para cada nucleosoma. Los tres nucleosomas poseen una acetilación basal en el control, la cual aumenta significativamente en los tres nucleosomas luego de la inducción con TPA 30 min, siendo los nucleosomas posicionados en el promotor, especialmente N-1, los que parecen presentar un mayor grado de acetilación. En la figura 47B se presentan los resultados del análisis de la trimetilación de H3K9, que a diferencia de lo observado en la figura 47A, aumenta considerablemente en el nucleosoma N+1 y se encuentra menos representada en los nucleosomas N-2 y N-1.

Estos resultados sugieren que la modificación de acetilación en H3K9 en el promotor de *Egr1*, facilita el acceso a los elementos *cis*, para que el complejo de la transcripción se forme y de esta manera la polimerasa comience la transcripción. La trimetilación de este residuo está descrita generalmente como asociada a represión, pero en el caso del gen *Egr1* parece estar relacionada con la activación génica.

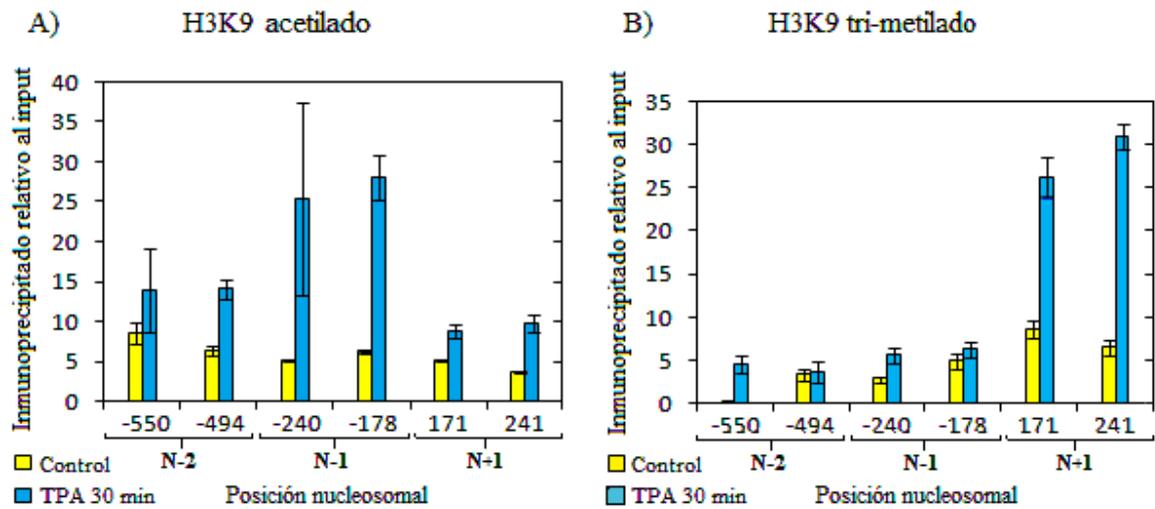


Figura 47. Cuantificación de Nu-ChIP de H3K9 Ac. y H3K9 met3 en los nucleosomas -2, -1 y +1 del gen *Egr1*, en el control y a los 30 min de inducción con TPA. Los histogramas representan la media de tres determinaciones de qPCR y las barras dan el correspondiente error estándar.

5.5.3 Enzimas acetiltransferasa y metiltransferasas de histonas. Se ha visto que la trimetilación de H3K9, es catalizada principalmente por la enzima metiltransferasa SUV39 (Inglis *et al.*, 2007; Selvi 2010). La acetiltransferasa GCN5 está relacionada con la acetilación de H3K9 (Allis *et al* 2007). Por lo anterior se estudió la posible participación de estas enzimas en las modificaciones de H3K9 observadas es los resultados anteriores.

En la figura 48A se presentan los resultados obtenidos del análisis realizado mediante qPCR de las muestras de Nu-ChIP, en ella se observa que la enzima acetiltransferasa GCN5 se encuentra más representada en el amplicon +241 del nucleosoma N+1, lo cual no concuerda con los resultados obtenidos para la acetilación, en donde se observa una mayor representación de la marca H3K9 Ac en los nucleosomas N-1 y N-2 (fig. 47A). Esto puede deberse a que GCN5 no sea la enzima responsable de la acetilación de H3K9, ya que además de ésta, en bibliografía también se describe a la acetiltransferasa PCAF como responsable de la acetilación de este residuo (Allis *et al* 2007).

Los resultados obtenidos en Nu-ChIP para la metiltransferasa SUV39, demuestran que la enzima se encuentra más representada en el amplicon +241, del nucleosoma N+1, esto coincide en parte, con los resultados observados para la marca de trimetilación del nucleosoma N+1 (fig. 47B) en H3K9.

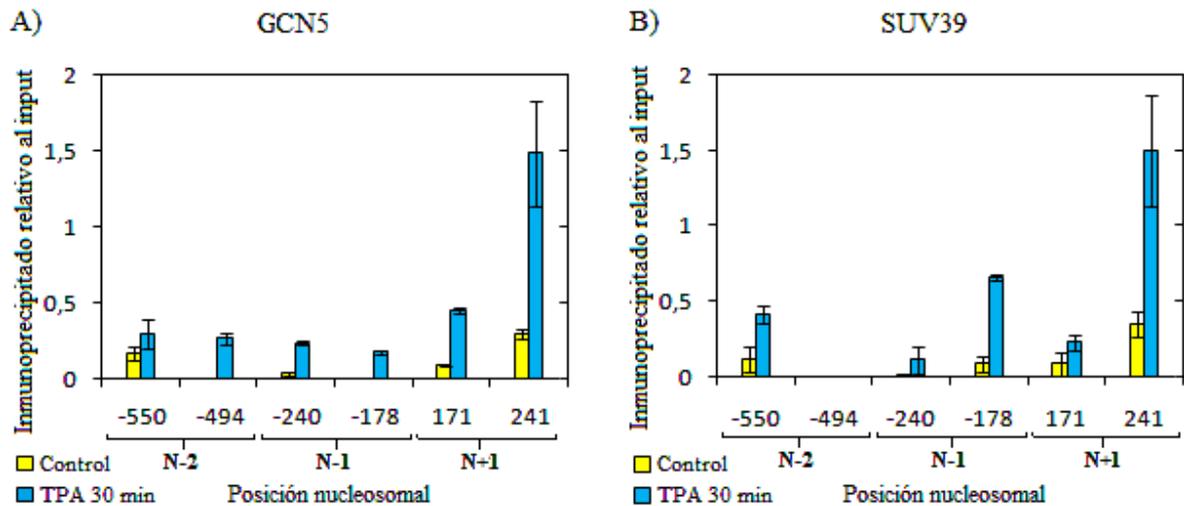


Figura 48. Cuantificación de Nu-ChIP para GCN5 y SUV39, en el control y a los 30 min de inducción con TPA. Los histogramas representan la media de tres determinaciones de qPCR y las barras dan el correspondiente error estándar.

5.6 Predicción del posicionamiento nucleosomal

Como ya se ha mencionado anteriormente, el nucleosoma tiene importancia no solo en la organización estructural de la cromatina, sino también en la regulación de la accesibilidad de los complejos proteicos implicados en la transcripción del ADN. Se sabe que la posición que ocupa el nucleosoma en el ADN no es al azar, sino que existen secuencias que experimentalmente se repiten preferentemente en el ADN nucleosomal y otras se observan con mayor frecuencia en el ADN entre nucleosomas. A partir de estos y otros parámetros que intervienen en la estructura del nucleosoma, se han elaborado programas bioinformáticos con el fin de predecir la posición dentro de la secuencia de ADN que tiene una mayor probabilidad de unirse al octámero de

histonas. La utilización de programas bioinformáticos permite optimizar tiempo y recursos en la investigación científica.

En primer lugar se comparó la posición de los nucleosomas obtenidos experimentalmente, con la predicción que realizan los programas bioinformáticos NXSensor, nuScore y NuPop. El programa NXSensor se basa en la composición de la secuencia para discriminar las zonas que por su contenido de CG, poli-A o poli-T, no pueden curvarse y por ende no forman parte de los nucleosomas (Luykx *et al.*, 2006). NuScore se basa en el costo energético que supone curvar el ADN alrededor del octámero de histonas, el cual depende de la secuencia de ADN y de la interacción de la secuencia con las histonas (Tolstorukov *et al.*, 2008). NuPop utiliza el modelo de cadenas de Markov (HMM) de primer y cuarto orden, teniendo en cuenta dos estados, el nucleosoma (N) y el enlazador (L) calculados en levadura, pero al escoger otra especie, el programa reestructura los datos (Xi *et al.*, 2010).

En la figura 49 se compara la posición obtenida experimentalmente de los nucleosomas N-2, N-1 y N+1 (recuadro amarillo) con las predicciones realizadas por NXSensor. Este programa predice tanto la zona del ADN con mayor probabilidad de que se forme un nucleosoma (fig. 49 a), como la zona entre nucleosomas, o región excluyente (fig. 49 b). Para el primer caso, que se indica en color gris, la predicción realizada por NXSensor no coincide con la posición obtenida experimentalmente en todos los casos, he incluso, la zona excluyente de nucleosomas (fig. 49 b, en rojo) es la que coincide, en mayor medida, con la posición de los nucleosomas N-2, N-1 y N+1. Estos resultados sugieren que basarse solamente en la composición del ADN, no es suficiente para predecir asertivamente la posición de los nucleosomas N-2, N-1 y N+1, identificados experimentalmente en el gen *Egr1*.

En la figura 50A, se muestran los resultados de la predicción realizada por NuPop, empleando la opción “Mouse” y el cálculo de cadenas de Markov de cuarto orden, presentando los resultados en una escala de probabilidad de 0 a 1. En la figura 50B se exhibe la predicción de la zona de inicio de un nucleosoma (en azul). Los recuadros amarillos representan la región experimental de los nucleosomas N-2, N-1 y N+1 del gen *Egr1*. De estos resultados se puede concluir que, para el caso de *Egr1*, NuPop no parece predecir, de manera precisa, la posición de los nucleosoma, ya que los lugares de mayor probabilidad de inicio de un nucleosoma no coinciden con lo observado experimentalmente.

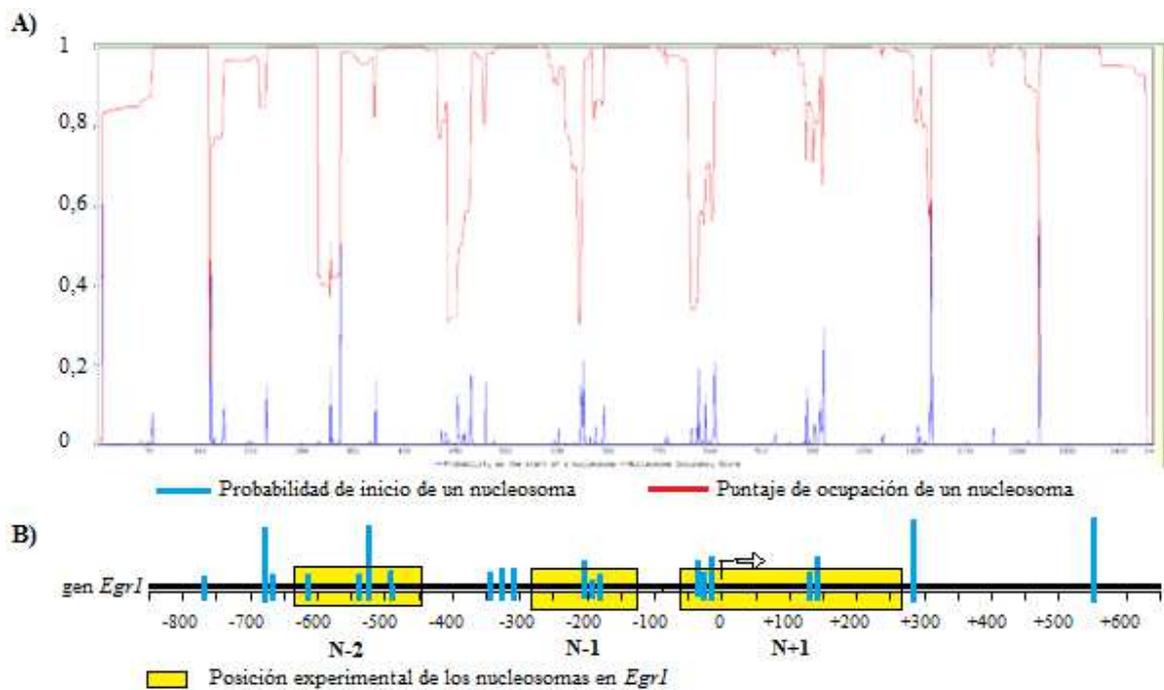


Figura 50. Resultados de la predicción realizada por NuPop en *Egr1*. A) Predicción de la secuencia nucleosomal (en rojo) y predicción del inicio de un nucleosoma (en azul), en escala de 0 a 1. B) Representación de las zonas de inicio, comparadas con los nucleosomas identificados experimentalmente para *Egr1* (recuadro amarillo).

Los resultados de la predicción realizada por nuScore para los nucleosomas en *Egr1* se muestran en la figura 51. La energía parece disminuir en la zona entre el -800 y el -600 y entre 0 y +300 (fig. 51), lo cual indicaría que en estas regiones del ADN existe una mayor probabilidad de que se forme un nucleosoma. NuScore parece predecir mejor la posición de los nucleosomas en cuanto al puntaje de energía.

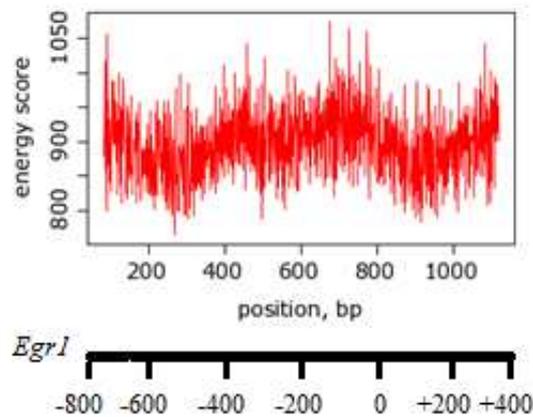


Figura 51. Resultados de la predicción realizada por nuScore en *Egr1*, en cuanto al puntaje de energía requerida para curvar el ADN. En la parte inferior se muestra el equivalente a la posición en *Egr1*.

La base de datos en la cual se sustentan los programas de predicción, es otro factor a tener en cuenta, ya que en gran medida, de ésta depende la capacidad de predicción que posean. Por ello se comparan los resultados de la posición de los nucleosomas identificados en el presente trabajo, con los datos almacenados en el mapa de la posición de nucleosomas del genoma de hígado de ratón proporcionada por GEO. En la presente tesis se trabaja con células progenitora de hepatocitos de ratón, por lo que la base de datos más adecuada para el diseño de un programa sería la que proporciona GEO.

Para realizar la comparación de los datos experimentales, con los que posee la base de datos, primero se identificó la región del genoma de ratón correspondiente a la secuencia del gen *Egr1* (fig. 52), en donde la posición 35020861 en el cromosoma 18 del ratón corresponde al +1 del gen *Egr1* (fig. 53).

Posteriormente se indica la longitud de los fragmentos protegidos frente a la MNasa almacenados en la base de datos y la longitud del espacio entre nucleosomas, tal como se muestra en la figura 52. En la mayoría de los casos la longitud no corresponde al tamaño nucleosomal. De la franja roja hacia abajo, se muestra la posición de los nucleosomas en la región transcribible, en el otro extremo se muestran los nucleosomas de la región promotora. Las coordenadas (en pb) que ocupa el gen *Egr1* dentro del genoma del ratón, fueron extraídas de la base de datos MGI, en donde se indica que la región codificante del gen *Egr1* es Chr18:35020861-35024638 bp, mientras que la región promotora se localiza 1300pb desde el 35020861, río arriba.

Región entre nucleosomas	desde	hasta	Nucleosoma
41 pb	chr18 35018748	35019006	name 19 + — 258 pb
37 pb	chr18 35019047	35019176	name 9 + — 129 pb
17 pb	chr18 35019213	35019428	name 17 + — 215 pb
96 pb	chr18 35019445	35019584	name 12 + — 139 pb
66 pb	chr18 35019680	35019758	name 6 + — 78 pb
598 pb	chr18 35019824	35020344	name 21 + — 520 pb
114 pb	chr18 35020942	35021171	name 9 + — 229 pb
63 pb	chr18 35021285	35021681	name 29 + — 396 pb
258 pb	chr18 35021744	35021881	name 9 + — 137 pb
14 pb	chr18 35022139	35022588	name 18 + — 449 pb
38 pb	chr18 35022602	35022799	name 15 + — 197 pb
101 pb	chr18 35022837	35022859	name 5 + — 22 pb
89 pb	chr18 35022960	35022965	name 5 + — 5 pb
22 pb	chr18 35023054	35023310	name 13 + — 256 pb
48 pb	chr18 35023332	35023337	name 4 + — 5 pb
217 pb	chr18 35023385	35023513	name 8 + — 128 pb
7 pb	chr18 35023730	35024204	name 31 + — 474 pb
370 pb	chr18 35024211	35024339	name 9 + — 128 pb
	chr18 35024709	35024848	name 9 + — 139 pb

Figura 52. Región de la base de datos de nucleosomas de hígado de ratón correspondiente al gen *Egr1*. En los costados se indica la longitud del ADN nucleosomal e internucleosomal. La línea roja identifica la región cercana al inicio de la región transcribible.

Como se observa en la figura 53, la zona transcribible de *Egr1* comienza en la posición 35020861 (en celeste) del cromosoma 18, de acuerdo a lo indicado en la base de datos 81 pb río abajo comienza una región nucleosomal, este fragmento de 229 pb corresponde seguramente al

nucleosoma N+1, cuya localización experimental se extiende aproximadamente entre el inicio de la transcripción y +300 pb (fig. 45). Desde la posición 35020942 hasta la posición 35020344 río arriba, se localiza un fragmento entre nucleosomas de 598 pb, al que le sigue una zona nucleosomal de 520 pb (desde 35020344, hasta 35019824 río arriba). En letras rojas subrayadas se indica la posición de los regiones nucleosomales identificados experimentalmente, las cuales se representan con los dos amplicones que poseen una mayor protección frente a MNasa. Destacadas en amarillo se muestran las regiones nucleosomales que se encuentran en la base de datos para la misma zona del gen *Egr1*.

Estos resultados sugieren que la base de datos más adecuada para el presente trabajo, no muestra una concordancia en la posición del nucleosoma N-1, presentando una concordancia parcial con la posición del nucleosoma N-2. En cambio para el nucleosomas N+1 la posición de la base de datos coincide con la posición determinada experimentalmente.

En nuestro caso concreto los programas actuales no han tenido éxito en la predicción del ADN que tiene mayor probabilidad de formar nucleosomas, ya que existen discrepancias con respecto a los resultados experimentales, es por esto que se requiere de nuevos programas que permitan predecir un posicionamiento nucleosomal de manera más exacta. El último objetivo de este trabajo es plantear la necesidad de crear un nuevo programa para la predicción de la posición de los nucleosomas.

La base de datos podría ser empleada como fuente de información que permita comparar la composición de la secuencia de los nucleosomas y a partir de esta información crear un programa para la predicción de la posición de los nucleosoma, pero se debe tener presente que la longitud de los fragmentos nucleosomales no corresponden a las 147 pb que posee un nucleosoma y que la concordancia con los datos experimentales es solo parcial.

GCTTGGGGTCCAGACTGCCAAAGCC AAGTCCCTCTTTAGGACAGGGCAGGGTTCGTGCCGACAGTCCCTGGCCT
GGATAAAAGTCAGGAAGTGTCTAACCATCACAAAGAACCAACAGATCCTGGCGGGGACTTAGGACTGACCTAGAACAAATC
AGGGTTCGC AATCCAGGTC CCCAAAGGTGGGATCCTCAACCGCAGGACGGAGGGAATAGCCTTTGGATTCTGGGTGGT
CATTGGAAGC CCCAGGCTCTAAAACCCCAACCTACTGACTGGTGGCCGAGTATGCACCOCGACTGCTAGCTAGGCAGTGT
CCCAAGAACCAGTAGCCAAATGTCTTGGCC TCAGTTTTCCCGGTGACACCTGGAAAAGTGACCC TGCCATTAGTAGAGGCTC
AGGTCAGGGCCCCGCCCTCTCC TGGGCGGCCCTGTGCCCTAGCCCGCCCTGCCGC TCCTCCTCCGCAGGCTCGC TOCCACG
GTCCCCGAGGTGGGCGGGTGAGCC CAGGATGAOCCGCTGTAGAACCCCGGCCCTGACTCGCC TCGCCCCCGGCCGGGCT
GGGCTTCCCTAGCCAGCTCGCACCCGGGGGCCGTCGGAGCCGC CGCGCGCCAGCTCTACGGCCCTGGCGCCCTCCCA
CGCGGGC GTCCCGACTCCCGCGCGCGCTCAGGCTCCAGTTGGGAACCAAGGAGG GGGAGGATGGGGGGGGGGTGT
GCGCCGACCCGAAACGCCATATAAGGAGCAGGAAGGATCCCCGCCGGAACAGACCTTATTGGGACAGCGCTTATATG
GAGTGGCCCAATATGGCCCTGCCCTTC CGGCTCTGGGAGGAGGGGC GAGCGGGGTTGGGGCGGGGGCAAGCTGGGAA
CTCAGGCGCTGGCCCGGAGGCCACTGCTGCTGTTCAGAACTACTAGGCTTTCCAGGAGCCTGAGCGCTCGCGATGCCCG
AGCGGGTCCAGGGTGGAGGTGCCACCACTCTTGGATGGGAGGGCTTCACGTC ACTCCGGTCTCCCGGCCGGTCTC
CATATTAGGGCTTCCTGCTTCCATATATGGCCATGTACGTCACGGCGGAGGGGGCCCGTGTGTTCAGACCTTGAAA
TAGAGGCCGATTCGGGAGTCCGAGAGATCCAGCGCGCAGAACTTGGGAGCGCCGCCCGC GATTGCGCGCCGCCCGCC
AGCTTCCGCCCGCAGATCGGCCCTGCCCCAGCCTCCGCGGACGCCCTGGTCC ACCACGGGCCCGGCTACCGCC
GCCTGGGGGCCCACTACACTCCCGCAGTGTGCCCTGCACCCCGCATGTAACCCGGCCAAACCCCGCGAGTGTGCCCT
CAGTAGCTTCGGCCCGGGCTGCGCCACCACCAACATCAGTTCOCAGCTCGCTGGTCCGGATGGCAGCGGCCAAGG
CCGAGATGCAATTGATGTC TCCGCTGCAGATCTGTGACCCGTTCCGGCTCCTTCC TCACTCACCCACCATGGACAACCTACCC
CAAAC TGAGGAGATGATGCTGCTGAGCAACGGGGTCC CAGTTCCTGGTGTGCCGGAACCCAGAGGGCAGGGCC
GTAATAGCAGCAGCACACAGCAGCGGGGGCGGTGGTGGGGCGGCAGCAACAGCGGCAGCAGCGCTTCAATCTCA
AGGGGAGCCGAGCGAACAACTATGAGCACTGACCACAGGTAAGCGGTGGTCTGCGCCGAGGCTGAATCCCCCTTCGT
GACGACCCAAACGTCAGTCTTTCCTGCACGGACCTGTTCTATCCCTTAGGGATGGGAC TGGGGTTTCCCTCTATTTCAC
ACAGCTCCAGGGTCTTGTGTTAGAGGGATGTCTAGGGACACCOCCTACCC TCCATCC TTGCGGGTGC CCGGAAGGCAGGA
CGTTTGT TTTGGATGGAGAATCAAGCTGCGTGGGTGGTGGAGGGGGTGGGGGGAGGGTTTGT TTTGATGAGCGGGGT
TGCCCCCTCCCGCGCGCGCGGTGTGCGCGCC TGTGTTGAGATGTTCCCAAGGCAGGGCTGAAATCTGTGAC CAGG

Región entre nucleosomas	desde	hasta	Nucleosoma
96 pb chr18	35019445	35019584	— 139 pb
66 pb chr18	35019680	35019758	— 78 pb
598 pb chr18	35019824	35020344	— 520 pb
114 pb chr18	35020942	35021171	— 229 pb
63 pb chr18	35021285	35021681	— 396 pb
chr18	35021744	35021881	— 137 pb

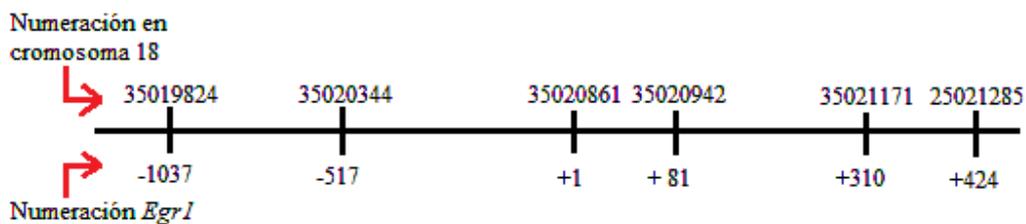


Figura 53. Esquema comparativo de la posición de los nucleosomas del promotor proximal y al inicio de la región transcribible, entre lo señalado en la base de datos (en amarillo) y los resultados obtenidos experimentalmente (en letra y subrayado rojo), en donde N-2 está representado por -550 y -498, N-1 por -240 y -178, y N+1 por los amplicones +171 y +241. En celeste el inicio de la región transcribible.

6. CONCLUSIONES

- Las condiciones determinadas para la extracción exclusiva de mononucleosomas son: Digerir 20 min con 500 U de MNasa y 3 mM de CaCl₂ en el tampón BG.
- La posición de los nucleosomas del promotor proximal y al inicio de la región transcribible queda definida entre los amplicones -666 y -370 para el nucleosoma N-2, entre los amplicones -307 y -126 para el nucleosoma N-1 y entre el -69 y +305 para el nucleosoma N+1.
- En la región ocupada por los nucleosomas N-2, N-1 y N+1 disminuye la protección frente a MNasa luego de inducir la expresión de *Egr1* con TPA 30 min. El nucleosoma N-2, además, se desplaza una 50 pb río abajo, dejando libre la región de acceso a EGR1 para su autorregulación. Esto demuestra la importancia que presenta el remodelamiento de la cromatina en la regulación de la transcripción.
- Los remodeladores BRM y BRG1 se encuentran presentes en el nucleosoma N+1 cuando *Egr1* se halla inactivo, sin embargo cuando el gen es activado con TPA 15 min, BRM abandona el nucleosoma y BRG1 se ve más representado. Por lo que este último podría estar involucrado en la remodelación del nucleosoma N+1, lo que causa la pérdida parcial de la estructura del nucleosoma, que se ve reflejado en la disminución de la protección frente a la digestión con MNasa.
- Los nucleosomas N-2 y N-1 se encuentran acetilados en H3K9, cuando el gen *Egr1* es activado con TPA 30 min, mientras que el nucleosoma N+1 se encuentra trimetilado.
- Los programas bioinformáticos para predicción de la posición nucleosomas NXSensor, NuPop y NuScore, en este caso, no fueron asertivos a la hora de predecir la posición de los nucleosomas del promotor proximal y al inicio de la región transcribible en *Egr1* (N-2, N-1,

N+1), en comparación a las posiciones identificadas experimentalmente. Por lo que se hace patente la necesidad de diseñar nuevos programas bioinformáticos que permitan predecir con mayor exactitud la zona del ADN que forma parte de un nucleosoma.

- La base de datos de nucleosomas de hígado de ratón, no mostró concordancia con la posición del nucleosoma N-1 y la coincidencia mostrada para la posición de N-2 es parcial, sin embargo la posición del nucleosoma N+1 concuerda con lo registrado en la base de datos.

7. LITERATURA CITADA

ALLIS CD, BERGER SL, COTE J, DENT S, JENUWIEN T, KOUZARIDES T, PILLUS L, REINBERG D, SHI Y, SHIEKHATTAR R, SHILATIFARD A, WORKMAN J, ZHANG Y. 2007. “New nomenclature for chromatin-modifying enzymes”. *Cell*. Nov 16; 131(4):633-6.

BANNISTER AJ, KOUZARIDES T. 2011. “Regulation of chromatin by histone modifications”. *Cell Res*. Mar; 21(3):381-95. Review.

BAO Y, SHEN X. 2007. “SnapShot: Chromatin Remodeling Complexes”. *Cell*, Volume 129, Issue 3, 4 May, Pages 632.e1-632.e2, ISSN 0092-8674.

BATTISTINI F, HUNTER CA, GARDINER EJ, PACKER MJ. 2010. “Structural mechanics of DNA wrapping in the nucleosome”. *J Mol Biol*. Feb 19; 396(2):264-79.

BHATTACHARYYA S, SARGENT JL, DU P, LIN S, TOURTELLOTTE WG, TAKEHARA K, WHITFIELD ML, VARGA J. 2011a. “Egr-1 induces a profibrotic injury/repair gene program associated with systemic sclerosis”. *PLoS One*. 6(9):e23082.

BHATTACHARYYA S, WU M, FANG F, TOURTELLOTTE W, FEGHALI-BOSTWICK C, VARGA J. 2011b. “Early growth response transcription factors: key mediators of fibrosis and novel targets for anti-fibrotic therapy”. *Matrix Biol*. May; 30(4):235-42.

BOONE DN, QI Y, LI Z, HANN SR. 2011. “*Egr1* mediates p53-independent c-Myc-induced apoptosis via a noncanonical ARF-dependent transcriptional mechanism”. *Proc Natl Acad Sci U S A*. Jan 11; 108(2):632-7.

- BOZON B, DAVIS S, LAROCHE S.** 2002. "Regulated transcription of the immediate-early gene Zif268: mechanisms and gene dosage-dependent function in synaptic plasticity and memory formation". *Hippocampus*. 12(5):570-7. Review.
- BRAUN AC.** 1981. "An epigenetic model for the origin of cancer". *Q Rev Biol*. Mar; 56(1):33-60. Review.
- CHRISTY B, NATHANS D.** 1989. "Functional serum response elements upstream of the growth factor-inducible gene zif268". *Mol Cell Biol*. Nov; 9(11):4889-95.
- CHUNG HR, VINGRON M.** 2009. "Sequence-dependent nucleosome positioning". *J Mol Biol*. Mar 13; 386(5):1411-22.
- CLAPIER CR, CAIRNS BR.** 2009. "The biology of chromatin remodeling complexes". *Annu Rev Biochem*. 78:273-304. Review.
- CUI P, ZHANG L, LIN Q, DING F, XIN C, FANG X, HU S, YU J.** 2010. "A novel mechanism of epigenetic regulation: nucleosome-space occupancy". *Biochem Biophys Res Commun*. Jan 1;391(1):884-9."
- DARLAND T, SAMUELS M, EDWARDS SA, SUKHATME VP, ADAMSON ED.** 1991. "Regulation of Egr-1 (Zfp-6) and c-fos expression in differentiating embryonal carcinoma cells". *Oncogene*. Aug; 6(8):1367-76.
- DEFRANCO C, RO M, GROSSEL M, ENGLISH MA, HANSEN UM, WAGNER JA, LICHT JD.** 1993. "NGFIA (EGR1) contains transcription activating domains in both the amino terminal and carboxyl terminal regions of the protein". *Biochem Biophys Res Commun*. Jul 15; 194(1):425-31.

- DEGUCHI T, FUJIMORI KE, KAWASAKI T, XIANGHAI L, YUBA S.** 2009. "Expression patterns of the Egr1 and Egr3 genes during medaka embryonic development". *Gene Expr Patterns*. Apr; 9(4):209-14.
- DILLON N.** 2004. "heterochromatin structure and function". *Biology of the cell* 96 (2004) 631-637.
- EBIHARA K, ISHIDA Y, TAKEDA R, ABE H, MATSUO H, KAWAI K, MAGATA Y, NISHIMORI T.** 2011. "Differential expression of FosB, c-Fos, and Zif268 in forebrain regions after acute or chronic L-DOPA treatment in a rat model of Parkinson's disease". *Neurosci Lett*. Jun 1; 496(2):90-4.
- EDELMAN GM.** 1987. "Epigenetic rules for expression of cell adhesion molecules during morphogenesis". *Ciba Found Symp*. 125: 192-216. Review.
- EINBOND LS, WU HA, SU T, CHANG T, PANJIKARAN M, WANG X, GOLDSBERRY S.** 2010. "Digitoxin activates EGR1 and synergizes with paclitaxel on human breast cancer cells". *J Carcinog*. Nov 18; 9:10.
- FERNANDEZ AG, ANDERSON JN.** 2007. "Nucleosome positioning determinants". *J Mol Biol*. Aug 17; 371(3):649-68.
- FUKUDA N.** 2010. "Cigarette smoking induces vascular proliferative disease through the activation of Egr-1". *Cardiovasc Res*. Nov 1;88(2):207-8. Epub 2010 Sep 10.
- GITENAY D, BARON VT.** 2009. "Is EGR1 a potential target for prostate cancer therapy?" *Future Oncol*. Sep; 5(7):993-1003. Review.

- GOMEZ-MARTIN D, DIAZ-ZAMUDIO M, GALINDO-CAMPOS M, ALCOCER-VARELA J.** 2010. "Early growth response transcription factors and the modulation of immune response: implications towards autoimmunity". *Autoimmun Rev.* Apr; 9(6):454-8.
- HAMADA N, MIYATA M, ETO H, IKEDA Y, SHIRASAWA T, AKASAKI Y, MIYAUCHI T, FURUSHO Y, NAGAKI A, ARONOW BJ, TEI C.** 2011. "Loss of clusterin limits atherosclerosis in apolipoprotein E-deficient mice via reduced expression of Egr-1 and TNF- α ". *J Atheroscler Thromb*. Mar 28; 18(3):209-16.
- HERMAN D, JENSSEN K, BURNETT R, SORAGNI E, PERLMAN SL, GOTTESFELD JM.** 2006. "Histone deacetylase inhibitors reverse gene silencing in Friedreich's ataxia". *Nat Chem Biol.* Oct; 2(10):551-8.
- HOLLIDAY R.** 1987. "DNA methylation and epigenetic defects in carcinogenesis". *Mutat Res.* Dec;181(2):215-7. Review.
- INGLIS J, ALLIS C.D, JENUWEIN T, REINBERG D.** 2007, "Epigenetics". Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. Printed in the United States of America. ISBN-13:978-0-87969-724-2. 195 p.
- KATAOKA F, TSUDA H, ARAO T, NISHIMURA S, TANAKA H, NOMURA H, CHIYODA T, HIRASAWA A, AKAHANE T, NISHIO H, NISHIO K, AOKI D.** 2011. "EGRI and FOSB gene expressions in cancer stroma are independent prognostic indicators for epithelial ovarian cancer receiving standard therapy". *Genes Chromosomes Cancer.* Nov 17. doi: 10.1002/gcc.21916.
- KHESIN RB, LEĪBOVICH BA.** 1976. "Chromosome structure, histones and gene activity in *Drosophila*". *Mol Biol (Mosk).* Jan-Feb;10(1):3-34. Review.

- KOUZARIDES T.** 2007a. “Chromatin modifications and their function”. *Cell*. Feb 23; 128(4):693-705. Review.
- KOUZARIDES T.** 2007b. “SnapShot: Histone-modifying enzymes”. *Cell*. Nov 16;131(4):822.
- LÄNGST G, BECKER PB.** 2001. “Nucleosome mobilization and positioning by ISWI-containing chromatin-remodeling factors”. *J Cell Sci*. Jul;114(Pt 14):2561-8. Review.
- LEFEVRE P, BONIFER C.** 2006. “Analyzing histone modification using crosslinked chromatin treated with micrococcal nuclease”. *Methods Mol Biol*. 325:315-25.
- LEVITSKY VG, KATOKHIN AV, PODKOLODNAYA OA, FURMAN DP, KOLCHANOV NA.** 2005. “NPRD: Nucleosome Positioning Region Database”. *Nucleic Acids Res*. Jan 1;33 (Database issue): D67-70.
- LI YJ, FU XH, LIU DP, LIANG CC.** 2004. “Opening the chromatin for transcription”. *Int J Biochem Cell Biol*. Aug; 36(8):1411-23. Review.
- LI Z, SCHUG J, TUTEJA G, WHITE P, KAESTNER KH.** 2011. “The nucleosome map of the mammalian liver”. *Nat Struct Mol Biol*. Jun; 18(6):742-6. Epub 2011 May 29.
- LUYKX P, BAJIĆ IV, KHURI S.** 2006. “NXSensor web tool for evaluating DNA for nucleosome exclusion sequences and accessibility to binding factors”. *Nucleic Acids Res*. Jul 1; 34(Web Server issue):W560-5.
- MELLO CV, CLAYTON DF.** 1994. “Song-induced ZENK gene expression in auditory pathways of songbird brain and its relation to the song control system”. *J Neurosci*. Nov;14(11 Pt 1):6652-66

MUTHUKKUMAR S, NAIR P, SELLS SF, MADDIWAR NG, JACOB RJ, RANGNEKAR VM. 1995. "Role of EGR-1 in thapsigargin-inducible apoptosis in the melanoma cell line A375-C6". *Mol Cell Biol.* Nov; 15(11):6262-72.

NAIR P, MUTHUKKUMAR S, SELLS SF, HAN SS, SUKHATME VP, RANGNEKAR VM. 1997. "Early growth response-1-dependent apoptosis is mediated by p53". *J Biol Chem.* Aug 8; 272(32):20131-8.

NELSON D.L. Y COX M. 2006, "Lehninger, principios de bioquímica", cuarta edición, ISBN: 8428214107. pag 940- 956.

PAGEL JI, DEINDL E. 2011. "Early growth response 1--a transcription factor in the crossfire of signal transduction cascades". *Indian J Biochem Biophys.* Aug; 48(4):226-35.

PIATTI P, ZEILNER A, LUSSER A. 2011. "ATP-Dependent Chromatin Remodeling Factors and Their Roles in Affecting Nucleosome Fiber Composition". *Int J Mol Sci.* 12(10):6544-65.

POIRIER R, CHEVAL H, MAILHES C, GAREL S, CHARNAY P, DAVIS S, LAROCHE S. 2008. "Distinct functions of egr gene family members in cognitive processes". *Front Neurosci.* Jul; 2(1):47-55.

RIVERA RM, BENNETT LB. 2010. "Epigenetics in humans: an overview". *Curr Opin Endocrinol Diabetes Obes.* Dec; 17(6):493-9. Review.

SACILTTO N. 2010. "The Growth Arrest-Specific gene 1 in Hepatic cell proliferation: Functional role and Epigenetic. Tesis presentada ante la Universidad de Valencia – España, como requisito para obtener el grado de PhD en Biología Molecular.

SAKAMOTO KM, BARDELEBEN C, YATES KE, RAINES MA, GOLDE DW, GASSON JC. 1991. “5' upstream sequence and genomic structure of the human primary response gene, EGR-1/TIS8”. *Oncogene*. May; 6(5):867-71.

SAKURAI K, FURUKAWA C, HARAGUCHI T, INADA K, SHIOGAMA K, TAGAWA T, FUJITA S, UENO Y, OGATA A, ITO M, TSUTSUMI Y, IBA H. 2011. “MicroRNAs miR-199a-5p and -3p target the Brm subunit of SWI/SNF to generate a double-negative feedback loop in a variety of human cancers”. *Cancer Res*. Mar 1; 71(5):1680-9.

SCHILTZ RL, MIZZEN CA, VASSILEV A, COOK RG, ALLIS CD, NAKATANI Y. 1999. Overlapping but distinct patterns of histone acetylation by the human coactivators p300 and PCAF within nucleosomal substrates. *J Biol Chem* 274: 1189–92.

SELVI BR, MOHANKRISHNA DV, OSTWAL YB, KUNDU TK. 2010. “Small molecule modulators of histone acetylation and methylation: a disease perspective”. *Biochim Biophys Acta*. Oct-Dec; 1799(10-12):810-28.

SHI Y, LAN F, MATSON C, MULLIGAN P, WHETSTINE JR, COLE PA, CASERO RA, SHI Y. 2004. “Histone demethylation mediated by the nuclear amine oxidase homolog LSD1”. *Cell*. Dec 29; 119(7):941-53.

SHIMIZU N, OHTA M, FUJIWARA C, SAGARA J, MOCHIZUKI N, ODA T, UTIYAMA H. 1992. “A gene coding for a zinc finger protein is induced during 12-O-tetradecanoylphorbol-13-acetate-stimulated HL-60 cell differentiation”. *J Biochem*. Feb; 111(2):272-7.

STERNER DE, BERGER SL. 2000. “Acetylation of histones and transcription-related factors”. *Microbiol Mol Biol Rev*. Jun; 64(2):435-59. Review.

STEWART N, SANO H. 2004. “Measuring changes in chromatin using micrococcal nuclease”. *Methods Mol Biol*. 287:65-75.

STRAHL BD, ALLIS CD. 2000. “The language of covalent histone modifications”. *Nature*. Jan 6; 403(6765):41-5.

SUKHATME VP, KARTHA S, TOBACK FG, TAUB R, HOOVER RG, TSAI-MORRIS CH. 1987. “A novel early growth response gene rapidly induced by fibroblast, epithelial cell and lymphocyte mitogens”. *Oncogene Res. Sep-Oct*; 1(4):343-55.

SUKHATME VP. 1990. “Early transcriptional events in cell growth: the Egr family”. *J Am Soc Nephrol. Dec*; 1(6):859-66. Review.

TEIF VB, RIPPE K. 2009. “Predicting nucleosome positions on the DNA: combining intrinsic sequence preferences and remodeler activities”. *Nucleic Acids Res. Sep*;37(17):5641-55.

TOLLEFSBOL T, 2011, Editor(s), *Handbook of Epigenetics*, Academic Press, San Diego, Pages i,iii, ISBN 9780123757098, 10.1016/B978-0-12-375709-8.00038-1.

TOLSTORUKOV MY, CHOUDHARY V, OLSON WK, ZHURKIN VB, PARK PJ. 2008. “nuScore: a web-interface for nucleosome positioning predictions”. *Bioinformatics. Jun* 15;24(12):1456-8.

TUR G. 2007. “Transcriptional regulation of egr-1 gene in murine cells. Towards the understanding of the role of chromatin”. Tesis doctoral para optar el grado de Dr. En Bioquímica, presentada ante la Universidad de Valencia – España.

TUR G, GEORGIEVA EI, GAGETE A, LÓPEZ-RODAS G, RODRÍGUEZ JL, FRANCO L. 2010. “Factor binding and chromatin modification in the promoter of murine Egr1 gene upon induction”. *Cell Mol Life Sci. Dec*; 67(23):4065-77.

WATSON JD, CRICK FH. 1953. “The structure of DNA”. *Cold Spring Harb Symp Quant Biol.* 18:123-31.

- XI L, FONDUFE-MITTENDORF Y, XIA L, FLATOW J, WIDOM J, WANG JP.** 2010. “Predicting nucleosome positioning using a duration Hidden Markov Mode”l. *BMC Bioinformatics*. Jun 24;11:346.
- XIE B, WANG C, ZHENG Z, SONG B, MA C, THIEL G, LI M.** 2011. “Egr-1 transactivates Bim gene expression to promote neuronal apoptosis”. *J Neurosci*. Mar 30; 31(13):5032-44.
- XING Y, ZHAO X, CAI L.** 2011. “Prediction of nucleosome occupancy in *Saccharomyces cerevisiae* using position-correlation scoring function”. *Genomics*. Nov; 98(5):359-66.
- YANG W, ZHAO JG, LI XY, GONG SL, CAO JP.** 2010. “Anti-tumor effect of 125I-UdR in combination with Egr-1 promoter-based IFN γ gene therapy in vivo”. *Radiat Environ Biophys*. Nov; 49(4):705-14.
- YOST JM, KORBOUKH I, LIU F, GAO C, JIN J.** 2011. “Targets in epigenetics: inhibiting the methyl writers of the histone code”. *Curr Chem Genomics*. 5(Suppl 1):72-84.
- YOUNG TJ, KIRCHMAIER AL.** 2011. “Cell cycle regulation of silent chromatin formation”. *Biochim Biophys Acta*. Oct 25.
- YU J, DE BELLE I, LIANG H, ADAMSON ED.** 2004. “Coactivating factors p300 and CBP are transcriptionally crossregulated by Egr1 in prostate cells, leading to divergent responses”. *Mol Cell*. Jul 2; 15(1):83-94.

8. ANEXOS

Anexo 1. Nueva nomenclatura de enzimas que modifican histonas, en cuadro 1, 2 y 3 (Allis *et al* 2007).

Cuadro 1. Nueva nomenclatura de las enzimas lisinas desmetilasas, sustrato en el actúan y organismo en el que se encuentra cada enzima, además de la función de la modificación.

New Name	Human	<i>D. melanogaster</i>	<i>S. cerevisiae</i>	<i>S. pombe</i>	Substrate Specificity	Function
KDM1	LSD1/BHC110	Su(var)3-3		SpLsd1/Swm1/Saf110	H3K4me1/2, H3K9me1/2	Transcription activation and repression, heterochromatin formation
KDM2			Jhd1		H3K36me1/2	Transcription elongation
KDM2A	JHDM1a/FBXL11				H3K36me1/2	
KDM2B	JHDM1b/FBXL10				H3K36me1/2	
KDM3A	JHDM2a				H3K9me1/2	Androgen receptor gene activation, spermatogenesis
KDM3B	JHDM2b				H3K9me	
KDM4			Rph1		H3K9/K36me2/3	Transcription elongation
KDM4A	JMJD2A/JHDM3A				H3K9/K36me2/3	Transcription repression, genome integrity
KDM4B	JMJD2B				H3K9/H3K36me2/3	Heterochromatin formation
KDM4C	JMJD2C/GASC1				H3K9/K36me2/3	Putative oncogene
KDM4D	JMJD2D				H3K9me2/3	
KDM5		Lid	Jhd2	Jmj2	H3K4me2/3	
KDM5A	JARID1A/RBP2				H3K4me2/3	Retinoblastoma-interacting protein
KDM5B	JARID1B/PLU-1				H3K4me1/2/3	Transcription repression
KDM5C	JARID1C/SMCX				H3K4me2/3	X-linked mental retardation
KDM5D	JARID1D/SMCY				H3K4me2/3	Male-specific antigen
KDM6A	UTX				H3K27me2/3	Transcription activation
KDM6B	JMJD3				H3K27me2/3	Transcription activation

Cuadro 2. Nueva nomenclatura de las enzimas lisina metiltransferasas, sustrato en el actúan y organismo en el que se encuentra cada enzima, además de la función de la modificación.

New Name	Human	<i>D. melanogaster</i>	<i>S. cerevisiae</i>	<i>S. pombe</i>	Substrate Specificity	Function
KMT1		Su(Var)3-9		Clr4	H3K9	Heterochromatin formation/silencing
KMT1A	SUV39H1				H3K9	Heterochromatin formation/silencing
KMT1B	SUV39H2				H3K9	Heterochromatin formation/silencing
KMT1C	G9a				H3K9	Heterochromatin formation/silencing
KMT1D	EuHMTase/GLP				H3K9	Heterochromatin formation/silencing
KMT1E	ESET/SETDB1				H3K9	Transcription repression
KMT1F	CLL8					
KMT2			Set1	Set1	H3K4	Transcription activation
KMT2A	MLL1	Trx			H3K4	Transcription activation
KMT2B	MLL2	Trx			H3K4	Transcription activation
KMT2C	MLL3	Trr			H3K4	Transcription activation
KMT2D	MLL4	Trr			H3K4	Transcription activation
KMT2E	MLL5				H3K4	Transcription activation
KMT2F	hSET1A				H3K4	Transcription activation
KMT2G	hSET1B				H3K4	Transcription activation
KMT2H	ASH1	Ash1			H3K4	Transcription activation
KMT3			Set2	Set2	H3K36	Transcription activation
KMT3A	SET2				H3K36	Transcription activation
KMT3B	NSD1				H3K36	
KMT3C	SYMD2				H3K36 (p53)	Transcription activation
KMT4	DOT1L		Dot1		H3K79	Transcription activation
KMT5				Set9	H4K20	DNA-damage response
KMT5A	Pr-SET7/8	PR-set7			H4K20	Transcription repression
KMT5B	SUV4-20H1	Suv4-20			H4K20	DNA-damage response
KMT5C	SUV4-20H2					
KMT6	EZH2	E(Z)			H3K27	Polycomb silencing
KMT7	SET7/9				H3K4 (p53 and TAF10)	
KMT8	RIZ1				H3K9	Transcription repression

Cuadro 3. Nueva nomenclatura de las enzimas lisina acetiltransferasas, sustrato en el actúan y organismo en el que se encuentra cada enzima, además de la función de la modificación.

New Name	Human	<i>D. melanogaster</i>	<i>S. cerevisiae</i>	<i>S. pombe</i>	Substrate Specificity	Function
KAT1	HAT1	CG2051	Hat1	Hat1/ Hag603	H4 (5, 12)	Histone deposition, DNA repair
KAT2		dGCN5/PCAF	Gcn5	Gcn5	H3 (9, 14, 18, 23, 36)/ H2B; yHtz1 (14)	Transcription activation, DNA repair
KAT2A	hGCN5				H3 (9, 14, 18)/H2B	Transcription activation
KAT2B	PCAF				H3 (9, 14, 18)/H2B	Transcription activation
KAT3		dCBP/NEJ			H4 (5, 8); H3 (14, 18)	Transcription activation, DNA repair
KAT3A	CBP				H2A (5); H2B (12, 15)	Transcription activation
KAT3B	P300				H2A (5); H2B (12, 15)	Transcription activation
KAT4	TAF1	dTAF1	Taf1	Taf1	H3 > H4	Transcription activation
KAT5	TIP60/PLIP	dTIP60	Esa1	Mst1	H4 (5, 8, 12, 16); H2A (yeast 4, 7; chicken 5, 9, 13, 15); dH2Av/yHtz1 (14)	Transcription activation, DNA repair
KAT6		(CG1894)	Sas3	(Mst2)	H3 (14, 23)	Transcription activation and elongation, DNA replication
KAT6A	MOZ/MYST3	ENOK			H3 (14)	Transcription activation
KAT6B	MORF/MYST4				H3 (14)	Transcription activation
KAT7	HBO1/MYST2	CHM		(Mst2)	H4 (5, 8, 12) > H3	Transcription, DNA replication
KAT8	HMOF/MYST1	dMOF (CG1894)	Sas2	(Mst2)	H4 (16)	Chromatin boundaries, dosage compensation, DNA repair
KAT9	ELP3	dELP3/ CG15433	Elp3	Elp3	H3	
KAT10			Hap2		H3 (14); H4	
KAT11			Rtt109		H3 (56)	Genome stability, transcription elongation
KAT12	TFIIIC90				H3 (9, 14, 18)	Pol III transcription
KAT13A	SRC1				H3/H4	Transcription activation
KAT13B	ACTR				H3/H4	Transcription activation
KAT13C	P160				H3/H4	Transcription activation
KAT13D	CLOCK				H3/H4	Transcription activation

Anexo 2. Protocolo de laboratorio para extracción de nucleosomas a partir de cultivos celulares

Se descarta el medio de cultivo, luego se lava con 10 ml de PBS por cada botella de 75 cm², posteriormente se añaden 10 ml de formaldehído al 1% en PBS y se incuba en agitación a temperatura ambiente (T.A.) durante 7 min. El entrecruzamiento se detiene con glicina 0,125 M (concentración final) y se incuba en agitación a T.A. durante 5 min. Se lava dos veces con 10 ml de PBS frío y a continuación se añaden 10 ml de PBS frío para rascar y recoger las células. Se lavan las botellas con 2 ml más de PBS frío para recuperar las células que hayan quedado y se centrifuga a 4500 g 5 min. Se elimina el sobrenadante y posteriormente las células se resuspenden en 6 ml de tampón de lisis celular frío (contiene espermina 0,15 mM, espermidina 0,5 mM, Tris-HCl pH 7.5 a 0,01 M, NaCl 0,01 M, MgCl₂ 0,003 M, sacarosa 0,03 M, EDTA 10 mM y 0,5% de NP-40) más inhibidores de proteasas (2 µl/mL), se traspasa a un homogeneizador de Potter, en donde las células se lisan. Luego se incuba en hielo de 10 a 15 min. Se centrifuga a 800 g por 5 min a 4°C para depositar los núcleos y se elimina el sobrenadante. Los núcleos se resuspenden en 1 mL de tampón de lavado RSB (contiene Tris-HCl pH 7.5 50 mM, NaCl 10 mM, MgCl₂ 3 mM, β-Mercaptoetanol 10 mM, espermina 0,15 mM, espermidina 0,5 mM y 2 µl/mL de inhibidores de proteasas) y centrifugar a 800g por 5 min a 4°C. Se elimina el sobrenadante y el residuo se resuspende en 500 µl de tampón de digestión RSB (contiene Tris-HCl pH 7.5 50 mM, NaCl 10 mM, MgCl₂ 3 mM, β-Mercaptoetanol 10 mM, espermina 0,15 mM, espermidina 0,5 mM y 1mM de CaCl₂) y añadir MNasa 100U/ml. Se incuba a 37°C durante 15 min y se detiene la digestión con 100 mM de EDTA. La cromatina digerida es tratada con 5 µg/ml de ARNasa durante 60 min a 65°C. Finalmente la muestra es tratada con 100 µg/ml de proteinasa K por 90 min a 65°C.

Anexo 3. Condiciones empleadas por Tur (2007) y Sacilotto (2010) en la extracción de nucleosomas.

En principio se prueba el protocolo de Tur (2007), quien fija la cromatina por 3 min con formaldehído 1% en PBS, emplea 100 U/ml de MNasa con CaCl₂ 0,1 mM por cada 300 µl de cromatina e incuba por 5 min a 37°C. También se probó el protocolo de Sacilotto (2010), quien parte de 7x10⁸ núcleos, los fija con formaldehído 1% por 15 min y digiere con 150 U de MNasa en tampón de digestión RSB, durante 30 min a 37°C. En ambos casos se purifican las bandas de ADN mononucleosomal a partir del gel de agarosa. Se comprobó que se obtienen mono, di, tri y polinucleosomas con los protocolos antes mencionados, se extrajo el fragmento mononucleosomal y se purificaron para comenzar el estudio de la protección a la digestión.