

LAS MOLÉCULAS DE LA HERENCIA BIOLÓGICA. DE LAS PROTEÍNAS AL DNA... Y VUELTA

Luis Franco Vera

Real Academia de Medicina de la Comunidad Valenciana

Universitat de València e Instituto de Investigación Sanitaria INCLIVA

UN POCO DE HISTORIA

La historia sobre la organización del material genético de eucariotas comienza a finales del siglo XIX. La investigación se realizó en dos frentes separados que solo llegaron a reunirse a mediados del siglo XX. Esos dos frentes se podrían denominar como la “línea celular”, consecuencia fundamentalmente de observaciones microscópicas, y la “línea molecular”, en la que químicos y fisiólogos se esforzaron por conocer la naturaleza química de los materiales implicados en la herencia biológica.

Los primeros resultados se obtuvieron en esta segunda línea, en la que Friedrich Miescher, entre 1869 y 1871 aisló, a partir de núcleos de leucocitos de pus, una sustancia que denominó *nucleína*, cuyo análisis elemental reveló la presencia de una elevada cantidad de nitrógeno y fósforo. Posiblemente, se trate del primer aislamiento de ácidos nucleicos mezclados con proteínas. Poco tiempo más tarde, entre 1872 y 1874, aisló *nucleína* de esperma de salmón y acuñó el término *protamina* para referirse a su principal componente proteico, aunque en aquella época la naturaleza de las proteínas era todavía desconocida.



Figura 1
Friedrich Miescher
(1844-1895)

En 1884, Albrecht Kossel fue capaz de aislar por primera vez unas proteínas nucleares que denominó *histonas*, nombre que sigue vigente en la actualidad. Kossel trabajó con eritrocitos nucleados de aves. Hay que tener en cuenta que, en aquellos años, el fraccionamiento celular se tenía que realizar por métodos muy rudimentarios, habida cuenta de la inexistencia de la centrifugación diferencial a la que hoy en día estamos habituados. Si Miescher había utilizado leucocitos, era porque se podían lisar fácilmente y, debido a su elevado coeficiente

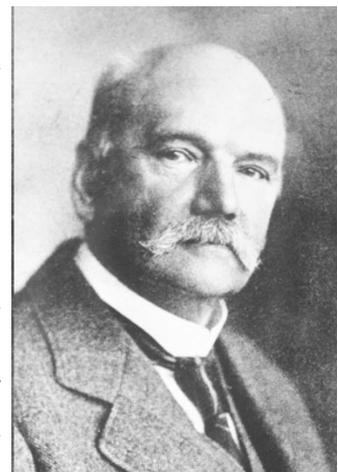


Figura 2
Albrecht Kossel
(1853-1927)

de sedimentación, sus núcleos pueden sedimentar simplemente por la acción de la gravedad. Otro tanto ocurre con los eritrocitos nucleados de ave. Kossel extrajo los núcleos en medio ácido y encontró que ese tratamiento solubilizaba las histonas. Años más tarde, el propio Kossel analizó su composición, encontró que en ellas abundaban los aminoácidos básicos y describió uno nuevo, desconocido hasta entonces, al que por proceder de las histonas, denominó histidina. La histidina no es precisamente uno de los aminoácidos más abundantes de las histonas, pero curiosamente debe su nombre a su presencia en ellas.

Aquí termina lo que se podría definir como la *prehistoria* de esa línea molecular dedicada al estudio del material genético eucariótico. Habrían de pasar aún bastantes años para que se conociera la estructura de las proteínas y del DNA, pero ya a finales del siglo XIX se sabía que en los núcleos eucarióticos existían histonas y DNA, aunque este último biopolímero aún no hubiera recibido la denominación con que se conoce en nuestros días.

Independientemente, se inició la “línea celular” con las investigaciones de Walther Flemming, a quien puede considerarse el padre de la Citogenética. En 1889, este autor observó que en los núcleos celulares se podían detectar dos tipos de estructuras cuando se teñían con colorantes básicos. Una, coloreada, que denominó por ese motivo *cromatina*, y otra a la que no se fijaba el colorante, que designó con los nombres de *acromatina* o *linina*. La cromatina, en determinados momentos, aparecía condensada en forma de filamentos, que más tarde serían designados por Waldeyer como *cromosomas*. Flemming no hizo ningún intento de describir esas estructuras en términos de su composición molecular, algo que hubiera estado fuera de lugar en la época, pero el nombre de cromatina sigue empleándose actualmente con un preciso significado molecular. Un año más tarde, el propio Flemming, simultánea aunque independientemente con Strasburger y Van Beneden describió la segregación cromosomal durante la división celular y acuñó para ese proceso el nombre de mitosis.



Figura 3
Walther Flemming
(1843-1905)

En ese contexto, Walter Sutton y Theodor Boveri, en 1903, postularon la teoría cromosómica de la herencia, según la cual, los cromosomas eran los portadores de la información genética. La citogenética y la genética clásicas progresaron considerablemente en la primera mitad del siglo XX. La aplicación a la genética de los trascendentales descubrimientos de Mendel, que inicialmente se realizaron en el ámbito de la hibridación de especies vegetales, fue decisiva. Desde el punto de vista molecular, los avances también fueron importantes. Se describió la naturaleza química de las proteínas y de los ácidos nucleicos, pero, no obstante, no tuvo lugar la integración de lo que venimos llamando “línea molecular” y “línea celular”.

El panorama cambió con la irrupción de la Biología Molecular en el estudio de la herencia biológica a mediados del siglo XX. Este nuevo enfoque de la Biología aspiraba a dar una explicación a todos los fenómenos biológicos —y no solo al metabolismo como, obviamente—

te, se había hecho hasta entonces— en términos de la estructura y cambios experimentados por las moléculas constituyentes de la materia viva. Además de interesarse por los metabolitos de pequeño tamaño molecular, se fue imponiendo la atención en las macromoléculas, singularmente en las proteínas y en los ácidos nucleicos. En 1941 ya se sabía que los cromosomas estaban constituidos principalmente por DNA e histonas, aunque había un consenso entre los investigadores que identificaba a las histonas como portadores de la información genética y al DNA se le atribuía un papel fundamentalmente estructural. Buena prueba de ello es que en el Simposio organizado ese año por el *Cold Spring Harbor Laboratory*, con el título *Genes and Chromosomes: Structure and Organization*, se aceptó pacíficamente esa idea.

DE LAS PROTEÍNAS AL DNA

Terminaba el apartado anterior viendo cómo hasta bien entrado el siglo XX se consideraba que las proteínas eran las portadoras de la información genética. Pero en 1944 se produjo un descubrimiento trascendental, que supuso un giro copernicano en la investigación, al asignar al

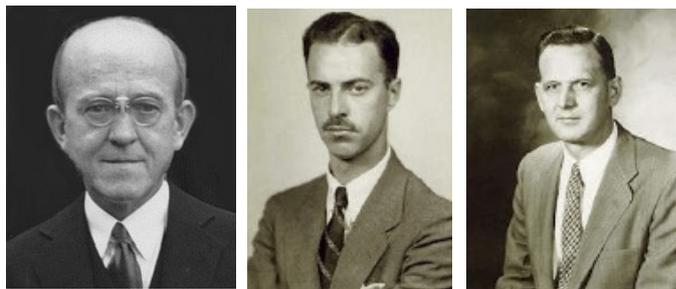


Figura 4

Oswald Avery
(1877-1995)

Colin MacLeod
(1909-1972)

Maclyn McCarty
(1911-2005)

DNA el papel que hasta entonces se atribuía a las proteínas. Fue Oswald Avery, en colaboración con Colin MacLeod, y Maclyn McCarty, quien comprobó, a través de unos experimentos de transformación de cepas bacterianas, que era el DNA el responsable de la herencia genética. Aunque los descubrimientos de Avery y sus colaboradores fueron recibidos con un cierto escepticismo, afortunadamente hubo destacados científicos que

aceptaron la idea revolucionaria y se dedicaron a dilucidar la estructura tridimensional del DNA, apoyándose en uno de los postulados fundamentales de la Biología Molecular, el de que la comprensión de la función y de la estructura son inseparables, como dos caras de una misma moneda. Entre ellos estaba Francis Crick, que junto con James Watson y apoyándose en experimentos realizados por otros autores llegaron a proponer la estructura del DNA en el que se ha calificado como el artículo más trascendental de la Biología del siglo XX.

Una vez aceptada la idea de que el DNA era el portador de la información genética, quedaba por decidir qué papel correspondía a las proteínas que le acompañan en la cromatina, especialmente a las histonas, las más abundantes de ellas. La primera hipótesis fue planteada en 1950 por Edgar Stedman y su esposa, Ellen. Según esa hipótesis, “las proteínas básicas del núcleo celular son inhibidores de los genes, de forma que cada histona o protamina es capaz de bloquear la actividad de un cierto número de genes”. Dada la diversidad y especificidad de los genes, la hipótesis implicaba que las histonas fueran muy variadas, específicas de especie y, posiblemente, también de tejido. Como se verá después, la hipótesis no resultó cierta, pero tuvo

la virtud de despertar el interés por la purificación y fraccionamiento de las histonas. Los primeros intentos, todavía infructuosos, se realizaron en 1954 por Cruft, pero hasta principios de la década siguiente no se comenzó a progresar en ese terreno.

El avance se debió fundamentalmente al trabajo de dos laboratorios: el de James Bonner en CalTech (Pasadena, California) y el de Ernt Johns en el Chester Beatty Research Institute de Londres. A mediados de la década de 1960 el fraccionamiento de histonas estaba prácticamente concluido. Johns había trabajado fundamentalmente con timo de ternera y en el laboratorio de Bonner habían prestado especial atención a las histonas de embrión de guisante. Para sorpresa de los investigadores que habían aceptado la hipótesis de los Stedman, en ambos organismos las histonas se fraccionaban en tan solo cinco clases que, con la nomenclatura actual, se designan como H1, H2A, H2B, H3 y H4. Además, el parecido en propiedades y composición de esas clases de histonas entre esos dos organismos era muy grande. Pero el resultado más llamativo fue consecuencia de la determinación, por el grupo de Bonner, de la estructura primaria de la histona H4, la más pequeña de las cinco, en timo de ternera y en embrión de guisante. En los dos casos, la histona tiene 102 residuos de los que solo dos son diferentes. Más aún, las dos diferencias son conservativas: la valina 60 de la histona de timo de ternera está sustituida por isoleucina en guisante, y la lisina 77 por arginina.

Pronto se obtuvieron más datos de estructuras primarias de histonas y se comprobó que la asombrosa conservatividad evolutiva observada en la H4 se daba también en la H3 y, aunque en menor medida, en las histonas H2A y H2B. Solo la histona H1 parecía exhibir una relativa especificidad de especie. Por otra parte, las histonas de los diversos órganos o tejidos de un mismo organismo, con alguna aislada excepción, son idénticas.

Evidentemente, esta extraordinaria semejanza entre las histonas echaba por tierra la

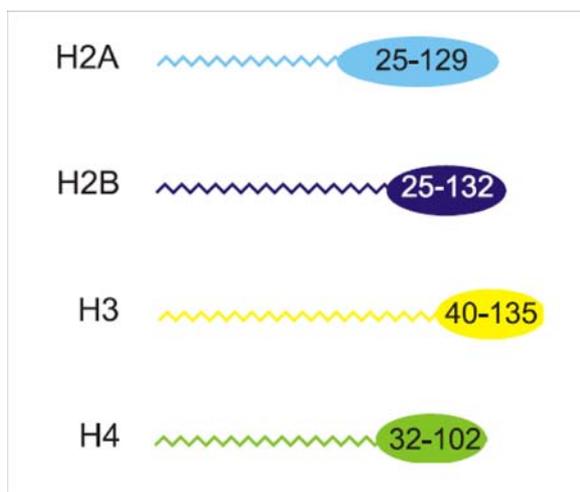


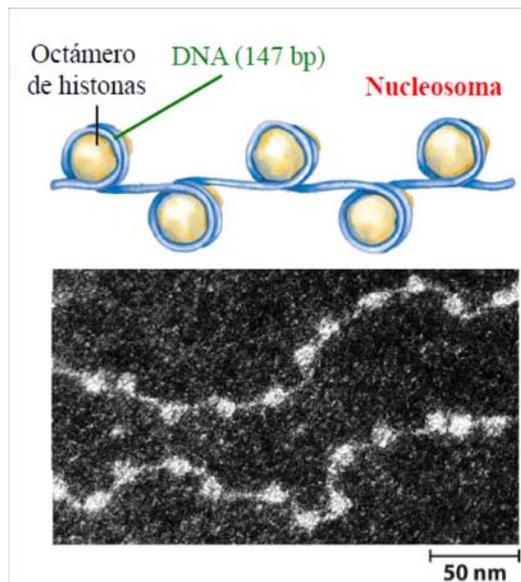
Figura 5. Modelo inicial para la estructura de las histonas H2A, H2B, H3 y H4.

posibilidad de que las histonas desempeñaran un papel de reguladores específicos de la expresión génica. Pero el hecho de que estuvieran tan conservadas hacía pensar que su papel en la cromatina debía ser muy importante. En la década de 1970 se comenzó a pensar que este papel era estructural y cobraron interés los estudios encaminados a dilucidar la estructura de la cromatina. Los primeros intentos resultaron fallidos; la propuesta más aceptada era que en la cromatina el DNA se enrollaba, con la ayuda de las histonas, formando una superhélice continua, con un

paso de rosca de unos 11 nm. Teniendo en cuenta que, con la excepción de la H1, la estructura primaria de las otras cuatro histonas mostraba en su tercio N-terminal una gran abundancia de aminoácidos polares y, especialmente, básicos, se veía muy improbable que esa región adoptara una estructura globular, la cual, a la vista de su composición, parecía posible en los dos tercios

restantes. De ese modo, se postuló el modelo general para la estructura de esas cuatro histonas que se recoge en la Fig. 5. En ella, se representa la presunta zona globular indicando la numeración de los residuos que comprendería, mientras que la cola N-terminal aparece como una línea en zigzag. En el modelo de superhélice continua para la cromatina se suponía que las colas, por su elevada densidad de carga positiva, serían los lugares de interacción con el DNA, mientras que los dominios globulares serían responsables de las interacciones histona-histona que mantendrían la superhélice.

El panorama cambió radicalmente a partir de 1973. Ese año Donald Olins y Ada Olins comunicaron que, observadas al microscopio electrónico, las fibras de cromatina aparecían con un aspecto arrosariado, como una serie de cuentas de collar ensartadas en un filamento que, por su diámetro, parecía corresponder al DNA desnudo. Las *cuentas* aproximadamente esféricas y con un diámetro de unos 11 nm (Fig. 6), se llamaron inicialmente cuerpos v, pero pronto se impuso la denominación de nucleosomas, con la que se siguen conociendo en la actualidad. En los dos años sucesivos se confirmó por diversos métodos esta estructura discontinua y repetitiva de la cromatina y se determinó que el nucleosoma está constituido por una



partícula núcleo y una longitud variable de DNA espaciador. La partícula núcleo contiene un octámero formado por dos moléculas de cada una de las histonas H2A, H2B, H3 y H4, alrededor del cual se enrolla una longitud fija, 147 pares de bases, de DNA, describiendo casi dos vueltas de superhélice. La histona H1 se asocia en algunas regiones de la cromatina, con el DNA espaciador.

Figura 6. Fibras de cromatina vistas al microscopio electrónico con contraste negativo. El esquema de la parte superior describe sucintamente la composición de los nucleosomas.

La importancia del nucleosoma como motivo estructural fundamental de la cromatina se comprendió inmediatamente y, como es lógico, pronto se iniciaron las investigaciones para resolver su estructura. Sin embargo, hubo numerosas dificultades para hacerlo con alta resolución y hasta 1991 no se consiguió un avance satisfactorio. En ese año, el grupo de Evangelos Moudrianakis describió la estructura del octámero de histonas con una resolución de 3,1 Å. De esa forma, se eludían los obstáculos que se habían encontrado para obtener cristales de nucleosomas de calidad suficiente, aunque, obviamente, solo con cristales de octámeros no se podía definir la trayectoria del DNA. No obstante, el conocimiento de la estructura del octámero aportó importantes datos. En primer lugar, se comprobó que posee una estructura tripartita, en la que un tetrámero central de dos moléculas de H3 y dos de H4 está flanqueado por dos dímeros de H2A y H2B. En segundo término, se vio que la estructura inicialmente propuesta para las histonas (Fig. 5) no era válida. Las histonas no tienen propiamente hablando una zona globular. Poseen una estructura que solo pueden adoptar cuando una molécula de histona

interacciona específicamente con otra: H2A con H2B y H3 con H4. Como consecuencia de esta interacción, las histonas adquieren un motivo estructural, que Moudrianakis denominó *histone fold*, constituido por una hélice α larga flanqueada por otros dos segmentos más cortos de hélice. Las dos histonas del par correspondiente se unen entre sí de modo que los dos *histone folds* se interpenetran mutuamente, formando el motivo que los autores denominaron gráficamente “apretón de

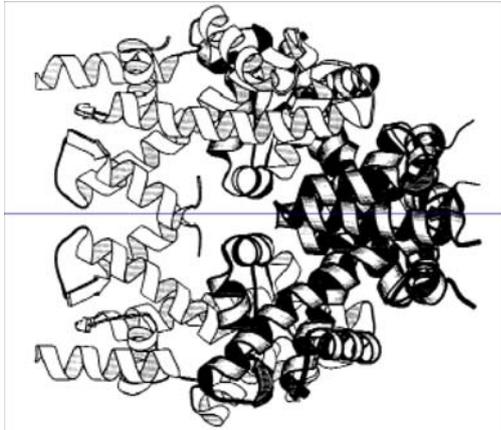


Figura 7. Estructura del octámero de histonas descrita por el grupo de Moudrianakis (Arents *et al.*, 1991). Las hélices del tetrámero H3-H4 aparecen en blanco y las de uno de los dímeros H2A-H2B en oscuro. El otro dímero, que no se ve en la figura, quedaría por debajo del plano.

manos”. En la figura 7 se recoge la estructura del octámero y en la 8 se apre-

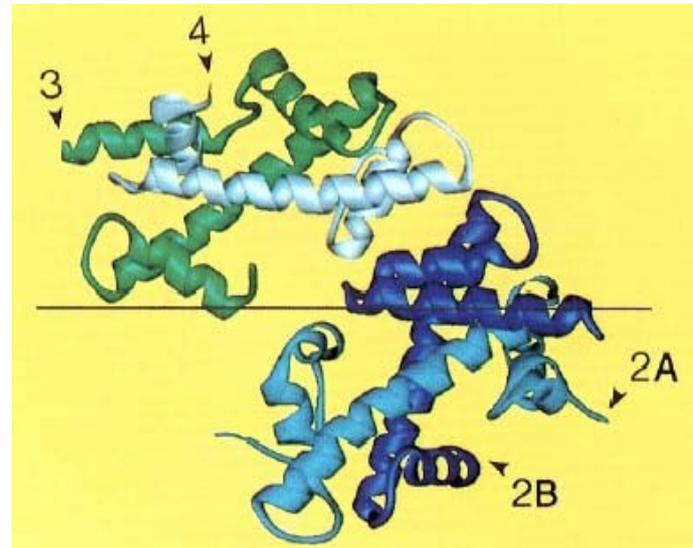


Figura 8. Se aprecia la mitad de un octámero, con la misma orientación que en la figura 7. Se puede observar el mismo dímero H2A-H2B y solo la mitad del tetrámero. La estructura del *histone fold* y la interpenetración de las histonas emparejadas es evidente.

ciendo con más claridad tanto el *histone fold* como el “apretón de manos”.

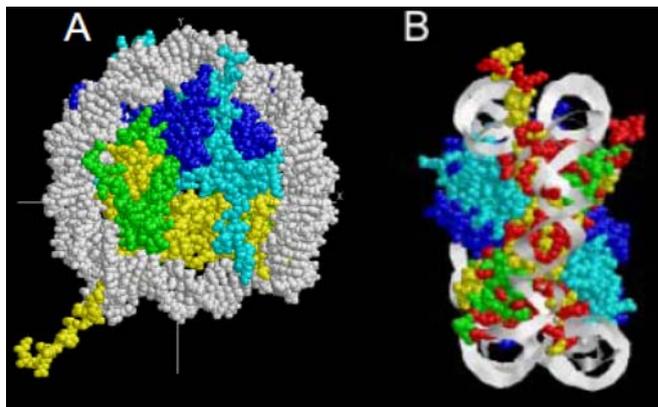


Figura 9. Estructura del nucleosoma. (A) Vista a lo largo del eje de la superhélice del DNA. (B) Vista desde el eje diádico de simetría. La imagen corresponde a la que se observa en A, vista desde abajo. El DNA aparece en blanco, la histona H3 en amarillo, la H4 en verde, la H2A en azul claro y la H2B en azul oscuro. Los residuos rojos en B corresponden a los aminoácidos básicos que interactúan con el DNA. Imagen realizada con el programa RasMol.

En 1997 el grupo de Richmond consiguió, mediante una ingeniosa combinación de técnicas, obtener cristales de nucleosomas y estudiar su estructura con alta resolución mediante difracción de rayos X (Fig. 9). La estructura del octámero coincide esencialmente con la descrita años atrás por Moudrianakis, pero se observan algunos residuos adicionales, que en los cristales de octámeros no ocupan posición fija y en el nucleosoma sí lo hacen por interactuar con el DNA del propio nucleosoma o, como es el

caso de la cola de H3 que se observa en la figura 9A, por interactuar con el DNA de un nucleosoma vecino en el cristal.

La estructura de la unidad básica de la cromatina quedaba, de esta manera, resuelta, aunque en años sucesivos se fueron perfilando algunos pequeños detalles, especialmente en lo que se refiere a la trayectoria del DNA, cuya superhélice presenta varias irregularidades. Las estructuras de orden superior de la cromatina, que en la mitosis tiene que alcanzar un grado máximo de compactación en los cromosomas, siguen sin estar bien conocidas.

En cualquier caso, el conocimiento profundo de la estructura del nucleosoma replanteó una cuestión que venía considerand-

do incluso desde antes del descubrimiento del propio nucleosoma: ¿cómo es posible que las RNA polimerasas puedan transcribir el DNA organizado en forma de cromatina? La pregunta puede ahora formularse desde una base más concreta. A la vista de la estructura del nucleosoma y de la RNA polimerasa II que se observan en la figura 10, resulta evidente que la enzima no puede recorrer el DNA si este se encuentra formando parte de un nucleosoma. En efecto, la RNA polimerasa II de levadura, que es la representada en la figura 10, es una molécula incluso mayor que el nucleosoma y el DNA debe pasar por el túnel que se observa en la parte superior

de la molécula, entre las subunidades Rpb1 y Rpb2. Es obvio que si el DNA está enrollado alrededor de un octámero de histonas, formando un nucleosoma, la enzima es incapaz de recorrerlo si no se disocia previamente. Ahora bien, desde la década de 1980 se sabe que los genes activamente transcritos poseen estructura nucleosomal en el cuerpo del gen.

Por otro lado, en las regiones reguladoras de los genes, como son los promotores, han de unirse los factores transcripcionales específicos, que muchas veces deben rodear totalmente la doble hélice del DNA. Además, en la secuencia de iniciación de la transcripción tiene que unirse el complejo de preiniciación que, además de la polimerasa, contiene numerosos factores transcripcionales basales. También en los años 80 se demostró que la RNA polimerasa no podía iniciar la transcripción en un nucleosoma. No obstante, también en muchos casos se ha comprobado que en numerosos genes potencialmente activos hay nucleosomas ocupando las secuencias

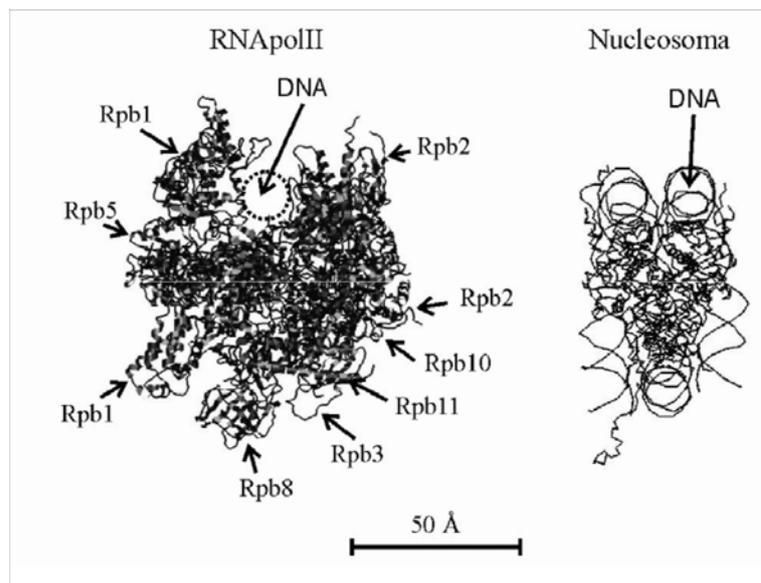


Figura 10. Estructura de la RNA polimerasa II de levadura y de un nucleosoma. Ambas estructuras están representadas a la misma escala con el programa RasMol. En la estructura de la polimerasa se señalan algunas de las subunidades de la enzima y el túnel de unos 20 Å de diámetro por el que ha de pasar el DNA para ser transcrito. El nucleosoma está dibujado en la orientación en que debe ser abordado por la enzima.

de reconocimiento de los factores específicos y que en las secuencias de iniciación con frecuencia se encuentran nucleosomas. Era, pues, necesario resolver esa dificultad que representaba la presencia de los nucleosomas para la actividad de la RNA polimerasa. Y otro tanto cabría decir para otras funciones del DNA. En efecto, tanto la replicación, como la reparación o la recombinación, son *a priori* incompatibles con la localización de nucleosomas en las secuencias en que han de intervenir las enzimas correspondientes.

La solución a este problema, que se antojaba imponente, vino en la década de 1990. De una manera un tanto fortuita se descubrieron los complejos de remodelación de la cromatina. Son estructuras formadas por varias subunidades proteicas, una de las cuales siempre tiene actividad de ATPasa. La energía liberada por la hidrólisis del ATP es suficiente para desestabilizar las fuertes interacciones DNA-histona que mantienen la estructura del nucleosoma. Y una vez debilitadas estas interacciones, los octámeros de histonas pueden deslizarse a lo largo del DNA, descubriendo secuencias clave para los diversos procesos, o bien pueden eliminarse totalmente —temporal o definitivamente— del DNA. También es posible que los nucleosomas se disocien parcialmente, o que se elimine una histona que se sustituya por otra variante que haga permanentemente más inestable el nucleosoma. En cualquier caso, la acción de los complejos de remodelación es capaz de solventar las dificultades que la presencia de nucleosomas implica para la iniciación de la transcripción o para la elongación de la misma, así como para los demás procesos en los que interviene el DNA.

Desde el descubrimiento de los mecanismos de remodelación, la estructura de la cromatina —y, en particular, la del nucleosoma— dejó de considerarse un simple obstáculo para la transcripción para pasar a contemplarse como un factor regulador de suma importancia. En efecto, un gen que contenga un nucleosoma ocupando el lugar de iniciación de la transcripción, no puede transcribirse. Pero si un complejo de remodelación elimina ese nucleosoma, la transcripción será posible. Además de la presencia o ausencia de activadores o represores transcripcionales, la presencia y localización exacta de los nucleosomas constituye un elemento regulador de la transcripción¹, tanto que, en la actualidad es imposible hablar de regulación de la transcripción en eucariotas sin tener en cuenta la estructura de la cromatina.

Pero, como ocurre con frecuencia cuando la investigación encuentra la respuesta a una pregunta, esta respuesta formula un amplio abanico de preguntas ulteriores. En este caso, esas preguntas podrían ser: ¿qué condiciona que un complejo de remodelación actúe en unos genes y no en otros?; ¿cuál es la causa de que se den diferencias en la estructura de la cromatina al comparar unos tipos celulares y otros de un mismo organismo? Y, sobre todo, ¿por qué mecanismos la célula distingue los genes que, aquí y ahora, debe transcribir? Las preguntas anteriores, y otras similares que se podrían plantear, parecen formidables. Hay que tener en cuenta que, en el hombre, por ejemplo, existen unos 22.000 genes, que son idénticos en todas las células somáticas —con la evidente excepción de los linfocitos B y T— que constituyen los 250 tipos

¹ Lo que se dice de la transcripción es extensible a otros procesos, pero en lo sucesivo se hará exclusivamente referencia a la transcripción.

celulares del adulto. Todos esos genes coinciden en todos los tipos celulares con los del cigoto a partir del cual se ha desarrollado el adulto. Pero, si bien el genoma del cigoto y el de una célula somática del adulto son idénticos, es obvio que su transcriptoma y su proteoma son distintos y que también son dispares cuando se comparan células de dos tipos diferentes. Si se consideran así las cosas, parece que la información genética por sí sola no es suficiente para justificar el desarrollo y diferenciación de los organismos superiores.

...Y VUELTA. LA EPIGENÉTICA

Que era insuficiente la información genética y había que añadirle algo fue ya intuido por Conrad Waddington, que en 1942 acuñó el término *Epigenética*, utilizando el prefijo griego ἐπί, que significa encima de, sobre, además de, añadido a. En la actualidad, el término se ha popularizado extraordinariamente, aunque el sentido que se le atribuye es algo diferente del propuesto por Waddington. Siguiendo a Esteller, se puede definir como el “estudio de los cambios en la expresión génica, hereditarios por medio de la mitosis y/o de la meiosis, que tienen lugar sin cambios en la secuencia del DNA”. Inmediatamente, surge una pregunta acerca de cuál puede ser la naturaleza de esos cambios. Los más clásicos son la metilación de citosinas del DNA y las modificaciones covalentes post-traduccionales de las histonas, ambos cambios conocidos desde hace más de 50 años. Actualmente, se añaden otros cambios covalentes, como la hidroximetilación de citosinas, y se suelen incluir como factores epigenéticos la presencia de RNAs no codificantes (ncRNAs) y la remodelación de la cromatina, de la que ya se ha hablado anteriormente y que, efectivamente, puede modular la actividad transcripcional sin implicar cambios en la secuencia de bases del DNA.

La metilación del DNA tiene lugar por una reacción que introduce restos metilo en el carbono 5 del anillo de citosina. Hay que destacar que esta metilación de la citosina no altera sus propiedades de formar enlaces de hidrógeno con la guanina. Por tanto, ni la organización de la doble hélice, ni su mensaje genético se alteran por la metilación. La reacción tiene lugar a expensas del donador universal de metilos, la S-adenosil-metionina, y está catalizada por DNA metiltransferasas. Aunque puede tener lugar en muchas citosinas del genoma, desde el punto de vista epigenético tiene interés cuando ocurre en las denominadas “islas CpG” localizadas en los promotores de los genes. Las islas CpG son secuencias en las que abunda este dinucleótido. Pues bien, la metilación de las citosinas en esas islas provoca el silenciamiento del gen. Recientemente se han descrito mecanismos capaces de desmetilar las citosinas, de modo que la modificación es reversible y constituye, de hecho, un importante mecanismo de regulación de la actividad génica. Existen también enzimas capaces de metilar el DNA hemimetilado, es decir, metilado en las citosinas de una sola de sus cadenas. De este modo, la metilación puede transmitirse en la división mitótica. En efecto, cuando en la replicación se construye una cadena nueva sobre cada una de las dos hebras metiladas, las moléculas hijas contienen una cadena metilada—la antigua— y otra desmetilada. Pero esta cadena de las moléculas hijas se metila a continuación por las DNA metiltransferasas que reconocen al DNA hemimetilado como sustrato. Así se

cumplen los requisitos que, según la definición de Esteller, debe poseer una modificación epigenética.

La modificación de las histonas se descubrió en 1964 por Allfrey y Mirsky. Estos autores encontraron que las histonas pueden acetilarse y metilarse en las cadenas laterales de sus residuos de lisina. La reacción de acetilación (Fig. 11) implica la formación de una acetamida en las cadenas laterales de lisina por transferencia de un resto acetilo procedente de la acetilcoenzima A y está catalizada por histona acetiltransferasas (HATs). La eliminación del resto acetilo se realiza por una reacción distinta, una hidrólisis catalizada por las histona desacetilasas

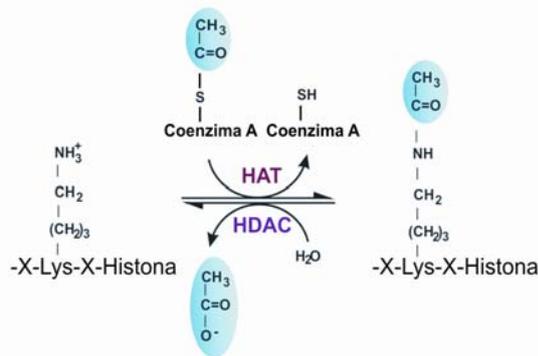


Figura 11. Reacciones químicas implicadas en la acetilación y desacetilación de histonas.

(HDACs), que libera acetato y restablece la estructura inicial de la lisina. Hay que destacar que la acetilación implica la pérdida de la carga positiva de las cadenas de lisina, por lo que desde el principio se sospechó que esa modificación debilitaría las interacciones electrostáticas de las histonas con el DNA y se postuló que estaría relacionada con la activación transcripcional. Aunque hubo que perfilar mucho esa hipótesis inicial, en líneas generales se puede decir que, efectivamente, hay una correlación entre acetilación de histonas y actividad génica.

El descubrimiento de la acetilación de histonas y su posible relación con la actividad transcripcional dio lugar a un cierto interés por la enzimología de esta modificación. Las primeras investigaciones permitieron la caracterización de dos HATs, denominadas inicialmente A y B. La B se encontraba en el citoplasma y parecía responsable de una modificación de la histona H4 necesaria para su posterior ensamblaje en el DNA recién replicado para organizarlo en forma de cromatina. La HAT A, nuclear, sería responsable de la acetilación relacionada con la actividad transcripcional.

Los estudios realizados fundamentalmente durante las décadas de 1970 y 1980 permitieron decidir que la acetilación tenía lugar mayoritariamente en las colas N-terminales de las histonas y se logró una inicial identificación de los sitios de acetilación, que en vertebrados correspondían a las lisinas 5 y 9 de H2A, 5, 12, 15 y 20 de H2B, 9, 14, 18 y 23 de H3 y 5, 8, 12, y 16 de H4. Aunque años más tarde se identificaron lisinas acetilables adicionales, algunas situadas en regiones más internas de las histonas, estos datos preliminares permitieron comprobar que las combinaciones posibles de acetilación de histonas eran inmensas. Habida cuenta de ello, en 1988 el investigador austriaco Peter Loidl propuso la hipótesis que se puede llamar de la señalización. Según ella, la acetilación de histonas no tendría simplemente una función de reducir la carga positiva de las colas N-terminales, sino que podría implicar “una señal específica para el mantenimiento o establecimiento de diferentes rasgos estructurales de la cromatina, aunque aún se desconozca el mecanismo molecular por el que actúa la acetilación de histonas”. Dicho de otro modo, una combinación de acetilaciones en unos residuos dados podría ser la

señal para la transcripción, otra combinación diferente actuaría para permitir la recombinación, etc.

La hipótesis implicaba algo muy importante: si la acetilación en un residuo dado podía determinar específicamente una función, era obvio que la acetilación de las histonas no podía ocurrir al azar. Pero eso conllevaba que la reacción de acetilación tenía que ser, al menos parcialmente, específica de histona y de residuo, lo que, a su vez, exigía que fueran específicas las enzimas —HATs y HDACs— responsables de mantener un estado dado de acetilación. En otras palabras, la existencia de una única HAT A nuclear era irreconciliable con la hipótesis de Loidl. Debían existir múltiples HATs. Nuestro laboratorio fue el primero en demostrar entre 1989 y 1991 la existencia de esa multiplicidad en las HATs de levadura y, poco después pudimos comprobar también la multiplicidad de HDACs. En trabajos subsiguientes en colaboración con el laboratorio de Peter Loidl se pudo comprobar que la multiplicidad de esas enzimas era una característica común de los eucariotas.

Estos hallazgos dispararon el interés por la investigación en el campo de la cromatina y, mediada ya la década de 1990, se logró la clonación de los primeros genes de HATs y HDACs, lo que convirtió a la Epigenética en un área emergente de creciente impacto (Fig. 12).

Además de la metilación del DNA y de la acetilación de histonas, se fue abriendo el abanico de las modificaciones epigenéticas. En el caso de las histonas, la metilación, que como se ha apuntado había sido descubierta en 1964, cobró un decisivo impulso al comprobarse que no solo existen histona metiltransferasas, sino también histona desmetilasas. Esta modificación puede afectar a cadenas laterales de lisina, que pueden mono- di- o trimetilarse, y a cadenas laterales de arginina, cuyos átomos de nitrógeno pueden metilarse. Caben tres posibilidades: la monometilación de un solo nitrógeno, la metilación de los dos o la dimetilación de uno de ellos. A diferencia de la acetilación, que está fundamentalmente relacionada con la activación de la transcripción, la metilación puede ser una señal para la actividad transcripcional o para la represión de la actividad génica, dependiendo de la histona y del residuo en que tenga lugar. Pero además, las histonas pueden ser objeto de fosforilación o de ubiquitinación —conjugación a través de cadenas laterales de lisina con la proteína ubiquitina a través de un enlace isopeptídico—, de modo que, en la actualidad, se sabe que un nucleosoma puede ser modificado en más de 140 sitios diferentes.

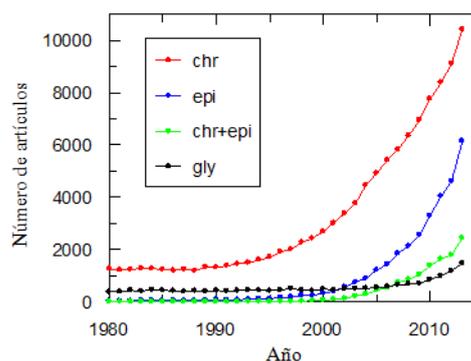


Figura 12. Evolución temporal del número de publicaciones que tratan de cromatina (chr), epigenética (epi) y de ambos temas a la vez. A efectos comparativos, se incluye el número de publicaciones sobre un campo clásico de la Bioquímica: la glicólisis (gly). Números tomados de la base de datos PubMed (www.ncbi.nlm.nih.gov/pubmed).

En este sentido, hay que señalar que la hipótesis de la señalización de Loidl, emitida para la acetilación de histonas, se extendió a las demás modificaciones por Allis, que propuso en el año 2000 la hipótesis del *código de las histonas*, según la cual “la modificación múltiple de las histonas, actuando de modo combinatorio o secuencial, en una o varias colas de las histonas, es capaz de especificar funciones singulares”.

La hipótesis de Allis fue el detonante final para que la investigación epigenética se disparara exponencialmente al comenzar el nuevo siglo (Fig. 12). Como consecuencia, en la actualidad se han descrito 10 HATs principales en el hombre, distribuidas en tres familias (tabla I), 18 HDACs, agrupadas en 8 clases (tabla II) y un número todavía no precisado de enzimas implicadas en la metilación de histonas.

Hay, por lo menos, 29 histona metiltransferasas que utilizan lisina como sustrato, y son específicas de histona, de residuo y de nivel de metilación; al menos 19 desmetilasas capaces de eliminar los grupos metilo de las cadenas laterales de lisina y se han descrito ya 11 arginina metiltransferasas, también específicas. También se conocen varias quinasas capaces de fosforilar histonas en residuos de serina o de treonina, así como las enzimas responsables de la introducción y eliminación de ubicuitina.

De muchas de estas modificaciones se ha conseguido establecer una funcionalidad, y se conocen proteínas capaces de reconocer específicamente las diferentes modificaciones,

Tabla I. Histona acetiltransferasas humanas

Familia	HAT	Especificidad
GCN5/PCAF	HAT1	H4
	Gcn5	H3/H4
	PCAF	H3/H4
MYST	MOF (MYST1)	H4K16
	HBO1 (MYST2)	H4>H3
	MOZ (MYST3)	H3
	MORF (MYST4)	H3
	Tip60	H4/H2A
p300/CBP	p300	Todas
	CBP	Todas

Tabla II. Histona desacetilasas humanas

Clase	Miembros
I	HDAC1, HDAC2, HDAC3, HDAC8
IIA	HDAC4, HDAC5, HDAC7, HDAC9
IIB	HDAC6, HDAC10
IV	HDAC11
Sirtuina clase I	SIRT1, SIRT2, SIRT3
Sirtuina clase II	SIRT4
Sirtuina clase III	SIRT5
Sirtuina clase IV	SIRT6, SIRT7

las propias histonas, los componentes de los complejos de remodelación, las enzimas

tanto de histonas como del DNA. De esta manera, la Epigenética ha vuelto a poner de relieve el papel de las proteínas en la expresión del mensaje genético. Este es el sentido de esa “vuelta” a las proteínas que aparece en el título de este artículo, vuelta que, por supuesto, no significa que se menosprecie el papel del DNA como portador de la información genética, sino que pone de relieve que ese papel no podría alcanzar su pleno sentido si no fuera por la participación de numerosas proteínas:

implicadas en la introducción y eliminación de las marcas epigenéticas y las proteínas encargadas de reconocer e interpretar estas marcas.

EL EJEMPLO DEL GEN *EGR1*

El gen *Egr1* (*early growth response 1*) es un gen inmediato-temprano que codifica un factor transcripcional. En nuestro laboratorio hemos estudiado profusamente los mecanismos de regulación de la transcripción de este gen y algunos de los resultados obtenidos pueden utilizarse como ejemplo de varios de los mecanismos comentados hasta ahora. En la figura 13 se esquematiza la estructura del promotor del gen en el ratón.

En la línea celular MLP29, de progenitores de hepatocitos, el gen se puede activar fácilmente por la adición de ésteres de forbol (TPA) al medio de cultivo. La activación es muy rápida y a los 5 min ya se observa transcripción del gen por encima del nivel basal. No obstante,

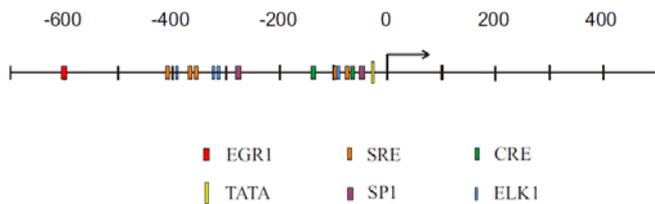


Figura 13. Promotor del gen *Egr1* de ratón. Se muestra, con una flecha acodada, el inicio de la transcripción, así como la posición de la caja TATA y de los diversos elementos de control del promotor. Puede apreciarse que existe un sitio de unión del propio producto del gen, varios elementos de respuesta a suero (SRE), y elementos de unión de los factores transcripcionales ELK1, SP1 y CREB.

La localización de nucleosomas en el promotor del gen se determinó por ensayos de protección a nucleasa de micrococo. Esta enzima digiere el DNA internucleosomal, pero no puede cortar en las regiones ocupadas por nucleosomas. De esta manera, si se utiliza DNA aislado de mononucleosomas como molde, puede evaluarse mediante PCR cuantitativa la protección de las diversas regiones, que corresponden a las zonas ocupadas por los nucleosomas. La figura 15 muestra esta protección a lo largo del promotor y zona proximal del cuerpo del gen, en función del tiempo tras la adición de TPA. Puede observarse que, en ausencia de estímulo, hay tres zonas protegidas, cuyo tamaño es compatible con la presencia de tres nucleosomas que, en lo sucesivo, se designan como N-2, N-1 y N+1, según se encuentren en el promotor o en el inicio de la región transcrita.

Comparando el curso temporal de la actividad

se trata de una activación transitoria, ya que 3 h después, la actividad transcripcional vuelve al nivel inicial (Fig. 14). La causa de este fenómeno es que en el promotor, en la posición -595, existe un sitio (Fig. 13) para la unión del propio producto del gen, que, si bien actúa inicialmente como activador, posteriormente recluta los represores NAB1 y NAB2, que causan la represión del gen.

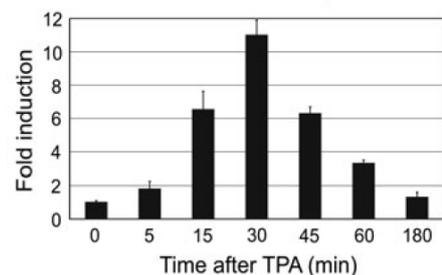


Figura 14. Transcripción del gen *Egr1* en células MLP29, tras la adición de TPA. La transcripción se ha determinado en tiempo real por RNAPol-ChIP. Datos de Tur *et al.* (2010) *Cell Mol. Life Sci.* **67**, 4065-4077.

transcripcional (Fig. 14), que está esquemáticamente reproducida en el encabezamiento de la

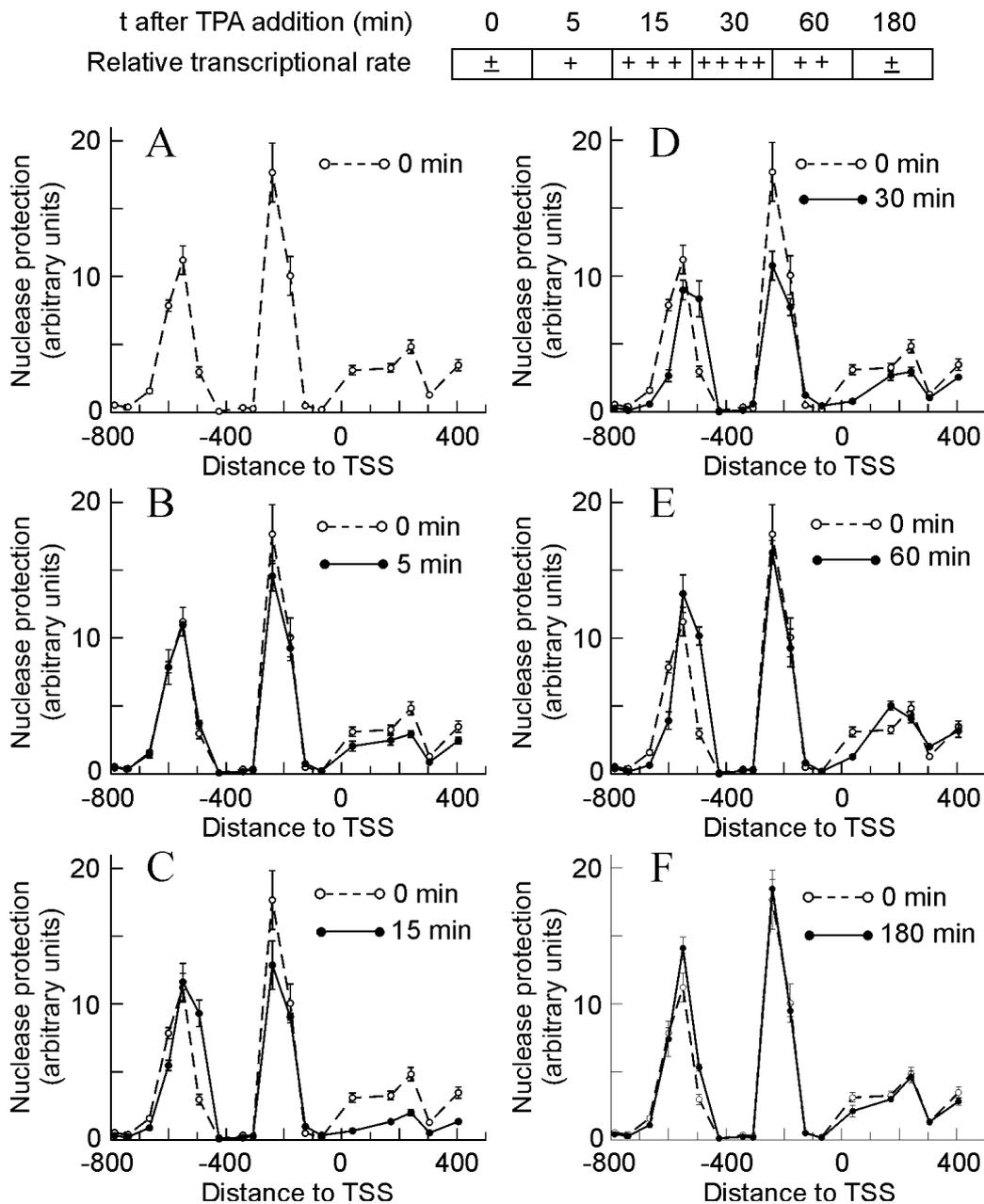


Figura 15. Cambios en la protección a nucleasa del promotor y región codificante proximal del gen *Egr1* tras la estimulación con TPA. A efectos comparativos se incluye en todos los paneles la protección basal ($t = 0$ min). Los puntos corresponden a la media de tres experimentos e incluyen el error estándar. Datos de Riffo-Campos *et al.* (2015) *J. Biol. Chem.* **290**, 197-208.

figura 15, con la localización de los nucleosomas, se puede apreciar claramente que la posición del nucleosoma N-2 cambia con la transcripción. El nucleosoma se desplaza hacia 3' y alcanza el máximo desplazamiento a los 30 min de estimulación (Fig. 15D), cuando es máxima la actividad de la transcripción (Fig. 14), para retornar a su posición basal una vez que la transcripción vuelve a su nivel basal (Fig. 15F). El nucleosoma está inicialmente localizado en una

posición de secuencia de DNA desfavorable gracias a la intervención de un complejo de remodelación de la familia SWI/SNF y en su desplazamiento interviene otro complejo de remodelación no identificado. Pero lo importante es que ese desplazamiento descubre el sitio de unión de EGR1 (Fig. 13), que está bloqueado en condiciones basales, y eso permite el juego de activación-represión antes mencionado que hace posible la expresión transitoria del gen.

Por otro lado, la protección a nucleasas en la región del nucleosoma N-1 disminuye en paralelo al aumento de la transcripción. Esto significa que el nucleosoma, más que cambiar de posición, se elimina parcialmente por la acción de otro complejo de remodelación. Posiblemente, esa desaparición parcial facilita la unión del complejo mediador que, como han demostrado otros autores, ocupa durante la transcripción una posición comprendida entre +1 y -400.

Cabe preguntarse qué cambios epigenéticos están asociados con estas alteraciones en la estructura de la cromatina. La cuestión se ha abordado en nuestro laboratorio mediante nuc-ChIP, una técnica de inmunoprecipitación de mononucleosomas con anticuerpos específicos para distintas modificaciones epigenéticas de las histonas. El método se ha empleado para ocho modificaciones de las histonas H3 y H4 a diversos tiempos tras la adición de TPA. De esta manera, se pueden correlacionar temporalmente los cambios epigenéticos a nivel de nucleosomas individuales con las alteraciones de su localización o estructura. Los resultados de estos experimentos se muestran en la figura 16.

Puede verse que algunas modificaciones son específicas de determinados nucleosomas. Por ejemplo, la fosfoacetilación de H3 ocurre mayoritariamente en el nucleosoma +1 y su aumento se advierte ya a los 5 min de la adición de TPA (Fig. 16A). Esta combinación de marcas en la H3 está seguramente asociada con la liberación de la RNA polimerasa II del promotor para comenzar la elongación productiva. Otras tres modificaciones, es decir, H3K4me3, H3K9me3 y H3K27me3², son también características, en mayor o menor medida, del nucleosoma +1 (Fig. 16, C-E). Por su parte, H3K9ac y H4K16ac parecen específicas del nucleosoma -1 y pueden estar relacionadas con su eliminación parcial durante la transcripción.

En resumen, el estudio del gen *Egr1* proporciona un excelente modelo para estudiar los diversos aspectos de la remodelación e introducción de marcas epigenéticas en la cromatina y la metodología empleada —detección de las marcas a nivel mononucleosomal y estudio de su evolución temporal— permite establecer correlaciones entre la modificación de histonas y su papel, lo que facilita la asignación de funciones dentro la hipótesis del código de las histonas.

EPIGENÉTICA Y ENFERMEDAD

² Se utiliza la convención de designar el residuo modificado con el código de una letra para los aminoácidos, seguido del número que ocupa en la secuencia. Las abreviaturas siguientes indican la modificación, acetilación (ac), metilación (me) y fosforilación (ph) y, en su caso, el número final indica el número de modificaciones sobre el residuo.

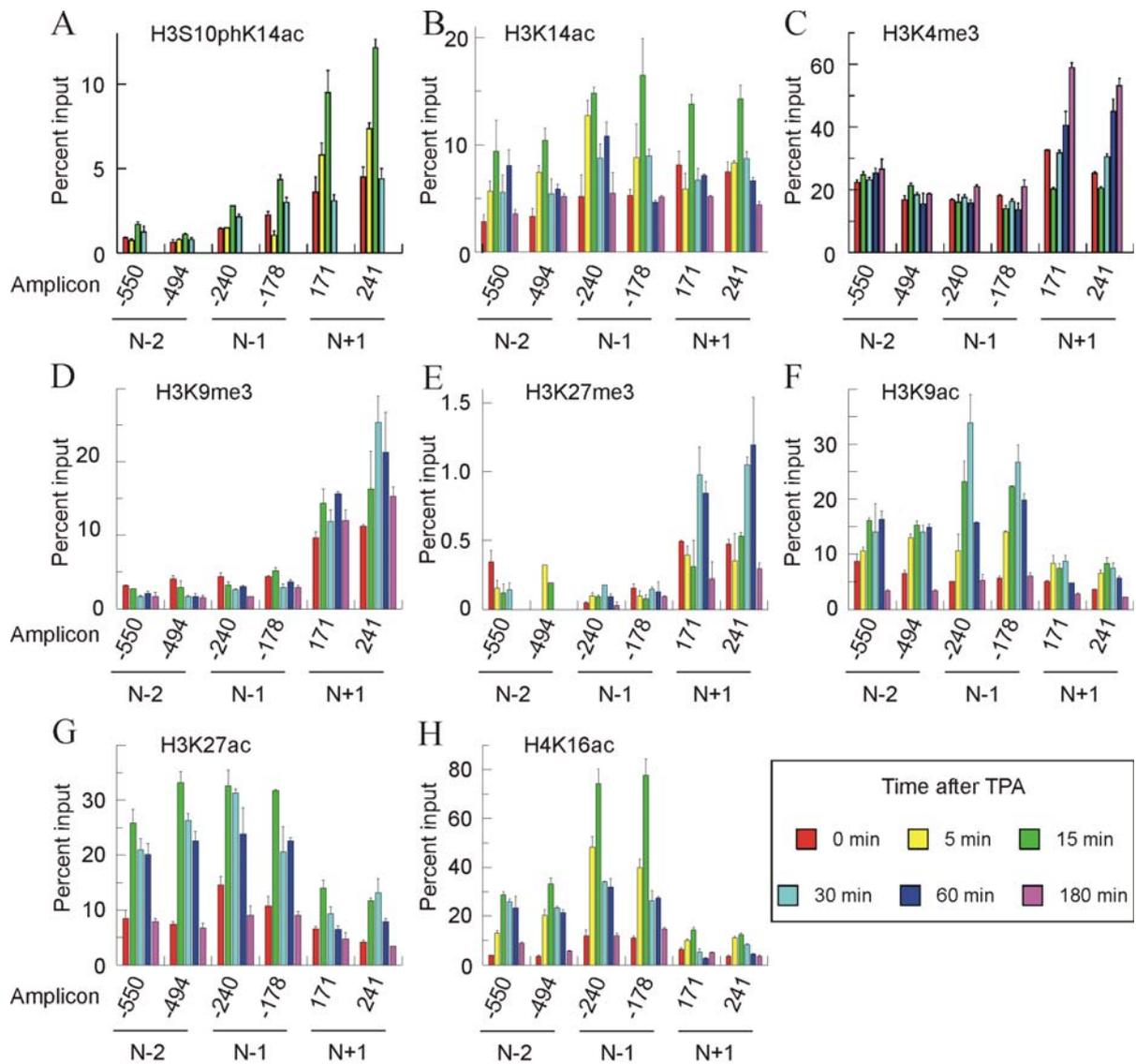


Figura 16. Cambios temporales de la modificación de histonas en los nucleosomas N-2, N-1 y N+1 tras la adición de TPA a un cultivo de células MLP29. Los paneles A-H muestran los niveles de las diferentes modificaciones estudiadas para dos amplicones en cada nucleosoma. Las barras de error representan la desviación estándar de tres determinaciones. Los tiempos tras la adición de TPA se dan a la derecha de la fila inferior. Datos tomados de Riffo-Campos *et al.* (2015) *J. Biol. Chem.* **290**, 197-208.

En las páginas precedentes ha quedado claro cómo las modificaciones epigenéticas pueden alterar la expresión del mensaje genético y, en este sentido, pueden tener unas consecuencias similares a las mutaciones. Una analogía puede ayudar a comprender la repercusión de los cambios en el patrón epigenético normal de una célula. De la misma manera que, en castellano, una frase se construye con las 27 letras de nuestro alfabeto y su sentido depende del orden en que las letras se encuentren, el mensaje genético depende del orden en que los cuatro nucleótidos que intervienen en la secuencia del DNA estén encadenados. Así como un cambio

de letras, una omisión o una alteración en su orden puede cambiar el sentido de una frase, una mutación, una deleción o una traslocación puede alterar el mensaje genético. En la figura 17 se ha reproducido el conocido inicio del Quijote. Un cambio de letras en la palabra Mancha, equivalente a una mutación, convertiría la frase en algo carente de sentido. Otras veces, como en el siguiente ejemplo, la omisión de una palabra, “que” en este caso, no tiene consecuencias

En un lugar de la Mancha, de cuyo nombre no quiero acordarme,
no ha mucho tiempo que vivía un hidalgo..
En un lugar de la Nqbxgq, de cuyo nombre no quiero acordarme,
no ha mucho tiempo que vivía un hidalgo..
En un lugar de la Mancha, de cuyo nombre no quiero acordarme,
no ha mucho tiempo vivía un hidalgo..
En un lagar de la Mancha, de cuyo nombre no quiero acordarme,
no ha mucho tiempo que vivía un hidalgo..
En un lugar de la Mancha, de cuyo nombre no quiero acordarme,
ha mucho tiempo que vivía un hidalgo..
En un lugar de la Mancha, de cuyo nombre no quiero acordarme,
no ha mucho tiempo que **vivía un hidalgo**..
En un lugar de la Mancha, de cuyo nombre no quiero acordarme,
no ha mucho tiempo que vivía un hidalgo..

Figura 17

importantes. La frase sigue teniendo el mismo sentido. Es lo que ocurre, por ejemplo, cuando se produce una deleción en una secuencia no codificante ni reguladora. Hay casos en que una mutación puede alterar la funcionalidad del producto de un gen. En el ejemplo del Quijote sería equivalente a la sustitución de la u de “lugar” por una a. Resulta la palabra lagar, que tiene sentido (¡y más en La Mancha!), pero distinto del original. En otras ocasiones, una deleción puede tener graves consecuencias, como ocurre en nuestro ejemplo al omitir la palabra “no”: se cambia totalmente el sentido de la frase.

Pero en los dos últimos ejemplos no se ha cambiado ni omitido ninguna letra. No obstante, o bien se han resaltado unas palabras cambiándolas de color, o bien se ha reducido el tamaño del tipo de otras, de manera que al leer el texto, pueden omitirse accidentalmente las palabras no resaltadas o escritas en tipo pequeño, con las mismas consecuencias que si se hubiera producido un cambio o una eliminación. Esto es lo que ocurre con los cambios epigenéticos que, como se ha visto, aunque no alteren la secuencia del DNA, pueden modificar la expresión de los genes.

De esta manera se comprende que hay cambios en el patrón de modificaciones epigenéticas que pueden tener las mismas consecuencias de una mutación. De hecho, algunos autores, como Manel Esteller, han comenzado a utilizar el neologismo “epimutación”. Y, por supuesto, de la misma manera que se habla de enfermedades de origen genético, se puede hablar de enfermedades epigenéticas, toda vez que sea cual sea el mecanismo por el que se produce una alteración de la expresión génica, si ocurre en una secuencia esencial, puede ser causa de un cambio patológico.

Así se explica que, en los últimos años, hayan proliferado las publicaciones sobre el origen epigenético de muchas enfermedades. En pocas ocasiones puede decirse que la etiología de una enfermedad sea exclusivamente epigenética, pero son, por contra, muy numerosos los casos en los que las epimutaciones influyen decisivamente. Entre estos casos, se encuentra el cáncer, enfermedades cardiovasculares y de vías respiratorias, la diabetes, la obesidad, algunas enfermedades neurodegenerativas, trastornos psíquicos, enfermedades autoinmunes, algunas alergias e, incluso, algunas enfermedades infecciosas. Como quiera que existe una relación indudable entre factores medioambientales y alteraciones epigenéticas, la investigación en este terreno está proporcionando datos valiosos para comprender los mecanismos moleculares causantes de muchas alteraciones patológicas. Como contrapartida, teniendo en cuenta que las alteraciones epigenéticas son reversibles, han aparecido “fármacos epigenéticos”, que tienden a restablecer el patrón epigenético correcto. Es de esperar que los próximos años sean de gran trascendencia en esta investigación.