

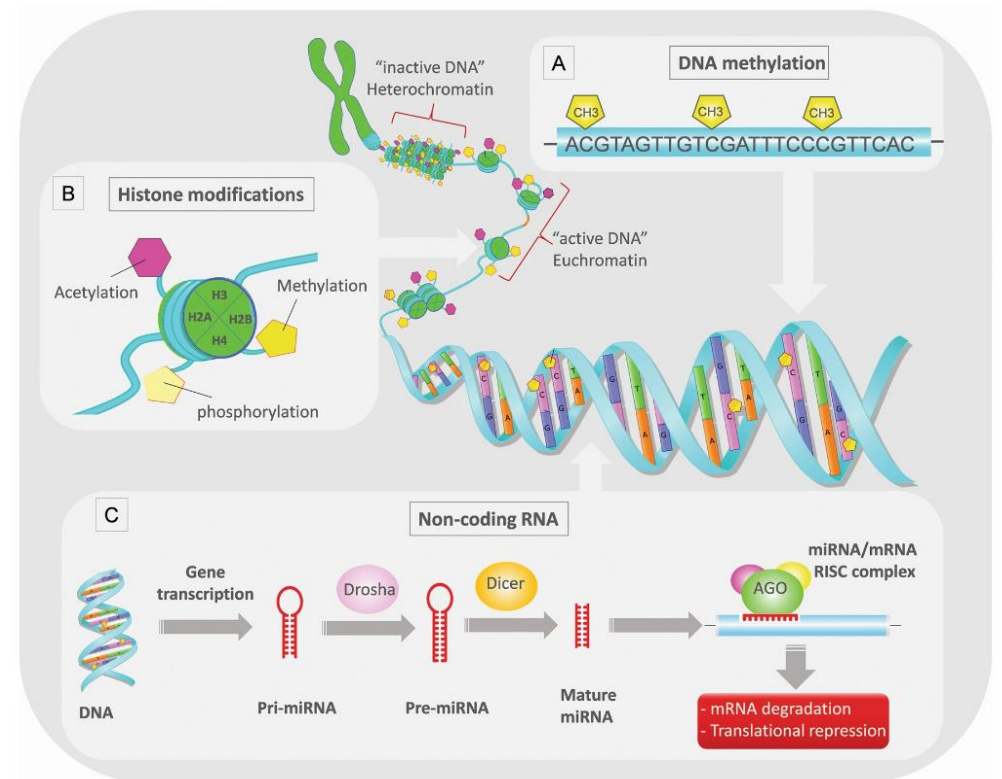
Regulación Génica 2

Guillermo Barturen Briñas
(gbarturen@ugr.es)

Regulación epigenética

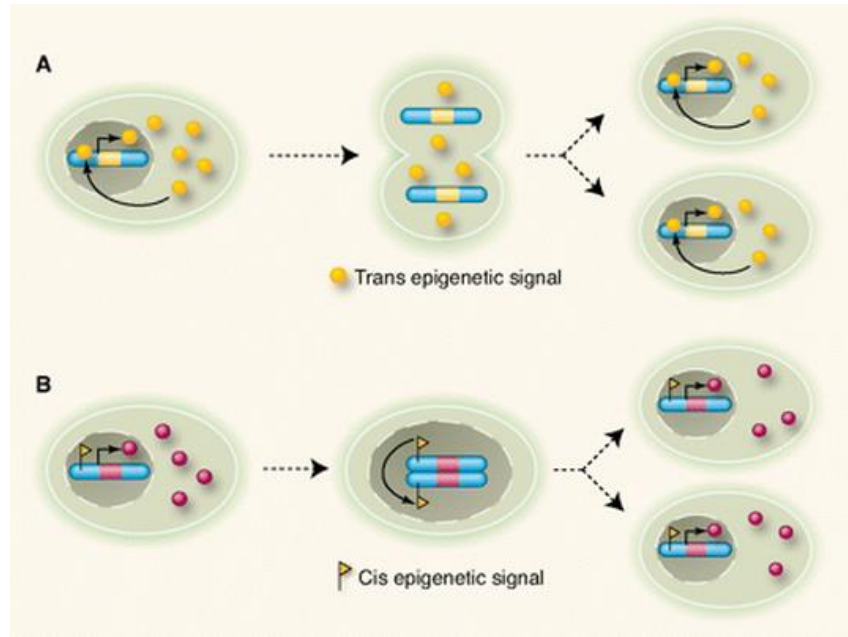
- La epigenética comprende todos los procesos que regulan la expresión de los genes por factores en los que no influya la composición de la secuencia de ADN.
- Controla la expresión génica de manera tejido y/o condición específica en un mismo individuo, independientemente de la secuencia de ADN.
- Principales intermediarios entre las condiciones ambientales y la regulación génica, lo que permite a los organismos responder a exposiciones ambientales concretas.

- Podemos clasificar la regulación epigenética en 3 factores:
 - Modificaciones del ADN
 - Modificaciones de las colas de las histonas
 - ARNs no codificantes

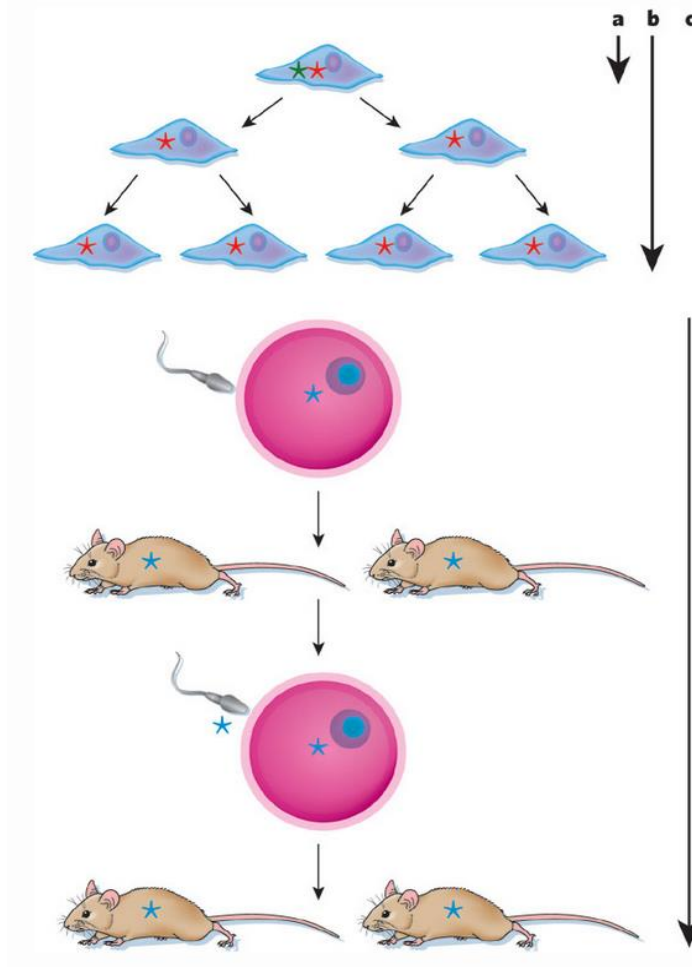


Características de los factores epigenéticos

- Regulan procesos genómicos
- No alteran la secuencia de ADN
- Heredables
- Auto-perpetuables
- Reversibles



Auto-Perpetuables

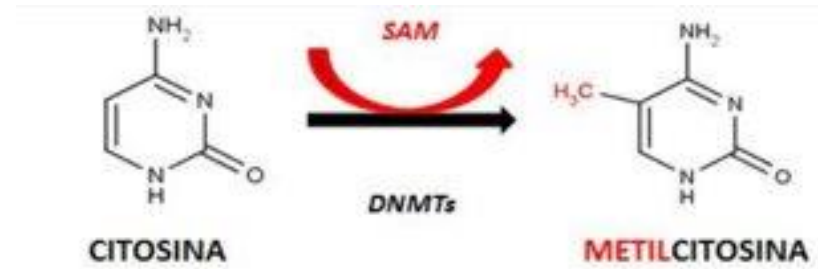


Heredables

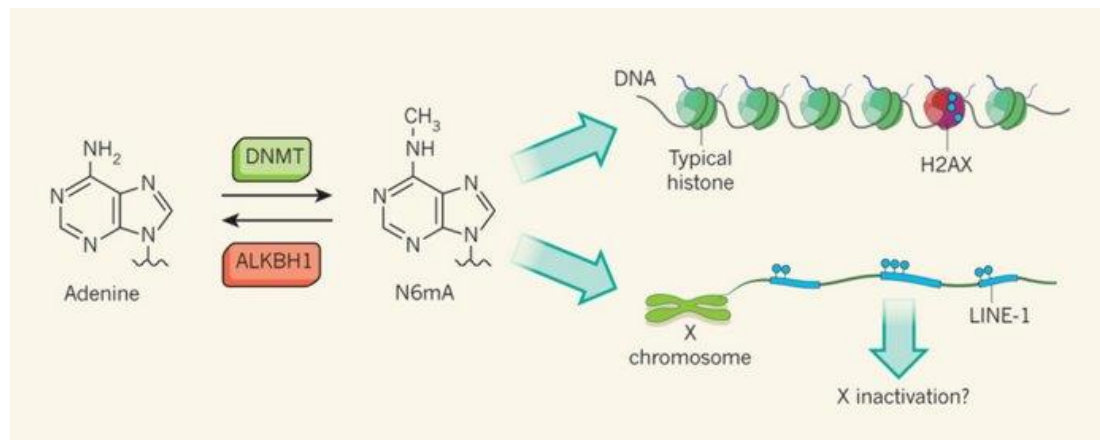
Metilación del ADN

- Es la adición química de un grupo metilo a ciertos nucleótidos del ADN
- En mamíferos suele encontrarse en la citosina de los dinucleótidos CpG

- La metilación es un proceso activo mediado por metiltransferasas específicas y requieren de S-adenosil metionina como sustrato de donde toma los grupos metilo.



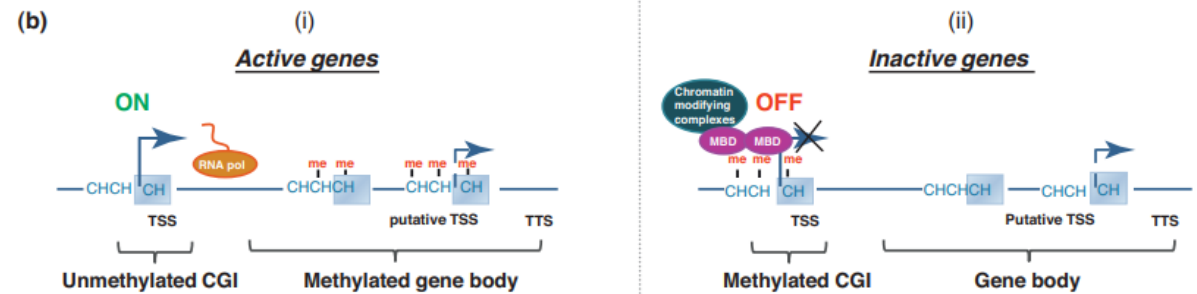
- Recientemente, se han identificado grupos metilo en las adeninas de células madre embrionarias



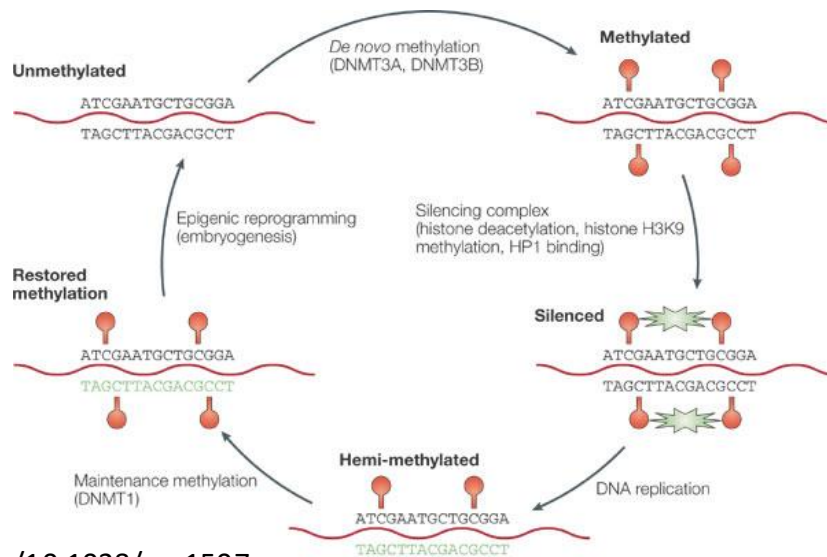
- Asociada muy específicamente a zonas de reparación del ADN (marcadas por la histona H2AX) y a la inactivación de elementos LINE-1 particularmente en el cromosoma X.

Características de la metilación del ADN

- La presencia de la metilación se relaciona con la ausencia de interacción de otras moléculas (por ejemplo, TFs) con el ADN, existen proteínas específicas que reconocen la metilación (MBD) e inducen la represión de la expresión de genes cercanos mediante modificaciones en las colas de las histonas
- La *metilación en el cuerpo génico* contribuye a la estabilidad del proceso y por lo tanto favorece la transcripción del gen
- La presencia de *metilación en el promotor* de los genes recluta MBDs que contribuyen a inhibir la transcripción de dicho gen



<https://doi.org/10.1016/j.tibs.2011.03.002>



- La metilación es heredable y perpetuable mediante dos complejos formados por metiltransferasas específicas, uno que define los **patrones de novo** (DNMT3A, DNMT3B) y otro de **mantenimiento** (DNMT1) que reconoce patrones de metilación hemi-metilados.

Características de la metilación del ADN

- Es un proceso reversible, incluso se han identificado enzimas específicas responsables de este proceso (TETs)
- 3 mecanismos de desmetilación en mamíferos:

A. Desmetilación pasiva

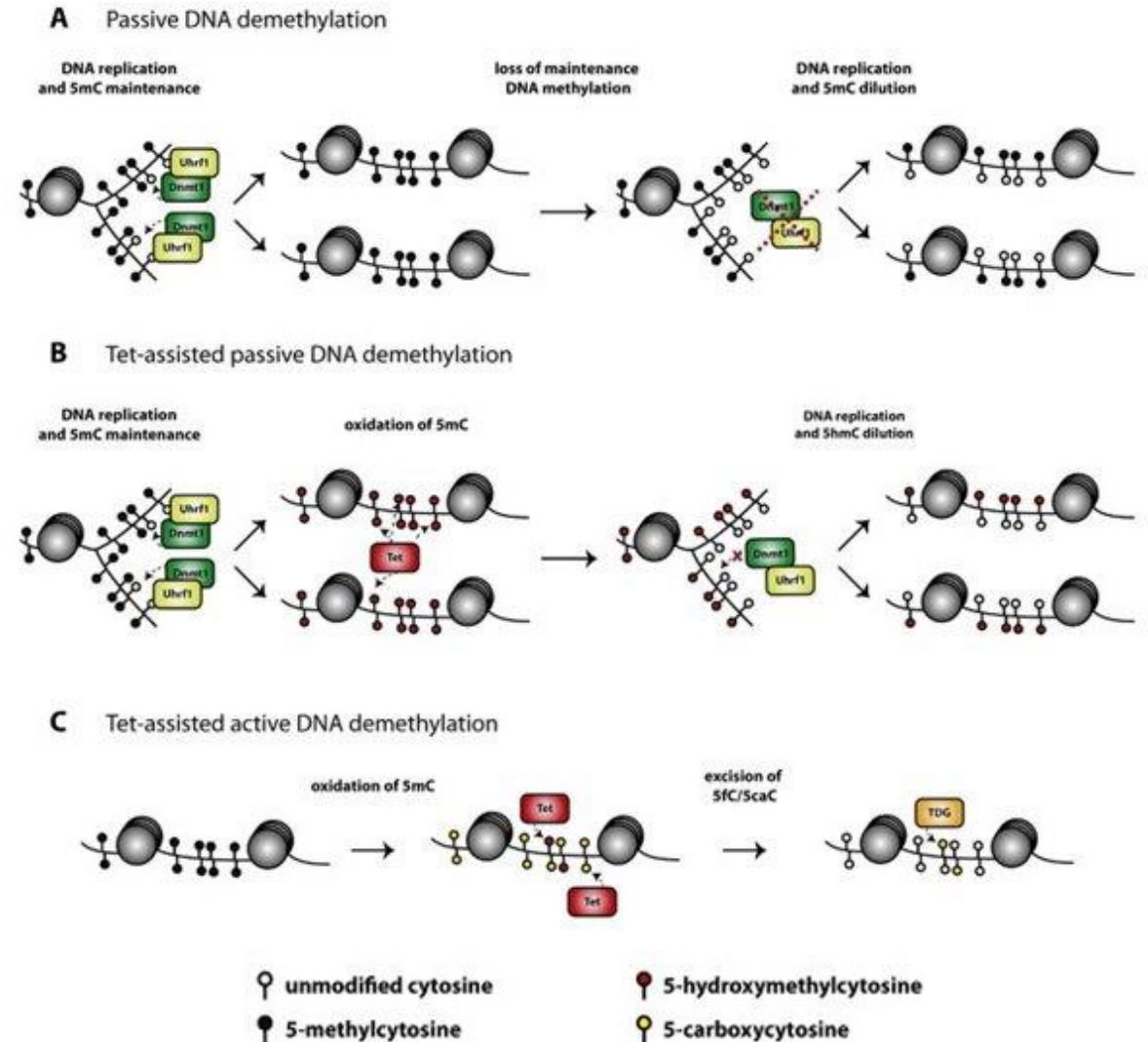
La ausencia de mantenimiento por la DNMT1 puede provocar paulatinamente la pérdida de la metilación en sucesivas replications

B. Desmetilación pasiva asistida por TETs

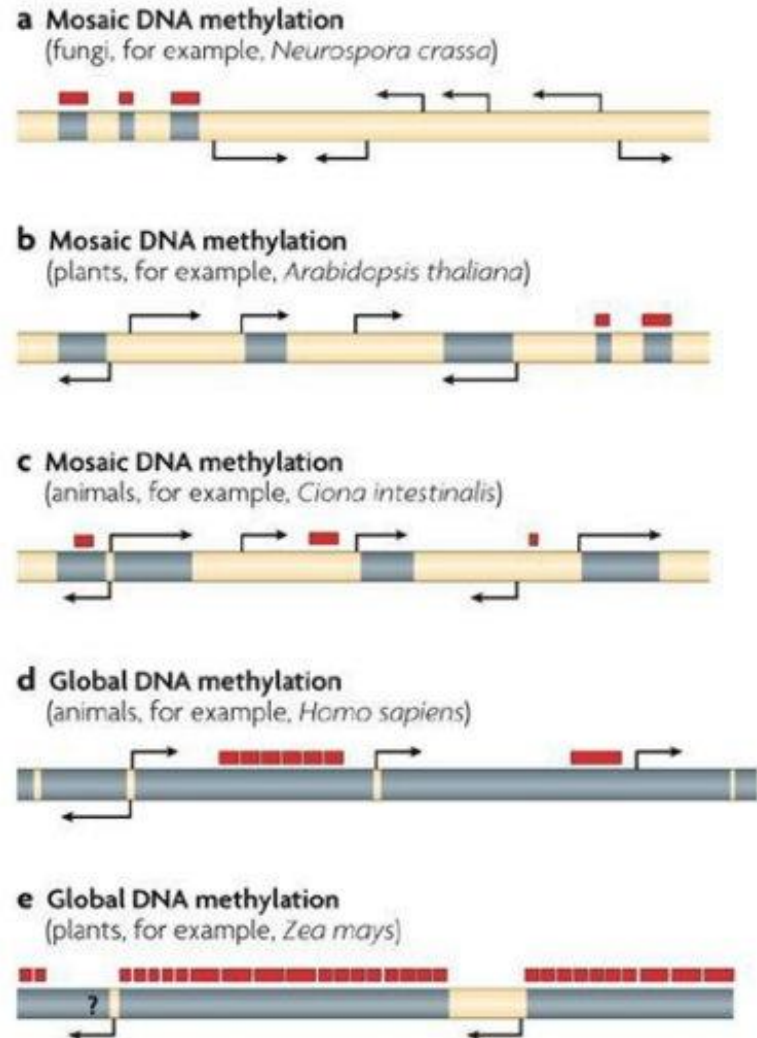
La oxidación activa por TETs de las metilcitosinas las convierte en hidroximetilcitosinas que no son reconocidas por el mecanismo de mantenimiento

C. Desmetilación activa asistida por TETs

Las hidroximetilcitosinas pueden ser de nuevo oxidadas por otras TETs produciendo formilcitosina y/o carboxilcitosina. Estas bases son reconocidos por mecanismos de reparación sustituyendo la citosina oxidada por una nueva citosina no metilada.



Distribución de la metilación



- **Patrón mosaico de metilación:** es el patrón mayoritario donde sólo ciertos elementos genómicos se metilan.
- **Patrón global de metilación:** es el patrón observado en vertebrados y plantas con genomas grandes. En este patrón la mayoría del genoma se encuentra metilado y sólo algunas regiones como los promotores de los genes y TFBSs no presentan metilación.
 - En plantas de genomas grandes este patrón ocurre debido a que la mayoría de su genoma se compone de retrotransposones.
 - En vertebrados se ha propuesto que la aparición de estos patrones globales está ligado a la aparición del TLR9.

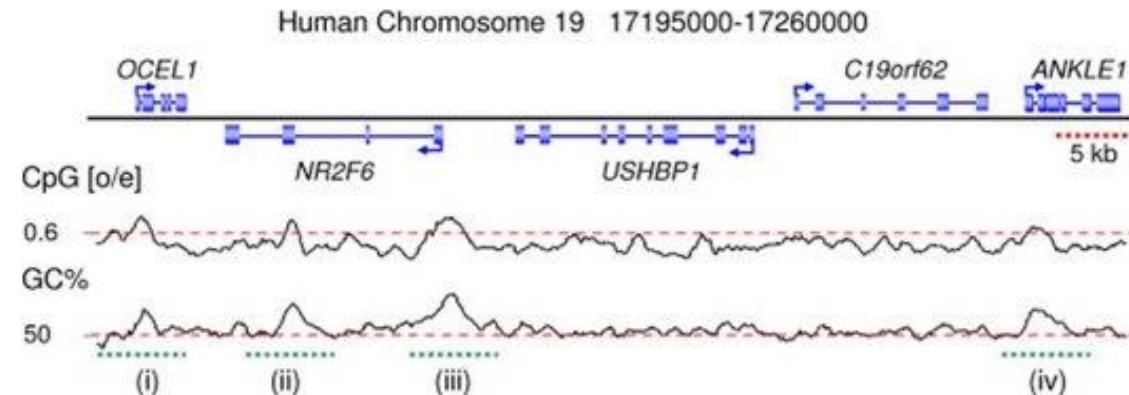
Islas CpG

- No provoca cambios en la secuencia. Sin embargo, en vertebrados la mayoría del genoma se encuentra metilado provocando un sesgo en la distribución de CpGs a lo largo del genoma a nivel evolutivo.

	Frecuencia nucleotídica observada	Frecuencia de CpGs esperada	Frecuencia de CpGs observada
Citosinas	0.2	0.2×0.2	0.008
Guaninas	0.2	0.04	

- La frecuencia observada de CpGs en el genoma es 5 veces menor que la esperada por azar

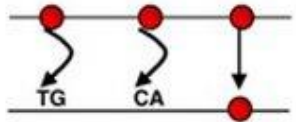
- Localizaciones concretas del genoma (principalmente promotores, no metilados) presentan enriquecimientos en dinucleótidos CpGs (Islas CpG)



Dinámica evolutiva de las islas CpG

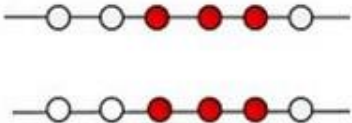
- Diferentes regímenes evolutivos mantienen los sesgos de distribución de CpGs en el genoma

A 5mCpG-Genomic background



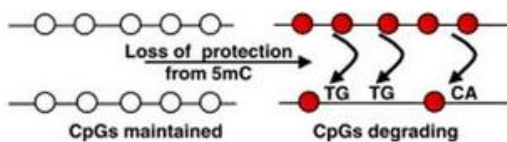
Rapid deamination
Low CpG content

C CpG selection



Low CpG loss rate
High CpG content

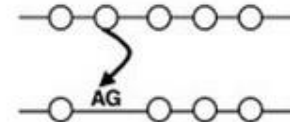
E CpG decay



Recent increase in deamination rate
CpG content is decreasing

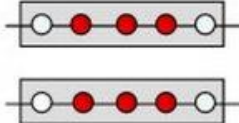
- La metilación global de citosinas aumenta la probabilidad de desaminación en la mayoría del genoma, eliminando los dinucleótidos CpGs por TpGs

B Bird's hypodeaminated regime



Low deamination
High CpG content

D Exon or general selection



Selection on non CpG + CpGs
High G+C, CpG content

- Existen regiones donde a pesar de la metilación los CpGs, estos son conservados

- Hay islas CpG recientemente metiladas en ciertas especies que acaban de iniciar su proceso de desaparición (desaminación)

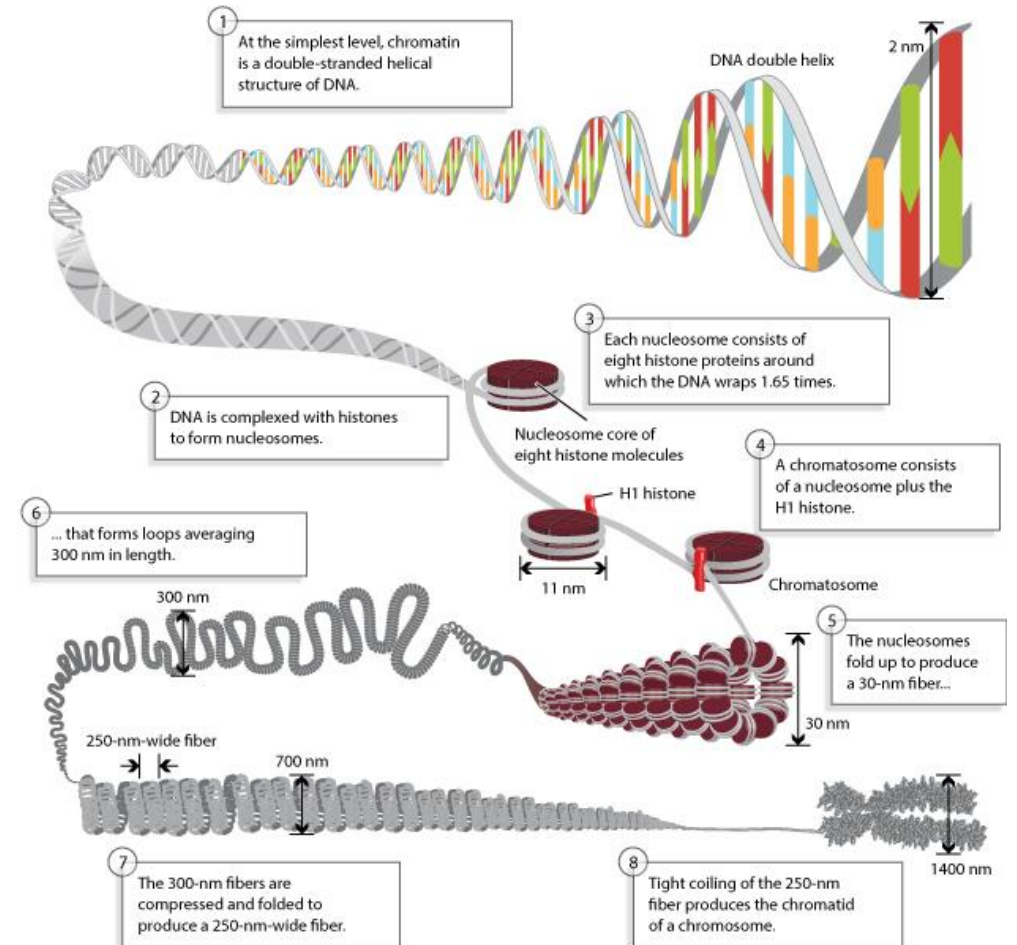
- Zonas hipometiladas como promotores no presentan esta tendencia elevada de desaminación, favoreciendo la conservación de CpGs

- Regiones donde se selecciona positivamente una elevada composición de GC y por lo tanto existe mayor probabilidad de formar CpGs

Histonas

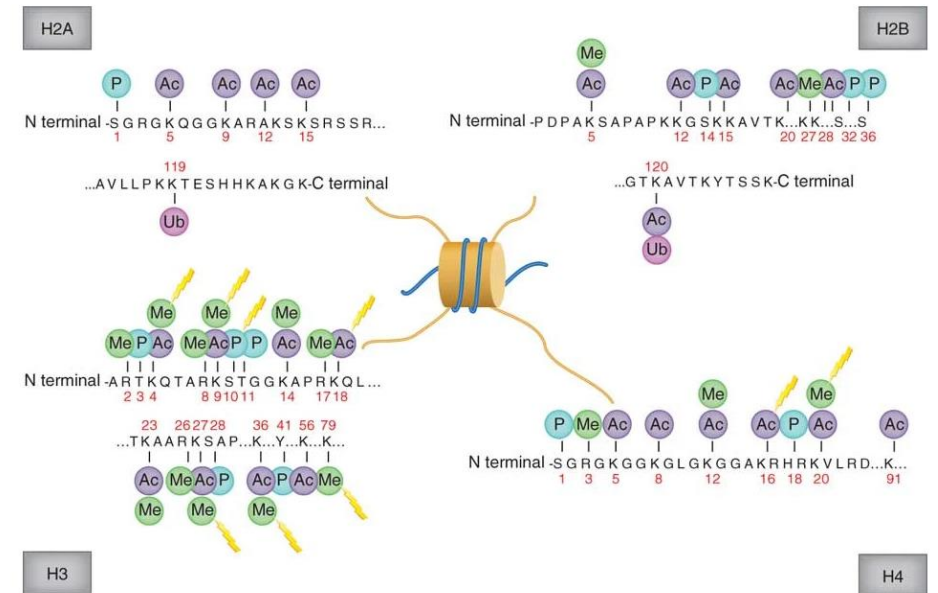
- Los humanos tenemos aproximadamente 3.000 millones de nucleótidos de ADN por célula.
- Cada nucleótido mide unos 0.34 nanómetros por lo que cada célula contiene unos 2 metros de ADN.
- Se estima que un humano está formado por unos 50 billones de células, por lo que aproximadamente tendríamos 100 billones de metros de ADN en nuestro interior.

- Las histonas son proteínas con carga positiva que forman los nucleosomas, estructuras que ayudan a compactar y plegar el ADN (carga negativa por sus grupos fosfato) dentro del núcleo y forman la cromatina.
- Un **nucleosoma** se compone de 8 histonas incluyendo 2 de cada tipo H2A, H2B, H3 y H4 que enrollan unas 146pbs (1.65 vueltas). Por otro lado, la histona H1 o de unión enrolla otras 20pbs y forma junto al nucleosoma una estructura conocida como **cromatosoma**.
- Estas estructuras se pliegan formando una fibra de 30 nm, la histona H1 juega un papel importante en el correcto plegamiento de la fibra de 30 nm.
- A su vez los pliegos de la fibra de 30 nm se comprimen y pliegan de manera más compacta para formar las **cromátidas** de los cromosomas.



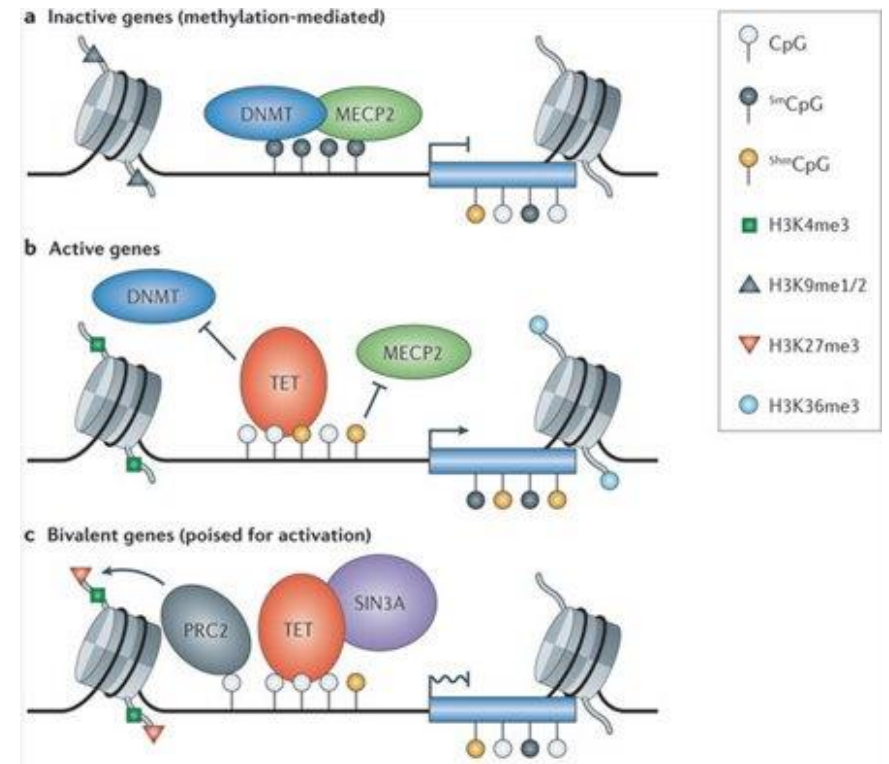
Modificaciones de las histonas

- La transcripción y la replicación del ADN requieren la separación de ambas hebras de ADN, y por lo tanto la compactación del ADN debe ser reversible para permitir el acceso de las polimerasas.
- Cuando el ADN se encuentra compactado (**Heterocromatina**) es difícilmente accesible, mientras que si no está enrollado será accesible y se considera que se encuentra activo (**Eucromatina**).
- Existen dos mecanismos que regulan la apertura temporal de la hebra de ADN y/o el desplazamiento de los nucleosomas:
 - i. Las **modificaciones enzimáticas de las colas de las histonas**, como la adición de grupos acetilo, metilo o fosfato que pueden cambiar la polaridad de las histonas y por lo tanto la configuración de la cromatina.
 - ii. Los nucleosomas pueden ser desplazados mediante los **complejos remodeladores de la cromatina**, exponiendo de esa manera partes del ADN previamente enrollados.
- Estas modificaciones son reversibles y pueden volver a su estado original una vez el proceso de transcripción o replicación haya terminado.



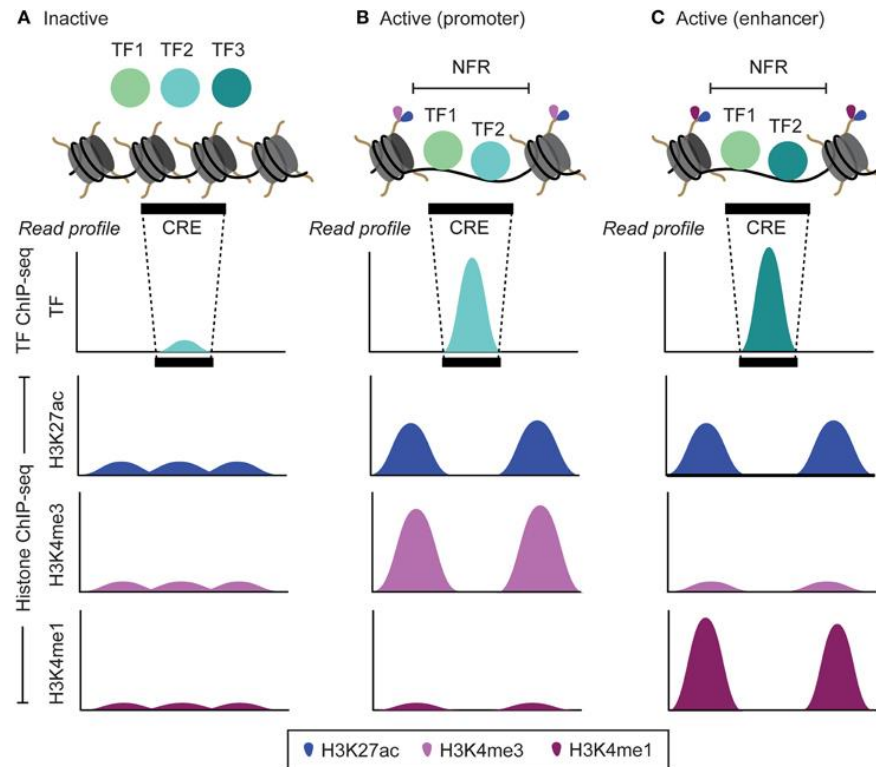
El código de histonas

- El código de histonas es el término que engloba toda la regulación de la expresión génica por medio de modificaciones postranscripcionales de las colas de las histonas.
- Estas modificaciones interactúan y definen la función de otros reguladores génicos como la metilación del ADN o los factores de transcripción.
- **Genes inactivos** presentan promotores metilados que suelen encontrarse flanqueados por **H3K9me1/2** y su cuerpo génico se encuentra parcialmente desmetilado.
- **Genes activos** presentan promotores desmetilados y flanqueados por histonas con **H3K4me3** y **H3K27ac**, mientras que sus cuerpos génicos presentan elevados niveles de metilación con histonas H3K36me3.
- **Gene bivalentes** presentan una configuración intermedia y particularmente histonas **H3K27me3** en su promotor.

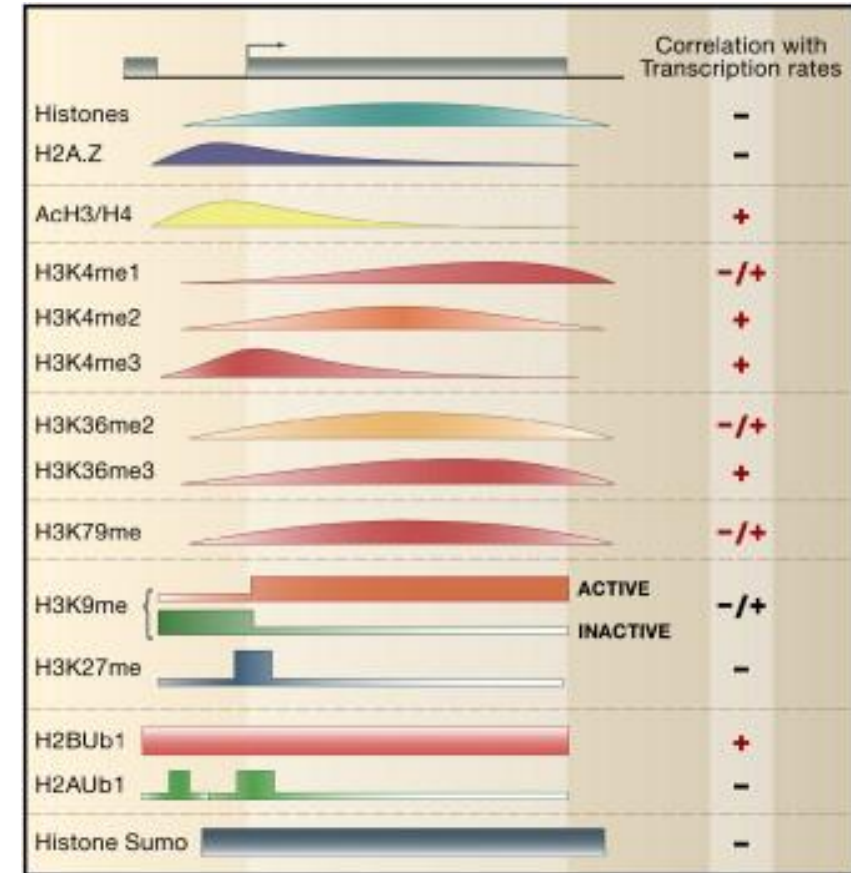


El código de histonas

- Las modificaciones de las colas definen la funcionalidad de los factores de transcripción. Promotores y potenciadores presentan H3K27ac, mientras que H3K4me3 y H3K4me1 son específicos de cada uno.



- El tipo de modificación y el número de moléculas de las mismas tienen propiedades diferentes.



<https://doi.org/10.1016/j.cell.2007.01.015>

<http://dx.doi.org/10.3389/fgene.2015.00188>

ARNs no codificantes

- La secuenciación masiva ha revelado que la gran mayoría de los genomas de eucariotas se transcriben (~75%), sin embargo, sólo el ~1-2% codifican proteínas. Por lo que la mayor parte del genoma se transcribe como ARNs no codificantes (ncRNAs).
- Existen numerosas evidencias de las funcionalidades de estos ncRNAs, que suelen expresarse a niveles muy inferiores a los ARNs codificantes lo que sugiere que deben tener funciones de regulación tejido/condición específicas.
- Se pueden clasificar en:
 - i. ARNs estructurales**
 - ARNs ribosómicos (rRNAs)
 - ARNs de transferencia (tRNAs)
 - ARNs pequeño nuclear (snRNAs)
 - ARNs pequeño nucleolar (snoRNAs)
 - ii. ARNs reguladores**
 - Micro-ARNs (miRNAs)
 - ARNs asociados a piwi (piRNAs)
 - ARNs pequeño de interferencia (siRNAs)
 - ARNs largos no-codificantes (lncRNAs)
 - Otros que son agrupados según su funcionalidad: eRNAs, PARs...

ARNs no codificantes estructurales

- Son ARNs fundamentales en el correcto funcionamiento de la célula, ya que intervienen en procesos vitales como la transcripción o la traducción de genes a proteínas.
- **rRNA:** es parte fundamental de los ribosomas, el orgánulo donde se sintetizan las proteínas. Cada ribosoma está formado por un 60% de rRNA (incluyendo diferentes subunidades) y un 40% de proteínas ribosomales. Este RNA supone aproximadamente el 80% del RNA transcrito en una célula.
- **tRNA:** hacen de unión entre la molécula mensajera de RNA y la cadena naciente de aminoácidos. La complementariedad entre los codones del mensajero (tripletes de nucleótidos) y los anti-codones de cada uno de los tRNAs existentes reclutan los aminoácidos específicos y por lo tanto median el proceso de traducción.
- **snRNA:** su principal función es procesar el RNA pre-mensajero (eliminando los intrones) y forma parte de los espliceosomas que regulan el proceso de barajamiento de exones (*alternative splicing*).
- **snoRNA:** dirigen las modificaciones químicas postranscripcionales de otros ARNs no codificantes, principalmente del resto de ARNs estructurales.

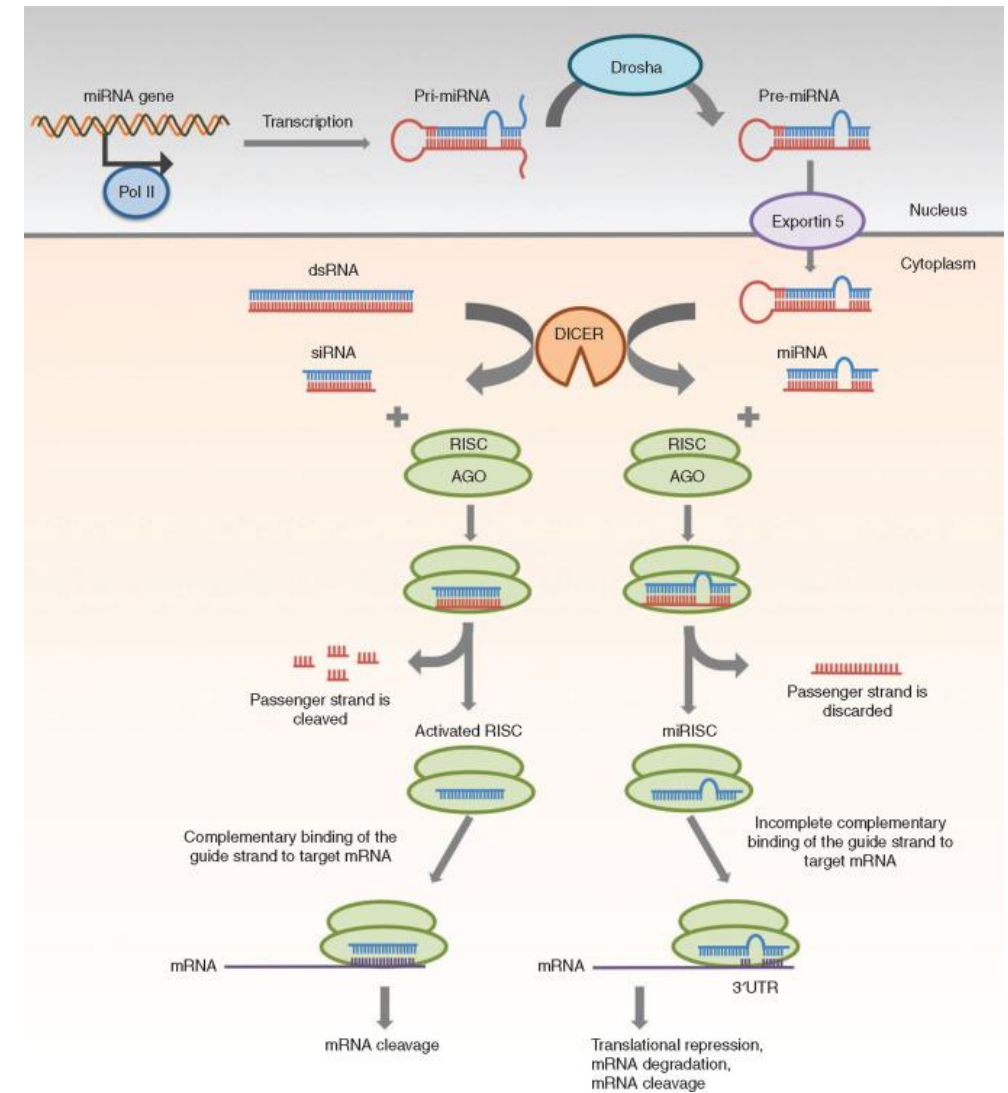
ARNs no codificantes reguladores

- Son ARNs funcionales que orquestan la regulación fina de la expresión de otros genes fundamentales para el correcto funcionamiento de la célula en tejidos y/o condiciones específicas.
- **miRNA:** son moléculas de ARN simple pequeñas (20-24 nts) que regulan la expresión de alrededor del 50% de los genes a nivel postranscripcional. Cumplen su función por complementariedad impidiendo la traducción de ARNs mensajeros que son degradados.
- **siRNA:** al igual que los miRNAs son moléculas cortas de RNA simples que regulan la expresión a nivel postranscripcional de manera similar a los miRNAs. Se diferencian de los miRNAs en su biogénesis y a nivel funcional suelen regular ARNs mensajeros específicos, mientras que un miRNA puede regular múltiples transcritos.
- **piRNA:** son ARNs pequeños (24-31 nts) que se unen a proteínas Piwi de la familia "Argonauta" e inhiben la expansión de elementos transponibles.
- **lncRNA:** son ARNs no codificantes largos (>200 nts), la mayoría de ARNs no codificantes pertenecen a esta categoría y algunos actúan como transcritos primarios para la producción de ARNs cortos no codificantes. Estos ARNs no codificantes se caracterizan por su localización nuclear, bajo nivel de expresión, bajo nivel de conservación de secuencia y el 70% de los mismos no contienen cola de poliA. Se suelen subdividir en función de su localización y dirección de transcripción con respecto a genes codificantes: anti-sentido, intrónicos, bidireccionales, intergénicos...

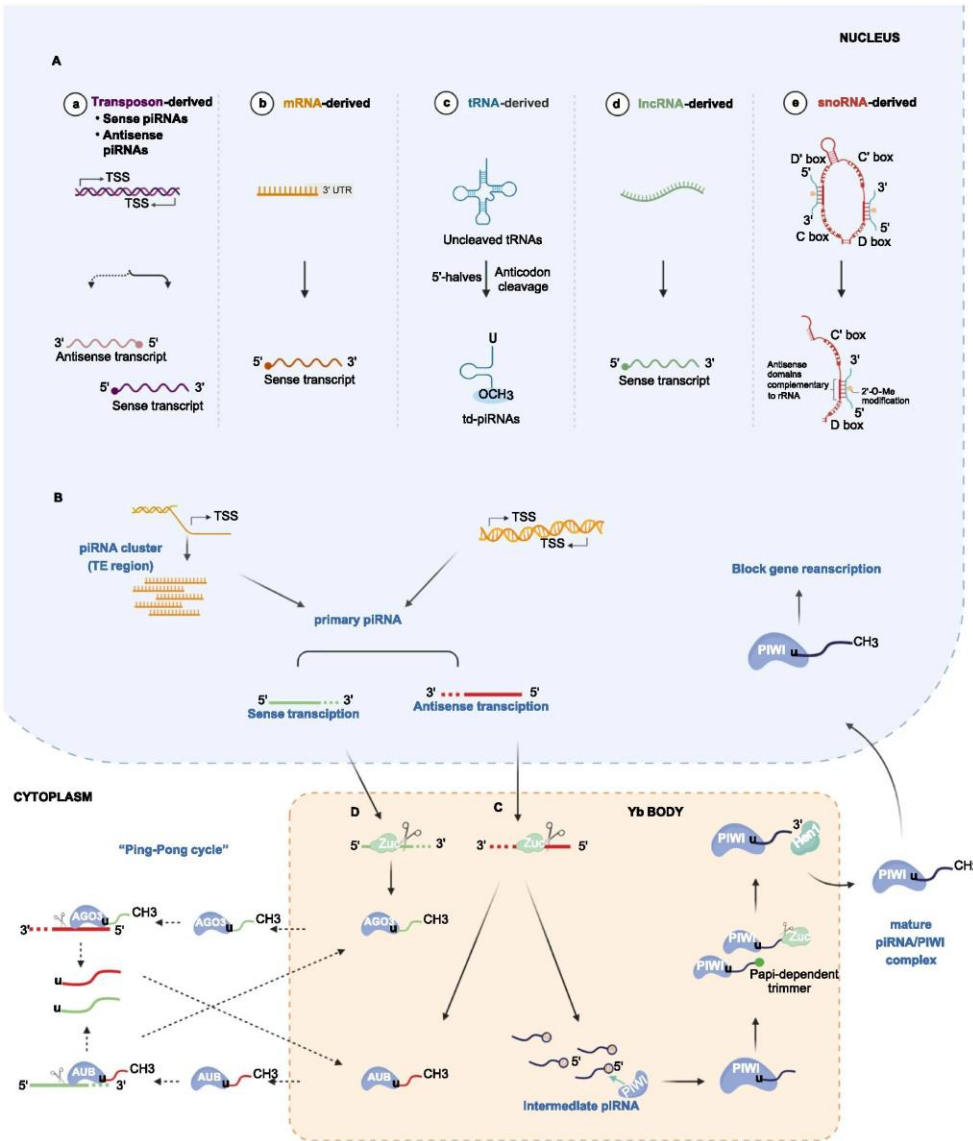
Biogénesis y funcionalidad de miRNAs y siRNAs

siRNA: un ARN de doble hebra (transcrito o artificial) es procesado en el citoplasma por Dicer para dar el siRNA que es cargado en el complejo RISC-AGO2, la hebra pasajera se trocea mientras que la hebra guía dirige el complejo RISC al mRNA diana. Los siRNAs requieren de emparejamiento perfecto para cumplir su función de inhibir el mRNA mediante rotura.

miRNA: los genes de miRNAs se transcriben en el núcleo mediante la RNA polimerasa II generando el pri-miRNA, que es procesado por Drosha separando el pre-miRNA de la secuencia transcrita completa. El pre-miRNA es transportado al citoplasma mediante la exportina 5 donde Dicer corta el tallo-asa (*stem-loop*) y el miRNA se carga en el complejo RISC. La hebra pasajera es descartada y la hebra guía dirige el complejo hacia su mRNA diana. En este caso no es necesaria una complementariedad completa de secuencias, sino que la complementariedad de la semilla (normalmente, nucleótidos 2-7 de la región 5') del miRNA con el mRNA diana es suficiente para inhibir la traducción mediante inhibición directa, degradación del mRNA diana o rotura como ocurre con los siRNAs.

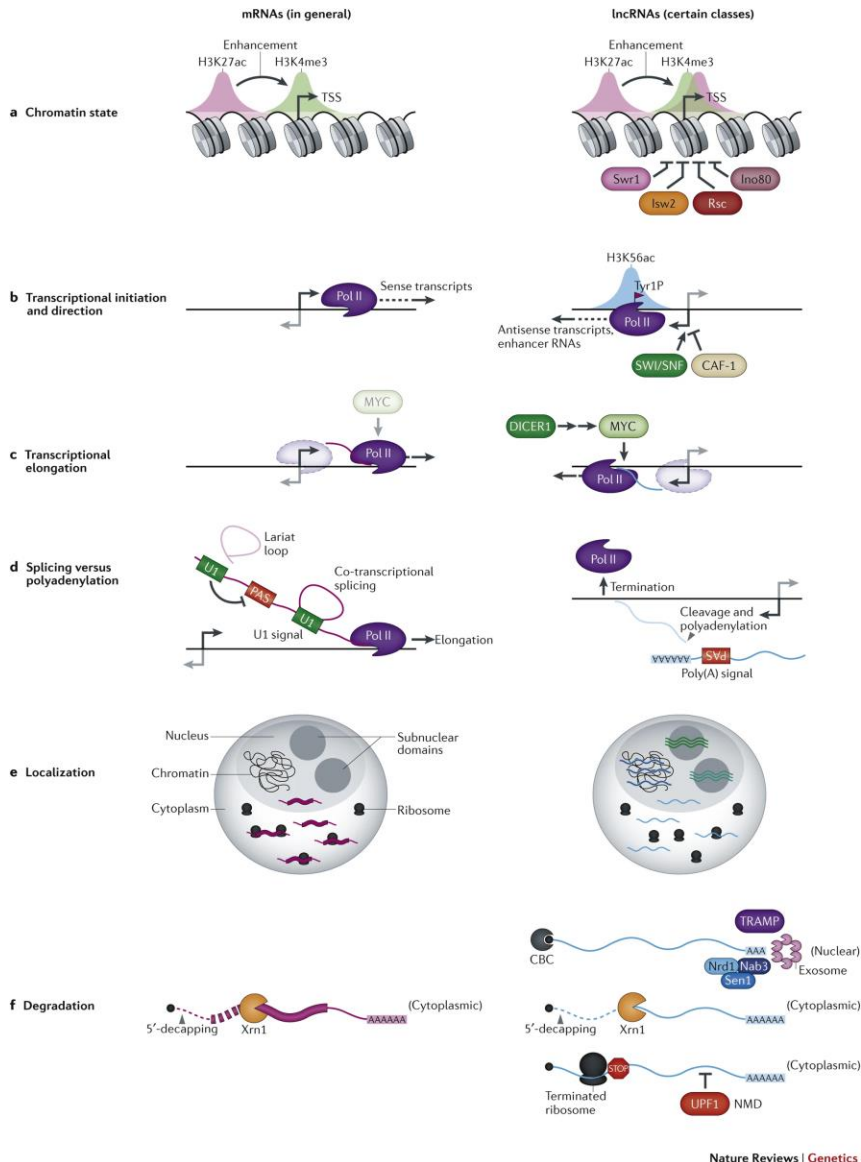


Biogénesis y funcionalidad de piRNAs



piRNA: (A/B) sus transcritos primarios pueden derivar de diferentes fuentes: transposones en regiones eucromáticas, mRNAs, tRNAs, lncRNAs, snoRNAs o loci que contienen múltiples secuencias de piRNAs (*clusters* de piRNAs). Los transcritos primarios se exportan a los cuerpos Yb en el citoplasma y son procesados por una endoribonucleasa llamada Zucchini, dando lugar a los intermediarios de piRNAs. (C) Los fragmentos 5' son cargados en el complejo Piwi donde son recortados en su extremo 3' hasta longitudes entre 24-31 nts, estos fragmentos se metilan en 3' mediante Hen1 dando lugar a su forma activa que es enviada de nuevo al núcleo para cumplir su función (principalmente silenciar elementos transponibles). (D) La ruta de amplificación "*ping-pong*" es un mecanismo secundario de producción de piRNAs en el citoplasma. Suele darse en líneas germinales, donde los intermediarios se unen a las proteínas AGO3 (principalmente intermediarios con A en posición 10) o AUB (principalmente intermediarios con U en 5'). Estos complejos contienen secuencias complementarias entre ellos, por lo que un complejo piRNA/AGO3 cortará su diana complementaria para generar secuencias sustrato para complejos piRNA/AUB y viceversa. Estos complejos actúan reclutando proteínas que inhiben la transcripción de elementos transponibles mediante cambios en las colas de las histonas o reclutamiento de metiltransferasas.

Biogénesis y funcionalidad de lncRNAs



lncRNA: son ARNs largos (>200nts) que no codifican proteínas, además su expresión es considerablemente menor que los genes codificantes pero también presentan un especificidad mucho mayor, por lo que se consideran especialistas en la regulación de estados concretos de la célula.

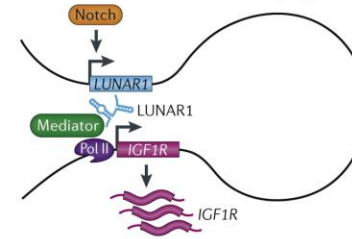
Existen otras diferencias en su regulación con respecto a los mRNAs:

- Estado de la cromatina similar pero mayor enriquecimiento de H3K27ac y mayor represión por parte de proteínas remodeladoras.
- lncRNAs antisentido en promotores bidireccionales presentan regulación diferente: activado por H3K56ac, complejo remodelador SWI/SNF y fosforilación del dominio Tyr1 de la RNAPol II.
- Elongación más dependiente del circuito de miRNAs y MYC que los mRNAs.
- En promotores bidireccionales enriquecimiento de maquinaria de splicing en los mRNAs y de poliadenilización en los lncRNAs antisentido (asegura su terminación eficiente y temprana).
- Localización diversa: citoplasma, cromatina, nucleosoma o dominios subnucleares.
- Degradación por múltiples vías: 5' decapping y digestión por exonucleasas 5'-3' (como los mRNAs), digestión mediante el complejo exosoma por exonucleasas 3'-5' o terminación prematura y vía de "non-sense mediated decay" (NMD).

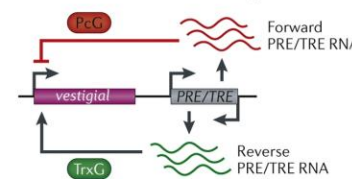
Biogénesis y funcionalidad de lncRNAs

- Muchos lncRNAs son post-procesados para dar lugar a otras clases de ARNs no codificantes más pequeños como: miRNAs, snoRNAs, tRNAs...
- Sin embargo, también cumplen importantes funciones como lncRNAs:
 - Actúan formando bucles que facilitan la interacción entre potenciadores y los genes que regulan (*cis*)
 - Regulan mediante el complejo PRE/TRE (elementos de respuesta Polycomb / Trithorax) modifican el estado de la cromatina de genes cercanos.
 - Regulan genes cercanos en respuesta a cambios de la metilación mediados por impronta.
 - Contribuyen a la compensación de dosis del cromosoma X (Xist y su regulación).
 - Inhiben en *cis* la expresión por reclutamiento de Polycomb.
 - Inhiben en *cis* la expresión por colisión de las RNAPol II en ambas hebras.
 - Inhiben la transcripción por complementariedad en el promotor.
 - Pueden autoregularse reclutando los TFs necesarios.

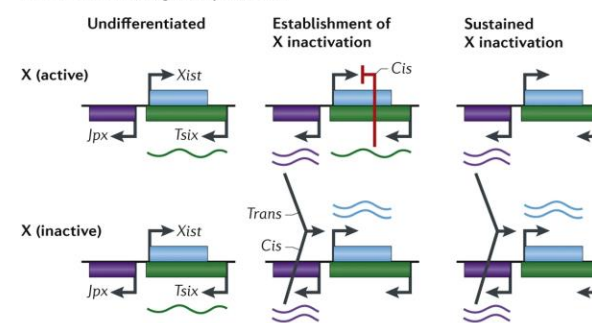
a Enhancer RNAs and chromosome looping



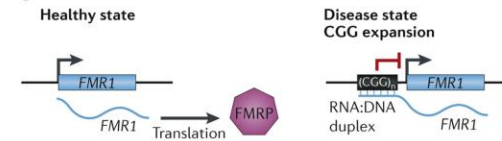
b PRE/TRE enhancer RNA switching



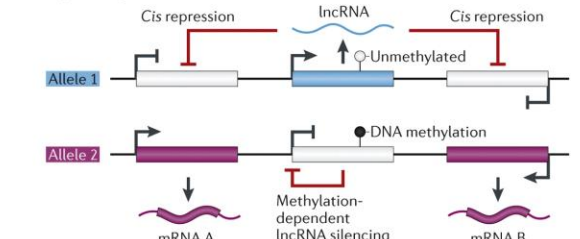
d Mammalian dosage compensation



g FMR1 auto-inhibition

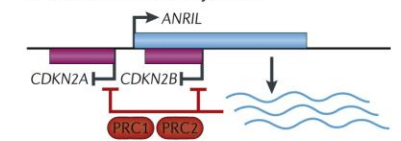


c Imprinted gene clusters

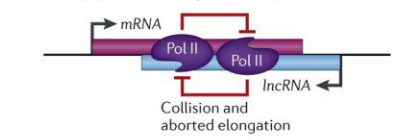


lncRNA	Expression	mRNAs	Expression	Refs
Air	Paternal	<i>Igf2r</i> , <i>Slc22a2</i> and <i>Slc22a3</i>	Maternal	136
Kcnq1ot1	Paternal	<i>Kcnq1</i> , <i>Cdkn1c</i> , <i>Ascl2</i> and others	Maternal	139
Nespas	Paternal	<i>Gnas</i> and <i>Nesp</i>	Maternal	140
H19	Maternal	<i>Igf2</i>	Paternal	132
Gtl2	Maternal	<i>Dlk1</i> and <i>Rtl1</i>	Paternal	141

e Antisense inhibition by ANRIL



f Antisense inhibition by Pol II collision



h roX1 autoregulatory loop

