

# ESTRUCTURA DEL GENOMA HUMANO Y MECANISMOS DE EXPRESIÓN GÉNICA

---

MODULO 1: CONCEPTOS GENERALES DE GENÉTICA HUMANA

CERTIFICADO DE GENÉTICA MÉDICA 2014 - 2015

**Material didáctico: Modulo1-Clase1**

**MODULO 1: CONCEPTOS GENERALES DE GENÉTICA HUMANA**

**ASIGNATURA 1.1- ESTRUCTURA DEL GENOMA HUMANO Y MECANISMOS DE EXPRESIÓN GÉNICA. Profesor: Dra. Nuria Paricio Ortiz**

**ÍNDICE**

1.-ESTRUCTURA DEL GENOMA HUMANO.....	1
1.1.- Genoma mitocondrial.....	1
1.2.- Genoma nuclear.....	2
1.3.- Tipos de secuencias del genoma humano.....	2
2.-MECANISMOS DE EXPRESIÓN GÉNICA.....	8
2.1.-Modificaciones a nivel de transcripción.....	9
2.2.-Modificaciones a nivel de procesado de RNA.....	10
2.3.-Modificaciones a nivel de traducción.....	12
2.4.-Mutaciones en regiones implicadas en la expresión de genes.....	13
3.-PROYECTO ENCODE.....	14
ANEXO I. Información adicional y preguntas.....	16

**NOTA: Las diapositivas asociadas al siguiente texto didáctico se encuentran disponibles en el Aula Virtual, Recursos/Módulo 1/Carpeta 1.1/PDF.**

## 1.-ESTRUCTURA DEL GENOMA HUMANO

El GENOMA es el contenido total de ADN presente en una célula. El genoma humano consta de 2 tipos, el genoma nuclear y el genoma mitocondrial.

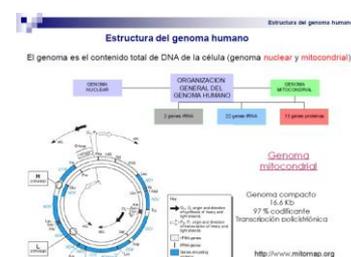
### 1.1.- Genoma mitocondrial

Este genoma incluido dentro de las mitocondrias está constituido por una molécula de ADN circular de cadena doble y de 16,6 kb de longitud (de tamaño muy pequeño comparado con el genoma nuclear), y además posee una capacidad codificante pequeña: sólo contiene 37 genes codificantes, de los cuales 24 contienen información para codificar RNA ribosómico (RNAr) y de RNA de transferencia (RNAt), y 13 genes que codifican proteínas necesarias para el funcionamiento de la mitocondria. El resto de proteínas necesarias para la mitocondria procede del genoma nuclear.

Es necesario comentar que el 97% del genoma mitocondrial es codificante y tiene una transcripción policistrónica, es decir, se transcribe en forma de un transcrito continuo, y su secuencia se conoció antes que el genoma humano, en el año 1981. Más información sobre el ADN mitocondrial se encuentra en la página web <http://www.mitomap.org>.

En el año 1988, se descubrió que en el genoma mitocondrial existían algunas mutaciones en genes causantes de enfermedades humanas y hereditarias, como por ejemplo la Neuropatía Óptica Hereditaria de Leber (LHON). Esta enfermedad es una enfermedad mitocondrial neurodegenerativa que afecta al nervio óptico y que se caracteriza por una pérdida súbita de la visión en los adultos jóvenes que son portadores. No se conoce con exactitud su prevalencia, pero está estimada en torno a 1/15.000 - 1/50.000 en todo el mundo. Como hemos comentado, la LHON está causada por mutaciones en el ADN mitocondrial (ADNmt), en las que más del 90% ocurren en las posiciones 11778, 3460 o 14484. Todas producen alteraciones en los genes *MT-ND1*, *MT-ND4* y *MT-ND6* del complejo I de la cadena respiratoria mitocondrial, y se transmite de forma materna, ya que el ADN mitocondrial exclusivamente se transmite de esta manera.

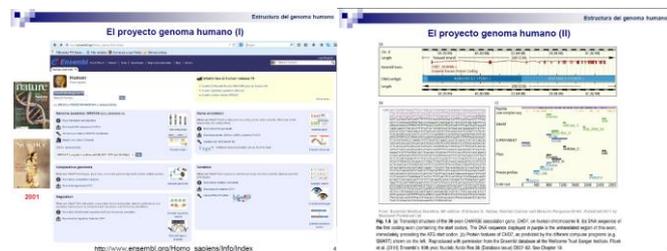
**CONSULTAR DIAPOSITIVA 3: Estructura del genoma humano.**



## 1.2.- Genoma nuclear

En los años 90 se inicia el Proyecto Genoma Humano, que persigue el objetivo de secuenciar el ADN nuclear, y que culminó en 2001 con el primer borrador de la secuencia completa del genoma humano. Desde entonces hasta ahora toda la información del genoma humano ha ido actualizándose constantemente. En la página web [http://www.ensembl.org/Homo\\_sapiens/index.html](http://www.ensembl.org/Homo_sapiens/index.html), está toda la información más actualizada sobre este proyecto, así como toda la información que actualmente se posee sobre el genoma humano, por ejemplo, la secuencia de todos los cromosomas humanos, información sobre genes concretos, información sobre gen/secuencia nucleotídica/estructura Exón-Intrón/nº de transcritos y cuáles son/codificación de proteínas y dominios estructurales, etc. La información sobre el genoma humano se renueva constantemente, la última versión es de junio de 2013.

**CONSULTAR DIAPOSITIVAS 4 y 5:  
El proyecto Genoma Humano.**



## 1.3.-Tipos de secuencias del genoma humano.

Como curiosidades dentro de las características generales del genoma humano, es que se está actualmente trabajando con la versión 73.37, y en total su longitud es de 3300Mb aproximadamente. Posee 20.769 genes codificantes para proteínas, genes que especifican RNA no codificantes (largos y cortos), 14.165 pseudogenes (genes inactivos) y 195.565 transcritos. El porcentaje de genoma humano que codifica proteína es sólo de un 1.5%.

Existen dos tipos de secuencias en el genoma humano:

- <50% secuencias de copia única, que son los genes que codifican proteínas, los codificantes, intrones, secuencias reguladoras y genes que contienen información específica sobre RNA no codificante.
- > 50% secuencias repetidas ó ADN repetido, en las cuales podemos encontrar dos categorías: 1) Secuencias agrupadas en tándem, que es una secuencia nucleotídica repetida en un lugar cromosómico un nº determinado de veces. Entre ellas se encuentran los ADN satélites, los ADN minisatélites, los microsátélites y los genes de ARNr; 2) Secuencias agrupadas de forma dispersa, en las cuales las repeticiones se encuentran en diferentes puntos del genoma. Entre ellos se encuentran los transposones ó elementos transponibles, denominados retrotransposones en el genoma.

Estas secuencias repetidas son importantes en estudios de genética poblacional y genética médica y además ofrecen una gran fuente de variación para producir la evolución. Encontramos:

- A) DNA repetido en tándem
- B) DNA repetido y disperso en el genoma

**CONSULTAR DIAPOSITIVAS 6 y 7:**

**Características generales del genoma humano y DNA repetido en el genoma humano.**

**A) DNA repetido en tándem:**

Dentro de los ADNs repetidos en tándem, vamos a centrar nuestros esfuerzos en explicar principalmente las tres clases más importantes, el ADN satélite, los minisatélites y los microsatélites, cuya diferencia viene determinada por la extensión de la región que contiene las secuencias repetidas en tandem:

- 1) ADN satélite: son sólo aquellos ADNs cuyas repeticiones abarcan regiones que tienen cientos de Kb.
- 2) ADN minisatélites: repeticiones que abarcan regiones mucho más pequeñas (0.1-20 Kb).
- 3) ADN microsatélites: repeticiones que abarcan regiones de 100 bp.

**CONSULTAR DIAPOSITIVA 8:**

**DNA repetido en tándem.**

1) ADN satélite:

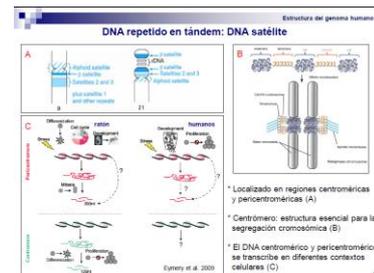
Es el ADN más variado ya que está constituido por muchas familias tales como  $\alpha$ ,  $\beta$ ,  $\delta$ , satélite 1, 2 y 3, y algunos satélites específicos del cromosoma Y. Estas secuencias tienen una localización bastante específica, son secuencias que se encuentran sobre todo en las regiones centroméricas y en las regiones pericentroméricas. En la diapositiva 8, vemos un fragmento del cromosoma humano 9, en donde la constricción representa el centrómero, y la mayor parte de la región centromérica está formada por

satélites tipo  $\alpha$  (alfoide). El resto de satélites  $\beta$ , 2 y 3, están localizados en la región pericentromérica. La importancia de las secuencias alfoides radica en que están implicadas de manera notable en la formación del centrómero, que es una estructura esencial para la segregación cromosómica durante la mitosis y la meiosis.

En las secuencias ADN alfoides se unen unas histonas específicas de los centrómeros que son las Cen H3 (variante de la Histona H3), que se unen al cinetocoro (estructura necesaria en el huso cromático para la separación de los cromosomas). En este aspecto se recalca que las regiones que contienen estas secuencias satélites son heterocromáticas y siempre esto se asocia con ADN silenciado (ADN que no se expresa). Pero recientemente se han encontrado transcritos procedentes de regiones centroméricas y pericentroméricas del ADN satélite, de hecho este ADN se transcribe en ciertas circunstancias, concretamente en el proceso de diferenciación, proliferación celular y estrés, por tanto la descondensación de estas regiones podría estar relacionada con la aparición de cáncer.

### CONSULTAR DIAPOSITIVA 9:

#### DNA repetido en tándem: DNA satélite.



## 2) ADN minisatélites

Existen dos familias, la familia de minisatélites que forma parte de los telómeros (extremos de los cromosomas), y las secuencias denominadas DNA minisatélites hipervariables o VNTR.

Los telómeros son estructuras del cromosoma formadas por una secuencia de 6 bp repetidas miles de veces, esto es el ADN minisatélite. Realiza principalmente dos funciones, la primera proteger a los extremos de los cromosomas de la degradación, a través de su reconocimiento por un complejo proteico llamado Selterina.

La estructura de los telómeros son secuencias de ADN de cadena doble hasta un cierto punto y luego continúa con una protuberancia de cadena sencilla que tiene 50-500 nucleótidos de largo. Todo el complejo proteico Selterina interacciona tanto con las secuencias de cadena doble como las de sencilla, y permite la formación de las estructuras Lazos T (o T-loops), que protegen al telómero de la degradación. Una segunda función de los telómeros es mantener la longitud de los cromosomas lineales. La enzima telomerasa alarga los cromosomas a través de estas secuencias porque contiene una secuencia complementaria a estos ADN



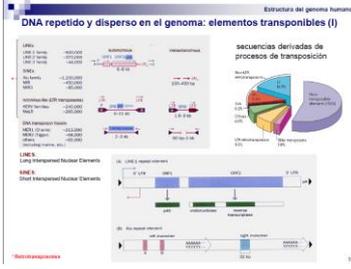


**B) DNA repetido y disperso en el genoma****Elementos transponibles:**

Son secuencias derivadas de la transposición, que son capaces de cambiar su posición en el genoma. En la diapositiva 15 vemos varios tipos de estos elementos transponibles que se han detectado en el genoma humano y el nº de copias que se han detectado de cada uno de ellos. La mayor parte de ellos pertenecen a la familia de los retrotransposones. Un transposón de ADN es una secuencia que se cambia de posición en el genoma, y un retrotransposón también cambia de posición pero a través de un intermediario de ARN, es decir, el retrotransposón se transcribe, da lugar a un transcrito, el transcrito se copia a ADN mediante la acción de una transcriptasa reversa codificada por el mismo, y ese ADN es el que se inserta en una nueva región del genoma.

**CONSULTAR DIAPOSITIVA 16:**

**DNA repetido y disperso en el genoma: elementos transponibles (I).**



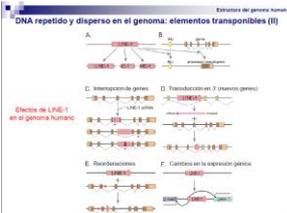
En el genoma humano, los retrotransposones más abundantes son LINE-1 (Long Interspersed Nuclear Elements) que son muy activos y se mueven constantemente en el genoma. Otra familia importante son Alu, retrotransposones que han perdido la capacidad de codificar la transcriptasa reversa, con lo que por ellos mismos no pueden moverse a lo largo del genoma, pero lo hacen con la maquinaria de los LINE-1.

Los transposones de ADN se les denominan fósiles porque son incapaces de replicarse para movilizarse en el genoma.

Los elementos transponibles como se insertan en regiones del genoma pueden interrumpir genes, pueden provocar reordenaciones cromosómicas, y son capaces de movilizar algunos exones de genes que están cercanos al sitio de inserción, por lo que el exón se inserta en un gen distinto, y por tanto, crean nuevos genes, y por último, pueden también influir en la expresión de genes que estén cerca del sitio de inserción.

**CONSULTAR DIAPOSITIVA 17:**

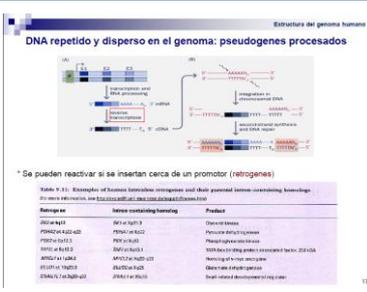
**DNA repetido y disperso en el genoma: elementos transponibles (II).**



**Pseudogenes procesados:** la transcriptasa reversa de los elementos transponibles LINE puede actuar sobre un transcrito de un gen ajeno a los retrotransposones, se genera una copia de ADN complementario (ADNc), y se introduce en otra región del genoma, por lo que se genera un pseudogen, ya que la copia no tiene promotor y por tanto no es activa, y está procesado porque no tiene ningún intrón. Si estos pseudogenes procesados se insertan cerca de un promotor de otro gen, se podrían reactivar dando lugar a lo que se llama Retrogenes.

**CONSULTAR DIAPOSITIVA 18:**

**DNA repetido y disperso en el genoma: pseudogenes procesados.**



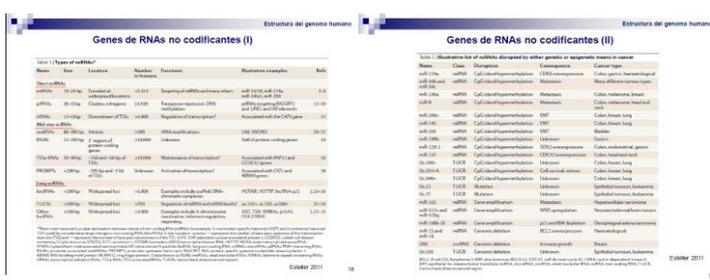
Retrogenes	Genes containing homolog	Product
ALU1-401	MYH11-101	Dimeric fibronectin
PER1-142-422	PER1-142-222	Proteinase 142-222
PER2-142-222	PER2-142-222	Proteinase 142-222
PER3-142-222	PER3-142-222	Proteinase 142-222
PER4-142-222	PER4-142-222	Proteinase 142-222
PER5-142-222	PER5-142-222	Proteinase 142-222
PER6-142-222	PER6-142-222	Proteinase 142-222
PER7-142-222	PER7-142-222	Proteinase 142-222
PER8-142-222	PER8-142-222	Proteinase 142-222
PER9-142-222	PER9-142-222	Proteinase 142-222
PER10-142-222	PER10-142-222	Proteinase 142-222
PER11-142-222	PER11-142-222	Proteinase 142-222
PER12-142-222	PER12-142-222	Proteinase 142-222
PER13-142-222	PER13-142-222	Proteinase 142-222
PER14-142-222	PER14-142-222	Proteinase 142-222
PER15-142-222	PER15-142-222	Proteinase 142-222
PER16-142-222	PER16-142-222	Proteinase 142-222
PER17-142-222	PER17-142-222	Proteinase 142-222
PER18-142-222	PER18-142-222	Proteinase 142-222
PER19-142-222	PER19-142-222	Proteinase 142-222
PER20-142-222	PER20-142-222	Proteinase 142-222

**Genes que especifican RNAs no codificantes**

Los más conocidos son los microRNAs y piwiRNAs de cadena corta (20-30 bp), pero también existen de cadena mediana (hasta 200 bp), y los de cadena larga (long non-coding RNA >200 bp). Cabe destacar que los microRNAs y piwiRNAs están implicados en la regulación de la expresión génica, y que se ha visto que una desregulación de estos genes está implicada en cáncer.

**CONSULTAR DIAPOSITIVAS 19 y 20:**

**Genes de RNAs no codificantes.**



Name	Class	Description	Consequence	Chromatin type
miR-101	miRNA	miR-101-1	miRNA	miRNA
miR-101	miRNA	miR-101-2	miRNA	miRNA
miR-101	miRNA	miR-101-3	miRNA	miRNA
miR-101	miRNA	miR-101-4	miRNA	miRNA
miR-101	miRNA	miR-101-5	miRNA	miRNA
miR-101	miRNA	miR-101-6	miRNA	miRNA
miR-101	miRNA	miR-101-7	miRNA	miRNA
miR-101	miRNA	miR-101-8	miRNA	miRNA
miR-101	miRNA	miR-101-9	miRNA	miRNA
miR-101	miRNA	miR-101-10	miRNA	miRNA
miR-101	miRNA	miR-101-11	miRNA	miRNA
miR-101	miRNA	miR-101-12	miRNA	miRNA
miR-101	miRNA	miR-101-13	miRNA	miRNA
miR-101	miRNA	miR-101-14	miRNA	miRNA
miR-101	miRNA	miR-101-15	miRNA	miRNA
miR-101	miRNA	miR-101-16	miRNA	miRNA
miR-101	miRNA	miR-101-17	miRNA	miRNA
miR-101	miRNA	miR-101-18	miRNA	miRNA
miR-101	miRNA	miR-101-19	miRNA	miRNA
miR-101	miRNA	miR-101-20	miRNA	miRNA
miR-101	miRNA	miR-101-21	miRNA	miRNA
miR-101	miRNA	miR-101-22	miRNA	miRNA
miR-101	miRNA	miR-101-23	miRNA	miRNA
miR-101	miRNA	miR-101-24	miRNA	miRNA
miR-101	miRNA	miR-101-25	miRNA	miRNA
miR-101	miRNA	miR-101-26	miRNA	miRNA
miR-101	miRNA	miR-101-27	miRNA	miRNA
miR-101	miRNA	miR-101-28	miRNA	miRNA
miR-101	miRNA	miR-101-29	miRNA	miRNA
miR-101	miRNA	miR-101-30	miRNA	miRNA
miR-101	miRNA	miR-101-31	miRNA	miRNA
miR-101	miRNA	miR-101-32	miRNA	miRNA
miR-101	miRNA	miR-101-33	miRNA	miRNA
miR-101	miRNA	miR-101-34	miRNA	miRNA
miR-101	miRNA	miR-101-35	miRNA	miRNA
miR-101	miRNA	miR-101-36	miRNA	miRNA
miR-101	miRNA	miR-101-37	miRNA	miRNA
miR-101	miRNA	miR-101-38	miRNA	miRNA
miR-101	miRNA	miR-101-39	miRNA	miRNA
miR-101	miRNA	miR-101-40	miRNA	miRNA
miR-101	miRNA	miR-101-41	miRNA	miRNA
miR-101	miRNA	miR-101-42	miRNA	miRNA
miR-101	miRNA	miR-101-43	miRNA	miRNA
miR-101	miRNA	miR-101-44	miRNA	miRNA
miR-101	miRNA	miR-101-45	miRNA	miRNA
miR-101	miRNA	miR-101-46	miRNA	miRNA
miR-101	miRNA	miR-101-47	miRNA	miRNA
miR-101	miRNA	miR-101-48	miRNA	miRNA
miR-101	miRNA	miR-101-49	miRNA	miRNA
miR-101	miRNA	miR-101-50	miRNA	miRNA
miR-101	miRNA	miR-101-51	miRNA	miRNA
miR-101	miRNA	miR-101-52	miRNA	miRNA
miR-101	miRNA	miR-101-53	miRNA	miRNA
miR-101	miRNA	miR-101-54	miRNA	miRNA
miR-101	miRNA	miR-101-55	miRNA	miRNA
miR-101	miRNA	miR-101-56	miRNA	miRNA
miR-101	miRNA	miR-101-57	miRNA	miRNA
miR-101	miRNA	miR-101-58	miRNA	miRNA
miR-101	miRNA	miR-101-59	miRNA	miRNA
miR-101	miRNA	miR-101-60	miRNA	miRNA
miR-101	miRNA	miR-101-61	miRNA	miRNA
miR-101	miRNA	miR-101-62	miRNA	miRNA
miR-101	miRNA	miR-101-63	miRNA	miRNA
miR-101	miRNA	miR-101-64	miRNA	miRNA
miR-101	miRNA	miR-101-65	miRNA	miRNA
miR-101	miRNA	miR-101-66	miRNA	miRNA
miR-101	miRNA	miR-101-67	miRNA	miRNA
miR-101	miRNA	miR-101-68	miRNA	miRNA
miR-101	miRNA	miR-101-69	miRNA	miRNA
miR-101	miRNA	miR-101-70	miRNA	miRNA
miR-101	miRNA	miR-101-71	miRNA	miRNA
miR-101	miRNA	miR-101-72	miRNA	miRNA
miR-101	miRNA	miR-101-73	miRNA	miRNA
miR-101	miRNA	miR-101-74	miRNA	miRNA
miR-101	miRNA	miR-101-75	miRNA	miRNA
miR-101	miRNA	miR-101-76	miRNA	miRNA
miR-101	miRNA	miR-101-77	miRNA	miRNA
miR-101	miRNA	miR-101-78	miRNA	miRNA
miR-101	miRNA	miR-101-79	miRNA	miRNA
miR-101	miRNA	miR-101-80	miRNA	miRNA
miR-101	miRNA	miR-101-81	miRNA	miRNA
miR-101	miRNA	miR-101-82	miRNA	miRNA
miR-101	miRNA	miR-101-83	miRNA	miRNA
miR-101	miRNA	miR-101-84	miRNA	miRNA
miR-101	miRNA	miR-101-85	miRNA	miRNA
miR-101	miRNA	miR-101-86	miRNA	miRNA
miR-101	miRNA	miR-101-87	miRNA	miRNA
miR-101	miRNA	miR-101-88	miRNA	miRNA
miR-101	miRNA	miR-101-89	miRNA	miRNA
miR-101	miRNA	miR-101-90	miRNA	miRNA
miR-101	miRNA	miR-101-91	miRNA	miRNA
miR-101	miRNA	miR-101-92	miRNA	miRNA
miR-101	miRNA	miR-101-93	miRNA	miRNA
miR-101	miRNA	miR-101-94	miRNA	miRNA
miR-101	miRNA	miR-101-95	miRNA	miRNA
miR-101	miRNA	miR-101-96	miRNA	miRNA
miR-101	miRNA	miR-101-97	miRNA	miRNA
miR-101	miRNA	miR-101-98	miRNA	miRNA
miR-101	miRNA	miR-101-99	miRNA	miRNA
miR-101	miRNA	miR-101-100	miRNA	miRNA

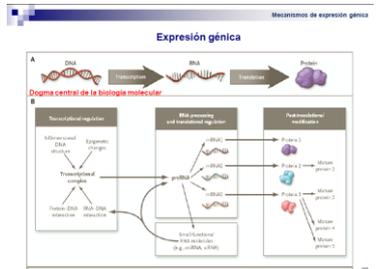
**2.-MECANISMOS DE EXPRESIÓN GÉNICA**

La regulación de la expresión génica es un proceso muy complejo en el cual se producen una serie de modificaciones a distintos niveles: a) a nivel de la transcripción, ya que tiene que modificarse la estructura tridimensional del ADN, para ello son necesarios algunos cambios epigenéticos, es decir, modificaciones covalentes de las proteínas y del ADN que forman parte de la cromatina. Una vez estas modificaciones se han producido, se tiene que dar una interacción correcta de factores de transcripción junto con la ARN polimerasa y el ADN, y en algunos casos también ocurren interacciones ADN-ARN; b) a nivel de procesamiento del ARN, ya que se genera un mensajero primario y no maduro, y c) a nivel de traducción,

ya que las formas de ARN mensajero generadas tienen que traducirse a proteínas. Además de que ciertas proteínas deben sufrir modificaciones post-traduccionales para ser funcionales.

**CONSULTAR DIAPOSITIVA 22:**

**Expresión génica.**

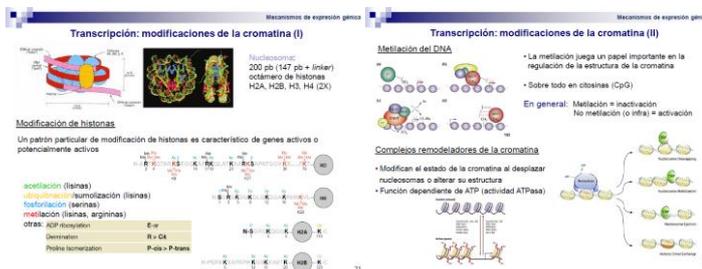


**2.1.-Modificaciones a nivel de la transcripción**

La transcripción comienza con una modificación de la cromatina, el ADN está unido a unas proteínas, las histonas, formando los nucleosomas. Ésta estructura es tan compacta que no permite la transcripción, y para que se puedan transcribir los genes, se necesita que se modifique esta estructura. Pero ¿cómo se modifica la cromatina en la transcripción? Por dos mecanismos, (1) el primero es mediante la modificación química de las histonas a través de una acetilación, fosforilación, metilación, etc en residuos de las histonas situados en las colas N-terminales. Esto produce la relajación en la interacción de proteína y ADN, lo que facilita la transcripción. Un patrón particular de modificación de histonas es característico de genes activos o potencialmente activos. (2) El segundo, mediante la modificación química del ADN mediante la metilación sobre las citosinas de la cromatina dentro del par CG (CpG). Cuando el ADN está muy metilado (en regiones promotoras), es inactivo y para que se produzca la activación de los genes se tiene que dar una inframetilación.

**CONSULTAR DIAPOSITIVAS 23 y 24:**

**Transcripción: modificaciones de la cromatina.**



A continuación actúan los Complejos Reguladores de la Cromatina, se ha relajado la estructura del nucleosoma, por lo que estos complejos proteicos mediante consumo de ATP modifican el estado de la cromatina desplazando a los nucleosomas o alterando su estructura. ¿Pero cómo? Desenrollando el ADN de los nucleosomas, movilizándolo a los nucleosomas o eliminándolos, introduciendo variantes de histonas que permiten que los nucleosomas estén relajados, etc. La importancia de todo esto es que secuencias que son

importantes para que los genes se transcriban estén accesibles para que la transcripción se pueda producir.

Estas secuencias importantes para la transcripción son las regiones Promotoras. El núcleo de promotor es una secuencia del ADN al cual se une la ADN polimerasa II en colaboración con factores de transcripción que la reclutan a la zona. También se necesitan Elementos Proximales al promotor, que son secuencias reguladoras que son reconocidas por proteínas que aumentan la eficacia del ensamblaje del complejo de transcripción. También son importantes los Elementos Distales en la transcripción, los Enhancers y Silenciadores. Son secuencias que influyen en la transcripción de los genes actuando a distancias muy amplias de dichos genes (del orden de Kb, hasta unos 50 kb de distancia hacia 3' o 5'). Los Enhancers o Amplificadores aumentan/promueven la expresión y la transcripción de los genes. Otras secuencias se denominan Silenciadores o Represores que actúan silenciando a los genes desde la distancia como en el caso anterior. Los Aislantes o Insulator son secuencias que sirven para restringir o limitar la acción de Enhancers y Silenciadores.

**CONSULTAR DIAPOSITIVA 25 y 26:**

**Transcripción: Promotores, Enhancers, Silenciadores y Aislantes.**

**Transcripción: promotores, enhancers, silenciadores y aislantes (I)**  
Inicio del promotor y elementos proximales

**Transcripción: promotores, enhancers, silenciadores y aislantes (II)**  
Elementos distales

**Enhancers o amplificadores:**

- Son elementos distales (10-50 kb)
- Carecen de actividad promotora
- Actúan la expresión de genes
- En 5', 3' y en ambas
- Pueden controlar grupos de genes
- Unión de proteínas específicas
- Los silenciadores o represores tienen las mismas características pero reprimen la expresión génica
- Los aislantes limitan la acción de enhancers y silenciadores

## 2.2.-Modificaciones a nivel de procesado de RNA

Al comienzo de la transcripción se producen unas modificaciones de los transcritos que se van generando en el extremo 5', 3' y en el interior de los genes.

La primera modificación es la adición de la caperuza en 5'. Es una estructura que es una guanina metilada que se añade al extremo 5' del transcrito. Tiene una función esencial para el ARN mensajero (ARNm) ya que permite que sea reconocido por el ribosoma y porque lo protege frente a nucleasas, además de aumentar su eficiencia de ser transportado al citoplasma para ser traducido. Esta adición se produce al principio cuando el transcrito tiene unos 25 nucleótidos de longitud y supone un check point de la transcripción, ya que si la caperuza no es correctamente añadida la elongación se interrumpe.

**CONSULTAR DIAPOSITIVA 27:**

**Procesado de los mRNAs: capping en 5'.**

**Procesado de los mRNAs: capping en 5'**

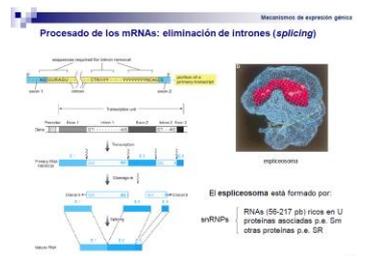
**Funciones de la caperuza en 5':**

- Reconocimiento por el ribosoma
- Protección contra nucleasas
- Eficiencia del transporte al citoplasma

La segunda modificación es la eliminación de los intrones, lo que se denomina *Splicing*. Los genes eucariotas se dividen en regiones codificantes (exones) y no codificantes (intrones). Los intrones tienen secuencias que han de ser eliminadas del transcrito primario para que el ARNm madure. Estas secuencias en 5', 3' e interior del intrón son reconocidas por un complejo llamado Espliceosoma, formado por ARNs muy cortos asociados a proteínas, concretamente, ribo-nucleoproteínas, y otras proteínas asociadas. Se elimina el intrón mediante el reconocimiento de esas secuencias, permitiendo la unión de los exones en el proceso de transcripción.

### CONSULTAR DIAPOSITIVA 28:

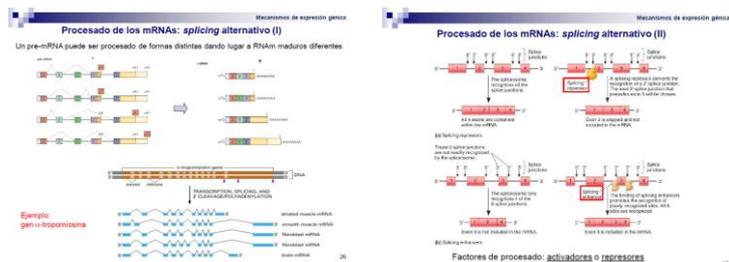
#### Procesado de los mRNAs: *eliminación de intrones (splicing)*.



Los intrones no siempre se eliminan de la misma forma, un ARNm primario puede dar lugar a varios y distintos ARNm maduros con diferentes combinaciones de exones, esto provoca la gran variedad de proteínas existentes y se conoce como *Splicing Alternativo*. El Splicing Alternativo se lleva a cabo principalmente por unas proteínas denominadas Factores de Procesado, que actúan tanto como Activadores como Represores y reconocen regiones exónicas e intrónicas. Un ejemplo se encuentra en la diapositiva 30.

### CONSULTAR DIAPOSITIVAS 29 y 30:

#### Procesado de los mRNAs: *splicing alternativo*.



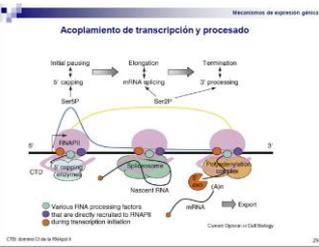
La tercera modificación es la poliadenilación en 3', que es la adición de una *cola polyA*, ya que se necesita dos secuencias fundamentales en el ARNm que se está transcribiendo, la señal de poliadenilación AAUAAA, y una secuencia rica de Uracilos y Guaninas. Las proteínas se unen a estas secuencias (complejo Factor de Especificidad de Corte y Poliadenilación -CPSF- y el factor de estimulación de corte -CstF-), esta unión produce un corte en una región intermedia entre las dos secuencias generándose dos transcritos. En el extremo 3' del fragmento del transcrito se añade la cola polyA por la acción de una polimerasa poli-A (PAP) para que se forme el transcrito maduro. El otro fragmento se degrada por la acción de las exonucleasas, y esto permite el desensamblaje del complejo de transcripción y que termine la misma. La poliadenilación está relacionada con el final

de la transcripción. La cola Poli-A añadida permite proteger al transcrito frente a la acción de las exonucleasa, regular la eficiencia de la traducción y además mejora la eficiencia del transporte del mRNA al citoplasma. No todos los mRNAs presentarán cola Poli-A, como por ejemplo los mRNAs que codifican para las proteínas Histonas.

Según lo comentado hasta ahora, queda evidente que existe un acoplamiento entre la transcripción y procesado de los transcritos para ser funcionales.

**CONSULTAR DIAPOSITIVAS 31 y 32:**

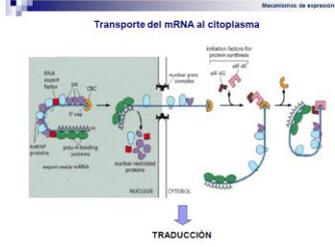
**Procesado de los mRNAs: poliadenilación en 3' y Acoplamiento de transcripción y procesado.**

En el transporte del transcrito (ARNm) al citoplasma, el transcrito generado está completamente cubierto por proteínas y entre ellas están los Factores de Exportación encargados de llevar del núcleo al citoplasma los transcritos para que sean traducidos. Una vez en el citoplasma, los ARNm son interceptados por unas proteínas esenciales para el inicio de la traducción (factores de iniciación).

**CONSULTAR DIAPOSITIVA 33:**

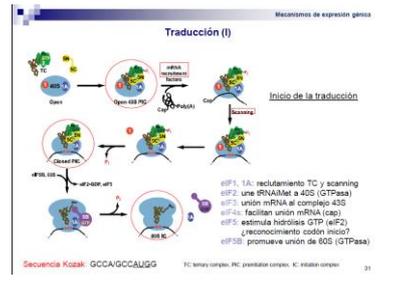
**Transporte del mRNA al citoplasma.**



**2.3.-Modificaciones a nivel de traducción**

Básicamente, el ribosoma recluta un ARNt que lleva la Metionina (1er aminoácido que se une a la proteína en eucariotas, además el codón de inicio es AUG que codifica la metionina). Este ARNt se une a la subunidad pequeña del ribosoma, y se recluta el ARNm con factores de iniciación que reconocen la caperuza 5' y comienza el proceso de SCANNING para la búsqueda del codón de inicio de la traducción. El codón de inicio está embebido en la secuencia Kozak GCCA/GCCAUGG y se genera el Complejo de preiniciación cerrado en el que tenemos el ARNt, la metionina colocada encima del codón de inicio y con la subunidad pequeña del ribosoma. Una vez se añade la subunidad grande del ribosoma comienza el proceso de traducción.

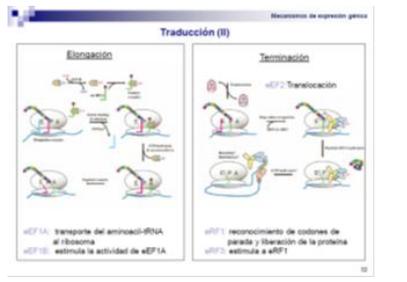
**CONSULTAR DIAPOSITIVA 34:  
Traducción (I).**



En la fase de elongación de la traducción se van incorporando nuevos ARNt cargados con los aminoácidos que van siendo reconocidos por los sucesivos codones después del de inicio. Los aminoácidos que se van incorporando se unen a los otros por enlaces peptídicos, el ribosoma se va trasladando al extremo 3' para ir reclutando aminoácidos a los demás codones. El ejemplo que ilustra los diferentes factores de elongación que participan se encuentra en la diapositiva 35.

Se termina la traducción cuando el siguiente codón que aparece dentro del ribosoma es un codón de parada, que no especifica ningún aminoácido, por tanto no va a ser reconocido por ningún ARNt sino por Factores de Terminación (-RF- *Release Factors*) de la traducción que permite el desensamblaje del ribosoma y la liberación del péptido que ya está sintetizado. Para profundizar en este tema podéis consultar libros generales de Genética y Biología Molecular.

**CONSULTAR DIAPOSITIVA 35:  
Traducción (II).**



Las proteínas recién sintetizadas pueden sufrir modificaciones postraduccionales para ser funcionales o para transportarlas a un compartimento celular donde ejerzan su función.

**CONSULTAR DIAPOSITIVA 36:  
Modificaciones postraduccionales de las proteínas.**

**Modificaciones postraduccionales de las proteínas**

Algunas proteínas parecen que ser modificadas para ser funcionales

Type of modification (group added)	Target amino acids	Comments
Hydroxylation (P-H)	Tyrosine, serine, threonine	Activated by iron & oxygen. Also is reversed by hydroxylases
Methylation (CH <sub>3</sub> )	Lysine	Activated by methylases and reversed by demethylases
Hydroxylation (OH)	Proline, lysine, aspartic acid	Hydroxyproline and L-hydroxylysine are particularly common in collagen
Acetylation (CH <sub>3</sub> CO)	Lysine	Activated by an acetylase and reversed by deacetylase
Carboxylation (CO <sub>2</sub> Me)	Glutamine	Activated by a carboxylase
Serine/threonine phosphorylation (PO <sub>3</sub> H)	Lysine	Activated by an acetylase and reversed by deacetylase
Hydroxylation (OH)	Asparagine, aspartic acid in the sequence Asn(Glu)X	Many glycoproteins, often embedded in membranes, have a sugar terminus acid chain (N-linked)
Hydroxylation (OH)	Serine, threonine, tyrosine, histidine	Many glycoproteins, often embedded in membranes, have a sugar terminus acid chain (O-linked)
Hydroxylation (OH)	Asparagine, Cysteine	Some enzymes, e.g. many antibodies, glycoproteins
Hydroxylation (OH), methyl (CH <sub>3</sub> )	Glutamine, Cysteine, Serine, Asparagine, Aspartic acid	Some enzymes, e.g. many antibodies, glycoproteins
Hydroxylation (OH), methyl (CH <sub>3</sub> )	Glutamine, Cysteine, Serine, Asparagine, Aspartic acid	Some enzymes, e.g. many antibodies, glycoproteins
Hydroxylation (OH), methyl (CH <sub>3</sub> )	Glutamine, Cysteine, Serine, Asparagine, Aspartic acid	Some enzymes, e.g. many antibodies, glycoproteins

### 3.4.-Mutaciones en regiones implicadas en la expresión de genes

Las mutaciones pueden ocurrir en regiones esenciales de los genes implicadas en su expresión, por ejemplo en el promotor (importante para la expresión), en la 5' UTR (importante para iniciar la traducción en el proceso de Scanning), en los exones (producirá una alteración en la capacidad codificante del gen), y en los intrones (se modificará su eliminación). Además estas mutaciones por tanto, pueden afectar a las distintas etapas de la expresión génica como se ilustra en la diapositiva 38.

**CONSULTAR DIAPOSITIVAS 37 y 38:**

**Mutaciones.**

**Las mutaciones pueden ocurrir en diferentes regiones de los genes**

**Posibles fallos**

**Las mutaciones pueden afectar diferentes etapas durante la expresión génica**

### 4.-Proyecto ENCODE

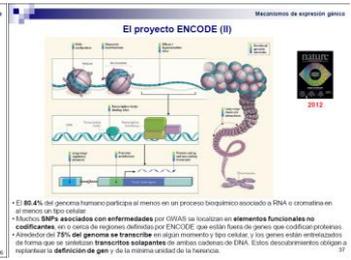
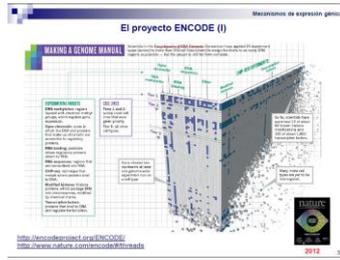
Este proyecto se inició en el año 2003 como continuación a los resultados obtenidos en el proyecto Genoma Humano donde se evidenciaba que sólo el 1% del genoma humano era codificante. Con el proyecto ENCODE se pretendía dilucidar qué función tenía el resto de genoma "basura". El objetivo era analizar 150 líneas celulares para identificar secuencias o elementos de ADN que fueran funcionales. Los experimentos de expresión de genes eran principalmente el mapeo de regiones en las que el ADN estaba metilado, regiones con modificaciones de la cromatina (histonas), regiones hipersensibles a la DNasa (desnudas de nucleosomas, que están siendo expresadas), factores de transcripción, elementos enhancers, promotores, y todos los transcritos que se podían generar en las líneas celulares, es decir, todos los ELEMENTOS FUNCIONALES DEL GENOMA.

Se extraen de este proyecto ambicioso, tres conclusiones principales:

- 1.- El 80.4% del genoma humano participa al menos en un proceso bioquímico asociado a ARN o a cromatina en al menos un tipo celular. En el 80% del genoma algún elemento funcional está presente, no todo es ADN basura.
- 2.- Muchos polimorfismos de nucleótido sencillo asociado a enfermedades por estudios a escala genómica que no estaban localizados en ningún gen concreto ni región codificante, los han mapeado en algunos de estos elementos funcionales del genoma.
- 3.- Alrededor del 75% del genoma se transcribe en algún momento y tipo celular, y los genes están entrelazados de forma que se sintetizan transcritos solapantes de ambas cadenas de DNA. Estos descubrimientos obligan a replantear la definición de gen y de la mínima unidad de la herencia.

**CONSULTAR DIAPOSITIVAS 39 y 40:**

**El proyecto ENCODE.**



**ANEXO I. Información adicional y preguntas.**

Podéis consultar las principales características del genoma humano en la diapositiva 41.

**CONSULTAR DIAPOSITIVA 41:**

**Principales características del genoma humano.**