

# PROCESOS GENÉTICOS DE LA SÍNTESIS DE PROTEÍNAS: LA TRANSCRIPCIÓN



Jacob



Monod



Cech



Pribnow

- [Fundamento Central de la Biología Molecular: "Dogma central de la Biología Molecular".](#)
- [Transcripción.](#)
- [Las ARN polimerasas y los diferentes tipos de ARN.](#)
- [Características de la transcripción: complementaridad, dirección y asimetría.](#)
- [Iniciación de la transcripción.](#)
- [Elongación de la transcripción.](#)
- [Terminación de la transcripción.](#)
- [Procesamiento del ARN.](#)
- [Procesamiento alternativo.](#)
- [Autoprocesamiento.](#)
- [Edición o corrección del ARN.](#)

## FUNDAMENTO CENTRAL DE LA BIOLOGÍA MOLECULAR

La aceptación de la hipótesis de la secuencia propuesta por Crick (1958) significaba por un lado la existencia de un código o clave que permitiera pasar de un lenguaje escrito con cuatro letras (las cuatro bases nitrogenadas) en el ADN a un lenguaje escrito con 20 letras (los aminoácidos) en las proteínas. Por otro lado, significaba la existencia de una serie de procesos que realizaran en la célula el trabajo de pasar de una estructura química como la de los ácidos nucleicos a una estructura química como la de las proteínas. Ya hemos visto en el capítulo anterior las características del código o clave genética y su desciframiento, ahora vamos a estudiar los procesos genéticos de la síntesis de proteínas.

Teniendo en cuenta que en las células eucariontes los cromosomas (material genético organizado) se encuentran en el núcleo y que la información que contienen los cromosomas (los genes) se expresa en el citoplasma, se supuso que tendría que haber alguna molécula intermediaria en el flujo de la información desde el ADN a las proteínas. Jacob y Monod en 1961 propusieron la Hipótesis del mensajero: "debe existir una molécula que transporte la información fielmente desde el ADN hasta las proteínas". En ese mismo año Brenner y colaboradores (1961) demostraron la existencia del este intermediario que resultó ser una

molécula de ácido ribonucleico que se denominó ARN mensajero (ARN-m). Posteriormente, el ARN-m tenía que ser traducido a proteína.



Jacob



Monod



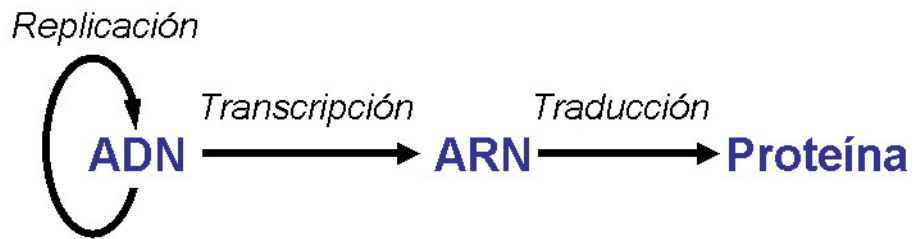
Brenner

Basándose en estos trabajos Crick (1970) alumbró la idea de un sistema fundamental de mantenimiento y flujo de la información genética en los organismos vivos que denominó *Dogma Central de la Biología Molecular*. De manera que la información genética contenida en el ADN se mantiene mediante su capacidad de *replicación*. La información contenida en el ADN se expresa dando lugar a proteínas, mediante los procesos de *transcripción*, paso por el que la información se transfiere a una molécula de ARN mensajero (ARN-m) y, mediante el proceso de la *traducción* el mensaje transportado por el ARN-m se traduce a proteína.



F. Crick

### Propuesta inicial de Crick (1970)



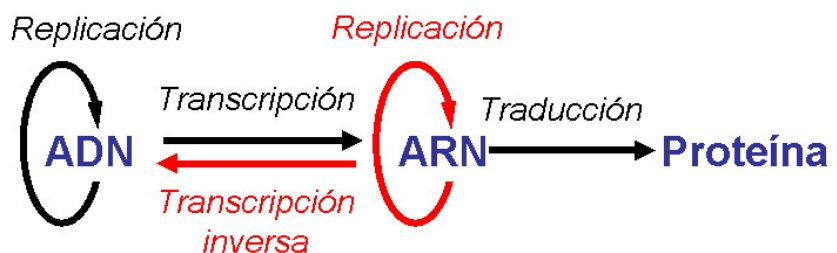
"Dogma Central de la Biología Molecular"

Este esquema central de flujo de la información pronto fue modificado, ya que en algunos virus cuyo material hereditario es ARN, la información se conserva o mantiene mediante replicación del ARN. Además, también se comprobó que la información no va siempre del ADN hacia el ARN (ADN→ARN), en algunos casos la información puede fluir del ARN hacia el ADN (ARN→ADN), es decir sintetizar ADN tomando como molde ARN, teniendo lugar el fenómeno de la *transcripción inversa*. H. Temin recibió el Premio Nobel en 1975 por sus descubrimientos en relación con la interacción entre los virus tumorales y el material genético en la célula, en particular por el descubrimiento de la transcripción inversa en virus ARN-ADN.



Temin

### Modificaciones posteriores



Modificaciones del "Dogma Central de la Biología Molecular"

Por tal causa, en una ciencia como la Genética no debería emplearse la palabra "dogma" ya que como acabamos de ver en muy poco tiempo el fundamento central de la Biología Molecular propuesto por Crick tuvo que ser modificado.

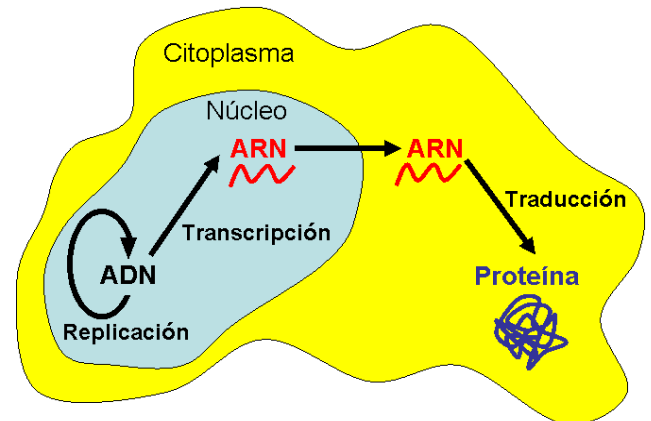
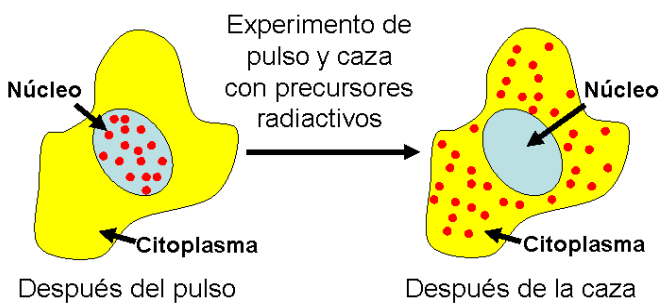
## TRANSCRIPCIÓN

La transcripción consiste en la síntesis de ARN tomando como molde ADN y significa el paso de la información contenida en el ADN hacia el ARN. La transferencia de la información del ADN hacia el ARN se realiza siguiendo las reglas de complementaridad de las bases nitrogenadas y es semejante al proceso de transcripción de textos, motivo por el que ha recibido este nombre. El ARN producto de la transcripción recibe el nombre de transcrito.

En las bacterias la transcripción y la traducción tienen lugar en el citoplasma bacteriano y al mismo tiempo, son simultáneas. Sin embargo, en eucariontes la transcripción tiene lugar en el núcleo y la traducción en el citoplasma.

Mediante experimentos de pulso y caza, se suministran precursores radiactivos que marcan específicamente el ARN (uridina tritiada) a las células durante un breve período de tiempo (pulso). Una vez que las células han incorporado la uridina tritiada se transfieren a un medio con precursores sin marcar (caza). De este forma es posible seguir el destino del ARN marcado durante el pulso, ya que la síntesis del nuevo ARN se produce con precursores sin marcar (uridina normal). Las muestras de células tomadas después de la caza, mostraban marcaje en el núcleo, indicando que ARN se sintetiza allí, sin embargo, las muestras de células tomadas después de la caza mostraban el marcaje radiactivo en el citoplasma. Por tanto, parece que el ARN se sintetiza en el núcleo y se transporta posteriormente al citoplasma.

Flujo de la información genética en eucariontes



Experimento de pulso y caza con precursores radiactivos (uridina tritiada)

Eucariontes: transcripción en el núcleo y traducción en el citoplasma

Volkin y Astrachan (1957) demostraron que después de la infección de *E. coli* por el fago T2 aumenta la síntesis de ARN y que este ARN inducido por la infección de T2 tiene una vida media muy corta. Infectaron bacterias con el fago T2 en un medio que contenía uridina tritiada, base nitrogenada específica del ARN (pulso), y después las pasaron a un medio con uridina normal no marcada (caza). El ARN recuperado después del pulso estaba marcado, pero el obtenido después de la caza no lo estaba, lo que indicaba que el ARN tenía una vida media muy corta. Por último cuando se comparó el porcentaje de los diferentes tipos de nucleótidos del ARN sintetizado inmediatamente después de la infección por T2 y del ADN del fago, se observó que era muy parecido, sugiriendo este resultado que el ARN sintetizado después de la infección por T2 procedía del ADN del fago.

# LAS ARN POLIMERASAS Y LOS DISTINTOS TIPOS DE ARN

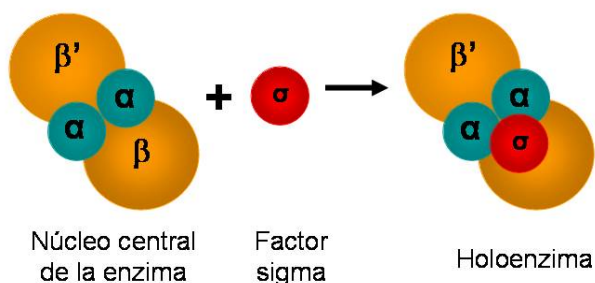
El ARN suele tener una sola hélice o cadena de nucleótidos pudiendo formar una amplia gama estructuras tridimensionales diferentes. El ARN contiene como azúcar a la ribosa y las bases nitrogenadas mayoritarias que posee son adenina (A), guanina (G), citosina (C) y uracilo (U). Por tanto, el uracilo (U) es característico del ARN. También se han encontrado bases poco frecuentes que forman parte del ARN, como son: la pseudouridina ( $\psi$ ), metilguanosa (mG), dimetilguanosa ( $m_2G$ ), metilinosina (mI) y dihidrouridina (DHU,  $UH_2$ ).

En las células existen diferentes tipos de ARN. Por un lado están los *ARN funcionales*, o ARN que tienen una función o actividad en la célula y que no se traducen a proteína. Dentro de esta categoría están el *ARN ribosómico* (ARN-r) que forma parte de los ribosomas que intervienen en la traducción, los *ARN transferentes* (ARN-t) cuya función es transportar a los aminoácidos durante el proceso de traducción, los *ARN nucleares pequeños* (ARN-np) que interactúan con proteínas formando los complejos de ribonucleoproteínas necesarios para el procesamiento de los transcritos en el núcleo y los *ARN citoplásmicos pequeños* (ARNcp) que intervienen en el transporte de los polipéptidos en las células eucarióticas.

Por otro lado, están los *ARN informativos* que son los que se van a traducir a proteínas. Estos *ARN informativos* son los *ARN mensajeros* (ARN-m). En bacterias el transcrito primario, tal y como se sintetiza ya es el ARN-m maduro que se traduce a proteínas. En eucariontes, el ARN recién transcrito se denomina *ARN heterogéneo nuclear* (ARN-hn) y es un pre-ARN-m que es procesado antes de convertirse en el ARN-m maduro que posteriormente se traducirá a proteína.

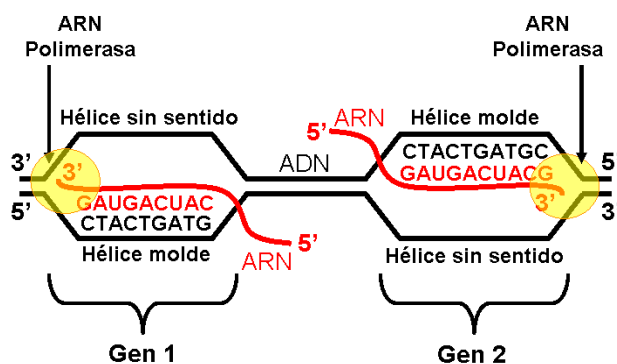
En bacterias, existe solamente una ARN polimerasa que es capaz de sintetizar todos los tipos de ARN, el ribosómico, el transferente y los mensajeros. La ARN polimerasa de *E. coli* está formada por varios polipéptidos, el núcleo central del enzima tiene dos cadenas de tipo  $\alpha$  (peso molecular de 40.000), una  $\beta$  (peso molecular de 155.000) otra  $\beta'$  (peso molecular de 165.000). Además, la enzima completa u holoenzima tiene la subunidad  $\sigma$  o factor  $\sigma$  (peso molecular de 95.000) que es necesario para iniciar la transcripción. La subunidad  $\sigma$  una vez iniciada la transcripción se libera y el núcleo central prosigue con la elongación del ARN.

## ARN polimerasa o Transcriptasa



ARN Polimerasa de *E. coli*: subunidades

## Asimetría de la Transcripción



ARN-m policistrónico: operón lactosa

En bacterias, con mucha frecuencia, los ARN-m son poligénicos o policistrónicos, de manera que un solo ARN-m contiene información para la síntesis de varios polipéptidos distintos. Habitualmente se trata de genes que comparten un sistema común que controla su expresión (operón).

Sin embargo, en eucariontes hay diferentes polimerasas encargadas de sintetizar distintos tipos

de ARN.

- **La ARN polimerasa I** sintetiza los precursores del ARN ribosómico (ARN-r).
- **La ARN polimerasa II** produce ARN heterogéneo nuclear (ARN-hn) que tras el procesamiento da lugar a los ARN mensajeros (ARN-m) que se traducen a proteínas.
- **La ARN polimerasa III** transcribe los precursores de los ARN transferentes (ARN-t), los ARN nucleares y citoplásmicos de pequeño tamaño y los genes para el ARN 5S que forma parte de la subunidad grande de los ribosomas.

Las ARN polimerasas eucarióticas son bastante más complejas que las de *E. coli*, poseen subunidades semejantes o correspondientes a las de *E. coli* y otras que no lo son.

Las ARN polimerasas o transcryptasas, a diferencia de lo que ocurre con las ADN polimerasas, carecen de función "*correctora de pruebas*". Esta diferencia se debe en primer lugar a que los transcritos son cortos y la probabilidad de que uno de los ARN posea una alteración es baja, y en segundo lugar a que la vida media de los ARN es corta y pronto se vuelve a sintetizar otro ARN nuevo. Por consiguiente el que exista un ARN con una alteración no es grave ya que durará poco y será remplazado pronto por otro nuevo sin la alteración. Sin embargo, un error en la replicación del ADN puede transmitirse a todas las células que deriven por división de la célula afectada.

En eucariontes los ARN-m son monogénicos o monocistrónicos, de manera que un ARN-m contiene información para sintetizar un solo polipéptido.



## CARACTERÍSTICAS DE LA TRANSCRIPCIÓN: COMPLEMENTARIDAD, DIRECCIÓN Y ASIMETRÍA

Seguidamente vamos a ver las principales características de la transcripción.

- **Complementaridad:** El parecido entre el ADN y el ARN sugiere que la transcripción probablemente esta basada en las reglas de complementaridad de las bases nitrogenadas al igual que la replicación del ADN. De manera que la ARN polimerasa o enzima encargada de llevar a cabo la transcripción toma como molde el ADN para sintetizar ARN y sigue las reglas de complementaridad, la A del ADN empareja con U del ARN, la G con C, la C con G y la T con A. Existen experimentos que demuestran que la proporción  $(A+U)/(G+C)$  del ARN es similar a la proporción  $(A+T)/(G+C)$  del ADN. En la siguiente tabla se muestra la proporción de nucleótidos de diferentes ADNs y de sus correspondientes transcritos:

Origen del ADN	$(A+T)/(G+C)$ del ADN	$(A+U)/(G+C)$ del ARN
Fago T2	1.84	1.86
Vaca	1.35	1.40
<i>Micrococcus</i> (bacteria)	0.39	0.49

En la siguiente figura, se ilustra que la secuencia de bases en el ARN es complementaria a la secuencia de la hélice codificadora e igual a la secuencia de la hélice estabilizadora, cambiando T por U.



# COMPLEMENTARIDAD

Hélice estabilizadora **CTGCCATTGTCAGACATGTATACCCCGTAC**

Hélice codificadora **GACGGTAACAGTCTGTACATATGGGGCATG**

↓ *Transcripción*

ARN **CUGCCAUUGUCAGACAUGUAUACCCCGUAC**

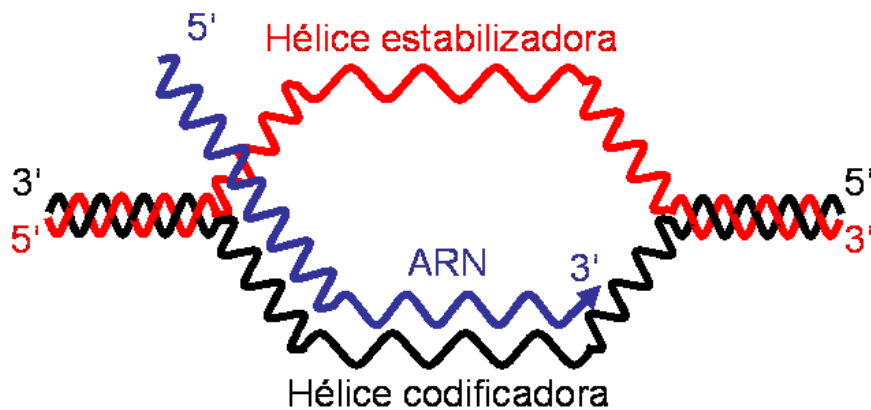
La secuencia de bases del ARN es complementaria a la secuencia de la cadena molde o codificadora. La secuencia de bases de la hélice estabilizadora es la misma que la del ARN, cambiando T por U.

La secuencia de bases del ARN es complementaria a la secuencia de la hélice codificadora

Otra manera de comprobar que existe complementaridad entre el ADN y el ARN es realizar experimentos de hibridación, de manera que el ADN se desnaturaliza y se mezcla con el ARN que se sintetiza en su presencia, cuando se lleva a cabo la renaturalización (mediante un enfriamiento lento), se producen híbridos ADN-ARN. Las moléculas híbridas ADN-ARN se pueden distinguir mediante centrifugación en gradiente de CsCl ya que presentan una densidad diferente a la de las dobles hélices ADN-ADN. La aparición de moléculas híbridas ADN-ARN solo es posible si existen segmentos largos de complementarios, por tanto estos resultados indican que el transcrito es complementario del ADN parental.

- **La dirección** en la que las ARN polimerasas sintetizan ARN es siempre 5'P→3'OH, es decir el ARN producto de la transcripción crece solamente en esta dirección. Recuerde que la dirección en la que las ADN polimerasas sintetizan ADN es también la misma 5'P→3'OH.
- **Asimetría de la transcripción:** la asimetría de la transcripción significa que solamente se transcribe para cada gen una de las dos hélices de ADN, la hélice que se toma como molde para producir el ADN se la denomina *hélice codificadora* o *hélice con sentido* y la otra hélice de ADN, la que no se transcribe, se la denomina *hélice estabilizadora* o *hélice sin sentido*.

## Dirección y Asimetría de la Transcripción



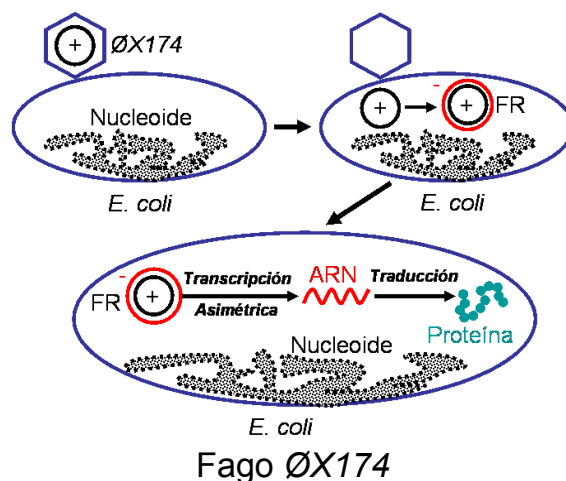
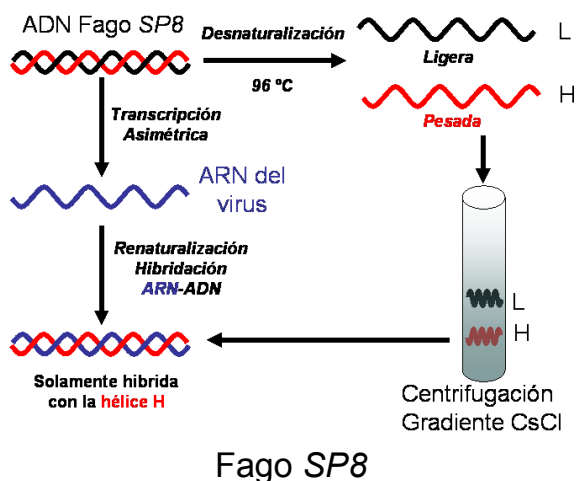
Hélice estabilizadora = Hélice sin sentido

Hélice codificadora = Hélice con sentido

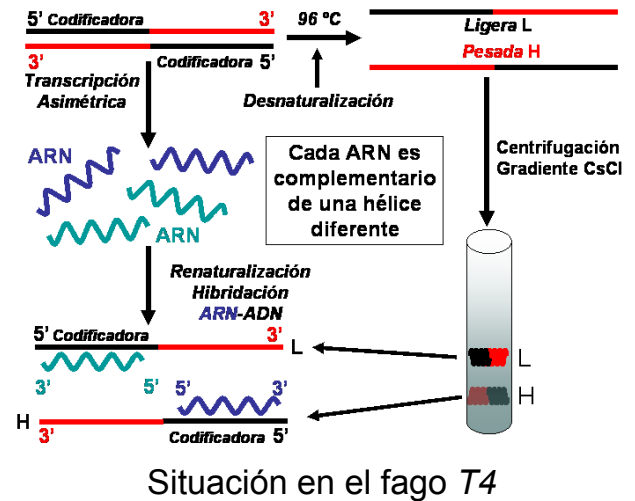
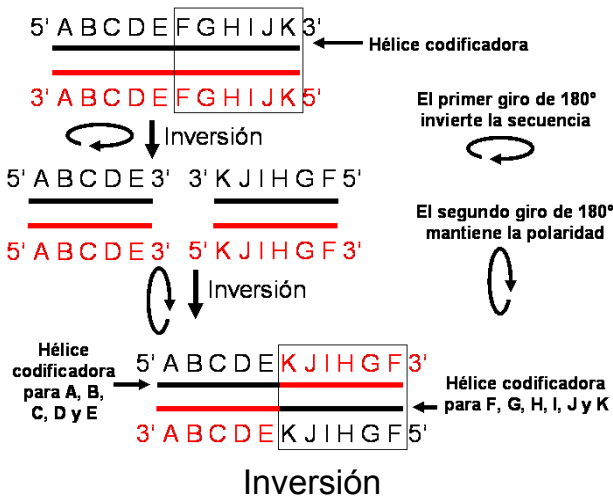
### Dirección y asimetría de la transcripción

Por ejemplo, Marmur y Greenspan (1963) comprobaron que en el fago SP8 que infecta a *Bacillus subtilis*, las dos hélices de su ADN se pueden separar mediante desnaturalización, enfriamiento rápido para mantenerlas separadas y posterior centrifugación en gradiente de densidad de CsCl. Por consiguiente, este virus posee una hélice ligera (L) y otra hélice pesada (H) con distinta relación de purinas/pirimidinas. Si se renaturaliza (enfriamiento lento) el ARN sintetizado después de la infección de *B. subtilis* por el virus y se hibrida por separado con el ADN de la hélice L y con el ADN de la hélice H, se puede comprobar que el ARN solamente hibrida con el ADN de la hélice H. Por tanto, en el virus SP8 todos los genes se transcriben utilizando como molde (hélice codificadora) la hélice H. La transcripción es asimétrica.

Hayashi y col. (1963) demostraron que el fago  $\phi X174$  cuyo material hereditario es ADN circular de una hélice, denominada hélice (+), cuando infecta a *E. coli*, pasa por una forma replicativa de doble hélice, la nueva hélice de ADN sintetizada después de la infección se la llama hélice (-). La hélice (-) sirve de molde durante las últimas etapas de la infección para producir mediante una *replicación asimétrica* nuevas hélices (+) del ADN del virus que se introducirán en el interior de las cápsidas ya formadas. A su vez, comprobaron que el ARN que se sintetiza inmediatamente después de infectar a *E. coli*, solamente hibrida con el ADN de la hélice (-). Por consiguiente, todos los genes de este virus se transcriben utilizando como molde (hélice codificadora) la hélice (-). La transcripción es asimétrica.

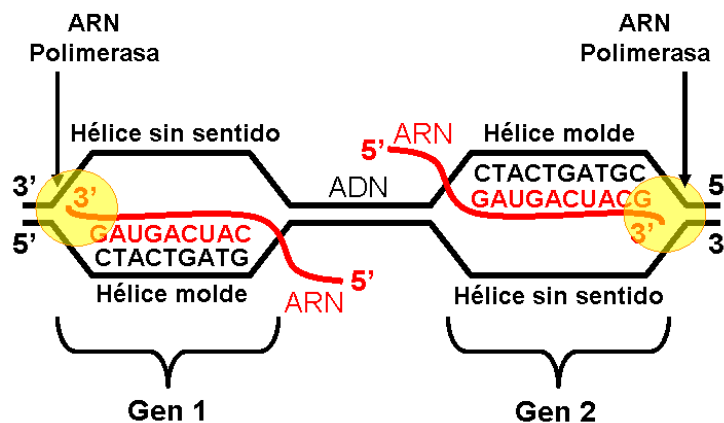


En otros fagos, como el bacteriofago *T4*, se ha comprobado que parte de los genes se transcriben a partir de una hélice y otro grupo de genes utiliza como codificadora a la otra cadena. Estos resultados se pueden explicar suponiendo que inicialmente todos los genes del fago se transcribían utilizando solamente una de las hélices y que posteriormente se produjo una inversión (se invierte el orden de un grupo de genes mediante dos giros perpendiculares de  $180^\circ$ , un giro que invierte la secuencia y otro giro perpendicular al anterior para mantener la polaridad de la hebras del ADN). Como consecuencia, después de la inversión el grupo de genes del segmento invertido utiliza como codificadora la otra hélice.



En los organismos eucariontes, para cada gen solamente se transcribe una hebra de ADN (la codificadora), de forma que genes distintos del mismo cromosoma pueden utilizar como codificadora una hélice diferente a la de otros genes.

### Asimetría de la Transcripción



### INICIACIÓN DE LA TRANSCRIPCIÓN

El inicio de la transcripción necesita que el factor  $\sigma$  esté unido al núcleo central de la ARN polimerasa. Existen unas secuencias de ADN específicas y necesarias para que la holoenzima reconozca el lugar de comienzo de la transcripción, dichas secuencias específicas se denominan secuencias promotoras. Pribnow (1975) aisló y secuenció las regiones de ADN que quedaban protegidas por la ARN polimerasa después de someterlas a digestión con ADNasa pancreática. Siguiendo este método se secuenciaron los primeros promotores de fagos como el T7 y I, y del promotor del operón lactosa. Pribnow observó una secuencia de unos 7 pares de



bases que se encontraba 10 bases antes de la primera que se transcribe y que aparecía en la mayoría de los fragmentos protegidos por la ARN polimerasa. Esta secuencia se la denominó *Caja de Pribnow* y es: 5' TATPuATPu 3'.

## SECUENCIAS CONSENSO PROMOTORAS

### BACTERIAS



### EUCARIONTES



Secuencias promotoras consenso de *E. coli* y eucariontes

Pribnow

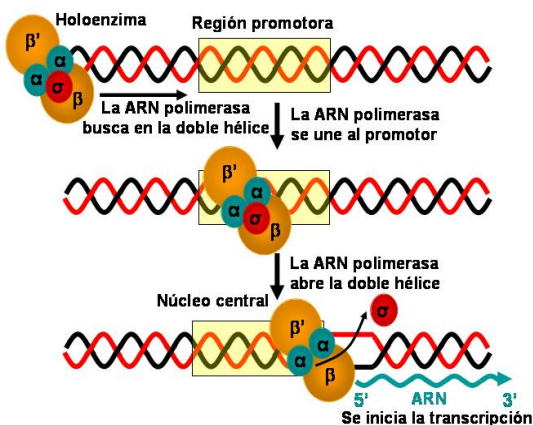
Cuando se analizan las secuencias de las regiones anteriores a la primera base que se transcribe aparecen dos regiones cortas que muestran un parecido o similitud bastante grande, la primera región se encuentra 10 bases antes de la primera que se transcribe (*Caja de Pribnow*) y la segunda se localiza 35 bases antes. El estudio de muchas regiones promotoras de diferentes genes permite determinar las secuencias que aparecen con mayor frecuencia y establecer lo que se denomina las *secuencias promotoras consenso o ideales*.

Los resultados obtenidos en *E. coli* después de analizar 100 regiones promotoras son los siguientes:

Secuencias consenso promotoras en <i>E. coli</i>	
5' .....	T <sub>82</sub> T <sub>84</sub> G <sub>78</sub> A <sub>65</sub> C <sub>54</sub> a <sub>45</sub> ..... T <sub>80</sub> A <sub>95</sub> t <sub>45</sub> A <sub>60</sub> a <sub>50</sub> T <sub>96</sub> ..... CAT ..... 3'
5' .....	-35 ..... -10 ..... 1ª bases transcrita

El subíndice indica el porcentaje de aparición de la base más frecuente de esa posición. Si la letra es mayúscula indica que el porcentaje de aparición es mayor del 54% y si es minúscula que es inferior al 54%.

El factor sigma es el responsable de que la ARN polimerasa reconozca estas secuencias y se una a ellas. La holoenzima recorre el ADN hasta encontrar las regiones -10 y -35 (secuencias promotoras) y se une a ellas, posteriormente comienza a separar las dos hélices por la región -10 (rica en pares AT). Cuando la holoenzima ha reconocido y separado las dos hélices se forma lo que se denomina "*complejo abierto con el promotor*".



Iniciación de la Transcripción

Secuencias de diferentes promotores de *E. coli*

tyr tRNA	TCTCAACGTAACACTTTTACAGGCGGCG • CGTCATTTGATATGATGC • GCCCGCTTCCCGATAAGGG
rrn D1	GATCAAAAAAATACTTGTGCAAAAAA • TTGGGATCCCTATAATGCGCCTCCGTTGAGACGACAACG
rrn X1	ATGCATTTTTCCGCTTGTCTCCTGA • GCCGACTCCCTATAATGCGCCTCCATCGACACGGCGGAT
rrn (DXE) <sub>2</sub>	CCTGAAATTCAGGGTTGACTCTGAAA • GAGGAAAGCGTAATATAC • GCCACCTCGCGACAGTGAGC
rrn E1	CTGCAATTTTTCTATTGCGGCGCTGCG • GAGAACTCCCTATAATGCGCCTCCATCGACACGGCGGAT
rrn A1	TTTTAAATTTCTCTTGTCAAGGCGGG • AATAACTCCCTATAATGCGCCACCCTGACACGGGAACAA
rrn A2	GCAAAAAATAATGCTTGACTCTGTAG • CGGGAAGGCGTATATGC • ACACCCCGCGCCGTGAGAA
λ PR	TAACACCGTGCGTGTGACTATTTTA • CCTCTGGCGGTGATAATGG • TTGCATGTAATAAGGAGGT
T7 A3	TATCTTGCGGCTGTGACATAAATA • CCAC TGGCGGTGATACTGA • GCACATCAGCAGGACGAC
T7 A1	GTGAACAAAAACGGTTGACAACATGA • AGTAAACACGGTACGATGT • ACCACATGAAACGACAGTGA
T7 A2	TATCAAAAAGAGTATGACTTAAAGT • CTAACCTATAGGATCTTA • CAGCCATCGAGAGGGACCG
fd VIII	ACGAAAAACAGGTATGACAACATGAAGTAACATGCAGTAAGATAC • AAATCGTAGGTAACACTAG
	GATACAAAATCTCCGTTGTACTTTGTT • TCGCGCTTGGTATAATCG • CTGGGGCTCAAAGATGAGT

Secuencia consenso de los promotores

-35 region      15-17 bp      -10 region

TTGACAT      TATAAT

Secuencias de regiones promotoras de *E. coli*

La actividad de los promotores puede modificarse por la presencia de otras *secuencias estimuladoras* o "*enhancers*" que aumentan la tasa de transcripción o por *secuencias atenuadoras* que disminuyen la tasa de transcripción. Estas secuencias estimuladoras y/o atenuadoras suelen estar cerca de las promotoras pero no a una distancia fija de estas y pueden ejercer su función sobre varios promotores diferentes.

En eucariontes la iniciación de la transcripción es más compleja que la de *E. coli* y depende de la presencia de *Factores proteicos de transcripción* (TF). La mayoría de las secuencias promotoras eucarióticas se encuentran antes (o "aguas arriba") de la primera base transcrita, aunque algunos promotores de la ARN Pol III se localizan después de la primera base transcrita ("aguas abajo"). Los TF interaccionan con las ARN polimerasas para iniciar la transcripción, los promotores de la ARN pol II son los que muestran mayor variación en su secuencia, motivo por el que hay muchos TF diferentes que interaccionan con la ARN pol II para iniciar la transcripción. Existen tres tipos de Factores proteicos de transcripción (TF):

- **Factores basales:** necesarios para el comienzo de la síntesis de ARN.
- **Factores que actúan aguas arriba:** reconocen secuencias consenso cortas situadas antes del punto de iniciación cuya actividad no está regulada, de manera que se unen a cualquier promotor que contenga estas secuencias.
- **Factores inducibles:** reconocen secuencias consenso cortas o *elementos de respuesta* situadas antes del punto de iniciación cuya actividad si está regulada. Estos factores inducibles son sintetizados o activados en diferentes tejidos y momentos del desarrollo controlando los patrones de transcripción en diferentes lugares (tejidos) y momentos del desarrollo en los organismos.

Los promotores que tienen secuencias que solamente son reconocidas por factores basales y factores que actúan aguas arriba, corresponden a genes que se expresan de manera constitutiva (se expresan siempre) en todos los tejidos. Se trata de los genes encargados del metabolismo básico de la célula involucrados en el buen funcionamiento celular. A estos genes se les ha denominado genes ("*Housekeeping genes*"), es decir, los genes que guardan o mantienen la casa.

Los promotores de la ARN pol I (transcribe los precursores del ARN ribosómico) tienen dos secuencias ricas en pares GC y que muestran una homología del 85%, una secuencia de 65 pb denominada núcleo que incluye la primera base que se transcribe (de -45 a +20) y una segunda secuencia unos 70 pb antes (aguas arriba) que actúa como *elemento de control* y que va desde la posición -107 a la -180.

Los promotores de la ARN pol III (transcriben ARN-t, ARN-5S y ARN de tamaño pequeño) contienen secuencias de reconocimiento situadas después de la primera base que se transcribe (aguas abajo), a este tipo de promotores se les denomina *promotores internos* (están dentro del gen). Se han encontrado en los promotores de ARN-t y ARN-5S pero no se han descrito en ARN de tamaño pequeño.

Los promotores de la ARN pol II son más variables que los de las ARN pol I y pol III pero a pesar de esta variación tienen una estructura general bastante conservada, de manera que se han reconocido al menos tres secuencias canónicas o consenso situadas en las regiones -25 (*caja TATA*, con la secuencia **TATAAAA**), -75 (*caja CAAT* con la secuencia **GGCCAATCT**) y -90 (*caja GC* con la secuencia **GGGCGG**). La *caja TATA* es equivalente a la *caja de Pribnow* de los promotores de bacterias y se la denomina *caja de Hogness*.

El punto de iniciación de la transcripción en eucariontes suele ser una adenina flanqueada por

pirimidinas, aceptándose la existencia de una región iniciadora (Inr) con el siguiente tipo de secuencia: Pir-Pir-C-A-Pir-Pir-Pir-Pir-Pir.

Resumen de las <i>secuencias consenso</i> promotoras en eucariontes									
5'	....	GGGCGG	.....	GGCCAATCT	.....	TATAAAA	....	Pir-Pir-C-A-Pir-Pir-Pir-Pir-Pir	.....
3'									
5'	.....	-90	.....	-75	.....	-25	.....		1ª bases transcrita



## ELONGACIÓN DE LA TRANSCRIPCIÓN

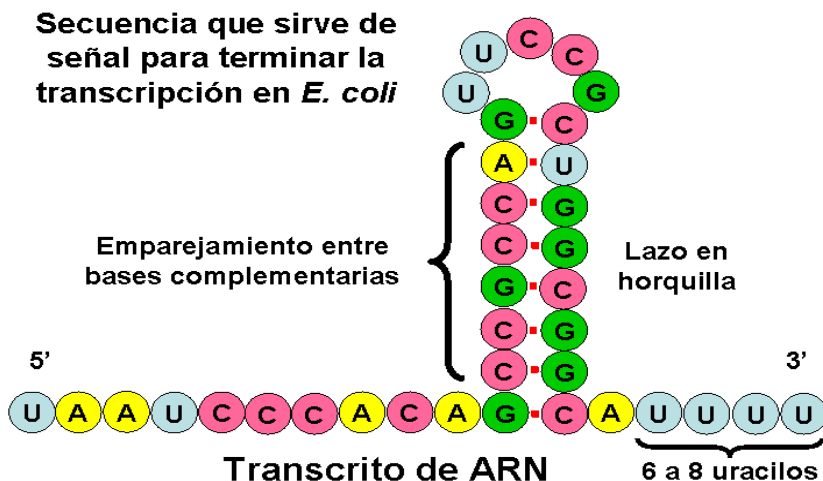
La elongación de la transcripción no necesita del factor  $\sigma$ , una vez iniciada la transcripción el factor sigma se suelta y el núcleo central de la ARN polimerasa comienza a sintetizar el ARN en la dirección 5' P  $\rightarrow$  3'OH a partir de ribonucleósidos trifosfato libres que sirven de sustrato al enzima. Parece ser que se van produciendo desenrollamientos parciales del ADN de la región que se está transcribiendo y que la velocidad de transcripción en *E. coli* a 37° C es de 2500 ribonucleótidos por minuto (aproximadamente 42 ribonucleótidos por segundo).



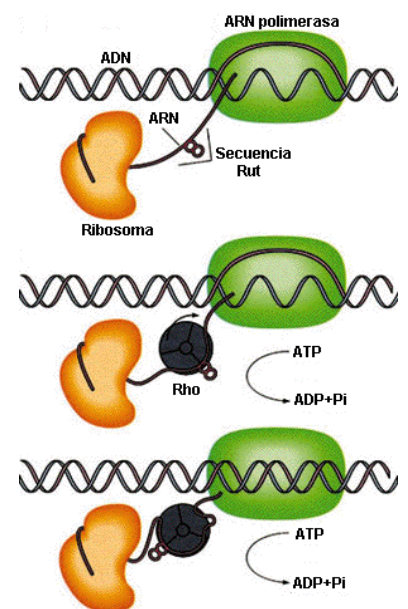
## TERMINACIÓN DE LA TRANSCRIPCIÓN

La terminación de la transcripción en *E. coli* puede tener lugar por dos mecanismos distintos:

- Existencia de unas secuencias terminadoras de unos 40 pares de bases que contienen una región rica en pares GC seguida por una serie de 6 o más adeninas (A). Cuando esta región del ADN se transcribe da lugar en el ARN a una secuencia rica en pares GC seguida de 6 o más uracilos (U), la región rica en pares GC forma una estructura en forma de horquilla (por autoapareamiento). Este lazo en horquilla seguido de uracilos actúa como señal para la separación de la ARN polimerasa del ADN y terminación de la transcripción.



Secuencias terminadoras ricas en pares G-C que forman un lazo en horquilla, seguidas de 6 a 8 uracilos.



Terminación dependiente de rho

- Terminación dependiente de la actuación del factor proteico rho ( $\rho$ ) que es un hexámero

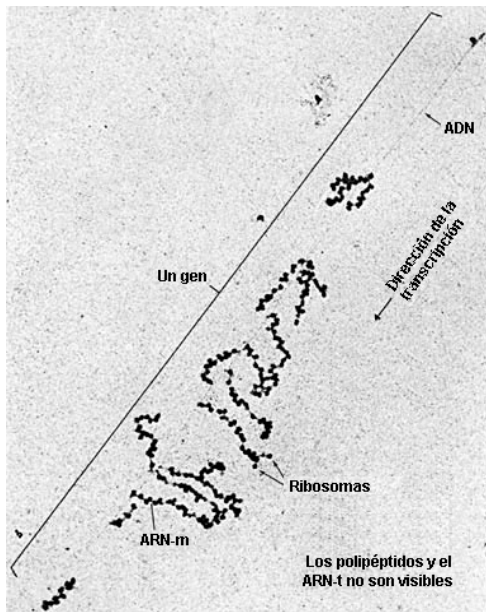
formado por 6 subunidades. Los ARN que presentan señales de terminación dependientes de rho ( $\rho$ ) no suelen presentar el lazo en horquilla seguido de uracilos. Par que se produzca la terminación de la transcripción, el factor rho ( $\rho$ ) reconocería una secuencia específica del ARN denominada sitio *rut* y se uniría a ella para posteriormente tirar del ARN y soltarlo de la ARN polimerasa. Las secuencias *rut* suelen estar un poco antes (aguas arriba) del lugar en el que la ARN polimerasa termina la transcripción.

En eucariontes la terminación de la transcripción es más compleja.



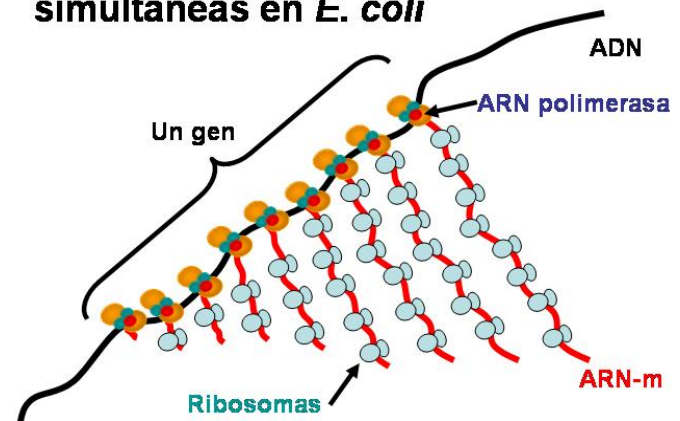
## PROCESAMIENTO DEL ARN

Los ARN recién sintetizados suelen sufrir un procesamiento posterior. Por ejemplo, en bacterias, el ARN precursor del ARN ribosómico (ARN-r) es de mayor tamaño, sucediendo algo parecido con los ARN precursores de los ARN transferentes (ARN-t). Sin embargo, el ARN que se va a traducir a proteínas no sufre procesamiento alguno en bacterias, tal y como se sintetiza comienza a traducirse. De hecho, en bacterias la transcripción y la traducción son simultáneas (al mismo tiempo), es más, aún no ha terminado de transcribirse un ARN mensajero y ya se le han unido ribosomas que están traduciéndolo.



Visualización de la transcripción y la traducción en la bacteria *E. coli*

### Transcripción y Traducción simultáneas en *E. coli*



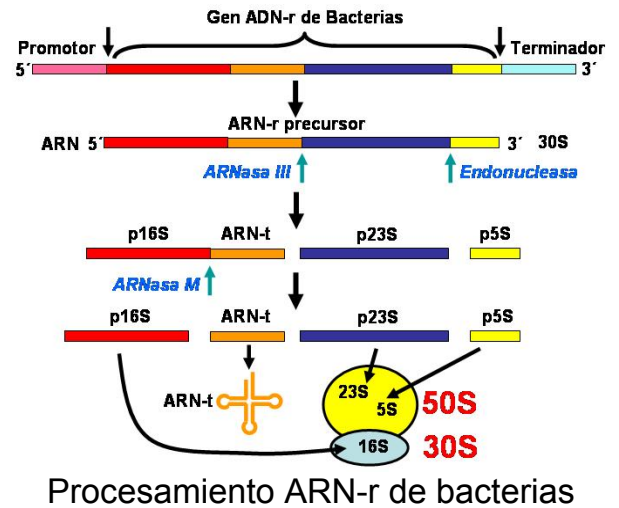
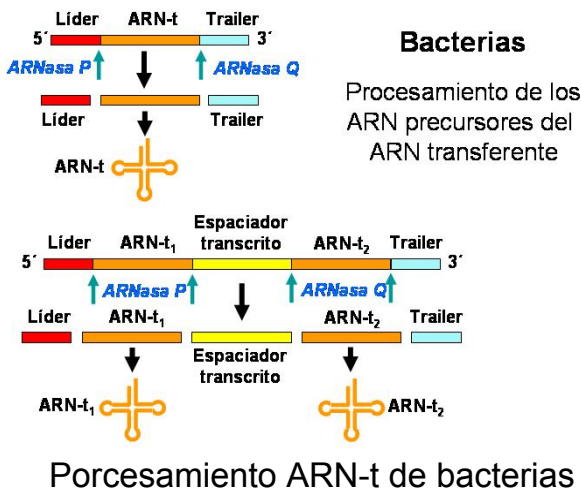
Esquema de la transcripción y la traducción en la bacteria *E. coli*

El ARN precursor del ARN transferente de bacterias tiene una región "líder" por el extremo 5' y una "trailer" por el extremo 3' que desaparecen durante el procesamiento gracias a la acción de la ARNasa P y de la ARNasa Q. A veces se encuentra la información para dos ARN-t distintos en el mismo ARN precursor, existiendo entre ambos una región espaciadora que se transcribe, de forma que la estructura del ARN precursor es la siguiente: 5' Líder - ARN-t<sub>1</sub> - Espaciador transcrito - ARN-t<sub>2</sub> - Trailer 3'. Posteriormente, la ARNasa P y la ARNasa Q cortan el precursor liberando los dos ARN transferentes.

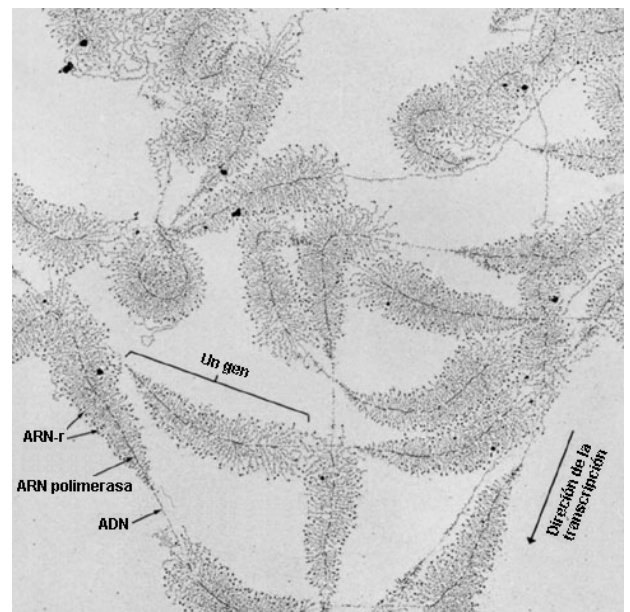
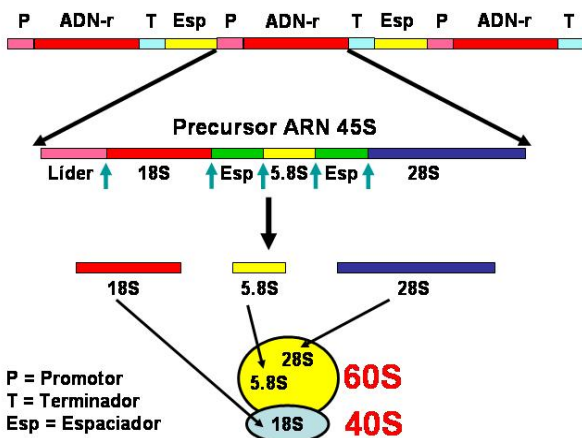
El número de copias de genes de ADN-r (ADN ribosómico) en bacterias es bajo entre 5 y 10 copias. Cada gen de ADN-r posee un promotor y una región terminadora que da lugar tras la transcripción a un ARN precursor del ARN ribosómico (ARN-r) de mayor tamaño que el ARN-r maduro. EL ARN precursor tiene un coeficiente de sedimentación 30S e incluye información para el ARN 16S, una molécula de ARN-t, el ARN 23S y el ARN 5S. Este ARN precursor es



cortado por la ARNasa III y una endonucleasa que generan tres piezas, una que contiene el precursor del ARN 16S (subunidad pequeña ribosoma) junto con un ARN-t, el ARN 23S (subunidad grande del ribosoma) y el ARN 5S (subunidad grande del ribosoma). Después actúa la ARNasa M que separa el ARN 16S del ARN-t.



Los genes para el ARN ribosómico en eucariontes se encuentran repetidos en tandem del orden de cien veces en la región del organizador nucleolar (NOR). En el caso de la especie humana existen varios cromosomas con regiones organizadoras nucleolares, en particular, cinco pares de cromosomas, los pares 13, 14, 15, 21 y 22. En cada región NOR cada gen ADN-r tiene un promotor y un terminador y está separado del siguiente por una región espaciadora que no se transcribe. El ARN precursor (recién transcrito) del ARN ribosómico tiene un coeficiente de sedimentación 45S y su estructura es la siguiente: 5' Líder- precursor del ARN 18S - Espaciador trasncrito - precursor del ARN 5,8S - Espaciador transcrito - precursor del ARN 28S 3'. Posteriormente, diferentes ARNasas y endonucleasas producen cortes en el ARN 45S que dan lugar al ARN 18S de la subunidad pequeña de los ribosomas y a los ARN 28S y 5,8S que forman parte de la subunidad grande de los ribosomas. Los genes para el ARN 5S de eucariontes se localizan en otra región diferente de los cromosomas eucarióticos.



Genes para ADN-r de la región NOR

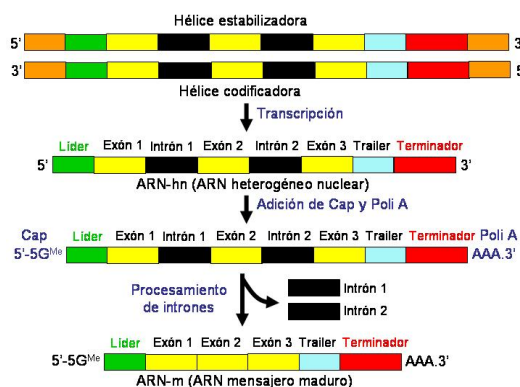
Sin embargo, en eucariontes además de sufrir procesamientos posteriores a la transcripción los *ARN funcionales* que van a dar lugar a los ARN-r y ARN-t, también sufren un procesamiento posterior a su transcripción los *ARN informacionales* que van a ser traducidos a polipéptidos. Este procesamiento consta de varios pasos y tiene como principal consecuencia la disminución



del tamaño del ARN-hn (heterogéneo nuclear) recién transcrito por pérdida de varios segmentos del interior. Los principales pasos del procesamiento del ARN-hn son los siguientes:

- Adición de una caperuza o protección por el extremo 5'P del ARN-hn que entre otras cosas lo protege de su degradación por dicho extremo. Esta caperuza ("capping" o abreviadamente tapón o "cap") consiste en la adición de un residuo de 7-metilguanosa mediante el establecimiento de un enlace trifosfato. Existen tres tipos de caperuzas, cap0 solamente está metilada la primera base (la guanina), cap1 está metilada también la segunda base (es el tipo más frecuente de caperuza) y cap2 en el que también está metilada la tercera base. Otra posible función de la caperuza del extremo 5'P es la suministrar una estructura que sea reconocible por el ribosoma para iniciar la traducción.
- Adición por el extremo 3'OH de una cola de Poli-A, esta cola de poli-A consiste en la adición de 150 a 200 residuos de adenina por el extremo 3'OH. En primer lugar una endonucleasa corta el ARN y después una Poli A polimerasa añade la cola de Poli A. La longitud de la cola de Poli A varía en eucariontes, es de 150 a 200 nucleótidos en células de mamíferos, y de 50 a 60 residuos en levaduras. Los ARN de histonas no tienen poliadenilación por el extremo 3'OH, sufren un procesamiento diferente por ese extremo. La función de la cola de Poli A se cree que puede servir para evitar la degradación por el ARN por ese extremo e intervenir en el transporte del ARN hacia el citoplasma.
- Eliminación de segmentos interiores del ARN-hn. Los genes en eucariontes están fragmentados, de manera que poseen regiones que se traducen a aminoácidos en la secuencia del polipéptido y que se denominan *Exones* y, regiones que no se traducen a aminoácidos en el polipéptido y que se denominan *Intrones*. Cuando la región de ADN correspondiente a un gen se transcribe, el ARN-hn recién transcrito contiene los *Exones* y los *Intrones*, y durante el procesamiento posterior se van a eliminar los *Intrones* o regiones que no se traducen a aminoácidos en el polipéptido. Esta eliminación de los *Intrones* que tiene lugar durante el procesamiento no suele afectar al orden relativo de los que tenían los *Exones* en el ADN.

En adenovirus se demostró por P. A. Sharp y R. J. Roberts (1977) que el ARN-m y el ADN del gen correspondiente no coincidían con exactitud al hibridar, de manera que aparecían lazos de ADN monocatenarios. Por tanto, había segmentos del ADN del gen que no estaban representados en el ARN-m maduro. Por consiguiente, los genes estaban fragmentados o en piezas. Por sus trabajos Sharp y Roberts recibieron el Premio Nobel en 1993. Posteriormente, se comprobó la estructura en Intrones y Exones (genes fragmentados) de los genes eucarióticos. La primera evidencia se obtuvo con los genes de la  $\beta$ -globina de conejo y de la ovoalbúmina de gallina. Los genes de la  $\beta$ -globina de conejo y ratón tienen tres exones y dos intrones, mientras que el de la ovoalbúmina de gallina tiene 7 intrones y 7 exones.



Esquema de procesamiento de un ARN-hn



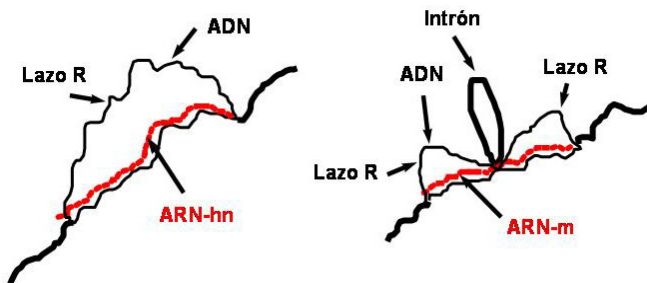
P. A. Sharp



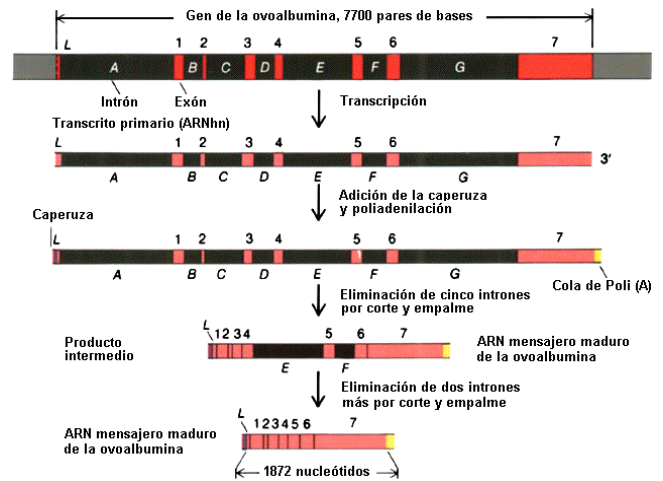
R. J. Roberts

A partir de eritrocitos de conejo se consiguió aislar el ARN mensajero maduro (ARN-m) que lleva información para la  $\beta$ -globina y el ARN recién transcrito o ARN heterogéneo nuclear (ARN-hn). También, fue posible aislar el segmento de ADN que lleva la información para la  $\beta$ -globina. Mediante la técnica de los "lazos R" desarrollada por White y Hogness (1977) es posible hibridar ARN con ADN de doble hélice a altas temperaturas y en presencia de formamida. En estas condiciones, los híbridos ADN-ARN son más estables que las moléculas de ADN doble hélice. Como consecuencia, el ARN hibrida con la secuencia complementaria de ADN y desplaza a la otra hélice de ADN que queda sin aparear formando un "lazo R". Si se hibrida, por un lado, el ARN-m de la  $\beta$ -globina con el segmento de ADN que lo ha originado, y por otro, el ARN-hn (recién transcrito) de la  $\beta$ -globina con el segmento de ADN que lo originó y se analizan los resultados al microscopio electrónico, se obtienen las siguientes imágenes:

Los resultados obtenidos indicaban que hay una zona del ARN-m maduro que no hibrida con el ADN del gen de la  $\beta$ -globina ya que en el centro se observa la aparición de un lazo de ADN doble hélice y a ambos lados la existencia de dos lazos R. Sin embargo, el ARN-hn (recién transcrito) hibrida en toda su longitud con el ADN del gen de la  $\beta$ -globina sin que aparezcan en el centro zonas o lazos de ADN doble hélice. Por tanto, parece que la diferencia existente en el tamaño del ARN-hn y del ARN-m se debe a que existiría un procesamiento que eliminaría una parte del centro del ARN-hn. Estas regiones del ADN del gen de la  $\beta$ -globina que no se traducen en aminoácidos en la proteína recibieron el nombre de Intrones y aquellas que sí se traducen en aminoácidos son los Exones. En el siguiente esquema se da una interpretación a los resultados obtenidos cuando se hibrida el ARN-m con el ADN del gen de la  $\beta$ -globina.

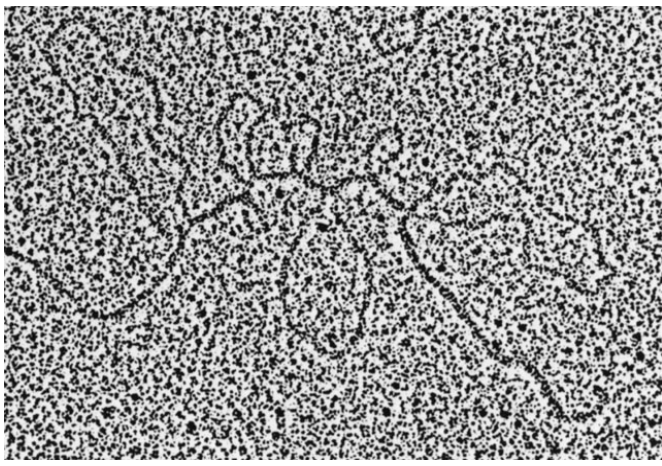


Lazos R: gen de la  $\beta$ -globina de conejo



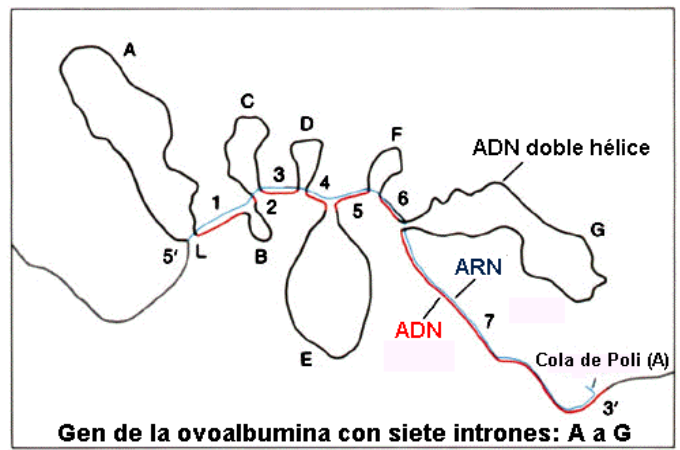
Procesamiento del ARNhn de la ovoalbumina

El gen de la seroalbumina de rata tiene 14 exones y 13 intrones y el factor VIII de coagulación humano posee 26 exones. El gen de la ovoalbumina de gallina tiene 7700 pares de bases entre la región líder, los siete intrones (1 al 7) y los siete exones (A al G) que posee. El ARN-m tiene 1872 nucleótidos, el resto hasta los 7.700 pares de bases se retira durante el procesamiento. El tamaño de los intrones de este gen varía desde los 251 pb del intrón B hasta los 1600 pb del intrón G.



Hibridación del ADN del gen de la ovoalbumina con el ARN mensajero maduro (ARN-m)

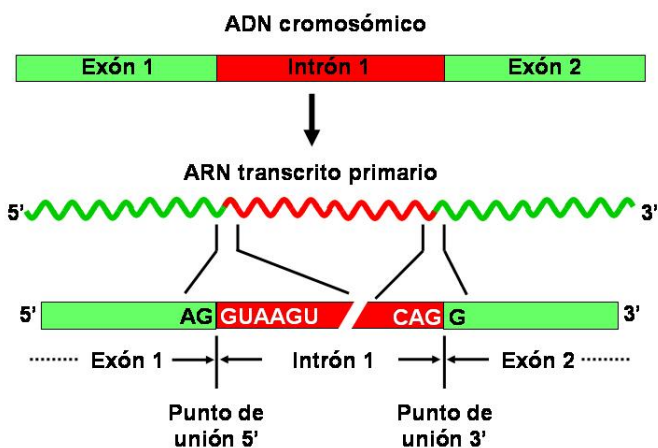
Hibridación del gen de la ovoalbumina



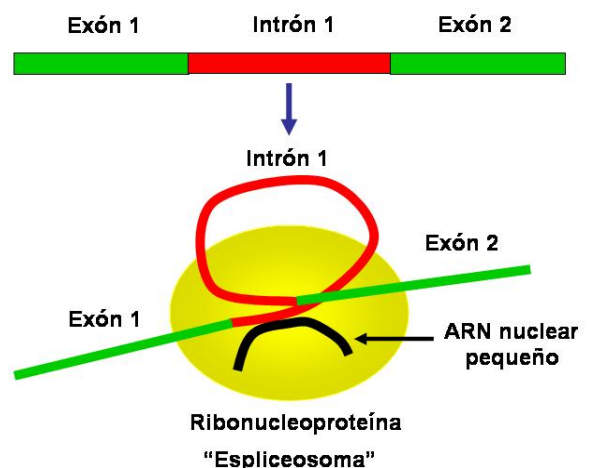
En el gen del factor VIII de coagulación humano el tamaño de los exones varía entre 69 y 3106 pb y posee intrones de hasta 32.400 pb. El número de intrones y exones y su tamaño es muy variable de unos genes a otros dentro de la misma especie. El número de genes que poseen un solo intrón varía mucho de unas especies a otras, en *Levadura (Saccharomyces cerevisiae)* es el 95%, en *D. melanogaster* el 17% y en mamíferos el 6%.

También se han detectado intrones en los genes que codifican para el ARN-r, por ejemplo en el ARN 28S en *D. melanogaster*. Igualmente, en ARN-t de Tyr en levadura se ha detectado un intrón de 14 bases justo antes del extremo 3' del triplete del anticodón. Posteriormente se han detectado intrones en posiciones semejantes en otros ARN-t (ser, phe, leu, etc).

La eliminación de los intrones se lleva a cabo mediante un mecanismo de corte en las regiones de paso de exón a intrón y de intrón a exón para eliminar los intrones y de unión posterior de los exones sucesivos. Este mecanismo de corte y unión ha recibido el nombre de "splicing" en inglés. Cuando se analizan las secuencias de las regiones de paso de intrón a exón y de exón a intrón se encuentra unas secuencias bastante conservadas, en casi todos los casos aparece la secuencia GU en el punto de corte 5' del intrón y AG en el punto de corte 3'. Estas secuencias son reconocidas por moléculas de ARN nucleares pequeñas (ARNnp) que interaccionan con proteínas formando partículas ribonucleoproteicas pequeñas que constituyen lo que se ha denominado el "espliceosoma".



Secuencias consenso exón-intrón-exón



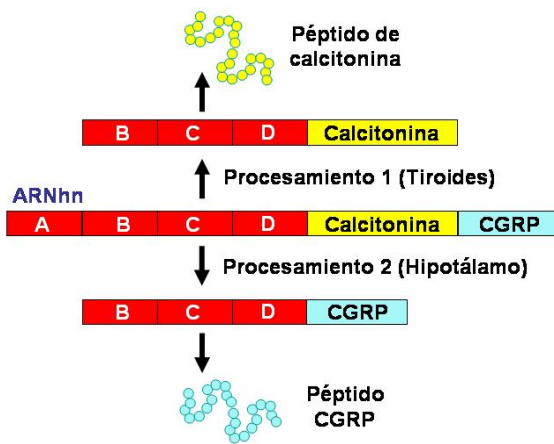
Espliceosoma

Las reacciones de corte y empalme en las que se elimina el intrón y se unen los exones consecutivos son dos reacciones transesterificación, estas reacciones consisten en el intercambio de un enlace fosfodiéster por otro, dando como resultado la unión de dos exones consecutivos.

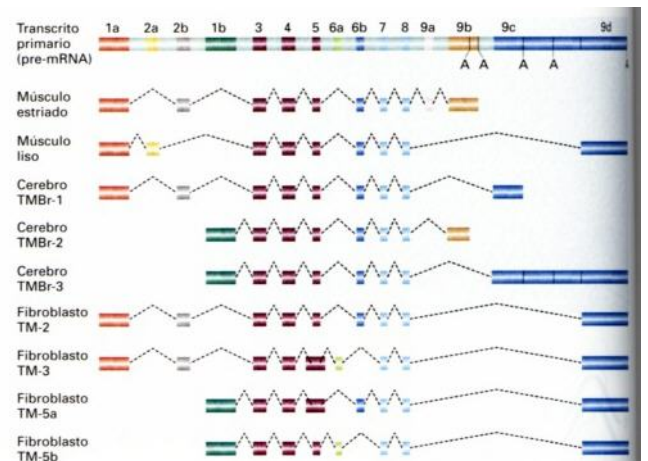


## PROCESAMIENTO ALTERNATIVO

También se ha observado que en diferentes tejidos de un mismo individuo o en el mismo tejido en diferentes momentos del desarrollo se pueden producir procesamientos distintos del mismo ARNhn, como consecuencia en cada tejido o momento del desarrollo se produce un ARN-m maduro diferente que da lugar a un polipéptido distinto. Un ejemplo de procesamiento alternativo es lo que sucede en el tiroides y en el hipotálamo en los que el mismo ARNhn da lugar con procesamientos diferentes al péptido de calcitonina en el tiroides y al péptido CGRP en el hipotálamo. Otro ejemplo de procesamiento alternativo es el caso del transcrito primario del gen de la  $\alpha$ -tropomiosina en músculo estriado, músculo liso, cerebro y fibroblasto. En los dos ejemplos citados de procesamiento alternativo se mantiene el orden relativo de los exones.



Procesamiento alternativo calcitonina y CGRP

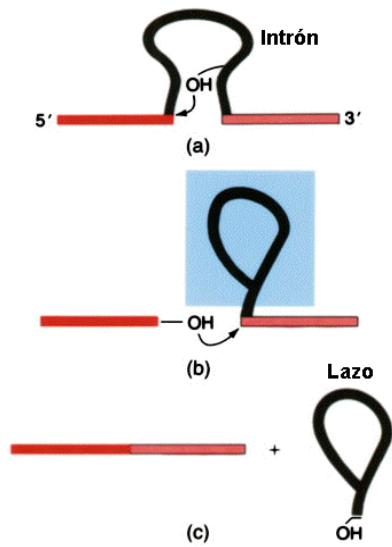


Procesamiento alternativo  $\alpha$ -tropomiosina

## AUTOPROCESAMIENTO

Existen muchos ejemplos de moléculas de ARN que pueden realizar el autoprocésamiento de sus intrones sin necesidad de la ayuda de proteínas. Estas moléculas de ARN que pueden actuar como una enzima con propiedades catalíticas se han denominado *ribozimas*. Esta capacidad de autoprocésamiento de algunas moléculas de ARN fue descubierta por Cech y colaboradores en 1981 estudiando el ARN 35S precursor del ARN-r 26S de *Tetrahymena thermophila* (ciliado). La molécula del ARN-r 26S contiene un intrón de 400 bases que se autoprocésa. T. Cech recibió en 1989 el Premio Nobel por el descubrimientos de los ARN con propiedades catalíticas.





Autoprosesamiento de un intrón



T. Cech

Existen dos tipos de intrones que se autoprosesan, los del grupo I en los que el producto de eliminación del intrón es una molécula lineal y los del grupo II en los que el producto de eliminación del intrón es una molécula en forma de lazo. En el grupo I se encuentran los transcritos primarios de algunos virus que infectan a *E. coli*, los transcritos primarios de *Tetrahymena*, en mitocondrias y cloroplastos, y en algunos ARN-t primarios de bacterias. En el grupo II se encuentran algunos ARN-t primarios y algunos transcritos primarios de mitocondrias y cloroplastos.

↑ Inicio

## EDICIÓN O CORRECCIÓN DEL ARN

Cuando se estudió la secuencia de bases del ARN-m de *cox II* que codifica para la subunidad II de la citocromo c oxidasa mitocondrial del *Trypanosoma brucei* encontraron que en la posición 520 había cuatro uracilos que no se encontraban presentes en la secuencia de bases del gen. Por tanto, estos cuatro uracilos habían sido añadidos después de la transcripción (Benne y col. 1986).

La edición o corrección de los ARN se ha descrito en diferentes especies (protozoos, hongos, plantas y mamíferos), en orgánulos (mitocondrias y cloroplastos) y en el núcleo. Además se produce tanto en regiones codificantes (exones) como no codificantes (intrones). La corrección o edición del ARN puede producirse por adición de bases nitrogenadas, pérdida de bases nitrogenadas o por sustitución de bases después de la transcripción.

↑ Inicio

