

Glosario de términos

Nota: Las palabras en **negritas** en una definición indican términos también definidos en este Glosario

término	Definición
3'	"3 primo"; Se refiere al carbono #3 del componente de azúcar del ácido nucleico (ya sea ribosa en RNA o desoxirribosa en el DNA) al que la polimerasa puede añadir nucleótidos adicionales, a menudo utilizados para referirse a ese extremo de una molécula de DNA o RNA de una sola cadena donde el carbono de 3' retiene su grupo hidroxilo (-OH) y no se unen más nucleótidos.
5'	"5 primo"; Se refiere al carbono #5 del componente de azúcar del ácido nucleico (ya sea ribosa en RNA o desoxirribosa en el DNA), al que se une el tripfosfato en un tripófato de nucleótido, a menudo utilizado para referirse a ese extremo de una molécula de DNA o RNA de una sola cadena donde el grupo de fosfato del carbono 5' no está asociado a un nucleótido anterior.
empalme alternativo (Alternative splicing)	La inclusión o exclusión de ciertos exones en las reacciones de empalme que determinan las secuencias incluidas en el producto final del mRNA . Este mecanismo se utiliza para generar una serie de isoformas proteicas estrechamente relacionadas, que difieren por la inclusión o exclusión de las regiones proteicas particulares codificados por esos exones. El empalme alternativo está dirigido por proteínas que unen RNA y que pueden bloquear o estimular la utilización de un sitio de empalme en particular.
Aminoácido (Amino acid)	El bloque básico de las proteínas, una molécula pequeña con un núcleo -C-C-, un grupo de amino (-NH ₂) en un extremo y un grupo de ácido carboxílico (-COOH) en el otro extremo. La estructura general se puede representar como NH ₂ -CHR-COOH, donde R puede ser cualquiera de los 20 grupos funcionales diferentes de carácter ácido, básico o no polar.
Anotación (annotation)	La anotación genética es el proceso de notar la ubicación, estructura e identidad de los genes en un genoma. A medida que los intentos iniciales pueden basarse en información incompleta, las anotaciones genéticas cambian constantemente a medida que se dispone de más datos. Las bases de datos de anotación genética se actualizan regularmente, y diferentes bases de datos pueden referirse al mismo gen/proteína por diferentes nombres, reflejando nuevos conocimientos y una mejor comprensión de la función proteica.
Base (Base)	Aunque formalmente incorrecto (la base nitrogenada que define A, C, G, T y U es sólo una parte de todo el nucleótido), esto se utiliza a menudo como sinónimo de "nucleótido" al referirse a los componentes A, C, G, T y U del DNA y RNA.
par de base/pareo de base	La unión de hidrógeno de una de las bases (A, C, G, T, U) con otra, dictada por la formación óptima de enlaces o puentes de hidrógeno en el DNA (A-T y C-G)

(base pair/ base pairing)	o en el RNA (A-u y C-G). Se dice que dos hebras de polinucleótidos, o regiones de estas, en las que todos los nucleótidos forman dichos pares base son complementarios. Al lograr la complementariedad, cada hebra de ADN puede servir como una plantilla para la síntesis de su hebra complementaria - el secreto de la precisión extremadamente alta de la replicación de DNA y por lo tanto de la herencia.
Canónico (canonical)	De acuerdo con los principios y estándares existentes generados a partir de datos y pruebas. Por ejemplo, la secuencia de donantes de empalme canónico es GT. En raras ocasiones, sin embargo, la secuencia GC se utiliza en su lugar; puesto que GC no es el "estándar", se denominaría no canónico .
cDNA	"DNA complementario"; una molécula de DNA de doble cadena preparada <i>in vitro</i> ("fuera del cuerpo"; es decir, en un tubo de ensayo) mediante el empleo de una molécula de RNA como plantilla para sintetizar DNA utilizando la transcriptasa al reverso. El componente de RNA del híbrido de RNA-DNA resultante se degrada enzimáticamente, y la hebra complementaria luego sintetizada por la polimerasa de DNA. El DNA de doble cadena resultante se puede utilizar para la clonación y el análisis.
CDS	"secuencia de codificación"; esa parte de la secuencia de DNA de un gen que se traduce en proteínas.
exón de codificación (coding exon)	En un gen, cualquier exon que contenga alguna parte de la CDS ; en cambio, un exon que no tiene ninguna parte traducida a proteína se llama un "exón no codificante".
hebra de codificación/ cadena positiva (coding strand/positive strand)	En un gen, la hebra de DNA que tiene la secuencia encontrada en la molécula de RNA. También se denomina hebra con sentido, positiva o cadena no-molde.
Codón (Codon)	La secuencia de tres nucleótidos en DNA o RNA que especifica un aminoácido en particular.
Coordenada (Coordinate)	Posición numérica dentro de una secuencia biológica; por ejemplo, la primera base de una secuencia de DNA tendría la coordenada "+1".
río abajo(downstream)	Se refiere a la región genómica que comienza en la posición +1. También se le conoce como característica .
Exón (Exon)	Un segmento contiguo de DNA eucariota que corresponde a una porción del producto de RNA maduro (procesado) de ese gen. Los exones se encuentran sólo en genomas eucariotas y están separados por intrones . Aunque los exones se transcriben con los intrones, estos últimos se empalman durante el procesamiento del ARN y se degradan.
Característica (Feature)	Cualquier región de estructura/secuencia definida en un fragmento genómico de DNA. Las características inherentes incluirían genes, pseudogéneos y elementos repetitivos. Una característica también puede ser predicha por

	algoritmos computacionales, como los dirigidos a identificar genes de codificación de proteínas.
Intrón (Intron)	Sección no codificante de una secuencia de ácido nucleico eucariota que se encuentra entre exones. Los intrones se eliminan ("empalmados") de la transcripción primaria/PRE-mRNA después de la transcripción y antes de que la molécula se exporte al citoplasma para su traducción.
Isoformas (Isoform)	Versiones potencialmente diferentes de una proteína codificada por un solo gen. Las isoformas son el resultado del empalme alternativo de un pre-mRNA en particular y/o el uso de un sitio diferente de inicio de transcripción.
mRNA	RNA mensajero maduro que ha sido completamente procesado y está listo para la traducción; tiene un gorro de 7-metilguanosa en su extremo 5' , una cola o rabo de poli(A) en su extremo 3' , y tiene todos sus intrones empalmados.
hebra no codificante / hebra negativa (non-coding strand/negative strand)	También se denomina hebra anti-sentido, molde o no codificante. Esta hebra de la secuencia de DNA de un solo gen es la complementaria a la hebra de DNA de 5' a 3' conocida como la hebra positiva, con sentido, no-molde o cadena codificante . El término pierde su significado para secuencias de DNA más largas con genes en ambas hebras.
Nucleótido (Nucleotide)	El bloque básico de DNA (A, C, G, T) y RNA (A, C, G, U). Los nucleótidos consisten en una base nitrogenada, un azúcar de 5 carbonos (ya sea ribosa en RNA o desoxirribosa en ADN) y grupos de fosfato.
ORF	"Marco de lectura abierto"; un largo tramo de codones en el mismo marco de lectura ininterrumpido por los codones de terminación ; un ORF puede reflejar la presencia de un gen.
Fase (Phase)	La fase describe el número de bases entre el final del exon (definido por el sitio de empalme) y el codón completo más cercano a ese sitio de empalme. El número de bases entre el codón completo adyacente y un sitio de exon/empalme puede ser 0, 1 o 2. La fase de un exón ascendente determinará qué marco se traduce en el exón rio abajo indicando cuántas bases después del sitio del aceptador de empalme son necesarias para crear un codón completo de 3 bases.
Rabo o cola de poli(A) (Poli (A) tail)	Alrededor de 250 nucleótidos de adenina que son añadidos post-transcripcionalmente por polimerasa de poli(A) en el terminal de 3' de transcritos eucariotas, después del corte del RNA recién sintetizado ~ 20 nucleótidos rio abajo de una secuencia de señal de poli adenilación AAUAAA.
pre-mRNA (transcripto principal) (Primary transcript)	La transcripción inicial de un gen de codificación de proteínas que contiene tanto intrones como exones . El PRE-mRNA requiere la adición de un gorro en el terminal 5' y una cola de poli (A) en el 3' y la eliminación de intrones para producir la molécula final de mRNA que contiene exones unidos.
Promotor (Promoter)	Un segmento de DNA al que la RNA polimerasa se une para iniciar la transcripción de los genes rio abajo (downstream) .

Putativo (Putative)	Algo que puede ser predicho o inferido, pero que requiere más evidencia para confirmar o refutar.
Lectura (Read)	Una secuencia de DNA cruda.
marco/marco de lectura (frame/reading frame)	Un marco es una sola serie de tripletes de nucleótidos adyacentes en el DNA o el RNA: un marco tendría bases en las posiciones 1, 4, 7, etc. como la primera base de codón secuencial. Hay tres marcos de lectura posibles en una hebra de mRNA y seis en una molécula de DNA de doble hélice debido a las dos hebras de las que la transcripción es posible. Diferentes programas informáticos numeran estos marcos de manera diferente, por lo que se debe tener cuidado al comparar marcos designados por diferentes programas. Una forma común es referirse a los tres posibles marcos de lectura de izquierda a derecha como +1, +2 y +3 y los tres posibles fotogramas de lectura de derecha a izquierda como -1, -2 y -3.
Empalme (Splicing)	El proceso por el cual se eliminan los intrones y los exones se unen para producir un RNA maduro y funcional (mRNA) a partir de un transcripto primario . Algunos RNAs se auto-empalman, pero la mayoría requieren un complejo específico de ribo nucleoproteínas para catalizar la reacción.
sitio de aceptación de empalme (splice acceptor site)	El sitio de empalme en el extremo 3' de un intrón , en el límite entre un intrón y el exón inmediatamente rio abajo . La secuencia de dinucleótidos del sitio del aceptador de empalme canónico es AG.
sitio donante de empalme (splice donor site)	El sitio de empalme en el extremo 5' de un intrón , en el límite entre un intrón y el exón inmediatamente rio arriba (upstream) . La secuencia de dinucleótidos del sitio donante de empalme canónico es GT; en raras ocasiones, se utiliza la secuencia no canónica GC en su lugar.
lugar de empalme (splice junction)	Sitio de aceptación de empalme o un sitio donante de empalme .
codón de iniciación (Start Codon)	El primer codón de un CDS . En eucariotas esto es casi siempre ATG, que codifica para la metionina (uno de los 20 aminoácidos).
codón de terminación (Stop Codon)	Un codón que especifica la terminación de la síntesis de proteínas; a veces llamado un "codón sin sentido" ya que no especifica un aminoácido .
Transcripción (Transcription)	El proceso de copiar una hebra de una doble hélice de DNA por la RNA polimerasa, creando una hebra complementaria de RNA llamada transcripto.
Traducción (Translation)	El proceso por el cual los codones en un mRNA son "leídos" por el ribosoma y los tRNAs para dirigir la síntesis de proteínas.
TSS- sitio de inicio de la transcripción (transcription start site)	La ubicación en el DNA generalmente rio arriba de la secuencia codificante de un gen, donde la RNA polimerasa comienza la transcripción .

UTR	"Región no traducible"; un segmento de DNA (o RNA) que se transcribe y está presente en el mRNA maduro, pero no se traduce en proteínas. Los UTR se pueden encontrar en cualquiera o en ambos de los extremos de 5' y 3' de un gen o transcripto.
río arriba (upstream)	Hace referencia a la región genómica antes de examinar la característica .

Traducido por: Dr. Enrique Rodriguez Borrero