

ADN y Genes

26 de abril de 2013

El material genético: ADN

Seres vivos tienen la capacidad de almacenar, usar y transferir información. Esta información se almacena a nivel molecular a través de ácidos nucleicos (ADN).

Los ADN son moléculas complejas producidas por la célula, esenciales para todos los organismos.

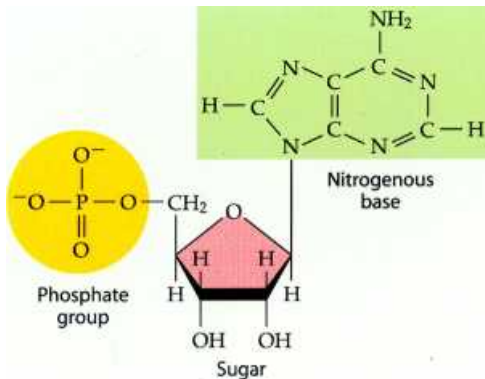
El ADN

- ▶ contiene la información que permite organizar moléculas en las células de un organismo,
- ▶ es capaces de regular su composición química interna, su crecimiento y su reproducción.
- ▶ es responsable para la herencia.

Nucleótidos

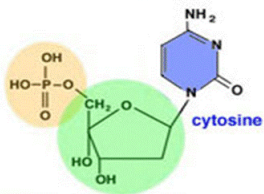
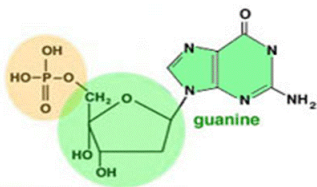
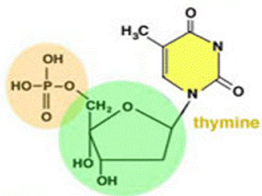
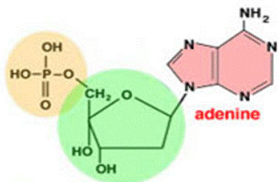
Las partes elementales de la ADN son los **nucleótidos**.

Hay 4 nucleótidos diferentes: **G**uanosina, **A**denina, **T**imina, **C**itosina.



Son hechos de un grupo fosfato, una ribosa (azúcar) y una base nitrogenada. La diferencia entre ellos es en la base.

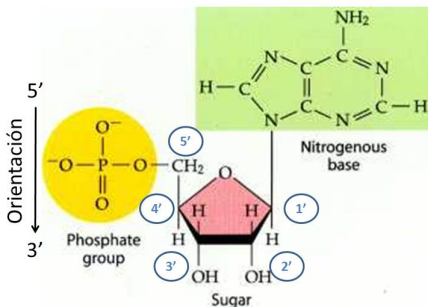
Las cuatro bases



Purinas

Pirimidinas

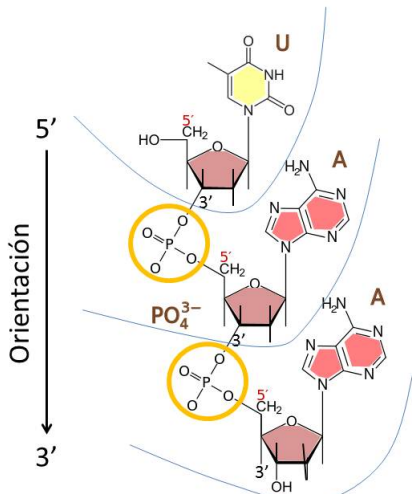
El azúcar



La **ribosa** contiene 5 carbonos que están numerados: carbono 1' está conectada la base nitrógena; carbono 4' está conectado con el grupo fosfato; los carbonos 3' y 5' se usan para encadenar nucleótidos.

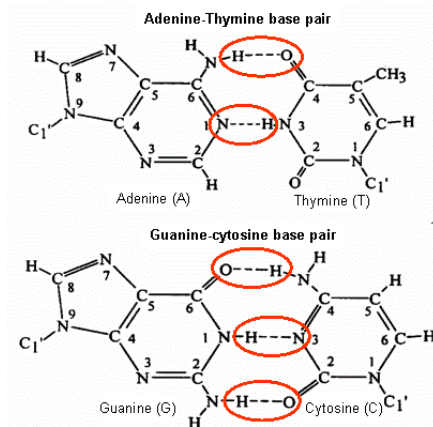
Orientación: 5' (el lado del grupo fosfato) → 3' (el lado del azúcar).

Cadenas de nucleótidos



Nucleótidos pueden formar cadenas: al carbono 3' de un nucleótido se agrega el grupo fosfato del siguiente. En la gráfica se ve la cadena 5'-UAA-3'. Los enlaces que se forman son covalentes (muy fuertes).

Nucleótidos complementarios

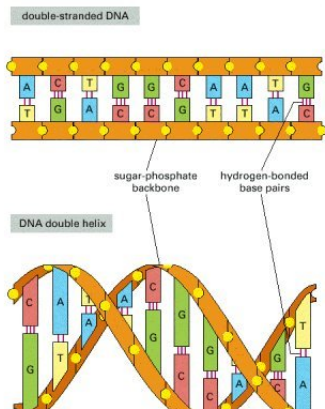


El ADN es una molécula formada de dos cadenas, estabilizadas por enlaces de hidrógeno (pares de bases):

- ▶ para cada A en una cadena se encuentra una T en la otra y viceversa (2 enlaces);
- ▶ para cada G en una cadena se encuentra una C en la otra y viceversa (3 enlaces).

Otras combinaciones de nucleótidos no se dan por razones de espacio o por incompatibilidad química.

ADN: estructura



- ▶ La información de las dos cadenas es redundante, sin embargo no son cadenas idénticas sino complementarias (conociendo una cadena se puede construir la otra).
- ▶ La orientación entre las dos cadenas se llama antiparalelo: El lado 5' de una cadena corresponde al lado 3' de la otra viceversa.

Por características de hidrofiliicidad, hidrofobicidad y electronegatividad en sus diferentes segmentos, la ADN se dobla y enrolla, formando así una estructura visible por el microscopio: una cromosoma.

ADN como portador de información (1)

La información contenida en el ADN está codificada por los 4 nucleótidos; es una cadena construida por 4 símbolos (A,T, C, G); por ejemplo:

ATTAGGCGCGTATATT

Compare con la forma como se almacena información en el computador: por cadenas de bits, usando los símbolos (0, 1), por ejemplo:

1101011100010011

ADN como portador de información (2)

Investigación Harvard University:

George M. Church, Yuan Gao, Sriram Kosuri. **Next-Generation Digital Information Storage in DNA.** Science 28 September 2012: Vol. 337 no. 6102 p. 1628

Digital information is accumulating at an astounding rate, straining our ability to store and archive it. DNA is among the most dense and stable information media known. The development of new technologies in both DNA synthesis and sequencing make DNA an increasingly feasible digital storage medium. We developed a strategy to encode arbitrary digital information in DNA, wrote a 5.27-megabit book using DNA microchips, and read the book by using next-generation DNA sequencing.

ADN como portador de información (3)

Investigación European Bioinformatics Institute (EBI):

Goldman, N.; Bertone, P.; Chen, S.; Dessimoz, C.; Leproust, E. M.; Sipos, B.; Birney, E. (2013). **Towards practical, high-capacity, low-maintenance information storage in synthesized DNA.** Nature 494 (7435): 77–80.

Digital production, transmission and storage have revolutionized how we access and use information but have also made archiving an increasingly complex task that requires active, continuing maintenance of digital media. This challenge has focused some interest on DNA as an attractive target for information storage¹ because of its capacity for high-density information encoding, longevity under easily achieved conditions and proven track record as an information bearer. Previous DNA-based information storage approaches have encoded only trivial amounts of information or were not amenable to scaling-up⁸, and used no robust error-correction and lacked examination of their cost-efficiency for large-scale information archival. Here we describe a scalable method that can reliably store more information than has been handled before. We encoded computer files totalling 739 kilobytes of hard-disk storage and with an estimated Shannon information of $5,2 \cdot 10^6$ bits into a DNA code, synthesized this DNA, sequenced it and reconstructed the original files with 100% accuracy. Theoretical analysis indicates that our DNA-based storage scheme could be scaled far beyond current global information volumes and offers a realistic technology for large-scale, long-term and infrequently accessed digital archiving. In fact, current trends in technological advances are reducing DNA synthesis costs at a pace that should make our scheme cost-effective for sub-50-year archiving within a decade.

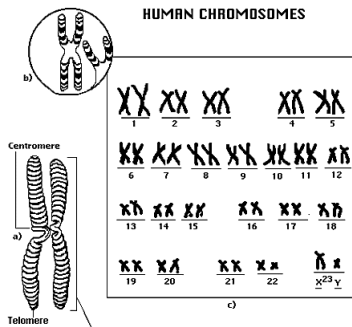
Genes

Las unidades de la ADN que contiene la información para una característica específica son los genes.

Determinan el desarrollo del cuerpo y todas sus características, para ello almacenan información hereditaria y dirigen la síntesis de proteínas.

- ▶ cada gen es la receta para formar una proteína determinada
- ▶ un gen complicado puede consistir de miles de nucleótidos.

Genoma



El conjunto de todas las cromosomas se llama **Genoma**.

- ▶ contiene todas las instrucciones genéticas de un organismo
- ▶ puede ser manejado por millones o miles de millones de nucleótidos

Dogma central de la biología molecular

El proceso de extracción de la información del gen y usar la información para producir una proteína se llama “Dogma central de la biología molecular”.



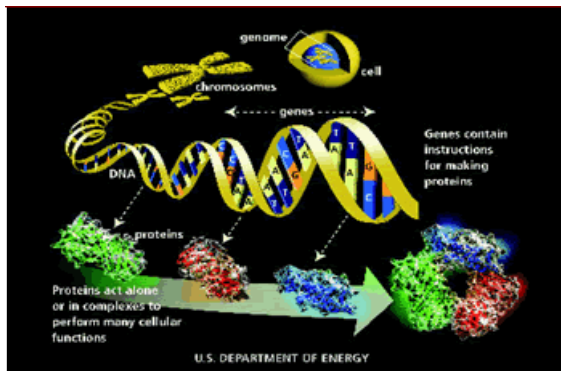
- ▶ Transcripción: Se hace una copia 1-a-1 de la ADN para obtener la ARN (una sola cadena), copiando G, A, T y C para obtener G, A, U y C (U = uracil)
- ▶ Traducción: La información en la ARN se convierte en una secuencia de aminoácidos que forma una proteína.

Cuando un gen produce una proteína, se dice que el gen **se expresa**.

Dogma central de la biología molecular

Describe el proceso de transición de **Información** (almacenada en el ADN) a **Acción** (ejecutada por las proteínas).

Analogía: un programa de computo y el efecto de su ejecución.



Universalidad

Los seres procariotas (como bacterias) y eucariotas (como levaduras, plantas, animales) usan:

- ▶ el mismo alfabeto (nucleótidos)
- ▶ el mismo formato de almacenamiento (ADN)
- ▶ la misma conversión a proteína (código genético)
- ▶ las mismas funcionalidades de las proteínas (proteínas homólogas)

Los genes

Los genes son trozos determinados de ADN, embebido en partes que no se expresan (zona intergénica).

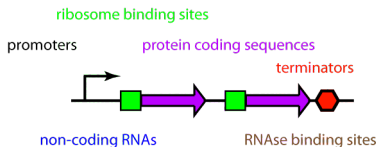
Pregunta biológica: ¿Cómo sabe una célula cual proteína falta y cual gen debe expresarse?

Respuesta: Existen mecanismos sofisticados de regulación, que determinan si se debe incrementar (control positivo) y disminuir (control negativo) la expresión de genes. Estos mecanismos están ligados a la estructura de un gen y las zonas adyacentes.

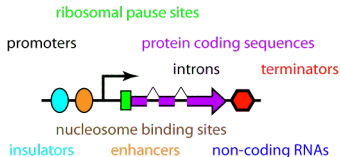
Pregunta computacional: ¿Cómo se pueden reconocer los genes en una secuencia de ADN?

Respuesta: Se modelan las partes de un gen y las zonas adyacentes (obteniendo un patrón) y se busca la secuencia de patrones que corresponde a la estructura de un gen.

Estructura de un gen



Prokaryotic Genetic Parts



Eukaryotic Genetic Parts

Promotores, terminadores: permiten al organismo responder a señales de otras partes del genoma o del ambiente celular, de modo que la unión de proteínas reguladoras codificadas por estas regiones favorece o evita la transcripción del gen y, por tanto, incrementa o disminuye la producción de proteínas en determinados momentos. En procariontes: Los elementos reguladores sirven para varios genes contiguos (operón); en eucariotes: cada gen está regulado.

Exones, intrones: muchos genes se encuentran constituidos por regiones codificantes (exones) interrumpidas por regiones no codificantes (intrones) que son eliminadas en la formación del ARN (en eucariotes).

Zona intergénica: los genes no están situados siempre de forma contigua en el ADN, sino que separando los genes existe una gran cantidad de ADN (variable según los distintos genomas). Las secuencias no codificadoras superan en número de diez o más a las codificadoras.