

Figura A: El código genético

| | U | C | A | G | |
|---|--------------------------|--------------------------|----------------------------|---------------------------|------------------|
| U | Phe Phe Leu Leu | Ser Ser Ser Ser | Tyr Tyr stop stop | Cys Cys stop Trp | U C A G |
| C | Leu Leu Leu Leu | Pro Pro Pro Pro | His His Gln Gln | Arg Arg Arg Arg | U C A G |
| A | Ile Ile Ile Met | Thr Thr Thr Thr | Asn Asn Lys Lys | Ser Ser Arg Arg | U C A G |
| G | Val Val Val Val | Ala Ala Ala Ala | Asp Asp Glu Glu | Gly Gly Gly Gly | U C A G |

Los genes se componen de una serie de cuatro bases de nucleótidos (A, C, G y T) que se alinean una tras otra en forma de cadena de nucleótidos. Para que un gen se pueda utilizar para formar una proteína, el ADN debe copiarse antes en un mensaje (una molécula relacionada llamada ARNm), que se puede entonces enviar fuera del núcleo de la célula a un ribosoma, donde se ensamblan las proteínas. Esta molécula de ARNm es similar al ADN, salvo que su estructura química se forma utilizando ribosa en lugar de desoxirribosa, y una de las bases, la timina (T), se sustituye por uracilo (U). El orden, o secuencia, de los nucleótidos del ARNm determina la manera en que se ensambla una proteína. Los aminoácidos son los elementos básicos de una proteína. Es decir, los diferentes aminoácidos se ensamblan, uno tras otro, en una cadena de "polipéptidos". Esta, a su vez, se dobla y se dispone en forma de proteína. Cada aminoácido añadido a una cadena creciente de polipéptidos se indica por medio de un conjunto específico de tres bases de nucleótidos dentro del ARNm. Cada triplete posible de bases de nucleótidos se llama codón.

Los aminoácidos especificados por cada secuencia de codón en el ARNm se indican en la siguiente tabla. La primera base de un codón se indica en el lado derecho de la tabla, la segunda en la parte superior y la tercera en el lado derecho.

Dado que el ARN se construye a partir de cuatro tipos de nucleótidos, hay 64 tripletes posibles para los codones (4x4x4). Tres de estos codones posibles especifican la terminación de la cadena de polipéptidos. Se llaman "codones de parada". Con esto, quedan 61 codones para especificar solamente 20 aminoácidos diferentes. Por lo tanto, la mayoría de los aminoácidos están representados por más de un codón. Por eso se dice que el código genético es degenerado.

Las claves de la tabla anterior:

Ala: Alanina; Cys: Cisteína; Asp: Ácido aspártico; Glu: Ácido glutámico;
 Phe: Fenilalanina; Gly: Glicina; His: Histidina; Ile: Isoleucina;
 Lys: Lisina; Leu: Leucina; Met: Metionina; Asn: Asparagina;
 Pro: Prolina; Gln: Glutamina; Arg: Arginina; Ser: Serina;
 Thr: Treonina; Val: Valina; Trp: Triptófano; Tyr: Tirosina

A = adenina; G = guanina; C = citosina; T = timina; U = uracilo