

IL PASSAGGIO DELL'INFORMAZIONE GENETICA

Con le conoscenze che abbiamo acquisito sul DNA, possiamo fornire una definizione più precisa di genotipo e fenotipo di un organismo: il genotipo è l'informazione ereditaria contenuta nel suo DNA; il fenotipo corrisponde, invece, alle sue caratteristiche fisiche. A livello molecolare, il collegamento tra genotipo e fenotipo è rappresentato dalle proteine. Il DNA che un organismo eredita dai genitori (genotipo), infatti, specifica quali proteine devono essere sintetizzate, e in quale momento. Il processo attraverso il quale il DNA "dirige" la sintesi delle proteine è chiamato espressione genica. Un gene, tuttavia, non sintetizza direttamente una proteina; il suo compito, infatti, è quello di fornire le istruzioni sotto forma di RNA, che a sua volta dirige la sintesi proteica. La serie di istruzioni parte dal DNA contenuto nel nucleo della cellula, è trasferita a una molecola di RNA ed è tradotta con la sintesi della proteina nel citoplasma. Le due fasi principali del processo sono: la trascrizione, ovvero il trasferimento dell'informazione genetica dal DNA a una molecola di RNA; la traduzione, cioè l'utilizzo dell'informazione contenuta nell'RNA per costruire una proteina. I geni forniscono le istruzioni per sintetizzare proteine specifiche; tuttavia, un gene non sintetizza direttamente la proteina. Il DNA viene trascritto nell'RNA, a sua volta tradotto nelle proteine. In altre parole, le cellule sono governate da quella che potremmo definire una "catena di comando molecolare": DNA-->RNA--> proteine

Per comprendere con quali modalità l'informazione genetica passa dal genotipo al fenotipo, dobbiamo capire in che modo il linguaggio chimico del DNA viene tradotto nel linguaggio chimico delle proteine. Che cos'è, esattamente, il linguaggio chimico degli acidi nucleici? Tanto il DNA quanto l'RNA sono polimeri costituiti da monomeri nucleotidici. Nel DNA troviamo 4 tipi di nucleotidi, che differiscono tra loro per le basi azotate (A,T,C e G); lo stesso vale per l'RNA, che però contiene l'uracile(U) al posto della timina(T). Come gli acidi nucleici, anche i polipeptidi sono polimeri, ma i monomeri che li compongono sono i 20 aminoacidi comuni a tutti gli organismi. Anche in questo caso, l'informazione è scritta in una sequenza lineare, e la sequenza dei nucleotidi della molecola di RNA indica l'esatta sequenza degli aminoacidi del polipeptide. In questo modo, l'RNA funge da messaggero che trasporta l'informazione genetica proveniente dal DNA. Questo tipo di molecola di RNA viene chiamato per questo mRNA o RNA messaggero. Durante la traduzione avviene, però, un cambiamento di linguaggio, dalla sequenza di nucleotidi dell'RNA alla sequenza di aminoacidi del polipeptide: in che modo? Ricordiamo che il DNA e l'RNA contengono soltanto 4 tipi diversi di nucleotidi: questi 4 nucleotidi devono in qualche modo specificare 20 aminoacidi. Se ogni base nucleotidica specificasse un aminoacido, potrebbero essere codificati soltanto 4 aminoacidi. Se invece le basi di un gene venissero lette due alla volta, la coppia AG per esempio potrebbe specificare un aminoacido, mentre AT potrebbe specificarne uno diverso. Ma combinando le 4 basi a coppie, otterremmo soltanto 16 (cioè 4^2) possibili combinazioni, un numero non sufficiente per specificare tutti i 20 aminoacidi che compongono le proteine. Supponiamo che nel DNA ogni parola in codice consista di una tripletta, cioè che ogni combinazione di tre basi consecutive codifichi per un aminoacido. Ci sarebbero allora 64 (cioè 4^3) possibili parole in codice, più che sufficienti per specificare i 20 aminoacidi. Le triplette di basi sono dunque le parole più brevi, di lunghezza uniforme, in grado di specificare tutti gli aminoacidi. In realtà, il numero di triplette è tale che a ogni aminoacido può corrispondere più di una. Per esempio le triplette AAT e AAC codificano entrambe per lo stesso aminoacido (leucina). Gli esperimenti hanno confermato che il flusso di informazioni dal gene alla proteina è effettivamente basato su un codice a triplette: le istruzioni genetiche per la sequenza aminoacidica di una catena polipeptidica sono scritte nel DNA e nell'RNA come una serie di parole di tre lettere, dette codoni. Negli anni Sessanta del secolo scorso, i Biologi molecolari, grazie a una serie di esperimenti, riuscirono a "decifrare" il codice genetico, cioè a individuare la traduzione di ciascuna tripletta di nucleotidi. Il primo codone fu decifrato nel 1961 dal biochimico statunitense Marshall Nirenberg, che sintetizzò una molecola di RNA artificiale costituita da una sequenza di un solo tipo di nucleotide, avente come base azotata l'uracile. Nirenberg introdusse questo RNA, che conteneva soltanto un tipo di codone, UUU, in una provetta contenente una miscela di ribosomi e altre sostanze necessarie per la sintesi di un

polipeptide. Dalla miscela ottenne un polipeptide contenente un unico amminoacido, la fenilalanina. Nirenberg ne deduce che il codone UUU dell'RNA specificava l'amminoacido fenilalanina(Phe). In breve tempo, variando questo metodo, furono individuati gli amminoacidi specificati da tutti gli altri codoni. Il codice genetico consiste di una serie di regole che stabiliscono la corrispondenza tra i codoni dell'RNA e gli amminoacidi delle proteine. Il codice genetico è universale, in quanto è condiviso da tutti gli organismi, dai batteri fino alle piante e agli animali più complessi. Ciò suggerisce che esso si sia evoluto in tempi molto antichi nella storia della vita ed è un'ulteriore testimonianza dell'origine di tutti i viventi da un antenato comune.