

## Il genoma umano

### Introduzione

Un gene è un frammento di DNA codificante un'informazione che fa stabilmente parte della struttura di una cellula e una copia della quale viene trasmessa, al momento della divisione cellulare, a ciascuna delle cellule figlie. Il genoma umano è costituito da 3 miliardi di paia di basi e contiene circa 35.000 geni. L'analisi della struttura di un gene dimostra che esso, oltre a specificare la sequenza aminoacidica della proteina codificata, contiene molti elementi regolatori che ne controllano l'espressione.

### Il messaggio genetico

In ogni cellula del corpo è contenuta l'informazione necessaria per specificare la struttura fisica di un intero essere umano. Questa informazione costituisce il genoma ed è scritta nel DNA in forma di una stringa di 3 miliardi di lettere, con un alfabeto costituito da soltanto quattro basi: A-adenina, T-timina, C-citosina e G-guanina. In una singola cellula il DNA è presente in una quantità pari a 7.1 picogrammi (10-12 grammi) ed è localizzato all'interno del nucleo sotto forma di un filamento lineare lungo 2 metri, con uno spessore di 0.2 nanometri.

Per apprezzare appieno le dimensioni e la complessità del genoma umano mi sembra utile considerare alcune analogie. Innanzitutto, immaginiamo la molecola del DNA portata alle dimensioni del binario di una ferrovia le cui rotaie siano costituite dallo scheletro dell'elica del DNA e le cui traversine siano le coppie di basi A-T o G-C: questo binario sarebbe lungo circa 1.700.000 chilometri. Le traversine contengono l'informazione genetica. Risalendo lungo una rotaia si potrebbe leggere la sequenza ATGGGTCTG, mentre scendendo lungo l'altra si leggerebbe la sequenza di basi complementare TACCCAGAC. Sequenziare il genoma umano significa percorrere 1.700.000 chilometri e registrare le lettere viste su 3 miliardi di traversine.

Attualmente, grazie alle moderne tecniche di sequenziamento e alla tecnologia dei chip a DNA, siamo in grado di decifrare 1.000 bp (paio di basi, base pairs) al secondo. Questa è una velocità sorprendente, ma ancora circa 10 volte inferiore alla velocità alla quale una cellula può replicare il proprio DNA. Durante il ciclo di divisione cellulare l'intero genoma diploide viene replicato nel giro di 12 ore.

### Contenuto informativo del genoma

Un gene è un segmento di DNA che codifica un'unità informativa che fa stabilmente parte della struttura di una cellula e una copia della quale viene trasmessa, al momento della divisione cellulare, a ciascuna delle cellule figlie. I tre elementi essenziali di questa definizione sono DNA, unità informativa e copia.

Un gene codifica per una specifica sequenza lineare di aminoacidi che viene assemblata sui poliribosomi della cellula. Gli eventi che controllano l'espressione del messaggio genetico controllano la funzione delle cellule nell'organismo mediante la modulazione della sintesi delle proteine.

Il messaggio genetico è espresso in “parole” di tre lettere denominate codoni. Ogni codone specifica uno dei 20 possibili aminoacidi che sono le unità costitutive di tutte le proteine. È importante notare che nell’RNA è presente la base uracile anziché la timina, il che implica la presenza di una U nell’RNA al posto di ogni T presente nel corrispondente DNA. Questo codice genetico è utilizzato da tutti gli organismi viventi presenti sulla terra! I codoni STOP, o di terminazione, UAA, UAG e UGA non specificano alcun aminoacido, ma sono usati come segni di interpunzione che stabiliscono la fine del messaggio. Si stima che i 3 miliardi di coppie di basi del genoma contengano circa 35.000 geni. È importante farsi un’idea della quantità di informazione che ciò rappresenta e anche in questo caso ci può aiutare un’analogia. I tre miliardi di lettere contenute nel genoma corrispondono, all’incirca, al numero di lettere contenuto in tutti i libri di una buona biblioteca di una facoltà di medicina. Nella biblioteca, le lettere vengono usate per specificare parole di varia lunghezza riunite in frasi con punteggiatura. Queste, a loro volta, sono organizzate in paragrafi, capitoli e libri. Nel genoma ci sono soltanto codoni di tre lettere e la punteggiatura è limitata ai segnali di inizio e di fine.

(tratto e adattato da Dennis W. Ross Introduzione alla MEDICINA MOLECOLARE pp.3-6)