

# Riassunto comandi latex

Matteo Mistri

5 aprile 2018

# Indice

<b>1</b>	<b>DNA</b>	<b>2</b>
<b>2</b>	<b>Geni</b>	<b>3</b>

# Capitolo 1

## DNA

Il DNA genomico è la lunga molecola composta da monomeri che contiene le info di ogni cellula. E' situato nel nucleo, è frammentato in cromosomi ed ha la forma di doppia elica. Quello umano è composto da 22 coppie di cromosomi, più i due cromosomi sessuali. La lunghezza non è costante, ma decresce. Il DNA è composto da tanti blocchi, come dei mattoncini. Ogni mattoncino è un nucleotide, composto a sua volta da uno zucchero, un gruppo fosfato e una base azotata(nitrogena). Ho 4 possibili basi azotate(Adenina, Citosina, Guanina ,Timina). La direzione 5' -> 3' (gruppo fosfato sotto, nucleotide in alto) è quella standard. Il collegamento tra i nucleotidi per creare la catena avviene dal gruppo fosfato allo zucchero. Per leggere la sequenza primaria di una seq di DNA, leggo le sue basi in direzione 5' -> 3'

L'unità di misura è il base pair, cioè il numero di basi di una catena(oppure un'elica). L'appaiamento è standard, avviene con una catena rovesciata e le basi si accoppiano così A+T(doppio legame idrogeno) C+G(triplo legame idrogeno). Per convenzione, si decide che la catena + è quella di sx, mentre quella - è quella di dx. La sequenza del genoma è una stringa sull'alfabeto  $\langle A, C, G, T \rangle$  L'operazione di reverseANDcomplement consiste, data una sequenza di DNA, di trovare la sua catena appaiata(ricordarsi che si legge in senso opposto, quindi prima la inverte e poi la complemento). RNA uguale a DNA, ma al posto della Timina ho l'Uracile.

## Capitolo 2

### Geni

Un gene è una sezione di DNA che codifica una proteina. E' quindi una sottostringa della catena del DNA. La sequenza primaria di una sequenza che codifica un gene è definita sequenza genomica. I geni hanno lo HUGO NAME. I geni sulla catena + sono diversi da quelli sulla catena -, non è detto in corrispondenti posizioni. Il formato FASTA(\*.fa, \*.fasta) è un plain text che codifica la sequenza primaria + alcune informazioni aggiuntive. La prima riga inizia con > e contiene, su quell'unica riga, informazioni a piacere (relase del genoma, genoma di provenienza, indice di inizio e indice di fine, catena di provenienza (+1 per la positiva, -1 per la negativa)). Essa è seguita dalla sequenza genica, su una sola riga o su più righe da 80 colonne. Ogni gene ha al suo interno delle regioni dette esoni (regioni codificanti) e introni (regioni non codificanti). Il confine tra esone e introni è detto 5' splice site, mentre tra introni e esoni è detto 3' splice site. La prima fase è quella della trascrizione, che crea il pre mRNA, nel quale è copiato il gene completo, sostituendo la timina con l'uracile. La seconda fase elimina gli introni e concatena gli esoni, dando vita all'mRNA trascritto. Qui vive la coding sequence, che inizia con AUG, finisce con UAG/UAA/UGA. Le parti prima e dopo della coding sequence sono dette UTR. Le triplette della coding sequence mappano un amminoacido, che poi si unirà in proteina. Ma i geni sono 25000, le proteine centinaia di migliaia, come è possibile? Il gene può produrre più proteine con lo splicing alternativo, combinando i suoi esoni in modo diverso. A seconda di dove si trova la cellula a cui appartiene il DNA, viene eseguito uno splicing diverso. Altri fattori sono le condizioni in cui si trova la cellula. Il sequenziamento significa determinare la sequenza primaria delle molecole biologiche. Non si può sequenziare tutto il DNA di una molecola.