

ESERCITAZIONE 4

Pagine internet

ExPaSY <http://www.expasy.org/tools>

SMS: <http://www.bioinformatics.org/sms2/>

CUTG: <http://www.kazusa.or.jp/codon>

Programmi locali

Notepad++: <http://notepad-plus.sourceforge.net/it/site.htm>

MS Excel

Obiettivi

1. Capire l'uso dei codoni
2. Identificare regioni idrofobiche in una proteina
3. Capire la frammentazione con proteasi in silico

Procedura

---Parte 1: codon usage ----

1. Cercare su CUTG la tavola dei codoni di escherichia coli (quella fatta con 8087 geni).
2. Copiarla e incollarla in Notepad++. Eseguire un sostituisci usando: cerca=" + " e sostituisci con="\". Copiare il risultato e incollarlo in Excel.
3. Cercare la sequenza codificante della proteina [P60422](#) (trovare il codice CAA26463). Copiarla e incollarla nel tool "Codon Usage" di SMS, dopodichè premere Submit. Copiare il risultato e ripetere la procedura di sostituzione del punto 2. Copiare il risultato in Excel (accanto alla tavola incollata in precedenza).
4. In excel, rimuovere le colonne relative al numero di codoni (Number). Ordinare separatamente le due tabelle usando le colonne "Codon", sistemare le tabelle una accanto all'altra, dopodichè inserire una nuova colonna C con la formula "=A2&"_"&B2. Ordinare di nuovo tutto secondo la nuova colonna.
5. Fare un grafico ad istogramma con i valori (/1000) del genoma di coli (serie1) e con quelli estratti dalla proteina L2. Osservare la disproporzione nell'uso dei codoni per gli aminoacidi glicina (G) e arginina (R).

--- Parte 2: sliding windows ---

1. Ottenere la sequenza della proteina PDGFR umana ([P16234](#)). Osservarne le caratteristiche nelle Features.
2. Andare su ProtScale ed incollare la sequenza, scegliere la scala di idrofobicità (Kyle & Doolittle) e lanciare l'analisi.
3. Scaricare il file numerico e lavorarlo come al punto 2 della parte 1: l'obiettivo è ottenere dei dati tab-separati per incollarli in excel.
4. Incollare i dati in excel e replicare il plot di ProtScale.
3. Evidenziare le regioni idrofobiche (> 1): c'è qualcosa di atteso in base alle features della proteina?

4. Ripetere le analisi con la propensione alpha (Chou-Fasman).

--- **Parte 3: proteolisi in silico** ---

1. Data la protein di prima, eseguire un taglio “in silico” con la tripsina con peptmass.
2. Collezionare la lista dei frammenti, seglierne 3 a caso
3. Aprire Mascot e incollarli nell’interfaccia. Riempire i campi richiesti e laciare l’analisi.
4. Osservare i risultati: è presente la proteina das cui si era partiti?