

Corso di Bioinformatica

http://compgen.bio.unipd.it/~stefania/Didattica/AA2010-2011/Bioinformatica_BTS/Bioinfo_BTS.html

Per ogni punto dell'esercitazione copiare ed incollare le informazioni richieste in un file sotto forma di breve testo con figure, in modo da assemblare un traccia di relazione con brevi commenti. Alla fine inviare via mail la relazione prodotta nel corso dell'esercitazione (stefania.bortoluzzi@unipd.it). Salvare il file della relazione ogni 5'!

III ESERCITAZIONE

Gli allineamenti multipli di proteine sono estremamente utili per ricavare informazioni sulla filogenesi, sulla struttura secondaria, sull'individuazione di domini strutturali e funzionali e sul folding proteico. Nel corso di questa esercitazione cercheremo di utilizzare la ricerca di similarità per identificare una proteina umana a partire da un frammento della sua sequenza e di evidenziare la sua eventuale appartenenza ad una famiglia di proteine. Le sequenze di alcuni altri membri della famiglia (proteine umane o di altri organismi, con elevatissima similarità con la proteina in analisi) verranno allineate con ClustalW. L'allineamento multiplo ed il cladogramma ottenuti saranno analizzati per interpretare i risultati.

- 1- Collegarsi all'home page del programma Blast dell'NCBI (<http://www.ncbi.nlm.nih.gov/BLAST/>).
- 2- Scegliere Protein-protein BLAST (blastp).
- 3- Incollare la sequenza seguente (cut and paste):
VQRPGRIPTSSRCGNTIPDDDNQVVLSLSPGSRVVLPRPVCFEKGTNYTVRLELPQYTSSSDSDVESPYTLIDSLVLMPYC
KSLDIFTVGGSGDGVVTVNSAWETFQRYRCLENSRSVVKTPMTDVCRNII FSI SALLHQ TGLACECDPQGSLSVCDPENG
QCQCRPNVVGRTCNRCAPGTFGFGPSGCKPCECHLQG
- 4- Scegliere il database RefSeq.
- 5- Inviare e formattare la ricerca.
- 6- Analizzare il risultato della ricerca di similarità, per capire a quale proteina umana corrisponde il frammento di sequenza in analisi, recuperarne la sequenza di riferimento, salvandola in un file di testo e seguire il collegamento a EntrezGene per conoscerne funzione, localizzazione ed eventuale coinvolgimento in malattie genetiche.
- 7- Ritornare a Protein-protein BLAST (blastp), usare come query la sequenza appena recuperata, limitando la ricerca a Reference sequences di proteine di mammiferi.
- 8- Inviare e formattare, analizzare il risultato della ricerca di similarità.
- 9- Selezionare le sequenze di uomo, topo e ratto della famiglia delle laminine beta (1, 2 e 3), recuperarle in formato FASTA e copiarle in un file di testo.
- 10- Modificare gli headers delle sequenze in modo da renderli brevi (max 8 caratteri), informativi ed univoci (ad es. hs1, hs2, mm1, ...).
- 11- Collegarsi all'home page di ClustalW <http://www.ebi.ac.uk/Tools/msa/clustalw2/>
- 12- Incollare le sequenze dal file di testo, avviare l'analisi con i parametri di default.
- 13- Analizzare l'allineamento e copiarlo
- 14- Collegarsi a <http://www.ebi.ac.uk/Tools/phylogeny/> Usando come input l'allineamento, ricostruire il dendrogramma che descrive le relazioni tra le sequenze considerate.
- 15- Allineare con ClustalW solo le tre sequenze delle laminine beta 2 e cercare la coppia di sequenze maggiormente simili.

CONSEGNARE UNA RELAZIONE CONTENENTE:

- Riassunto dei contenuti dell'esercitazione (massimo 150 parole).
- Sequenza di riferimento della proteina identificata, punto 6 ed informazioni sul gene.
- Sequenze delle proteine selezionate al punto 9.
- Allineamento ottenuto con ClustalW (punto 13).
- Risposta al punto 15.