

Βιοπληροφορική I

Εργαστηριακές Ασκήσεις – Μέρος A

Άσκηση 1: Να γίνει αναζήτηση στη βάση δεδομένων UniProt με σκοπό την ανεύρεση ανθρώπινων υποδοχέων συζευγμένων με G-πρωτεΐνες οι οποίοι να έχουν γνωστή (προσδιορισμένη) τρισδιάστατη δομή:

keyword:"G-protein coupled receptor [KW-0297]" AND organism:"Human [9606]" AND existence:"evidence at protein level" AND database:(type:pdb)

Πόσες από τις πρωτεΐνες έχουν δομή στη βάση δεδομένων PDB; Για δέκα από τις παραπάνω πρωτεΐνες (ομάδα A 1^η-10^η, ομάδα B 11^η-20^η, ομάδα Γ 21^η-30^η, ομάδα Δ 31^η-40^η) να σημειωθεί ποιες αντιστοιχούν στην περιοχή της πρωτεΐνης στην οποία βρίσκονται το σύνολο των διαμεμβρανικών τμημάτων, ποιες αντιστοιχούν σε μέρος της περιοχής των διαμεμβρανικών τμημάτων και ποιες δεν περιλαμβάνουν κανένα τμήμα της ακολουθίας το οποίο να αντιστοιχεί σε ακολουθία διαμεμβρανικών τμημάτων.

Άσκηση 2: Να γίνει αναζήτηση στις βάσεις δεδομένων UniProt, PDB, CATH, SCOP, Pfam για την πρωτεΐνη DHFR του ανθρώπου και του βακτηρίου *Lactobacillus casei* και να καταγραφούν δομικές διαφορές μεταξύ των δύο πρωτεϊνών.

Άσκηση 3: Δίνονται οι ακολουθίες: GAWGHEE και WHEAE. Να γίνει ολική στοίχιση και τοπική στοίχιση των δύο ακολουθιών με τη χρήση του πίνακα BLOSUM 50 και ποινή για το κενό 8. Να γίνει επαλήθευση των αποτελεσμάτων με τη χρήση του αλγορίθμου στοίχισης που βρίσκεται στη διεύθυνση <http://www.itu.dk/~sestoft/bsa/graphalign.html>.

Άσκηση 4:

- A. Χρησιμοποιήστε την πρωτεΐνη Hemoglobin subunit beta του ανθρώπου με accession number P68871 για να πραγματοποιήσετε αναζήτηση με το πρόγραμμα blastp έναντι της βάσης UniProt/SwissProt αποκλείοντας την ταξινομική βαθμίδα Mammalia (taxid:40674) και αυξάνοντας τον αριθμό αποτελεσμάτων σε 1000. Παρατηρήστε και σχολιάστε σύντομα τα αποτελέσματα.
- B. Από την αναζήτηση που μόλις κάνατε και μελετώντας τα αποτελέσματα στο κομμάτι Taxonomy, επιλέξτε 9 ακολουθίες που η κάθε μια να ανήκει σε διαφορετική ταξινομική βαθμίδα με βάση το blast name και κρατήστε τις σε fasta format σε ένα αρχείο μαζί με την ανθρώπινη. Στη συνέχεια χρησιμοποιήστε τις 10 ακολουθίες για να κατασκευάσετε πολλαπλές στοιχίσεις με τα προγράμματα Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) και T-COFFEE (<https://www.ebi.ac.uk/Tools/msa/tcoffee/>). Σχολιάστε τις πολλαπλές στοιχίσεις. Αποθηκεύστε τοπικά τις στοιχίσεις σε ClustalW format.
- C. Στον web server του HMMER <https://www.ebi.ac.uk/Tools/hmmer/> χρησιμοποιήστε την ανθρώπινη πρωτεΐνη για να κάνετε αναζήτηση μέσω του Search/rhmmr έναντι της UniProt/SwissProt αποκλείοντας την ταξινομική

βαθμίδα Mammalia (taxid:40674). Συγκρίνετε τα αποτελέσματα με αυτά της αντίστοιχης αναζήτησης με το πρόγραμμα blastp.

- D.** Στον web server του HMMER <https://www.ebi.ac.uk/Tools/hmmer/> χρησιμοποιήστε την πολλαπλή στοίχιση που κατασκευάσατε με το εργαλείο T-COFFEE για να κάνετε αναζήτηση μέσω του Search/hmmsearch έναντι της UniProt/SwissProt αποκλείοντας την ταξινομική βαθμίδα Mammalia (taxid:40674). Συγκρίνετε τα αποτελέσματά σας με τις αντίστοιχες αναζητήσεις που κάνατε προηγούμενα.