**ΕΡΓΑΣΤΗΡΙΟ ΒΙΟΠΛΗΡΟΦΟΡΙΚΗ**

**Άσκηση 4:**

Χρησιμοποιήστε την πρωτεΐνη **Epidermal growth factor receptor** του ανθρώπου με accession number P00533 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.

1. 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 (download 🡪 fasta (complete sequence)).
2. 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.
3. Οπτικοποιείστε τα αποτελέσματα με το Jalview: <https://www.jalview.org/jalview-js/JalviewJS/>
4. Οπτικοποιείστε τα αποτελέσματα με το Weblogo: <http://weblogo.berkeley.edu/>
5. Στον web server του HMMER <https://www.ebi.ac.uk/Tools/hmmer/> χρησιμοποιήστε την ανθρώπινη πρωτεΐνη για να κάνετε αναζήτηση μέσω του Search/phmmer έναντι της UniProt/SwissProt αποκλείοντας την ταξινομική βαθμίδα Mammalia (taxid:40674). Συγκρίνετε τα αποτελέσματα με αυτά της αντίστοιχης αναζήτησης με το πρόγραμμα blastp.
6. 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.

**Άσκηση 5:**

Κάντε φυλογενετική ανάλυση χρησιμοποιώντας το πρόγραμμα MEGA.

A. Επισκεφτείτε τον ιστότοπο της ΒΔ Conserved Database (<http://www.ncbi.nlm.nih.gov/cdd>) στο

NCBI.

B. Αναζητήστε την οικογένεια G-alpha (ή όποια άλλη σας ενδιαφέρει να μελετήσετε).

C. Στο πεδίο Στοίχιση αλληλουχιών επιλέξτε να εμφανίσετε τη στοίχιση από τα πιο διαφοροποιημένα μέλη (most diverse members) και αντιγράψτε την σε mFasta format (πατήστε Reformat αφού επιλέξετε το αντίστοιχο format). Αφού επικολλήσετε τη στοίχιση σε έναν κειμενογράφο (π.χ. Notepad++) απλοποιήστε τα ονόματα χρησιμοποιώντας μόνο το accession number και το κοινό όνομα του οργανισμού.

D. Εισάγετε το αρχείο στο MEGA όπως στην εικόνα, στοιχίστε τις ακολουθίες και αποθηκεύστε σε .mas και .meg formats.

E. Επιλέξτε Phylogeny > Construct/Test για να κατασκευάσετε τα φυλογενετικά δέντρα με τις μεθόδους neighbor-joining, maximum likelihood και UPGMA.

F. Για κάθε φυλογενετικό δέντρο, διαβάστε τις πληροφορίες της λεζάντας και δοκιμάστε τα εργαλεία που προσφέρονται (π.χ. placing a root, flipping nodes, showing or hiding branch lengths, interconverting display formats). Συγκρίνετε τα δέντρα που πήρατε και αναφέρετε σύντομα τις διαφορές τους.

G. Κατασκευάστε ξανά τα δέντρα εφαρμόζοντας σε καθένα τη μέθοδο του bootstrapping. Εντοπίστε τους κλάδους του δέντρου που εμφανίζουν μικρό βαθμό bootstrap (low bootstrap value). Γιατί συμβαίνει αυτό; Εξηγήστε.

