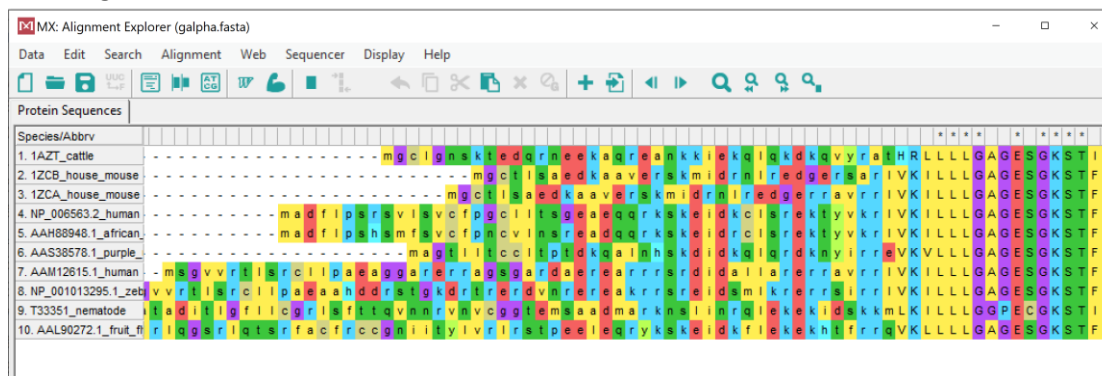


# Βιοπληροφορική Ι

## Εργαστηριακές Ασκήσεις – Μέρος Β

**Άσκηση 5:** Κάντε φυλογενετική ανάλυση χρησιμοποιώντας το πρόγραμμα MEGA.

1. Επισκεφτείτε τον ιστότοπο της BD Conserved Database (<http://www.ncbi.nlm.nih.gov/cdd>) στο NCBI.
2. Αναζητήστε την οικογένεια G-alpha (ή όποια άλλη σας ενδιαφέρει να μελετήσετε).
3. Στο πεδίο Στοιχισμός 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.
4. Εισάγετε το αρχείο στο MEGA όπως στην εικόνα, στοιχίστε τις ακολουθίες και αποθηκεύστε σε .mas και .meg formats.



5. Επιλέξτε Phylogeny > Construct/Test για να κατασκευάσετε τα φυλογενετικά δέντρα με τις μεθόδους neighbor-joining, maximum likelihood και UPGMA.
6. 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.
7. 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.

**Άσκηση 6:** Να γίνει πρόβλεψη της δευτεροταγούς δομής των πρωτεϊνών ompA Ecol (P0A910) και NAD-dependent malic enzyme (P27443) με τη χρήση των προγραμμάτων **Jpred**, **Porter** και **Psipred**. Στη συνέχεια να υπολογιστεί ο δείκτης Q3 για την πρόβλεψη της δευτεροταγούς δομής της πρωτεΐνης για καθένα από τα προγράμματα.

<http://www.compgen.org/material/courses/bioinformatics1/uoa/membrane-proteins>

(Πρόγνωση Δευτεροταγούς Δομής)

**Άσκηση 7:** Να γίνει πρόβλεψη των διαμεμβρανικών τμημάτων της πρωτεΐνης Rhodopsin (P02699) με τη χρήση των προγραμμάτων:

- TMHMM,
- MEMSAT,
- PHOBIUS.
- HMMTOP,
- TOPCONS και

Στη συνέχεια, σε αρχείο excel να καταχωρήσετε την ακολουθία της πρωτεΐνης και στις επόμενες γραμμές την αντίστοιχη πρόβλεψη του κάθε προγράμματος για κάθε αμινοξύ. Αν για κάθε αμινοξύ η πρόβλεψη συμφωνεί για τέσσερα προγράμματα πρόβλεψης και πάνω τότε να καταχωρείτε ότι το αμινοξύ αυτό την αντίστοιχη πρόβλεψη (π.χ. συμβολίστε με 1 αν ανήκει σε διαμεμβρανικό και 0 όταν δεν ανήκει). Τέλος, να απεικονιστεί γραφικά η τελική πρόβλεψη για τα διαμεμβρανικά τμήματα της πρωτεΐνης η οποία θα έχει προκύψει από το συνδυασμό των παραπάνω προγραμμάτων πρόβλεψης τοπολογίας διαμεμβρανικών πρωτεϊνών.