



Πανεπιστήμιο Ιωαννίνων

Σχολή Επιστημών Υγείας

Τμήμα Ιατρικής

Μεταπτυχιακό Πρόγραμμα «Βασικές Βιοϊατρικές Επιστήμες»

Μεταπτυχιακό Δίπλωμα Ειδίκευσης

**«Μοριακή ταυτοποίηση και φυλογενετική ανάλυση απομονωθέντος
παρασίτου *Dirofilaria repens*»**

Πέρτσαλης Αθανάσιος

Επιβλέπων: Δρ. Μποζίδης Πέτρος

Ιωάννινα 2019



Πανεπιστήμιο Ιωαννίνων

Σχολή Επιστημών Υγείας

Τμήμα Ιατρικής

Μεταπτυχιακό Πρόγραμμα «Βασικές Βιοϊατρικές Επιστήμες»

Μεταπτυχιακό Δίπλωμα Ειδίκευσης

**«Μοριακή ταυτοποίηση και φυλογενετική ανάλυση απομονωθέντος
παρασίτου *Dirofilaria repens*»**

Πέρτσαλης Αθανάσιος

Επιβλέπων: Δρ. Μποζίδης Πέτρος

Ιωάννινα 2019

«Η έγκριση του Μεταπτυχιακού Διπλώματος Ειδίκευσης από το Τμήμα Ιατρικής του Πανεπιστημίου Ιωαννίνων δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα Ν.5343/32, άρθρο 202, παράγραφος 2 (νομική κατοχύρωση του Ιατρικού Τμήματος)».

Όνοματεπώνυμο: *Πέρτσαλης Αθανάσιος*

Τίτλος του Μεταπτυχιακού Διπλώματος Ειδίκευσης: *Βασικές Βιοϊατρικές Επιστήμες – Κατεύθυνση: Εφαρμοσμένη Μικροβιολογία*

«*Μοριακή ταυτοποίηση και φυλογενετική ανάλυση απομονωθέντος παρασίτου *Dirofilaria repens**»

Ημερομηνία Παρουσίασης: *13/03/2019*

Επιβλέπων: *Δρ. Μποζίδης Πέτρος – ΕΔΙΠ Μικροβιολογίας*

Εξεταστική Επιτροπή:

- 1. Μποζίδης Πέτρος – ΕΔΙΠ Μικροβιολογίας*
- 2. Παπαδοπούλου Χρυσάνθη – Καθηγήτρια Μικροβιολογίας*
- 3. Σακκάς Ηρακλής – Λέκτορας Μικροβιολογίας*

Στην οικογένειά μου

ΠΡΟΛΟΓΟΣ

Η παρούσα διπλωματική εργασία πραγματοποιήθηκε στα πλαίσια του Μεταπτυχιακού Προγράμματος Σπουδών «Βασικές Βιοϊατρικές Επιστήμες», του Τμήματος Ιατρικής, της Σχολής Επιστημών Υγείας, του Πανεπιστημίου Ιωαννίνων, κατά το ακαδημαϊκό έτος 2018-2019.

Σε αυτό το σημείο θα ήθελα να εκφράσω τις ειλικρινείς μου ευχαριστίες σε όλους όσους συνέβαλαν στο να φέρω σε πέρας την παρούσα Μεταπτυχιακή Διπλωματική Εργασία. Ιδιαίτερα, θα ήθελα να ευχαριστήσω τον επιβλέποντα της εργασίας αυτής και Διδάσκοντα του Μ.Π.Σ. κ. Μποζίδη Πέτρο, για την πολύτιμη βοήθειά του και τη διαρκή υποστήριξή του, τόσο κατά τη διεξαγωγή του πρακτικού μέρους όσο και κατά τη συγγραφή της παρούσας εργασίας, καθώς και τα μέλη της εξεταστικής επιτροπής, την Καθηγήτρια Μικροβιολογίας κ. Παπαδοπούλου Χρυσάνθη και τον Λέκτορα Μικροβιολογίας κ. Σακκά Ηρακλή για τις συμβουλές τους και τη βοήθειά τους. Τέλος, θα ήθελα να ευχαριστήσω όλους τους δικούς μου ανθρώπους, οικογένεια και φίλους, για την αμέριστη συμπαράσταση, ποικιλότροπη βοήθεια και προ πάντων κατανόηση και ανοχή καθ' όλο το χρονικό διάστημα των σπουδών μου, γεγονός που οδήγησε στην πραγματοποίηση αυτού του μεταπτυχιακού προγράμματος.

ΠΕΡΙΕΧΟΜΕΝΑ

1. ΕΙΣΑΓΩΓΗ	8
1.1. Φυλογενετική: Ορισμοί και Ιστορική Αναδρομή.....	8
1.2. Φυλογενετική Ανάλυση	10
1.2.1. Εισαγωγικά Στοιχεία και Ορισμοί.....	10
1.2.2. Μέθοδοι Ευθυγράμμισης Αλληλουχιών.....	12
1.2.2.1. Ευθυγράμμιση Αλληλουχιών Ανά Ζεύγος.....	12
1.2.2.2. Πολλαπλή Ευθυγράμμιση Αλληλουχιών	14
1.2.3. Κατασκευή Φυλογενετικών Δέντρων	16
1.2.3.1. Ορισμοί.....	16
1.2.3.2. Τύποι Φυλογενετικών Δέντρων	17
1.2.3.3. Μέθοδοι Κατασκευής Φυλογενετικών Δέντρων.....	20
1.2.3.4. Επιπλέον Παράμετροι Κατασκευής Φυλογενετικών Δέντρων	24
1.2.3.5. Βάσεις Δεδομένων	26
1.2.3.6. Επιλογή Τρόπου Κατασκευής Φυλογενετικού Δέντρου	27
1.3. Παρασιτισμός και Ελμινθολογία.....	28
1.4. Διροφιλαρίαση και <i>Dirofilaria repens</i>	30
1.5. Φυλογενετική και Παράσιτα	32
1.6. Η Φυλογενετική στο Μέλλον.....	32
1.7. Σκοπός.....	33
2. ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ	34
2.1. Πειραματικό Παράσιτο	34
2.2. Αλληλουχίες.....	34
2.3. Διαδικασία Κατασκευής Φυλογενετικού Δέντρου	35
2.3.1. Συνοπτική	35
2.3.2. Αναλυτική.....	36
3. ΑΠΟΤΕΛΕΣΜΑΤΑ	50
4. ΣΥΖΗΤΗΣΗ	57
5. ΠΕΡΙΛΗΨΗ	68
6. ABSTRACT	69
7. ΒΙΒΛΙΟΓΡΑΦΙΑ	70

1. ΕΙΣΑΓΩΓΗ

1.1. Φυλογενετική: Ορισμοί και Ιστορική Αναδρομή

Ο όρος φυλογένεση αναφέρεται στην εξελικτική ιστορία των ειδών. Όταν δε, γίνεται λόγος για «φυλογενετική ανάλυση» (phylogenetic analysis) εννοούμε τη διασύνδεση αυτών των ιστοριών ή αλλιώς τη μελέτη της εξελικτικής συσχέτισης των ειδών με τη βοήθεια μαθηματικών μοντέλων. Η «μοριακή φυλογενετική» (molecular phylogenetics) χρησιμοποιεί τα δεδομένα που προκύπτουν από την αλληλούχιση των οργανισμών, προκειμένου να οδηγηθεί σε συμπεράσματα αναφορικά με τις σχέσεις μεταξύ των εξεταζόμενων οργανισμών αλλά και των γονιδίων τα οποία αυτοί φέρουν. Χάρη στην εκτεταμένη διαθεσιμότητα πολλών δεδομένων αλληλούχισης και τη χρήση υπολογιστικών αλγορίθμων και προγραμμάτων, η φυλογενετική ανάλυση αποδεικνύεται σταδιακά υψηλής σημασίας σε πολλούς τομείς της βιολογίας. Το πλέον καθιερωμένο αποτέλεσμα της φυλογενετικής ανάλυσης είναι ένα διάγραμμα με διακλαδώσεις, το λεγόμενο «δενδρόγραμμα» (dendrogram), ή ακόμη πιο συνηθισμένα «φυλογενετικό δέντρο» (phylogenetic tree) [7, 22, 55].

Ιστορικά, η χρήση των φυλογενετικών δέντρων, περιοριζόταν, κυρίως, στη μελέτη της εξελικτικής βιολογίας και σε τομείς όπως η ταξινόμηση των ειδών. Μάλιστα, μέχρι περίπου τριάντα χρόνια πριν, αυτό γινόταν κυρίως με συγκρίσεις βασιζόμενες στη μορφολογία των ειδών. Δύο κλασσικές προσεγγίσεις ταξινόμησης υπήρξαν: α) η «φυσική ταξινόμηση» (natural classification) και β) η «φαινετική (ή φαινοτυπική) ταξινόμηση» (phenetics), ενώ θεωρείται ότι ο Αριστοτέλης, ακόμη πιο παλιά (384-322 π.Χ.) είχε ήδη προχωρήσει στην ταξινόμηση έμβιων όντων, φτάνοντας μάλιστα στο ορθό συμπέρασμα ότι τα δελφίνια δεν συμπεριλαμβάνονται στην ίδια μεγάλη ομάδα στην οποία ανήκουν οι ιχθύες. Η πρώτη αναπτύχθηκε από τον Κάρολο

Λινναίο τον 18^ο αιώνα και βασιζόταν σε ανατομικά χαρακτηριστικά. Όντας μια ομαδοποίηση παρόμοιων χαρακτηριστικών δεν παρείχε πάντοτε εξελικτική συγγένεια. Έτσι, οι νυχτερίδες θα μπορούσαν να ομαδοποιηθούν μαζί με πολλά έντομα μόνο και μόνο επειδή πετούν. Η δεύτερη αναπτύχθηκε από τους Peter Sneath και Robert R. Sokal τον 20^ο αιώνα και ομαδοποιεί τους οργανισμούς βάσει συνολικής ομοιότητας, συνήθως, χρησιμοποιώντας μορφολογικά παρατηρήσιμα χαρακτηριστικά. Μπορεί να ειπωθεί, ορθά, ότι αποτελεί ένα είδος αναβαθμισμένης φυσικής ταξινόμησης, μιας και εδώ λαμβάνει χώρα η σύγκριση περισσότερων του ενός χαρακτηριστικών. Σαφώς και δεν χρησιμοποιείται σήμερα ως μέθοδος, όπως ούτε και η πρώτη. Ωστόσο, είναι σημαντικό να γίνει αντιληπτό το ότι οι σημερινές μέθοδοι είναι αποτέλεσμα της ιστορίας τους και είναι γεγονός ότι κάποιες από τις αρχές της φαινετικής ταξινόμησης έχουν διατηρηθεί. Παράδειγμα αποτελεί η χρήση των αλγορίθμων Σύνδεσης-Γειτονίας (Neighbor-joining ή NJ) ως πρώτο βήμα στη κατασκευή φυλογενετικών δέντρων όταν είναι πολύ χρονοβόρα η χρήση αποκλειστικά μοντέρνων μεθόδων, δεδομένο που θα αναλυθεί πιο κάτω. Επίσης, μία ακόμη σημαντική ορολογία είναι η «γονοτυπική ταξινόμηση» (genotypic classification), η οποία αναφέρεται στη σύγκριση γενετικής ομοιότητας μεταξύ οργανισμών χρησιμοποιώντας αλληλουχίες ολόκληρου γονιδιώματος ή μέρος αυτού [11, 40, 41, 42, 43, 51].

Η ευθυγράμμιση και η ανάλυση των αλληλουχιών DNA εμπίπτει στο ευρύτερο πεδίο της βιοπληροφορικής. Ο όρος «βιοπληροφορική» (bioinformatics) εισήχθηκε το 1970 από τους Paulien Hogeweg, Ben Hesper και David Searles οι οποίοι και χρησιμοποίησαν τον όρο αυτό εννοώντας «τη μελέτη των υπολογιστικών διεργασιών σε βιοτικά συστήματα». Η βιοπληροφορική είναι ο δεσμός της βιολογίας, της επιστήμης των υπολογιστών και των/της μαθηματικών/στατιστικής. Ειδικότερα, όμως, μπορεί να αναφερθεί ότι είναι ένας τρόπος να λαμβάνει χώρα η επεξεργασία αλληλουχιών, να εντοπίζονται μοτίβα, όπως επίσης και να

αναλύονται τα μεγάλα ποσά των δεδομένων που παράγονται από τα σύγχρονα πειράματα μοριακής βιολογίας και τη χρήση συστημάτων επόμενης γενιάς αλληλούχισης (Next Generation Sequencing ή NGS) [14, 15, 16, 23].

1.2. Φυλογενετική Ανάλυση

1.2.1. Εισαγωγικά Στοιχεία και Ορισμοί

Αναφορικά με τη διαδικασία κατασκευής φυλογενετικών δέντρων υπάρχει πληθώρα μεθόδων που μπορούν να χρησιμοποιηθούν. Αξίζει, ωστόσο, να σημειωθεί ότι όλες αυτές οι μέθοδοι ακολουθούν σε γενικές γραμμές την ίδια προσέγγιση, η οποία και θα μπορούσε να παρουσιαστεί με τα εξής πέντε βασικά βήματα: (α) Συλλογή των αλληλουχιών, (β) Ευθυγράμμιση των αλληλουχιών, (γ) Καθορισμός των διαφορών μεταξύ των αλληλουχιών, (δ) Έκφραση των διαφορών αυτών αριθμητικά και τέλος (ε) Χρήση του μέτρου αυτού των διαφορών προκειμένου να κατασκευαστεί το φυλογενετικό δέντρο. Τα τέσσερα πρώτα βήματα αφορούν στη φυλογενετική ανάλυση που πρέπει να προηγηθεί της κατασκευής του φυλογενετικού δέντρου [13].

Κατά τη διαδικασία της φυλογενετικής ανάλυσης λαμβάνει χώρα σύγκριση δύο ή περισσότερων αλληλουχιών DNA, RNA, ή πρωτεΐνης για τον εντοπισμό περιοχών ομοιότητας που μπορεί να είναι συνέπεια λειτουργικών, δομικών ή εξελικτικών σχέσεων [29, 37].

Προκειμένου να γίνουν κατανοητά τα βήματα αυτής της διαδικασίας θα πρέπει να δοθεί η ερμηνεία των παρακάτω όρων:

α) Ευθυγράμμιση αλληλουχιών (sequences alignment): Αφορά στη διαδικασία διαμόρφωσης των αλληλουχιών DNA, RNA ή πρωτεΐνης κατά τέτοιο τρόπο ούτως ώστε να καθίσταται εφικτός ο εντοπισμός περιοχών ομοιότητας. Η προκύπτουσα ευθυγράμμιση τυπικά έχει νουκλεοτίδια ή αμινοξέα ευθυγραμμισμένα σε διαδοχικές στήλες με κενά ή μη μεταξύ τους, γεγονός που σε μεγάλο βαθμό εξαρτάται από τη μέθοδο ευθυγράμμισης που χρησιμοποιείται. Τις μεθόδους ευθυγράμμισης θα μπορούσαμε να τις διακρίνουμε σε δύο μεγάλες κατηγορίες: i) εκείνη της **ευθυγράμμιση αλληλουχιών ανά ζεύγος** και ii) εκείνη της **πολλαπλής ευθυγράμμισης αλληλουχιών** [29].

β) Ταυτότητα (identity) αλληλουχιών: Με τον όρο «ταυτότητα» αναφερόμαστε στην έκταση εκείνη μεταξύ των δύο αλληλουχιών, όπου παρατηρείται πλήρης ταύτιση [29, 37].

γ) Ομοιότητα (similarity) και ομολογία (homology) αλληλουχιών: Ο πρώτος όρος σχετίζεται με την ποιότητα ευθυγράμμισης των αλληλουχιών, όπως για παράδειγμα ο αριθμός και το μήκος των κενών, ενώ ο δεύτερος χρησιμοποιείται για αλληλουχίες που έχουν συγχρόνως υψηλή «ταυτότητα» και «ομοιότητα». Ο όρος ομολογία χρησιμοποιείται μόνο όταν ο κοινός πρόγονος είναι αρκετά πρόσφατος, ώστε οι εκάστοτε αλληλουχίες να έχουν διατηρήσει αρκετή ομοιότητα για να χρησιμοποιηθούν σε φυλογενετική ανάλυση. Πιο συγκεκριμένα, θα μπορούσε να ειπωθεί ότι ο καλύτερος τρόπος για να συγκρίνουμε δύο αλληλουχίες είναι μέσω της στοίχισής τους. Η στοίχιση δύο αλληλουχιών αποδίδει εκτός από την καλύτερη δυνατή ευθυγράμμισή τους κι ένα μέτρο της ομοιότητάς τους που μπορεί να εκφραστεί είτε αρνητικά ως «απόσταση», είτε θετικά ως «βαθμός ομοιότητας» που συνήθως εκφράζεται με τη μορφή ποσοστού. Η ομοιότητα, λοιπόν, ορίζεται ως ένα αντικειμενικά μετρήσιμο μέγεθος που λαμβάνει αριθμητικές τιμές, κυρίως με μορφή ποσοστού, και η ομολογία ως μια ποιοτική ιδιότητα που αναφέρεται στην εξελικτική

σχέση δύο αλληλουχιών χωρίς να ποσοτικοποιείται αριθμητικά. Δύο αλληλουχίες είτε είναι ομόλογες είτε δεν είναι [29, 37].

1.2.2. Μέθοδοι Ευθυγράμμισης Αλληλουχιών

1.2.2.1. Ευθυγράμμιση Αλληλουχιών Ανά Ζεύγος

Οι μέθοδοι με τις οποίες επιτυγχάνεται ευθυγράμμιση αλληλουχιών ανά ζεύγος περιλαμβάνουν τους εξής τέσσερις κλασικούς τύπους :

α) Η μέθοδος «**Brute Force**», η οποία εμπεριέχει τη σύγκριση όλων των δυνατών νουκλεοτιδικών αντικαταστάσεων προκειμένου να εντοπίσει τη καλύτερη δυνατή ευθυγράμμιση [54].

β) Η μέθοδος «**Dot-Matrix**», η οποία εμπεριέχει την εύρεση όλων των αντιστοιχιών των νουκλεοτιδίων μεταξύ δύο αλληλουχιών προσανατολισμένων σε δύο κάθετους μεταξύ τους ΧΥ άξονες, όπου και στη συνέχεια τα ταιριάσματα, δηλαδή τα σημεία όπου ενώνονται δύο νοητές ευθείες που προέρχονται από ίδιου τύπου νουκλεοτίδια διαφορετικών αλληλουχιών, κάθετων ως προς τον άξονα της αλληλουχίας προέλευσης, καταγράφονται στον πίνακα αυτό με τη μορφή τελειών και αυτές ενώνονται διαγώνια προκειμένου να εντοπιστεί το βέλτιστο μονοπάτι [36].

γ) Η μέθοδος «**Dynamic Programming**», η οποία περιλαμβάνει τη σύγκριση κάθε ζεύγους χαρακτήρων μεταξύ δύο αλληλουχιών για τη δημιουργία μιας ευθυγράμμισης που μεγιστοποιεί τον αριθμό των αντιστοιχιών μεταξύ πανομοιότυπων ή σχετικών χαρακτήρων. Διακρίνεται δε, σε «global» και «local». Η πρώτη κατηγορία δημιουργεί μια ευθυγράμμιση δύο αλληλουχιών όπου μεγιστοποιεί τον αριθμό των ταιριασμάτων και ελαχιστοποιεί τον αριθμό των κενών,

χρησιμοποιώντας μια γραμμική λειτουργία ποινής κενού και ενδείκνυται για όμοιες αλληλουχίες ίδιου μήκους μιας και η ευθυγράμμιση γίνεται από το ένα άκρο έως το άλλο. Η δεύτερη δημιουργεί μια ευθυγράμμιση δύο αλληλουχιών που μεγιστοποιεί τις περιοχές ομοιότητας χρησιμοποιώντας μία αφινική λειτουργία ποινής κενού και ενδείκνυται για ανόμοιες αλληλουχίες με διαφορετικά μήκη μιας και η ευθυγράμμιση δε γίνεται υποχρεωτικά από τη μία άκρη έως την άλλη. Πιο συγκεκριμένα, στην πρώτη η βαθμολογία των ποινών που απονέμεται αυξάνεται όταν αυξάνεται ο αριθμός των κενών στην ευθυγράμμιση, ενώ στη δεύτερη εφαρμόζεται ποινή τόσο για την εμφάνιση ενός νέου κενού όσο και για το μήκος του κενού αυτού· μεγαλύτερο μήκος σημαίνει υψηλότερη βαθμολογία ποινής [36].

δ) Τέλος, οι «**Word Methods**» από τη πλευρά τους, οι οποίες αναπτύχθηκαν ως εναλλακτική του «Dynamic Programming» ως πιο υπολογιστικά αποδοτικές, λειτουργούν ψάχνοντας όμοιες «λέξεις» με αυτές που εμπεριέχονται σε ότι τους προσφέραμε, οι οποίες δεν είναι παρά τμήματα αλληλουχιών. Βέβαια, το αποτέλεσμα που προκύπτει δεν είναι απαραίτητως το βέλτιστο. Γνωστά είδη των μεθόδων αυτών είναι τα «BLAST» (Basic Local Alignment Search Tool) και «FASTA», τα οποία και είναι προτιμητέα σε περιπτώσεις που υπάρχει μία αλληλουχία και ψάχνονται παρόμοιες αλληλουχίες σε μια βάση δεδομένων. Το BLAST, ουσιαστικά, διαβάζει την αλληλουχία του οργανισμού ενδιαφέροντος, τις παραμέτρους που ο αναλυτής έχει θέσει και τη βάση δεδομένων όπου γίνεται η αναζήτηση. Παράγει, έπειτα, ένα σύνολο λέξεων που είναι σύντομες και των οποίων το μήκος έχει καθοριστεί από το χρήστη. Χρησιμοποιώντας μια «λέξη» κάθε φορά, αναζητούνται ταιριάσματα στην αντίστοιχη βάση δεδομένων. Μόλις βρεθεί, ένα ταιρίασμα επεκτείνεται έπειτα μέχρι να βρεθεί ένα κενό. Οι τοπικές (local) αυτές ευθυγραμμίσεις βαθμολογούνται. Οι υψηλότερες ευθυγραμμίσεις βαθμολογικά επεκτείνονται στη συνέχεια με κενά. Από αυτές τις μεγαλύτερες ευθυγραμμίσεις, σώζονται μόνο εκείνες μιας

συγκεκριμένης βαθμολογίας. Στο τέλος, όταν η βαθμολογία πέσει σε επίπεδα που δεν κρίνεται ορθό να συνεχιστεί η ευθυγράμμιση των αλληλουχιών αυτών εξαιτίας των κενών ή και αναντιστοιχιών, η επέκταση τερματίζεται και η ευθυγράμμιση αναδιπλώνεται για να αντικατοπτρίζει τη βαθμολογία όταν ήταν υψηλότερη, δηλαδή όταν υπήρχε το ζεύγος τμημάτων με την υψηλότερη βαθμολογία, ενώ λαμβάνει χώρα και ο υπολογισμός της σημαντικότητας για κάθε ζεύγος τμημάτων. Ο αλγόριθμος FASTA αρχικά βρίσκει όλες τις πιθανές αλληλοεπικαλυπτόμενες «λέξεις» της αλληλουχίας-ερωτήματος, το μήκος της οποίας καθορίζεται από τον χρήστη. Στη συνέχεια αυτές συνταιριάζονται με ίδιου μήκους «λέξεις» στις αλληλουχίες που βρίσκονται στη βάση δεδομένων αναζήτησης. Οι αλληλουχίες της βάσης δεδομένων που έχουν κοινές «λέξεις» με την αλληλουχία ερωτήματος επεκτείνονται κατά μήκος της αλληλουχίας-ερωτήματος και λαμβάνει χώρα η βαθμολόγηση [12, 24, 29, 36, 52].

1.2.2.2. Πολλαπλή Ευθυγράμμιση Αλληλουχιών

Αναφορικά, τώρα, με τη δεύτερη μεγάλη κατηγορία ευθυγράμμισης αλληλουχιών, εκείνη της πολλαπλής ευθυγράμμισης αλληλουχιών, μπορεί να αναφερθεί επιγραμματικά ότι είναι ακριβώς αυτό που δηλώνει, δηλαδή η ευθυγράμμιση τριών ή περισσότερων αλληλουχιών αρκεί να είναι όλες μεταξύ τους το ίδιο είδος. Στο σημείο αυτό κρίνεται ορθό να γίνει αναφορά στις υβριδικές μεθόδους που χρησιμοποιούνται, ως επί το πλείστον, στην πολλαπλή ευθυγράμμιση και οι οποίες είναι ο συνδυασμός «global» και «local», γνωστές ως «semi-global» ή «glocal». Προκειμένου να δημιουργηθεί μια πολλαπλή ευθυγράμμιση, πρώτα απ' όλα, πρέπει να οργανωθούν οι αλληλουχίες που είναι επιθυμητό να ευθυγραμμιστούν, να ευθυγραμμιστούν χρησιμοποιώντας μια μέθοδο πολλαπλής ευθυγράμμισης, με πιο δημοφιλείς σε αυτή την

κατηγορία να είναι οι «Clustal» και «Muscle», και ενδεχομένως κάποιες φορές, έπειτα από επιθεώρηση των αποτελεσμάτων, να εντοπιστούν τα υπάρχοντα κενά, να αφαιρεθούν τμήματα ή και αλληλουχίες και να επαναληφθούν, ενδεχομένως, τα προηγούμενα δύο βήματα, μέχρι να εντοπιστεί αυτό που ονομάζεται διατηρημένο κομμάτι αλληλουχίας. Ένα επιπλέον βήμα που ακολουθείται συχνά σε αυτή τη κατηγορία ευθυγραμμίσεων είναι αυτό της αφαίρεσης τμημάτων αλληλουχίας που επαναλαμβάνονται προκειμένου να μη δημιουργηθεί κακής ποιότητας αποτέλεσμα-συμπέρασμα. Ένα από τα πλέον συνηθισμένα εργαλεία που χρησιμοποιούνται για αυτή τη διαδικασία είναι το «RepeatMasker». Όσον αφορά τώρα μεθόδους πολλαπλής ευθυγράμμισης που εμπεριέχουν «Dynamic Programming», εν αντιθέσει με την ευθυγράμμιση ζεύγους που είχε λειτουργία βαθμολόγησης βάσει ταιριάσματος, αναντιστοιχίας και κενού, χρησιμοποιούν λειτουργίες βαθμολόγησης που καταγράφουν τη συχνότητα των προαναφερθέντων κατά μήκος των διάφορων αλληλουχιών που ευθυγραμμίζονται [5, 13, 17, 36, 46].

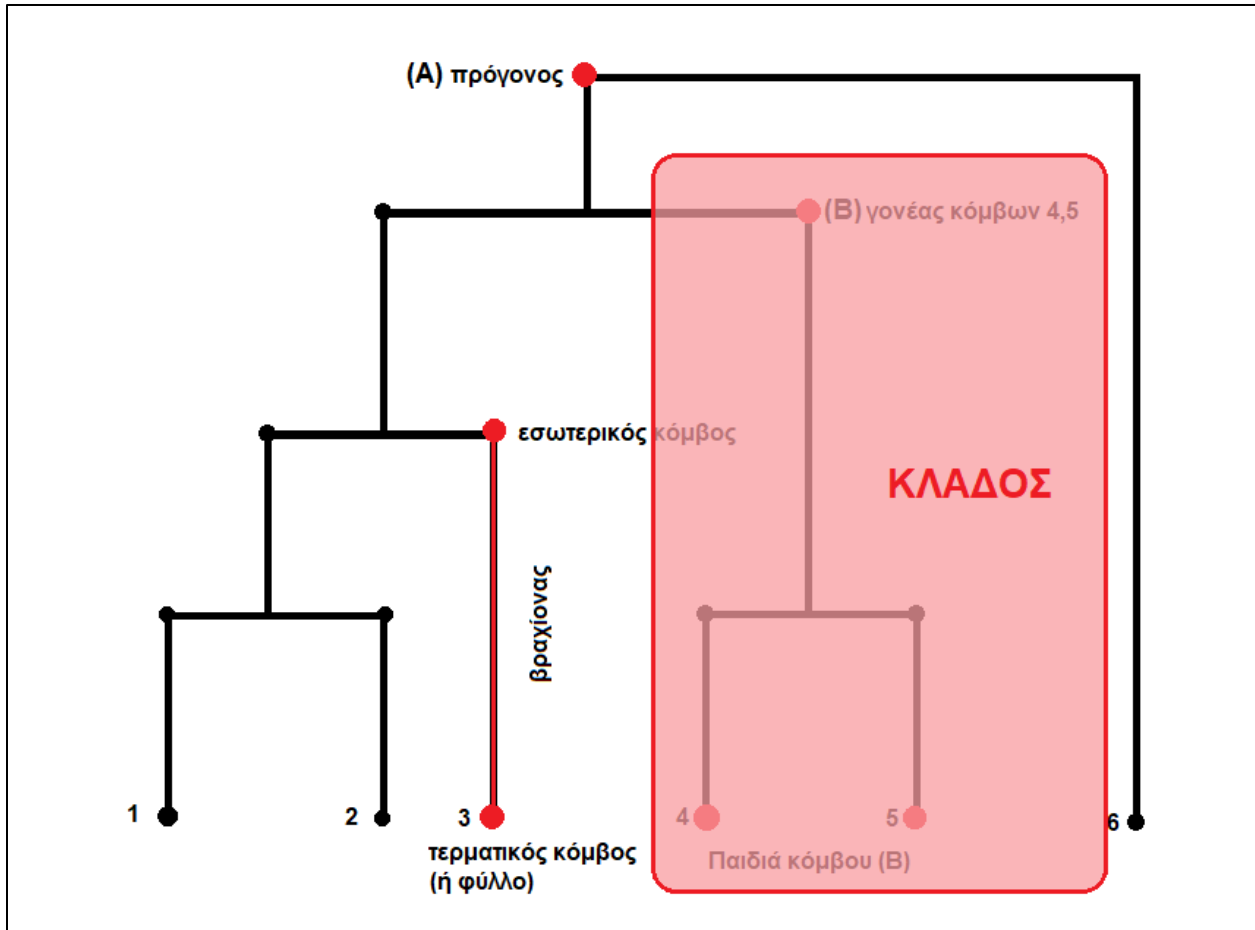
Βέβαια, στην περίπτωση των πολλαπλών ευθυγραμμίσεων έχει πλέον εισαχθεί και το «Progressive Dynamic Programming», το οποίο έχει την ιδιότητα να μειώνει την πολυπλοκότητα σε σχέση με το «Dynamic Programming» και να μειώνει τον απαιτούμενο χρόνο με αυτό τον τρόπο. Το «Progressive Dynamic Programming», λοιπόν, ξεκινά με ευθυγράμμιση όλων των δυνατών ζευγών παρέχοντας μια βαθμολογία ομοιότητας για κάθε επιμέρους ευθυγράμμιση. Στη συνέχεια αυτές οι βαθμολογίες αξιοποιούνται προς κατασκευή ενός «δέντρου οδηγού» (guide tree), ενώ στο τελικό στάδιο η συστηματική κατασκευή των πολλαπλών ευθυγραμμίσεων λαμβάνει χώρα με χρήση «Dynamic Programming», ξεκινώντας με την ευθυγράμμιση του πιο κοντινού βαθμολογικά ζεύγους. Γενικότερα, το «Progressive

Dynamic Programming» ενδεχομένως να μπορούσε εννοιολογικά να ενταχθεί στη κατηγορία του λεγόμενου «Progressive Alignment» [45].

1.2.3. Κατασκευή Φυλογενετικών Δέντρων

1.2.3.1. Ορισμοί

Προκειμένου να γίνεται σαφής η οποιαδήποτε αναφορά σε τμήματα του φυλογενετικού δέντρου χρησιμοποιούνται ευρέως κάποιες ορολογίες. Έτσι, οι γραμμές του δέντρου ονομάζονται βραχίονες (branches) και τα σημεία που ενώνονται οι βραχίονες αλλά και οι αλληλουχίες που απεικονίζονται ονομάζονται «κόμβοι» (nodes). Πιο συγκεκριμένα τα πρώτα ονομάζονται «εσωτερικοί κόμβοι» (internal nodes) και τα δεύτερα «τερματικοί κόμβοι» (end-nodes), συχνά αποκαλούμενα και «φύλλα» του φυλογενετικού δέντρου. Ένας κόμβος ο οποίος είναι υψηλότερα στην ιεραρχία κατά ένα βήμα σε σχέση με έναν άλλο κόμβο είναι ο «γονέας» (parent) του δεύτερου, ενώ οι κόμβοι που μοιράζονται τον ίδιο γονέα ονομάζονται «παιδιά» (children). Ο «κλάδος» (clade) περιλαμβάνει έναν «εσωτερικό κόμβο» και όλους τους «τελικούς κόμβους» που προκύπτουν από αυτόν. Οι «τελικοί κόμβοι» δεν έχουν «παιδιά», ενώ ο «εσωτερικός κόμβος» εκείνος, εάν υπάρχει, που ενώνεται με όλους τους κατώτερους ιεραρχικά «κόμβους» ονομάζεται «πρόγονος» (ancestor). Όλα τα προαναφερθέντα σχετίζονται με τη δομή και οργάνωση του εκάστοτε φυλογενετικού δέντρου, γνωστή και ως «τοπολογία» (topology) [11, 13].



Εικόνα 1.1.: Τοπολογία Φυλογενετικού Δέντρου

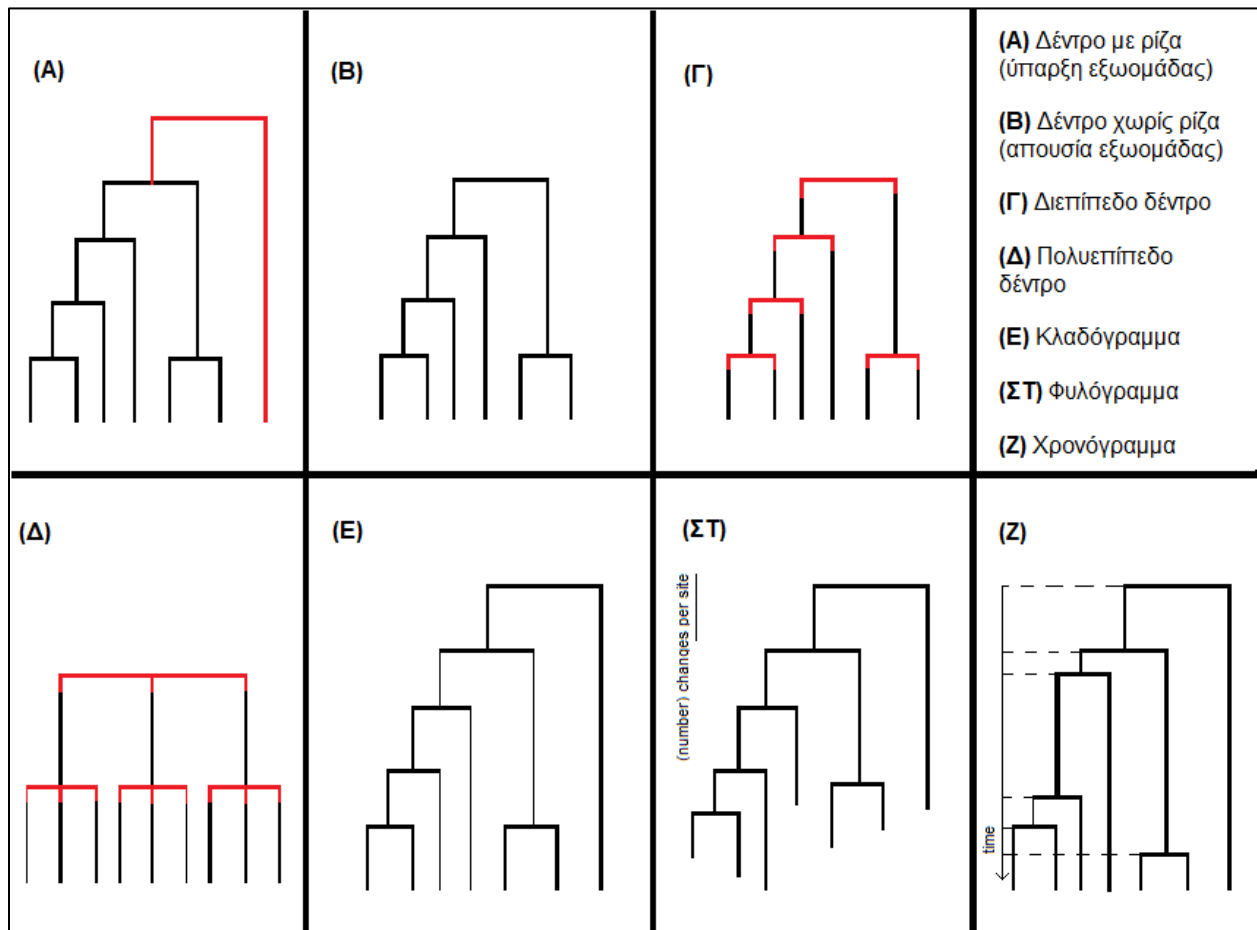
1.2.3.2. Τύποι Φυλογενετικών Δέντρων

Υπάρχουν διάφοροι τύποι φυλογενετικών δέντρων και μπορούν να ομαδοποιηθούν αναφορικά με τα χαρακτηριστικά τους. Πρώτη και βασική διαφοροποίηση είναι η ύπαρξη ή μη «ρίζας» (root). Είθισται συχνά σε κάποια φυλογενετικά δέντρα να συμπεριλαμβάνεται μια επιπλέον ταξινομική μονάδα σε σχέση με εκείνες που είναι άμεσου ενδιαφέροντος με σκοπό τη δημιουργία μιας ρίζας στο δέντρο. Η ταξινομική αυτή μονάδα καλείται «εξωομάδα» (outgroup). Η προσθήκη της προκαλεί την εμφάνιση ενός επιπλέον κόμβου στο φυλογενετικό δέντρο. Ο

κόμβος αυτός είναι η λεγόμενη «ρίζα», δηλαδή ο κοινός πρόγονος των οργανισμών που έχουν χρησιμοποιηθεί για την κατασκευή του φυλογενετικού δέντρου. Πιο συγκεκριμένα, αναφορικά με την επιλογή της εξωομάδας, μπορεί να αναφερθεί ότι είναι μια ταξινομική μονάδα που είναι γνωστό ότι απέχει εξελικτικά αρκετά από όλα τα στοιχεία που αναλύονται στην εκάστοτε περίπτωση, αλλά ταυτόχρονα είναι αρκετά κοντά ούτως ώστε να βρίσκεται κάτω από έναν κοινό πρόγονο. Η τακτική αυτή, με την προσθήκη εξωομάδας ακολουθείται συχνά στα NJ δέντρα. Όσα δεν έχουν ρίζα, ουσιαστικά αναπαριστούν τις σχέσεις μεταξύ των τελικών κόμβων χωρίς να γίνονται εικασίες σχετικά με την καταγωγή και προφανώς δε συμπεριλαμβάνουν κάποιο πρόγονο. Όσα, πάλι, έχουν ρίζα είναι καθορισμένες δομές δεδομένων οι οποίες ακολουθούν ένα ανιχνεύσιμο μονοπάτι μεταξύ όλων των κόμβων και ενός κοινού πρόγονου. Τα φυλογενετικά δέντρα μπορούν, επίσης, να είναι διεπίπεδα ή πολυεπίπεδα, με ετικέτες ή χωρίς, καθώς επίσης και να έχουν μία απεικονίσιμη κλίμακα ή όχι. Ένα διεπίπεδο δέντρο έχει ακριβώς δύο απογόνους οι οποίοι προκύπτουν από κάθε εσωτερικό κόμβο, ενώ ένα πολυεπίπεδο δέντρο έχει περισσότερους από δύο. Όσα δέντρα έχουν ετικέτες έχουν ακριβείς πληροφορίες, ή έστω ένα σημείο αναφοράς (π.χ. αριθμό) σε κάθε τελικό ή και εσωτερικό κόμβο, σε αντίθεση με όσα δεν έχουν. Τα δέντρα με κλίμακα έχουν μήκη κλαδιών αναλογικά με μια συγκεκριμένη μονάδα, η οποία συχνά είναι ο εξελικτικός χρόνος [3, 13, 18, 40].

Σήμερα, βέβαια, με την εμφάνιση των τεχνικών αλληλούχισης και την ευρεία χρήση της κλαδιστικής, η χρήση φυλογενετικών δέντρων έχει διεισδύσει σε πολλούς κλάδους της βιολογίας και πέραν αυτής. Με τον όρο «κλαδιστική» (cladistics) γίνεται αναφορά στην ταξινόμηση οργανισμών βάση κοινών χαρακτηριστικών τα οποία προέρχονται από έναν κοινό πρόγονο και δεν είναι παρόντα πριν από αυτόν. Τα φυλογενετικά δέντρα που προκύπτουν με τη χρήση της κλαδιστικής ονομάζονται «κλαδογράμματα» (cladograms) και ουσιαστικά είναι μία εκτίμηση

της φυλογένειας των συγκρινόμενων αλληλουχιών. Το μήκος των κλαδιών είναι ίδιο για τις διάφορες αλληλουχίες. Αναπαριστάται, λοιπόν, ένας κοινός πρόγονος, αλλά όχι η σύγκριση εξελικτικού χρόνου μεταξύ των αλληλουχιών. Εάν, πάλι γίνεται αναφορά σε «φυλόγραμμα» (phylogram), τότε και πάλι ουσιαστικά είναι μία σύγκριση των φυλογενειών των συγκρινόμενων αλληλουχιών, με τη διαφορά ότι τα μήκη των διακλαδώσεων είναι ανάλογα με το ποσό της εκάστοτε παρατηρούμενης εξελικτικής αλλαγής, ενώ συχνά σε αντίστοιχα διαγράμματα τα οποία αντιπροσωπεύουν εξελικτικό χρόνο χρησιμοποιείται και ο όρος «χρονόγραμμα» (chronogram) [11, 40, 41, 42, 43, 51].



Εικόνα 1.2.: Είδη Φυλογενετικών Δέντρων

Επίσης, αξίζει να σημειωθεί ότι κάθε τοπολογία στην οποία καταλήγουμε μετά τη χρήση συγκεκριμένων κριτηρίων, την οποία και κρίνουμε αληθή, δεν είναι παρά μια υπόθεση μεταξύ πολλών, ενώ το πλήθος των πιθανών υποθέσεων μπορεί να προσδιοριστεί εάν γνωρίζουμε τον αριθμό των ειδών που θέλουμε να συμπεριλάβουμε και εάν το δέντρο είναι με ρίζα ή χωρίς. Επομένως, δεν είναι εφικτό να γνωρίζουμε με σιγουριά εάν ένα δέντρο που κατασκευάσαμε είναι το σωστό, τουλάχιστον εξελικτικά [11, 49].

Επί προσθέτως, μια κύρια παραδοχή πολλών φυλογενετικών αναλύσεων είναι ότι ο ρυθμός με τον οποίο συμβαίνουν οι αντικαταστάσεις είναι σταθερός. Η υπόθεση αυτή είναι γνωστή και ως «υπόθεση του μοριακού ρολογιού» (molecular clock hypothesis). Η ιδέα του σταθερού ρυθμού αντικαταστάσεων βρίσκεται στη βάση τόσο των υπολογισμών αποστάσεων, όσο και της δημιουργίας φυλογενετικών δέντρων με διάφορες μεθόδους. Ωστόσο, η χρήση της χρειάζεται προσοχή, καθώς δεν ισχύει σε απόλυτο βαθμό [36].

1.2.3.3. Μέθοδοι Κατασκευής Φυλογενετικών Δέντρων

Οι μέθοδοι κατασκευής φυλογενετικών δέντρων μπορούν να διαχωριστούν στις εξής δύο κατηγορίες:

α) Αυτές που βασίζονται στην «απόσταση» (distance): σε αυτές οι πιο γνωστές που συμπεριλαμβάνονται είναι οι UPGMA και η Neighbor-Joining [13, 36].

β) Αυτές που βασίζονται στο «χαρακτήρα» (character): σε αυτές συμπεριλαμβάνονται η Μέγιστη Φειδωλότητα (Maximum Parsimony ή MP) και η Μέγιστη Πιθανοφάνεια (Maximum Likelihood ή ML) [13, 36].

Στις αναφερθείσες μεθόδους της πρώτης κατηγορίας χρησιμοποιείται μια ομάδα κριτηρίων γνωστή ως «Clustering» και σε εκείνες της δεύτερης κατηγορίας χρησιμοποιείται η ομάδα κριτηρίων «Optimality Criterion» [13, 36].

Όταν γίνεται αναφορά στο «Clustering», ουσιαστικά γίνεται αναφορά σε μια μέθοδο η οποία οδηγεί στη δημιουργία φυλογενετικού δέντρου συνδυάζοντας ζεύγη αλληλουχιών τα οποία έχουν τη μικρότερη απόσταση μεταξύ τους, γεγονός που γίνεται σε κάθε βήμα ξανά και ξανά. Το θετικό σε αυτή τη μέθοδο είναι ότι υπολογιστικά είναι αποδοτική ως προς το χρόνο τον οποίο χρειάζεται. Ωστόσο, δε σημαίνει ότι οδηγεί στη βέλτιστη λύση, καθώς η τοπολογία του δέντρου που κατασκευάζεται βασίζεται στη σειρά με την οποία λαμβάνει μέρος η σύγκριση των αλληλουχιών. Από την άλλη πλευρά, η μέθοδος «Optimality Criterion» εμπεριέχει τη δημιουργία όλων των πιθανών δέντρων έμμεσα και στη συνέχεια την επιλογή εκείνου του δέντρου το οποίο ικανοποιεί καλύτερα το επιλεγμένο σύστημα βαθμολόγησης-αξιολόγησης. Αν και η αξιολόγηση όλων των πιθανών δέντρων δείχνει, σε πρώτη φάση, να της δίνει προβάδισμα ως προς τη μέθοδο «Clustering», οδηγεί συχνά σε ένα μη αποδοτικό υπολογιστικά σύστημα, υπερβολικά αργό ακόμα και σε σχετικά μικρό αριθμό αλληλουχιών [36].

Οι περισσότερες από τις μεθόδους που αναφέραμε υπάρχουν διαθέσιμες σε πακέτα λογισμικού, τα οποία είναι ελεύθερα διαθέσιμα στον τελικό χρήστη. Από τα πιο γνωστά και παλιά είναι το PAUP και το PHYLIP. Ωστόσο, τα τελευταία χρόνια υπάρχουν νέες προσθήκες με ιδιαίτερα ευκολόχρηστα πακέτα λογισμικού με μεγάλες ικανότητες ανάλυσης κάτω από διαφορετικά μοντέλα και προϋποθέσεις. Το MEGA (Molecular Evolutionary Genetic Analysis) είναι ένα από αυτά και θεωρείται από τα πλέον χρησιμοποιημένα τα τελευταία χρόνια πακέτα φυλογενετικής ανάλυσης. Ενσωματώνει μεθόδους αποστάσεων, μέγιστης πιθανοφάνειας και

φειδωλότητας, ενώ ένα από τα χαρακτηριστικά του που το καθιστούν εύχρηστο είναι το γεγονός ότι λειτουργεί σε περιβάλλον Windows με παραθυρική διεπαφή [13, 26, 36].

Οι μέθοδοι που βασίζονται στην «απόσταση» είναι εκείνες που χρησιμοποιούν την απόσταση μεταξύ των αλληλουχιών για να κατασκευάσουν το φυλογενετικό δέντρο. Οι αποστάσεις αυτές προκύπτουν κατά την ευθυγράμμιση των αλληλουχιών, ενώ πιο συγκεκριμένα η απόσταση ανάμεσα σε δύο αλληλουχίες είναι η ποσοτική έκφραση του πλήθους των νουκλεοτιδίων τους που δεν ταυτίζονται έπειτα της ευθυγράμμισης. Στον υπολογισμό των αποστάσεων μπορούν να χρησιμοποιηθούν μοντέλα αξιολόγησης υποκατάστασης, με ένα από τα πιο γνωστά να είναι το PAM, ενώ κάποια από τα πιο γνωστά προγράμματα που χρησιμοποιούνται σε αυτή την κατηγορία για πολλαπλή ευθυγράμμιση αλληλουχιών DNA είναι το DNADIST, το οποίο και μπορεί να βρεθεί στο διαδίκτυο ως μέρος του πακέτου PHYLIP. Η μέθοδος UPGMA, ή αλλιώς «Unweighted Pair Group Method with Arithmetic Averages», ξεκινά από έναν πίνακα αποστάσεων. Έχει τη δυνατότητα να παράγει γρήγορα, ικανοποιητικά δέντρα για μεγάλο αριθμό αλληλουχιών, αλλά εμπεριέχει ορισμένους αρκετά δεσμευτικούς περιορισμούς. Ειδικότερα, ισχύει η υπόθεση του μοριακού ρολογιού, ενώ τα δέντρα που παράγονται είναι με ρίζα. Προκειμένου, λοιπόν, να αποφεύγονται λάθη στις περιπτώσεις εκείνες όπου δεν ισχύει η υπόθεση του μοριακού ρολογιού εΐθισται να χρησιμοποιείται η μέθοδος αποστάσεως Neighbor Joining (NJ) η οποία και οδηγεί στη δημιουργία αντίστοιχων δέντρων. Η μέθοδος NJ σε αντίθεση με την UPGMA αποδίδει δέντρα χωρίς ρίζα [13, 36, 47].

Όσον αφορά τις μεθόδους που βασίζονται στο «χαρακτήρα», προφανώς και δε χρησιμοποιούν την απόσταση, αλλά κάνουν απευθείας χρήση της πολλαπλής ευθυγράμμισης αλληλουχιών, αξιοποιώντας τα σημεία που δεν υπάρχει ταύτιση αλληλουχιών προκειμένου να κατασκευάσουν το πιο πιθανό δέντρο. Έχοντας λύσει, λοιπόν, το πρόβλημα του υπολογισμού

των αποστάσεων, στρεφόμεστε τώρα στο πρόβλημα της ανασύστασης των εξελικτικών σχέσεων μεταξύ αλληλουχιών. Μια φυλογενετική ανάλυση σε μοριακό επίπεδο αντιμετωπίζει το πρόβλημα της προσπάθειας ανασύστασης της πιθανότερης εξελικτικής διαδρομής βιομορίων με βάση τη σύγκριση της πρωτοταγούς τους αλληλουχίας. Δυστυχώς για τους αναλυτές, η απλή παρατήρηση των αντικαταστάσεων σε μια στοίχιση δεν αντιστοιχεί στον αριθμό των πραγματικών αντικαταστάσεων που έχουν όντως συμβεί. Σημαντικότερη παράμετρος είναι επίσης και το είδος των αντικαταστάσεων (ουδέτερες, συνώνυμες, μη-συνώνυμες (με λάθος νόημα (mis-sense) και χωρίς νόημα (non-sense))) που παρατηρούνται και οι επιπτώσεις που αυτές επιφέρουν στον εκάστοτε οργανισμό. Ωστόσο, για την παρούσα εργασία δε κρίνεται σκόπιμο να επεκταθούμε σε αυτό. Αξίζει, ωστόσο, να σημειωθεί ότι από τους διάφορους μηχανισμούς που δρουν στους οργανισμούς στην πορεία του χρόνου συμπεριλαμβανομένου μεταλλάξεων, διπλασιασμό γονιδίων και αναδιοργάνωση γονιδιώματος, και οι οποίοι οδήγησαν σε αυτό που σήμερα καλείται βιοποικιλότητα, οι σημειακές αυτές μεταλλάξεις είναι αυτές που χρησιμοποιούνται ως επί το πλείστον για να εξηγήσουν τις σχέσεις μεταξύ διαφόρων γονιδίων. Έτσι, θεωρούνται η βάση της εξέλιξης των αλληλουχιών και κατ' επέκταση όλων των ζωντανών οργανισμών. Γενικότερα αυτές είναι τυχαίες, αλλά, όπως έμμεσα προαναφέρθηκε, οι θέσεις στις οποίες θα συμβούν και με βάση τις επιπτώσεις στο προϊόν της αλληλουχίας στην οποία συμβαίνουν, υπάρχει διαφορετική πιθανότητα διατήρησής τους [13, 36, 48].

Λαμβάνοντας υπόψη τα είδη των αντικαταστάσεων αλλά και την πιθανότητα παρατήρησής τους μέσω μιας πολλαπλής στοίχισης, γίνεται σαφές ότι η εκτίμηση της εξελικτικής απόστασης δύο αλληλουχιών απαιτεί μαθηματικά μοντέλα που μας επιτρέπουν να υπολογίσουμε την απόσταση αυτή, με την καλύτερη δυνατή εκτίμηση του πραγματικού και όχι του φαινομενικού αριθμού αντικαταστάσεων [13, 36].

Οι μέθοδοι «optimality criterion» αποτελούν μια κατηγορία πιο εκλεπτυσμένων μεθόδων που βασίζονται στην προσπάθεια εύρεσης μιας βέλτιστης λύσης σε ένα αντίστοιχο πρόβλημα. Οι πιο ευρέως χρησιμοποιούμενες είναι οι μέθοδοι «Maximum Parsimony» και «Maximum Likelihood». Σε αντίθεση με τα ονόματά τους είναι στην ουσία μέθοδοι που προσπαθούν να ελαχιστοποιήσουν μια ποσότητα. Στη μέθοδο MP η ποσότητα αυτή είναι ο αριθμός των αντικαταστάσεων καταλοίπων, ενώ στην ML η ποσότητα που πρέπει να ελαχιστοποιηθεί είναι η διαφορά μεταξύ των παρατηρούμενων θέσεων στο δέντρο και μιας πιθανότητας που προκύπτει από ένα μοντέλο. Σαφέστατα και είναι εξαιρετικά απαιτητικές υπολογιστικά, γι' αυτό και χρησιμοποιούν αλγόριθμους όπως οι «branch and bound» και «divide and conquer» προκειμένου να απορρίψουν ευκολότερα κάποιες πιθανές τοπολογίες. Ωστόσο είναι γεγονός ότι είναι και πιο άρτιες από βιολογικής άποψης [9, 13, 33, 36].

1.2.3.4. Επιπλέον Παράμετροι Κατασκευής Φυλογενετικών Δέντρων

Επίσης, αναφορικά με τα μοντέλα νουκλεοτιδικών αντικαταστάσεων που προσφέρονται στο πρόγραμμα MEGA-X για τη μέθοδο ML, των οποίων και θα γίνει χρήση για τους σκοπούς της παρούσας εργασίας, αυτά είναι το General Time Reversible (GTR), Tamura-Nei, Hasegawa-Kishino-Yano, Tamura Three-Parameter, Kimura Two-Parameter, Tajima-Nei και Jukes-Cantor, με το General Time Reversible μοντέλο να προτείνεται ως επί το πλείστον από τη βιβλιογραφία για αντίστοιχα ML φυλογενετικά δέντρα [26].

Το τελικό στάδιο μιας φυλογενετικής ανάλυσης είναι αυτό της εκτίμησης των φυλογενετικών δέντρων, δηλαδή του ελέγχου της πιστότητάς τους. Έλεγχοι μπορούν προφανώς να διενεργηθούν και στα υπόλοιπα στάδια μιας ανάλυσης, όπως είναι η στοίχιση και ο

υπολογισμός των αποστάσεων, αλλά η κατασκευή του δέντρου είναι αυτό που περιέχει τη μεγαλύτερη πιθανότητα σφάλματος, ακριβώς εξαιτίας της επαγωγικής διαδικασίας και της πολυπλοκότητας των υπολογισμών που περιλαμβάνει. Οι έλεγχοι πιστότητας σε δέντρα μπορούν να είναι αναλυτικοί ή στατιστικοί. Ως αναλυτικούς ορίζουμε αυτούς που βασίζονται σε αντικειμενικά κριτήρια όπως το ελάχιστο μήκος κλάδων ή ο ελάχιστος αριθμός αντικαταστάσεων. Γενικά, οι αναλυτικοί έλεγχοι είναι εξειδικευμένοι ανάλογα με το είδος της φυλογενετικής ανάλυσης που έχει προηγηθεί. Οι στατιστικοί έλεγχοι, από την άλλη, είναι όλοι εκείνοι που προσπαθούν να αποτιμήσουν τη σημασία ενός δέντρου συγκρίνοντάς το με δέντρα που θα μπορούσαν να έχουν προκύψει τυχαία. Η κυριότερη από τις μεθόδους στατιστικού ελέγχου είναι η μέθοδος bootstrap. Ακριβώς, λοιπόν, επειδή είναι συνήθης η έλλειψη ακράδαντων στοιχείων έχει αναπτυχθεί η λεγόμενη bootstrap μέθοδος, η οποία ουσιαστικά προσπαθεί να αξιολογήσει πιθανότητες. Η μέθοδος αυτή αρχικά κατασκευάζει ένα δέντρο και κρατάει την τοπολογία που προκύπτει. Στη συνέχεια από την αρχή κατασκευάζει ξανά ένα φυλογενετικό δέντρο μόνο που σε αυτή την περίπτωση έχει βγάλει εκτός κάποια αλληλουχία. Αυτό επαναλαμβάνεται για όλες τις συγκρινόμενες αλληλουχίες και συλλέγονται οι τοπολογίες που προκύπτουν, οι οποίες και στη συνέχεια συγκρίνονται με την αρχική. Σε τελικό στάδιο, ως πιο αξιόπιστο δέντρο ορίζεται εκείνο του οποίου η τοπολογία αναπαράγεται στην πλειοψηφία των περιπτώσεων [13, 36].

Στις περιπτώσεις, λοιπόν, που έχει χρησιμοποιηθεί η μέθοδος bootstrap, στο πακέτο λογισμικού MEGA-X, το αποτέλεσμα της διαδικασίας κατασκευής των ML φυλογενετικών δέντρων είναι ένα «Original tree» και ένα «Bootstrap consensus tree». Τα Original trees συχνά δημοσιεύονται από μοριακούς βιολόγους ως φυλογενετικά δέντρα. Επειδή τα δέντρα αυτά, ωστόσο, αναπαριστούν απλώς τη μοναδική τοπολογία που κρίνεται ως καλύτερη, χωρίς να έχει

γίνει χρήση του προαναφερθέντος στατιστικού τεστ (bootstrap), κρίνεται ορθό να μην αποτελούν αυτά το τελικό αποτέλεσμα της έρευνας, αλλά τα Bootstrap consensus trees [13, 27, 36].

Επίσης, αναφορικά με την ευθυγράμμιση των αλληλουχιών που έχουν επιλεγεί ως εκείνες από τις οποίες θα δημιουργηθεί το φυλογενετικό δέντρο, αυτή γίνεται, όπως έχει προαναφερθεί, με χρήση συνήθως είτε του εργαλείου «Clustal» είτε του «Muscle». Αυτά, τουλάχιστον είναι τα πιο διαδεδομένα και καθιερωμένα για το συγκεκριμένο σημείο της φυλογενετικής ανάλυσης. Εντοπίζονται είτε μόνα τους, είτε συνοδευόμενα από τον όρο «codons». Η χρήση της τελευταίας αναφερθείσας εκδοχής περιορίζεται στις περιπτώσεις που υπάρχουν αλληλουχίες που κωδικοποιούν για πρωτεΐνες, προκειμένου να μη δημιουργηθεί οποιοδήποτε κενό κατά την ευθυγράμμιση χωρίς να λαμβάνεται υπόψη η εκάστοτε τριπλέτα που κωδικοποιεί για αμινοξύ. Στις υπόλοιπες περιπτώσεις χρησιμοποιείται η απλή εκδοχή [36].

1.2.3.5. Βάσεις Δεδομένων

Τέλος, η χρήση της βάσης δεδομένων NCBI είναι από τις πλέον καθιερωμένες σε αντίστοιχες εργασίες, χωρίς αυτό να σημαίνει ότι δεν υπάρχουν και άλλες εξίσου αξιόπιστες, όπως η European Molecular Biology Laboratory (EMBL) και η DNA Databank of Japan (DDBJ). Επίσης, δε σημαίνει το ότι επειδή μια βάση δεδομένων εστιάζει παραδείγματος χάρη σε συγκεκριμένες αλληλουχίες είναι καλύτερη σε σχέση με κάποια άλλη. Η χρήση της εκάστοτε βάσης δεδομένων σχετίζεται με το τι είναι πιο βολικό για τον χρήστη, κυρίως με γνώμονα το είδος της αλληλουχίας που τον ενδιαφέρει και την συμβατότητά της με το αντίστοιχο πρόγραμμα. Σε περίπτωση που είναι επιθυμητή η κατάθεση κάποιας αλληλουχίας, το NCBI

διαθέτει τα ευκολόχρηστα εργαλεία «BankIt» για μικρές αλληλουχίες και «Sequin» για μεγάλες. Υπάρχει βέβαια και η δυνατότητα κατάθεσης μέσω email, ενώ είθισται να χρησιμοποιείται ο κώδικας IUPAC. Οι πληροφορίες αποθηκεύονται σε μορφή που κάνει χρήση ASCII (American Standard Code for Information Interchange). Υπάρχουν περισσότερες της μιας μορφές αποθήκευσης. Οι πιο δημοφιλείς για αλληλουχίες DNA είναι τα GenBank και FASTA, ενώ αν χρειάζεται να γίνει μετατροπή από μια μορφή αποθήκευσης σε μια άλλη μπορεί να χρησιμοποιηθεί το πρόγραμμα «READSEQ». Είναι γεγονός, ωστόσο, ότι ο χρήστης οφείλει να είναι ιδιαίτερα προσεκτικός ως προς το περιεχόμενο το οποίο καταθέτει, καθώς δεν υφίστανται λεπτομερείς έλεγχοι διασφάλισης ποιότητας και πρόκειται για ιστοσελίδα που χρησιμοποιείται από το διεθνές επιστημονικό κοινό, με μεγαλύτερους κινδύνους να εμφανίζονται σε περιπτώσεις λανθασμένης εισαγωγής ή διαγραφής βάσεων σε συγκρίσεις αλληλουχιών σε πρωτεϊνικό επίπεδο. Ένα από τα θετικά του NCBI είναι ότι προσφέρει πολλές εναλλακτικές αναζήτησης σε διάφορες επιμέρους βάσεις δεδομένων με δυνατότητα εστίασης στον άνθρωπο ή σε διάφορα ζώα, ενώ κάθε δημοσίευση παρέχει πληθώρα χρήσιμων πληροφοριών με κυριότερα τον γενετικό τόπο, το μήκος της αλληλουχίας σε ζεύγη βάσεων, τον τύπο της αλληλουχίας, την ημερομηνία ανάρτησης και άλλα πολλά [4, 8, 10, 30, 31].

1.2.3.6. Επιλογή Τρόπου Κατασκευής Φυλογενετικού Δέντρου

Γενικότερα, αξίζει να σημειωθεί, πως δεν είναι ορθό να ειπωθεί ότι μια μέθοδος είναι η καλύτερη για όλες τις περιστάσεις. Η μέθοδος επιλογής εξαρτάται τόσο από το τι είναι επιθυμητό να γίνει γνωστό, όσο και από το μέγεθος και την πολυπλοκότητα του συνόλου δεδομένων. Στην πράξη, εξαρτάται επίσης από την ταχύτητα του υπολογιστή και την ευκολία

εφαρμογής της συγκεκριμένης μεθόδου. Οι μέθοδοι που αναφέρθηκαν στην ενότητα αυτή καλύπτουν την πλειοψηφία των πιο δημοφιλών, ευρέως χρησιμοποιούμενων, έγκυρων, ευκολόχρηστων μεθόδων. Φυσικά υπάρχουν και άλλες, όπως η Μπεϋζιανή μέθοδος (Bayesian method), η οποία βασίζεται στο «χαρακτήρα» και μάλιστα κερδίζει συνεχώς έδαφος στον συγκεκριμένο τομέα βιοπληροφορικής. Ωστόσο, δεν εξυπηρετούσε ικανοποιητικά τους σκοπούς της εργασίας αυτής με βάση το πλάνο το οποίο είχε αποφασιστεί να ακολουθηθεί και θεωρείται μία εκ των πιο πολύπλοκων ως προς τις διάφορες παραμέτρους που απαιτεί να συμπληρωθούν κατά τη διάρκεια της φυλογενετικής ανάλυσης. Αυτό σε έναν άριστο χειριστή-γνώστη των αντίστοιχων μεθόδων κρίνεται ως ιδιαίτερα χρήσιμο. Ωστόσο, στο ευρύτερο ενδιαφερόμενο κοινό ενδέχεται να οδηγήσει στη δημιουργία λανθασμένων φυλογενετικών δέντρων και κατ' επέκταση διάδοση λανθασμένων πληροφοριών, εντός και εκτός επιστημονικού κοινού, γι' αυτό και για τον παρόν σκοπό αποφεύχθηκε [13, 36, 46].

1.3. Παρασιτισμός και ελμινθολογία

Ξεκινώντας με την έννοια που συνοψίζεται στον όρο «παρασιτισμός» μπορεί να αναφερθεί ότι είναι η συμβίωση δύο οργανισμών, υπό την ιδιαιτερότητα ότι ο ένας εκ των δύο συμμετεχόντων οργανισμών βλάπτεται από τις ενέργειες του άλλου. Ο δεύτερος οργανισμός, ο οποίος ζει σε βάρος του πρώτου, καλείται «παράσιτο» και ενδέχεται να είναι οποιοσδήποτε μικροοργανισμός, συμπεριλαμβανομένων ιών, μυκήτων, βακτηρίων και ελμίνθων. Ο οργανισμός στον οποίο, λοιπόν, το παράσιτο φιλοξενείται ονομάζεται «ξενιστής» και διακρίνεται σε «τελικό», όπου και εντοπίζεται η γεννητικά ώριμη μορφή του παρασίτου, και «ενδιάμεσο» όπου εντοπίζεται η γεννητικά άωρη μορφή του. Το σύνολο των διάφορων σταδίων ανάπτυξης του

παρασίτου αποτελεί τον «βιολογικό κύκλο» του. Μία ακόμη σημαντική έννοια είναι αυτή του «μεταφορέα» ή αλλιώς «μεταδότη», ο οποίος και είναι ο οργανισμός εκείνος, ή το αντικείμενο, που μεταδίδει, όπως άλλωστε φανερώνει και το όνομά του, το παράσιτο από ένα ξενιστή σε έναν άλλο. «Βιολογικός» χαρακτηρίζεται ο μεταδότης εκείνος που δίνει τη δυνατότητα στο παράσιτο να εξελιχθεί σε στάδιο δυνητικά μολυσματικό για τον τελικό ξενιστή, το λεγόμενο «μολύνον στάδιο», ενώ «μηχανικός» ονομάζεται ο μεταδότης εντός του οποίου δεν υφίσταται η οποιαδήποτε εξέλιξη του παρασίτου. Ο τελικός ξενιστής, ο οποίος μολύνεται από το παράσιτο και το διατηρεί, καλείται «αποθήκη ή δεξαμενή», ενώ εκείνος που δεν εμφανίζει συμπτώματα μόλυνσης και το διασπείρει «φορέας». Ο άνθρωπος ενδέχεται να μολυνθεί από ένα παράσιτο κατά κύριο λόγο μέσω του δέρματος, του πεπτικού, αναπνευστικού ή γεννητικού συστήματος, γεγονός που μπορεί να οδηγήσει σε καταστάσεις όπως η μηχανική έμφραξη των οργάνων, η καταστροφή κυττάρων και ιστών, ο σχηματισμός όγκων, η ελλιπής ικανότητα του οργανισμού να απορροφήσει θρεπτικές ουσίες, τα τραύματα, η απομύζηση οργανικών υγρών, καθώς επίσης και η μετάδοση αιτιολογικών παραγόντων άλλων λοιμωδών νοσημάτων [56].

Αναφορικά με τη λέξη «έλμινθας», κοινώς σκώληκας, η χρήση της στην παρασιτολογία ισούται με την έννοια «παρασιτικός έλμινθας». Η μελέτη, λοιπόν, αυτών των σκωλήκων που παρασιτούν στον άνθρωπο αποτελεί κομμάτι της Ιατρικής Ελμινθολογίας. Οι έλμινθες αυτοί διακρίνονται σε Πλατυέλμινθες και Νηματέλμινθες. Στην πλειοψηφία τους οι περισσότεροι από αυτούς δεν απειλούν την ανθρώπινη υγεία και είναι παθογόνοι μόνο σε περιπτώσεις υψηλού παρασιτικού φορτίου. Βέβαια, αξίζει να σημειωθεί ότι λόγω της εκτεταμένης διάδοσης αυτού του είδους των νοσημάτων, υφίσταται ένα πρόβλημα μέγιστης ιατροοικονομικής σημασίας [56].

1.4. Διροφιλαρίαση και *Dirofilaria repens*

Η διροφιλαρίαση είναι μια ελμινθική παρασίτωση που προκαλείται από παράσιτα του γένους *Dirofilaria*, της οικογένειας *Onchocercidae*, των Νηματώδων. Είναι μία συνήθης ζωνοσός που επηρεάζει διάφορα ζώα όπως ο σκύλος, ο λύκος, η αλεπού και η γάτα. Μεταδίδεται μετά από νύγμα μολυσμένων αιμομυζητικών αρθροπόδων (κουνούπια των γενών *Aedes*, *Anopheles*, *Culex*, *Mansonia*) ενώ στον άνθρωπο που αποτελεί ενδιάμεσο ξενιστή, σπάνια εκδηλώνεται με νόσο. Ωστόσο, είναι ενδημική στις χώρες της Μεσογείου, συμπεριλαμβανομένης και της Ελλάδας. Έχουν υπάρξει, βέβαια, αναφορές εμφάνισής της σε διάφορα σημεία του κόσμου σε όλες τις Ηπείρους και μάλιστα σε περιοχές που δεν θεωρούνταν ενδημικές. Οι κλινικές μορφές της στον άνθρωπο συμπεριλαμβάνουν την πνευμονική, υποδόρια και οφθαλμική διροφιλαρίαση και σπάνια τη μόλυνση του εγκεφάλου, των όρχεων και των μαστών [20].

Η *Dirofilaria repens* είναι υπεύθυνη για υποδόρια και οφθαλμική διροφιλαρίαση. Η οφθαλμική διροφιλαρίαση γενικά προκαλεί αίσθημα δυσφορίας, οφθαλμικό πόνο και ερυθρότητα, ενώ υπάρχουν αναφορές ότι επηρεάζει κυρίως μεσήλικες με ελαφρά υπεροχή των αντρών ως προς τον αριθμό των κρουσμάτων. Θεραπεία εκλογής αποτελεί η χειρουργική αφαίρεση ολόκληρου του παρασίτου [20].

Πιο συγκεκριμένα, αναφορικά με τον κύκλο ζωής της *Dirofilaria repens*, κατά τη διάρκεια του γεύματός του ένα μολυσμένο κουνούπι, εισάγει τρίτου σταδίου προνύμφες στο δέρμα του ξενιστή (π.χ. σκύλος), οι οποίες και βρίσκουν δίοδο μέσω της πληγής που δημιουργήθηκε από το νύγμα του κουνουπιού. Στον ξενιστή αυτό, οι προαναφερθείσες προνύμφες περνάνε δύο επιπλέον στάδια και γίνονται ενήλικες. Στο τελευταίο στάδιο εντοπίζονται στους υποδόριους

ιστούς. Τα ενήλικα θηλυκά έχουν συνήθως μήκος 100-170 mm και πλάτος 460-650 μm, τα αρσενικά μήκος 50-70 mm και πλάτος 370-450 μm και είναι και τα δύο συνήθως λεπτοί κυλινδρικοί έλμινθες χρώματος κοντά στις αποχρώσεις του λευκού. Τα ενήλικα μπορούν να ζήσουν για 5-10 χρόνια. Στον υποδόριο ιστό, οι θηλυκοί έλμινθες είναι ικανοί να παράγουν μικροφιλάρια κατά τη διάρκεια ζωής τους. Οι μικροφιλάρια αυτές εντοπίζονται στο περιφερικό αίμα. Ένα κουνούπι καταπίνει τις μικροφιλάρια κατά τη διάρκεια ενός γεύματος αίματος. Μετά την κατάποση, οι μικροφιλάρια που πλέον βρίσκονται στο κουνούπι αναπτύσσονται σε προνύμφες πρώτου σταδίου και στη συνέχεια σε μολυσματικές προνύμφες τρίτου σταδίου σε χρονικό διάστημα το οποίο εξαρτάται και από τις περιβαλλοντικές συνθήκες. Οι μολυσματικές προνύμφες τρίτου σταδίου μεταναστεύουν στην προβοσκίδα του κουνουπιού και μπορούν να μολύνουν έναν άλλο ξενιστή όταν αυτός δέχεται νύγμα [6, 20, 39].

Πίνακας 1.1.: Ταξινομική κατάταξη ειδών χρησιμοποιούμενων στις φυλογενετικές αναλύσεις της παρούσας εργασίας [50]

ΣΥΝΟΜΟΤΑΞΙΑ (ΦΥΛΟ)	Nematoda
ΟΙΚΟΓΕΝΕΙΑ	Anisakidae
	- <i>Anisakis spp</i>
ΟΙΚΟΓΕΝΕΙΑ	Cystidicolidae
	- <i>Salmonema sp.</i>
ΟΙΚΟΓΕΝΕΙΑ	Dracunculidae
	- <i>Dracunculus medinensis</i>
ΟΙΚΟΓΕΝΕΙΑ	Filariidae
	- <i>Filarioid sp.</i>
ΟΙΚΟΓΕΝΕΙΑ	Gongylnematidae
	- <i>Gongylnema nepalensis</i>
ΟΙΚΟΓΕΝΕΙΑ	Onchocercidae
	- <i>Brugia malayi</i>
	- <i>Dirofilaria immitis</i>
	- <i>Dirofilaria repens</i>
	- <i>Dirofilaria sp. hongkongensis</i>
	- <i>Elaeophora schneideri</i>
	- <i>Loa loa</i>
	- <i>Mansonella ozzardi</i>
	- <i>Mansonella perstans</i>
	- <i>Onchocerca cervipedis</i>
	- <i>Rumenfilaria andersoni</i>
	- <i>Wuchereria bancrofti</i>

1.5. Φυλογενετική και Παράσιτα

Τις τελευταίες τρεις δεκαετίες, η συγκριτική ανάλυση μεταξύ των ειδών έχει χρησιμοποιηθεί ευρέως για να αποκαλύψει μοτίβα συσχετισμένης εξέλιξης μεταξύ των χαρακτηριστικών ή μεταξύ φαινοτυπικών χαρακτηριστικών και περιβάλλοντος, δηλαδή του φαινομένου της προσαρμογής. Σε αυτό έχει διατελέσει πρωτεύοντα ρόλο η φυλογενετική. Τα παράσιτα αντιπροσωπεύουν τους ιδανικούς στόχους για συγκριτικές μελέτες λόγω των προφανών προσαρμοστικών χαρακτηριστικών τους και της σχέσης τους με τους ξενιστές τους [28].

Έτσι, η ανάπτυξη μοριακών τεχνικών αύξησε τις γνώσεις αναφορικά με την ποικιλομορφία των παρασίτων και τους φορείς που τα μεταδίδουν. Η ποικιλομορφία και η διαφοροποίηση των παρασίτων συγκεντρώνει δύο ενεργούς τομείς έρευνας, εκείνον της φυλογενετικής και εκείνον της εξελικτικής οικολογίας, για να αποκαλύψουν και να εξηγήσουν τα μοτίβα της ποικιλομορφίας των παρασίτων και τη διαφοροποίηση των ξενιστών τους. Οι γνώσεις που αποκτούνται με τη χρήση της φυλογενετικής αναφορικά με την επιστήμη της παρασιτολογίας βρίσκουν εφαρμογή σε πολλούς τομείς, συμπεριλαμβανομένης της υγείας των ζώων και του ανθρώπου [28].

1.6. Η Φυλογενετική στο Μέλλον

Η δημιουργία φυλογενετικών δέντρων είναι μια ισχυρή και οπτικά ευκολόχρηστη προσέγγιση για την εξαγωγή συμπερασμάτων αναφορικά με τις εξελικτικές σχέσεις μεταξύ και των μικροβιακών αλληλουχιών. Η συνεχής βελτίωση των τεχνολογιών αλληλούχισης, μαζί με την αυξανόμενη εξάρτηση από μεθόδους που βασίζονται στην χρήση αλληλουχιών

ευθυγράμμισης για μοριακή ταυτοποίηση, διασφαλίζουν ότι η φυλογενετική προσέγγιση θα εξελιχθεί σε ένα σημαντικότατο εργαλείο, ιδίως σε μελέτες μοριακής επιδημιολογίας [53].

Τέλος, η ικανότητα να προβλέπονται υπολογιστικά οι βιολογικές λειτουργίες βιομορίων μειώνει σημαντικά το κόστος και αυτός είναι ένας ακόμη παράγοντας που συμβάλλει στην ύπαρξη και σταδιακή προσπάθεια για βελτίωση των εργαλείων και μέσων που χρησιμοποιούνται στις ποικίλες φυλογενετικές αναλύσεις.

1.7. Σκοπός

Σκοπός της παρούσας μελέτης είναι, κυρίως, η ταυτοποίηση, μέσω φυλογενετικής ανάλυσης, του παρασίτου ενδιαφέροντος το οποίο απομονώθηκε από ανθρώπινο οφθαλμό. Αυτό θα επιτευχθεί με φυλογενετική ανάλυση των αλληλουχιών του 12S rRNA, 18S rRNA και αλληλουχιών της υπομονάδας 1 της μιτοχονδριακής οξειδάσης του κυτοχρώματος c (COI), μέσω σύγκρισής τους με άλλες καταγεγραμμένες αλληλουχίες και της ένταξής τους, εν τέλει, σε φυλογενετικά δέντρα, κάνοντας χρήση των κατάλληλων παραμέτρων. Τα κεντρικά πακέτα λογισμικού και βάσεις δεδομένων που θα χρησιμοποιηθούν για το σκοπό αυτό είναι τα MEGA-X και NCBI.

2. ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ

2.1. Πειραματικό Παράσιτο

Στα πλαίσια της πειραματικής διαδικασίας ελήφθη χειρουργικά παράσιτο από ανθρώπινο οφθαλμό. Το παράσιτο εξετάστηκε φαινοτυπικά και διατηρήθηκε σε διάλυμα αιθανόλης.

2.2. Αλληλουχίες

Οι αλληλουχίες που χρησιμοποιήθηκαν για την ταυτοποίηση του παρασίτου, προήλθαν από ενίσχυση με αλυσιδωτή αντίδραση πολυμεράσης (PCR) χρησιμοποιώντας DNA του μικροοργανισμού που είχε προηγουμένως απομονωθεί. Τα τέσσερα αμπλικόνια που παρήχθησαν αντιστοιχούσαν σε καλά συντηρημένες περιοχές των γενετικών τόπων 12S rRNA (1), 18S rRNA (2), και COX1 (1). Ακολούθησε απομόνωση των DNA θραυσμάτων και αλληλούχισή τους κατά Sanger. Οι αλληλουχίες που προέκυψαν κατατέθηκαν στη βάση δεδομένων GenBank (NCBI) κι έλαβαν τους εξής κωδικούς πρόσβασης (accession numbers): MK192091 (12S rRNA), MK192092 (18S rRNAF), MK192093 (18S rRNAF2) και MK210632 (COX1).

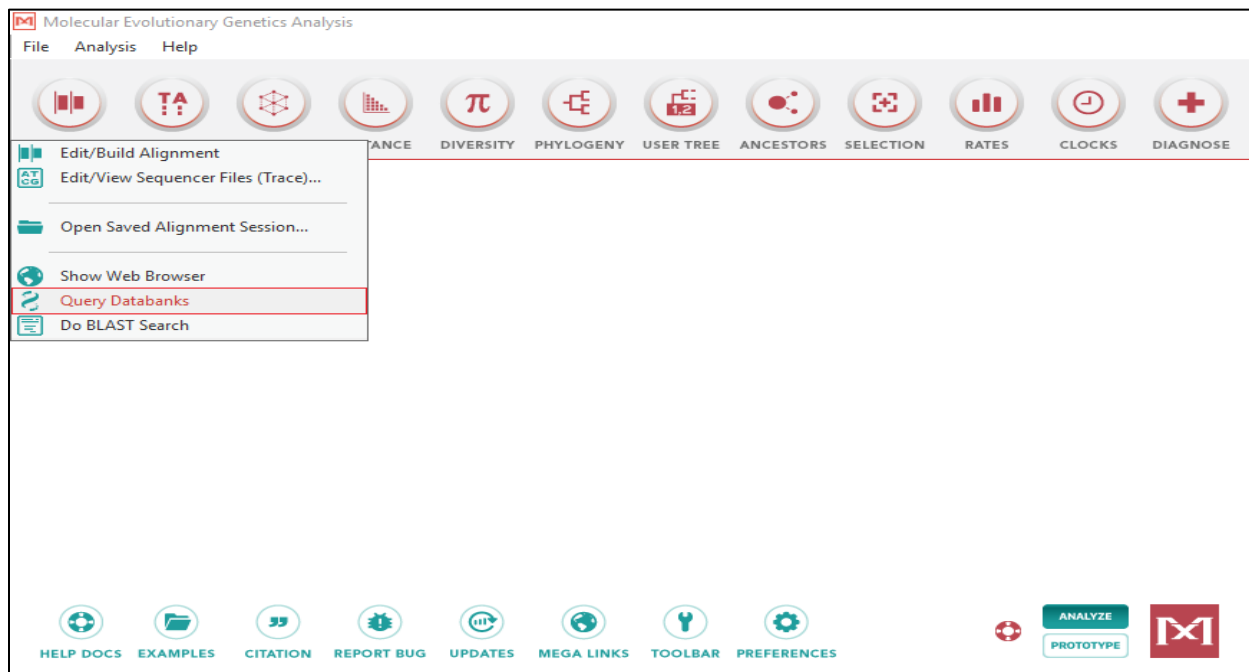
2.3. Διαδικασία Κατασκευής Φυλογενετικού Δέντρου

2.3.1. Συνοπτική

Προκειμένου να επιτευχθεί ο τελικός στόχος, ο οποίος είναι η δημιουργία του φυλογενετικού δέντρου, πρώτα απ' όλα ελήφθησαν οι αλληλουχίες του παρασίτου ενδιαφέροντος σε μορφή «*.fasta». Στη συνέχεια, πραγματοποιήθηκαν οι φυλογενετικές αναλύσεις των αλληλουχιών αυτών με χρήση πολλαπλών ευθυγραμμίσεων μέσω του πακέτου λογισμικού MEGA-X το οποίο διατρέχει τη βάση δεδομένων NCBI. Με τη βοήθεια της διαδικτυακής πλατφόρμας BLAST (NCBI) συγκρίνεται η αλληλουχία ενδιαφέροντος με αλληλουχίες του ίδιου είδους ή άλλων ειδών με τις οποίες εμφανίζει ομολογία και βρίσκονται στην προαναφερθείσα βάση δεδομένων. Προκύπτει έτσι μια σειρά ταξινόμησης σύμφωνα και με το συντελεστή ομοιότητας. Κατόπιν επιλέγονται κάποιες αλληλουχίες που διαθέτουν συντελεστή ομοιότητας πάνω από 85% κι αποθηκεύονται για τη δημιουργία του φυλογενετικού δέντρου. Στη συνέχεια γίνεται ευθυγράμμιση με το ClustalW, αποθήκευση των αποτελεσμάτων σε ένα αρχείο το οποίο και στη συνέχεια πρέπει να ανοιχθεί στο MEGA-X προκειμένου να κατασκευαστεί το φυλογενετικό δέντρο χρησιμοποιώντας «Bootstrap Method» με «1000 replicates», «General Time Reversible Model», «Gamma Distributed with Invariant (G + I)» και «No. Of Discrete Gamma Categories - 5». Τέλος, ακολουθεί η μορφοποίηση των Original Trees και Bootstrap Consensus Trees με προσθήκη δεδομένων όπως η αριθμητική έκφραση του μήκους των κλαδιών, ή ορθότερα του αναπαριστώμενου εξελικτικού χρόνου.

2.3.2. Αναλυτική

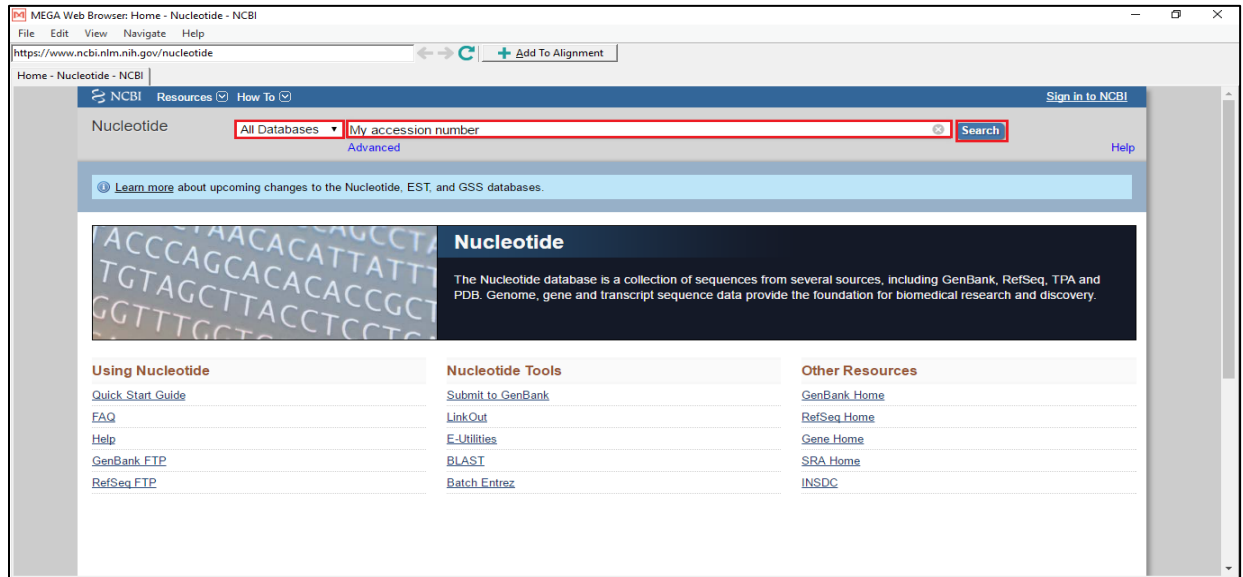
ΒΗΜΑ 1: Γίνεται εκκίνηση του πακέτου λογισμικού MEGA-X (εγκαταστημένο σε υπολογιστή) και ακολουθείται η διαδρομή «Align → Query Databanks» προκειμένου το λογισμικό να αποκτήσει πρόσβαση στη βάση δεδομένων του NCBI στην οποία και έχουν αναρτηθεί οι αλληλουχίες ενδιαφέροντος. Με τη βοήθεια της βάσης δεδομένων αυτής θα γίνει η αναζήτηση των αλληλουχιών, οι οποίες και θα χρησιμοποιηθούν για την κατασκευή του φυλογενετικού δέντρου.



Εικόνα 2.1.: Υλικά και Μέθοδοι - Βήμα 1

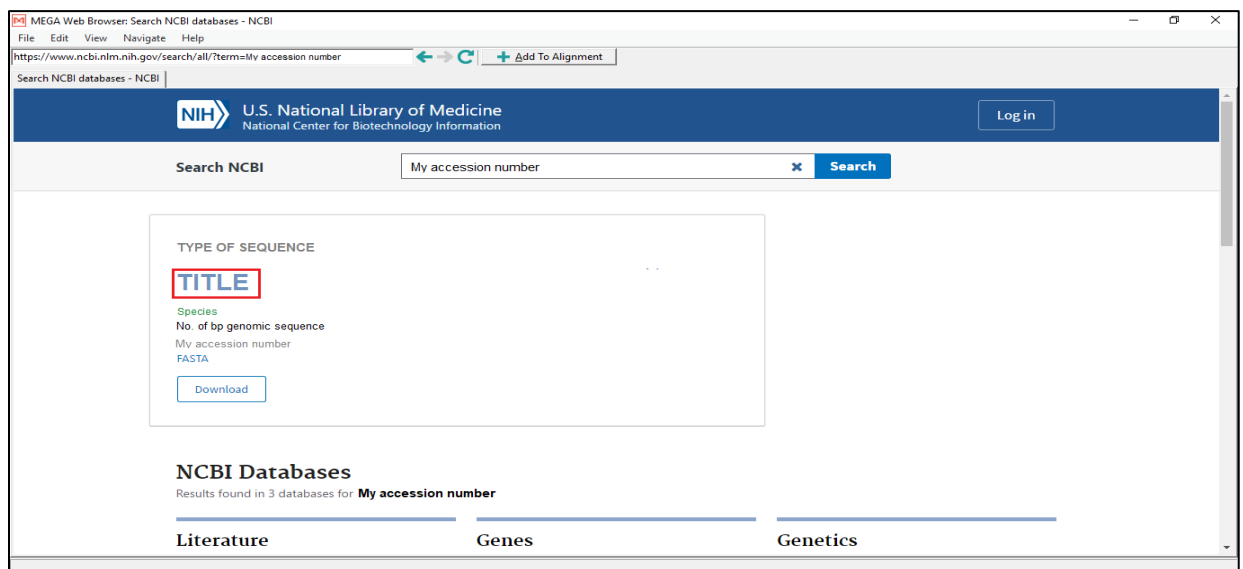
ΒΗΜΑ 2: Μετά τη διαπαραθυρική σύνδεση με τη βάση δεδομένων NCBI, εμφανίζεται ιστοσελίδα που στην κορυφή της φέρει δύο πλαίσια εντολών αριστερά του πλαισίου που αντιστοιχεί στην εντολή «Search». Το πρώτο πλαίσιο αριστερά ρυθμίζεται στην επιλογή «All Databases», ενώ στο δεύτερο πλαίσιο, δεξιά αυτού, το οποίο και αποτελεί πεδίο αναζήτησης

εισάγεται ο αριθμός πρόσβασης (accession number) της αλληλουχίας που έχει κατατεθεί στη βάση δεδομένων.



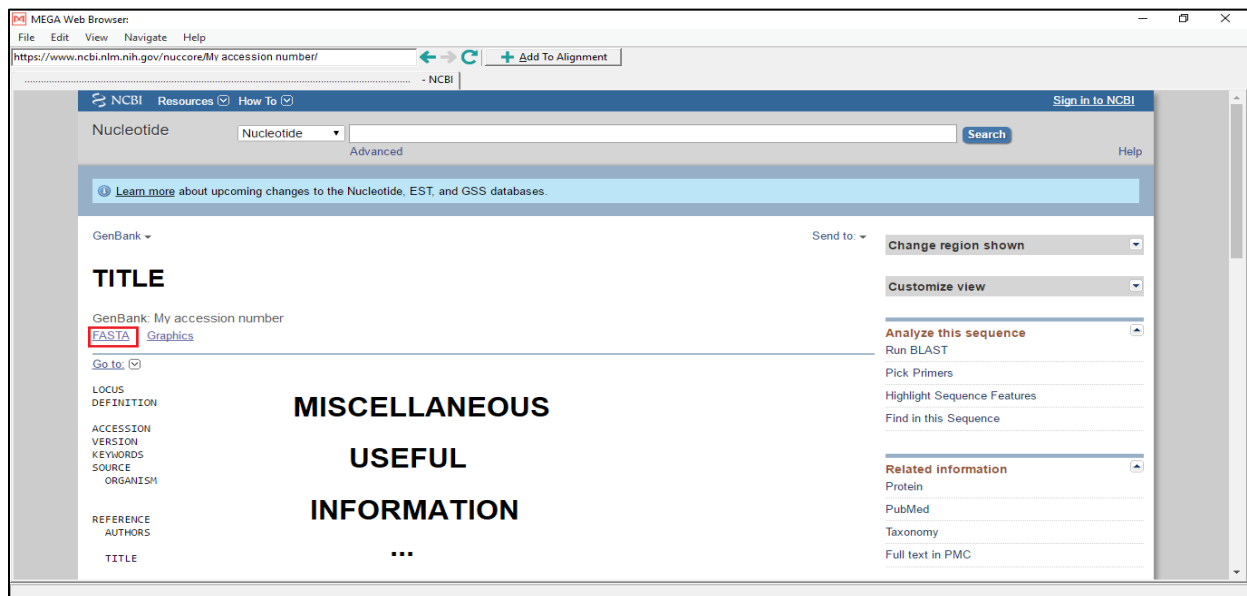
Εικόνα 2.2.: *Υλικά και Μέθοδοι - Βήμα 2*

ΒΗΜΑ 3: Η αναζήτηση οδηγεί στην ανάρτηση ενδιαφέροντος, η οποία και περιέχει την αντίστοιχη νουκλεοτιδική (ή αμινοξική) αλληλουχία. Επιλέγεται ο τίτλος.



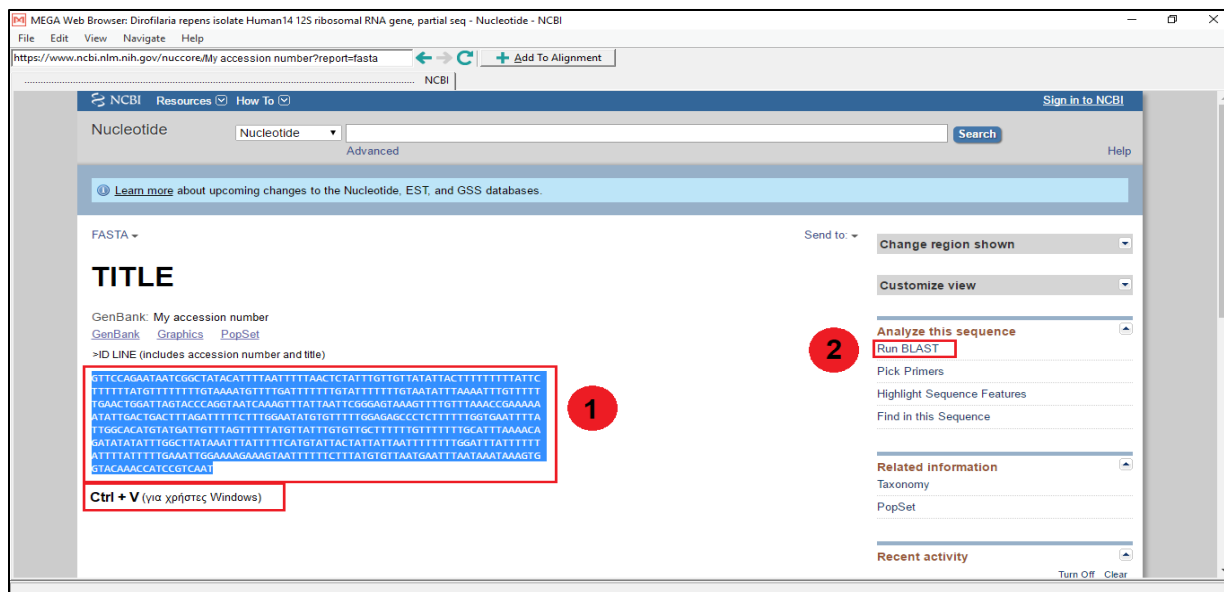
Εικόνα 2.3.: *Υλικά και Μέθοδοι - Βήμα 3*

ΒΗΜΑ 4: Στη σελίδα αυτή αποκτάται πρόσβαση τόσο στην ίδια την αλληλουχία όσο και σε αναλυτικές και χρήσιμες πληροφορίες που αφορούν σε αυτή. Ωστόσο, προκειμένου να συνεχιστεί η διαδικασία και να καταστεί εφικτή η ανάλυση, θα πρέπει η αλληλουχία να ληφθεί σε μορφή FASTA (*.fasta). Για το λόγο αυτό, στην κορυφή της σελίδας, ακριβώς κάτω από την επιλογή «Genbank accession number», και κάτω από τον τίτλο, επιλέγεται το «FASTA».



Εικόνα 2.4.: Υλικά και Μέθοδοι - Βήμα 4

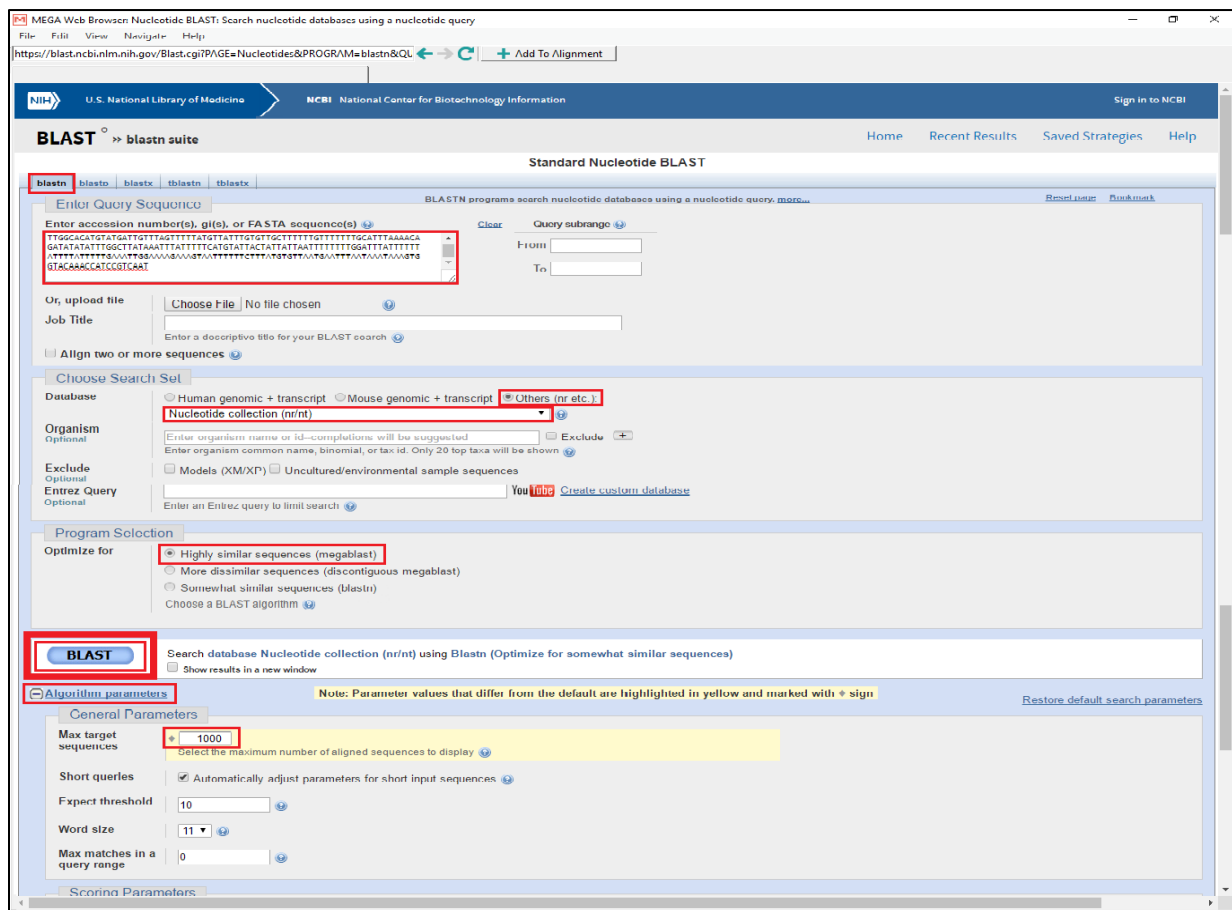
ΒΗΜΑ 5: Εμφανίζεται, λοιπόν, η πλήρης αλληλουχία στη μορφή «*.fasta». Σε αυτή τη μορφή, στην πρώτη σειρά υπάρχει μια σύντομη περιγραφή της αλληλουχίας (id line) κι ακολουθεί η νουκλεοτιδική αλληλουχία. Ξεκινώντας από την πρώτη βάση επιλέγεται ολόκληρη η αλληλουχία (αφήνοντας, έτσι, εκτός το id line) και αντιγράφεται. Στη συνέχεια, επιλέγεται η εντολή «Run BLAST», που βρίσκεται στη δεξιά στήλη της σελίδας ως υποενότητα του συνόλου εντολών «Analyze your sequence».



Εικόνα 2.5.: Υλικά και Μέθοδοι - Βήμα 5

ΒΗΜΑ 6: Στο βήμα αυτό παρουσιάζεται η σελίδα του λογισμικού σύγκρισης αλληλουχιών BLAST. Ανάμεσα στις εμφανιζόμενες καρτέλες διατηρούμε την εντολή «blastn», εφόσον θα γίνει σύγκριση νουκλεοτιδικών αλληλουχιών. Στην ενότητα εισαγωγής αλληλουχιών με τίτλο «Enter Query Sequences» γίνεται η εισαγωγή της αλληλουχίας που πρόκειται να συγκριθεί με αυτές της βάσης δεδομένων. Δύο τρόποι εισαγωγής είναι δυνατοί: α) με τη βοήθεια του κωδικού πρόσβασης της αλληλουχίας (accession number) και β) με αντιγραφή και επικόλλησή της στο αντίστοιχο πεδίο εισαγωγής. Στη συνέχεια, στην ενότητα με τίτλο «Choose Search Set» επιλέγεται στο «Database» το «Others (nr etc.)» και ακριβώς από κάτω το «Nucleotide collection (nr/nt)». Επίσης, στο «Program Selection» πρέπει να είναι επιλεγμένο το «Highly similar sequences (megablast)». Ακόμη πιο κάτω, στην ίδια σελίδα, στα αριστερά, βρίσκεται το «+Algorithm parameters». Γίνεται κλικ του υπολογιστή πάνω του και στην πρώτη ενότητα με τίτλο «General Parameters» επιλέγεται στο «Max target sequences» ο αριθμός «1000». Τέλος, γίνεται η εκκίνηση του προγράμματος σύγκρισης με κλικ του υπολογιστή στο μεγάλο γαλάζιο

κουμπί που βρίσκεται τέρμα κάτω, στα αριστερά και μέσα του γράφει «BLAST». Ακολουθεί αναμονή για την εμφάνιση των αποτελεσμάτων σύγκρισης. Ας σημειωθεί ότι ορισμένα από τα προαναφερθέντα κριτήρια ενδέχεται να είναι ήδη προεπιλεγμένα. Ωστόσο, κρίνεται ορθός ο έλεγχός τους.



Εικόνα 2.6.: Υλικά και Μέθοδοι - Βήμα 6

ΒΗΜΑ 7: Στην καινούρια σελίδα που εμφανίζεται παρατηρούνται κάποιες παράλληλες κόκκινες γραμμές. Όσο περισσότερες είναι, τόσο το καλύτερο. Κάθε γραμμή αντιπροσωπεύει μια ομόλογη αλληλουχία που συνήθως αντιστοιχεί σε διαφορετικό είδος. Ενδεχομένως, κάποιες από τις γραμμές να μην έχουν το ίδιο μήκος με τις υπόλοιπες κι έτσι εμφανίζονται κάποια λευκά κενά. Τα κενά αυτά σημαίνουν ότι σε αυτά τα τμήματα δεν έχει γίνει αλληλούχιση για την

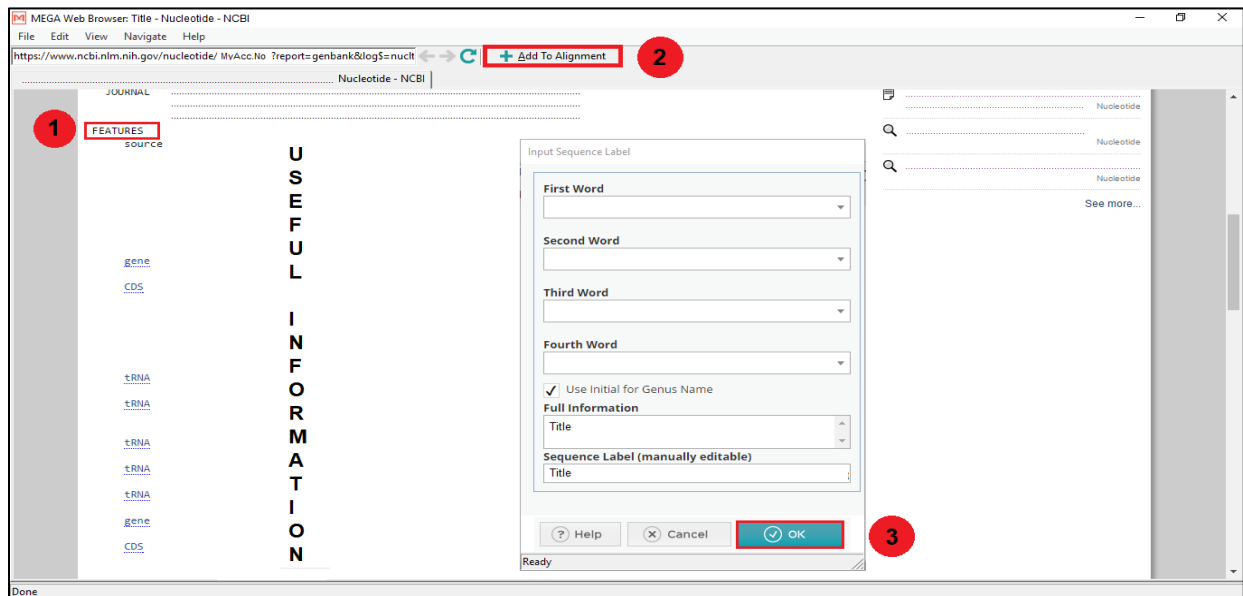
εκάστοτε αλληλουχία ή έπαψαν τα τμήματα αυτά να εμφανίζουν στατιστικά σημαντική ομοιότητα. Η πρώτη γραμμή είναι της αλληλουχίας ενδιαφέροντος. Παρακάτω, στην ενότητα με τίτλο «Descriptions», το «Ident» δείχνει ομοιότητα εκφρασμένη ποσοστιαία επί τοις εκατό. Η πρώτη γραμμή και πάλι είναι η αλληλουχία ενδιαφέροντος. Επιλέγονται ορισμένα από αυτά τα είδη με ποσοστό «Ident» μεταξύ 85% και 100% που εμφανίζουν συγκεκριμένη τουλάχιστον κάλυψη («Query cover») της αλληλουχίας του παρασίτου ενδιαφέροντος, διαφορετική για τα εκάστοτε είδη αλληλουχιών. Αυτό επιτυγχάνεται για το κάθε ένα ξεχωριστά, κάνοντας δεξί κλικ του υπολογιστή επάνω στον κωδικό πρόσβασης του (accession number) ο οποίος ανευρίσκεται στην τελευταία δεξιά στήλη και επιλέγοντας στη συνέχεια την εντολή «Open link in new window».

The screenshot displays the NCBI Blast Nucleotide Sequence (509 letters) web interface. The top section shows a 'Graphic Summary' with a color key for alignment scores and a horizontal bar chart. The bottom section shows 'Descriptions' with a table of sequences producing significant alignments. A context menu is open over the 'Accession' column, highlighting 'Open Link in New Window'.

Description	Max score	Total score	Query cover	E value	Ident	Accession
Sequence 1	919	919	100%	0.0		
Sequence 2	919	919	100%	0.0		
Sequence 3	919	919	100%	0.0		
Sequence 4	919	919	100%	0.0	100%	KX265053.1
Sequence 5	919	919	100%	0.0	100%	KX265052.1
Sequence 6	919	919	100%	0.0	100%	KX265048.1
Sequence 7	919	919	100%	0.0	100%	KM205403.1
Sequence 8	914	914	100%	0.0	99%	KX265089.1
Sequence 9	914	914	100%	0.0	100%	KX265095.1
Sequence 10	914	914	100%	0.0	100%	KX265096.1

Εικόνα 2.7.: Υλικά και Μέθοδοι - Βήμα 7

ΒΗΜΑ 8: Στο καινούριο παράθυρο που αναδύεται σημειώνονται οι χρήσιμες πληροφορίες για την αντίστοιχη αλληλουχία. Σημαντικότερο απ' όλα είναι να σημειωθούν τα τμήματα των αλληλουχιών που θα χρησιμοποιηθούν στην ανάλυση παρακάτω, όπως επίσης και να σημειωθεί σε ποιο τμήμα της αλληλουχίας αντιστοιχεί η κωδικοποιούσα περιοχή (coding sequence ή Cds), εφόσον αυτή περιλαμβάνεται στην αλληλουχία. Όλα αυτά ανευρίσκονται στην ενότητα με τίτλο «Features» με αριθμούς που δηλώνουν τις νουκλεοτιδικές βάσεις. Στη συνέχεια, από τη γραμμή διευθύνσεως επιλέγεται η εντολή «Add to Alignment» και μετά στο αναδύόμενο παράθυρο το «OK».

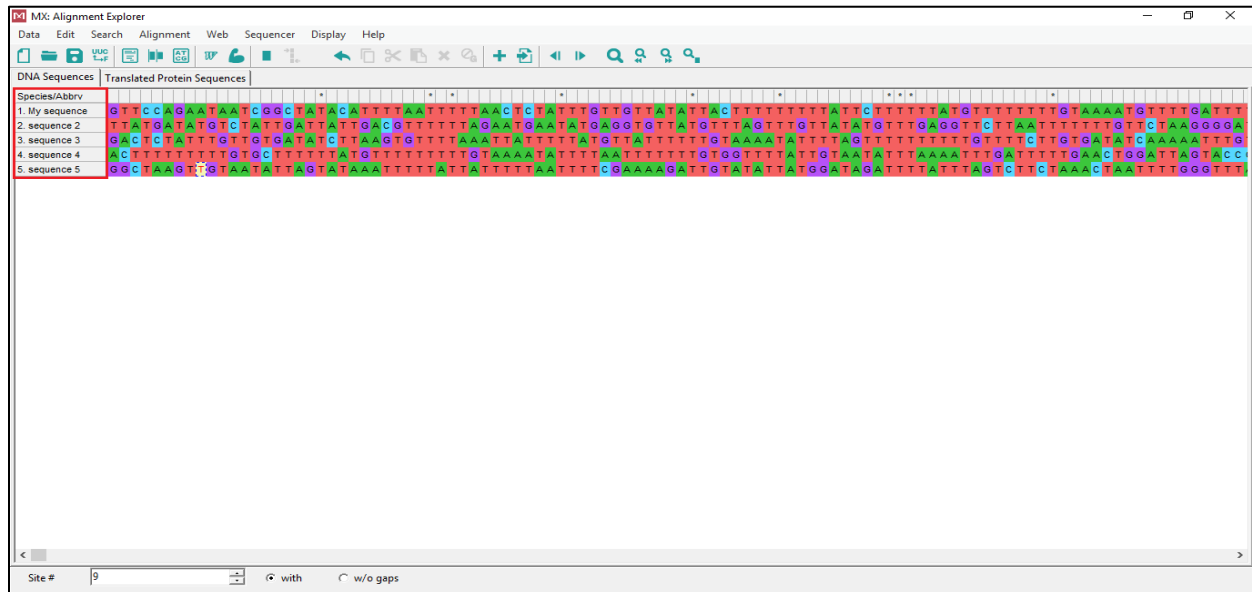


Εικόνα 2.8.: Υλικά και Μέθοδοι - Βήμα 8

ΒΗΜΑ 9: Τώρα, έχει αναδυθεί το παράθυρο «Alignment Explorer». Εδώ είναι δυνατόν να προσαρμοστεί η πρώτη στήλη, η οποία και δίνει την περιγραφή της αλληλουχίας ενδιαφέροντος, στο μέγεθος που είναι επιθυμητό και να δοθεί η επιθυμητή ονομασία.

ΒΗΜΑ 10: Αφού ολοκληρωθούν οι προαναφερθείσες ενέργειες κλείνεται το παράθυρο από το οποίο επιλέχθηκε το «Add to Alignment».

ΒΗΜΑ 11: Επαναλαμβάνεται η ίδια διαδικασία και για τις υπόλοιπες αλληλουχίες που αξιολογούνται ως κατάλληλες να συμμετέχουν στη φυλογενετική ανάλυση.



Εικόνα 2.9.: Υλικά και Μέθοδοι - Βήματα 9, 10, 11

ΒΗΜΑ 12: Αφού ολοκληρωθεί η διαδικασία για όλες τις συμμετέχουσες αλληλουχίες, λαμβάνει χώρα η αξιοποίηση των πληροφοριών που καταγράφησαν, προηγουμένως, από την ενότητα «Features», με αποτέλεσμα να διατηρούνται τα τμήματα της αλληλουχίας που κρίνεται σκόπιμο. Εφόσον υπάρχουν τμήματα, λοιπόν, που δε σχετίζονται με την ανάλυση ενδιαφέροντος, γίνεται η επιλογή τους, με την χρήση των εντολών «Shift + Home» ή «Shift + End» ή «Shift + βελάκι» και η απαλοιφή τους είτε με τη χρήση του πλήκτρου «Delete», είτε με την εντολή «cut the selected block» (εικονίδιο με ψαλίδι). Η διαδικασία αυτή είναι ίδια για όλες τις αλληλουχίες που συμπεριλαμβάνονται στο «Alignment Explorer».

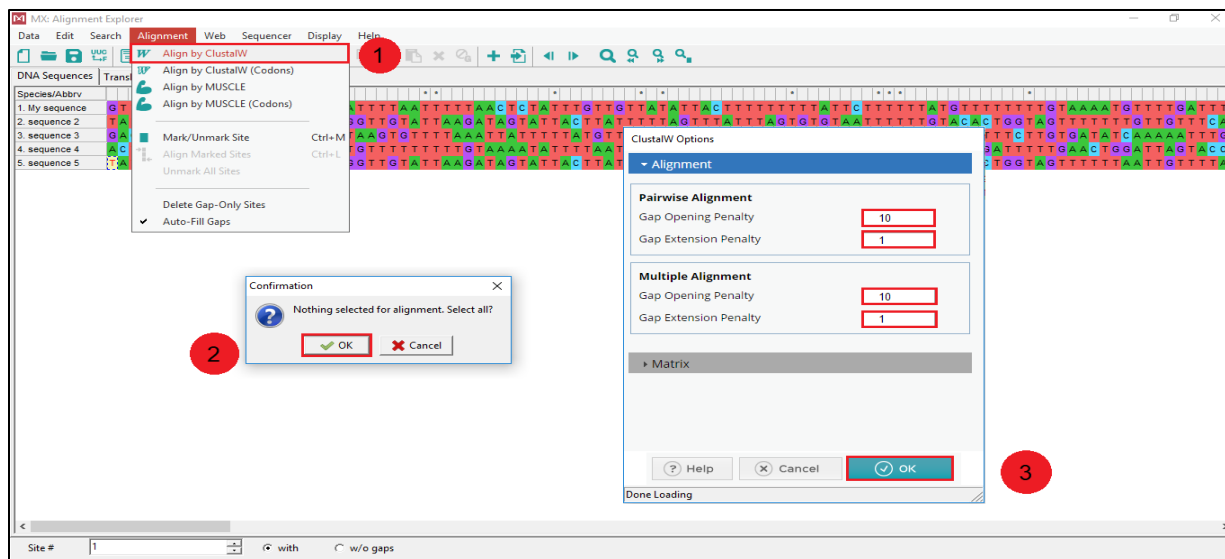


Εικόνα 2.10.: Υλικά και Μέθοδοι - Βήμα 12

ΒΗΜΑ 13: Στο «Alignment Explorer» επιλέγεται «Alignment → Align by ClustalW»

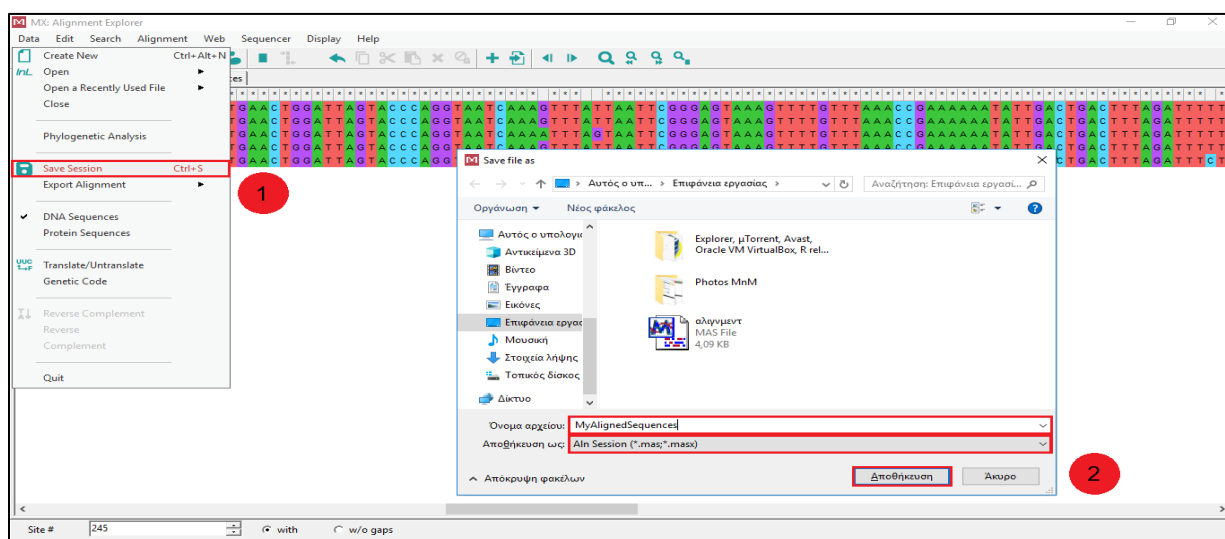
ΒΗΜΑ 14: Στο μήνυμα που εμφανίζεται επιλέγεται το «OK».

ΒΗΜΑ 15: Στο παράθυρο που εμφανίζεται αμέσως μετά, στα πλαίσια που αντιστοιχούν στο «Gap Opening Penalty» τόσο στην ενότητα «Pairwise Alignment», όσο και στην ενότητα «Multiple Alignment» επιβάλλεται η τιμή 10, ενώ αντίστοιχα στα πλαίσια που αντιστοιχούν στο «Gap Extension Penalty» η τιμή 1. Έπειτα, επιλέγεται το «OK».



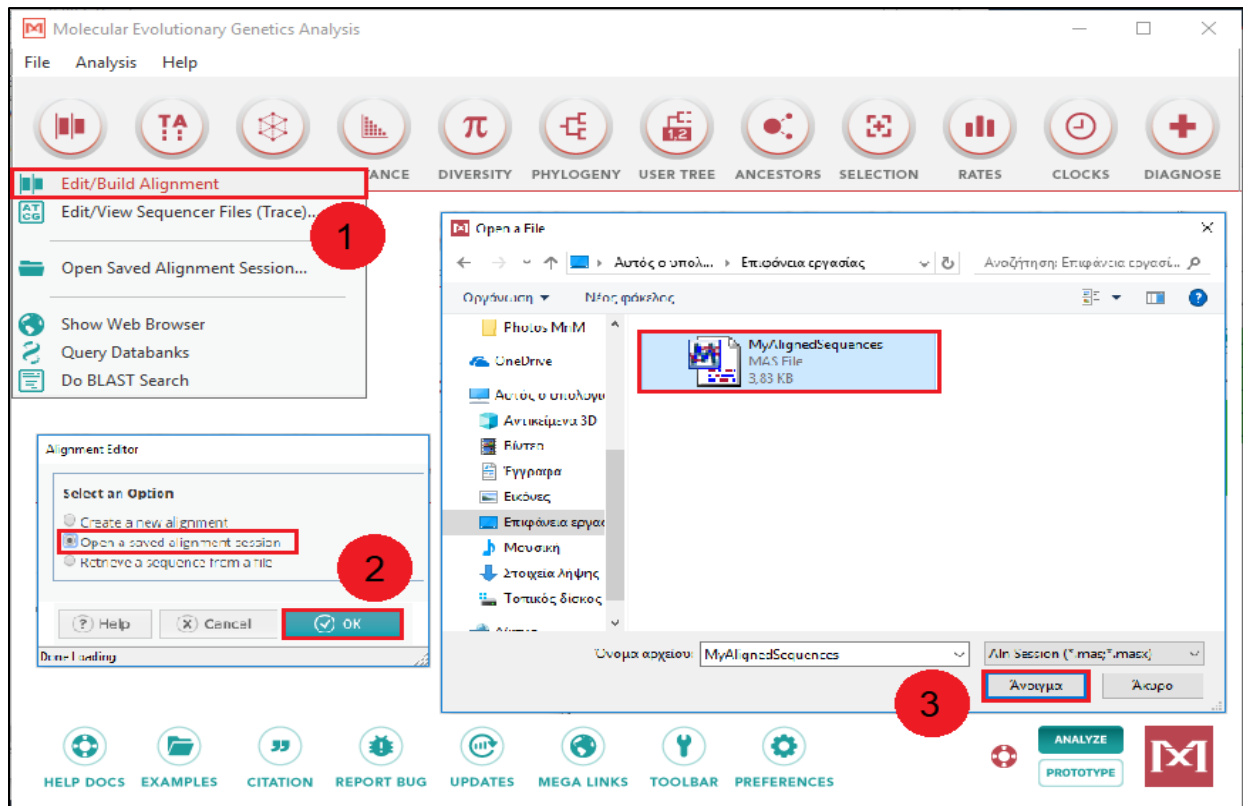
Εικόνα 2.11.: Υλικά και Μέθοδοι - Βήμα 13, 14, 15

ΒΗΜΑ 16: Μόλις ολοκληρωθεί η διεργασία, και αφού διαγράψουμε τμήματα άλλων αλληλουχιών τα οποία ενδεχομένως περισσεύουν της δικής μας αλληλουχίας, λαμβάνει χώρα η αποθήκευση της μέχρι τώρα εργασίας ακολουθώντας τη διαδρομή «Data → Save Session» και επιλέγοντας την κατάληξη «Aln Session (*.mas; *.masx)» για το αρχείο προς αποθήκευση. Έπειτα, κλείνεται το «Alignment Explorer».



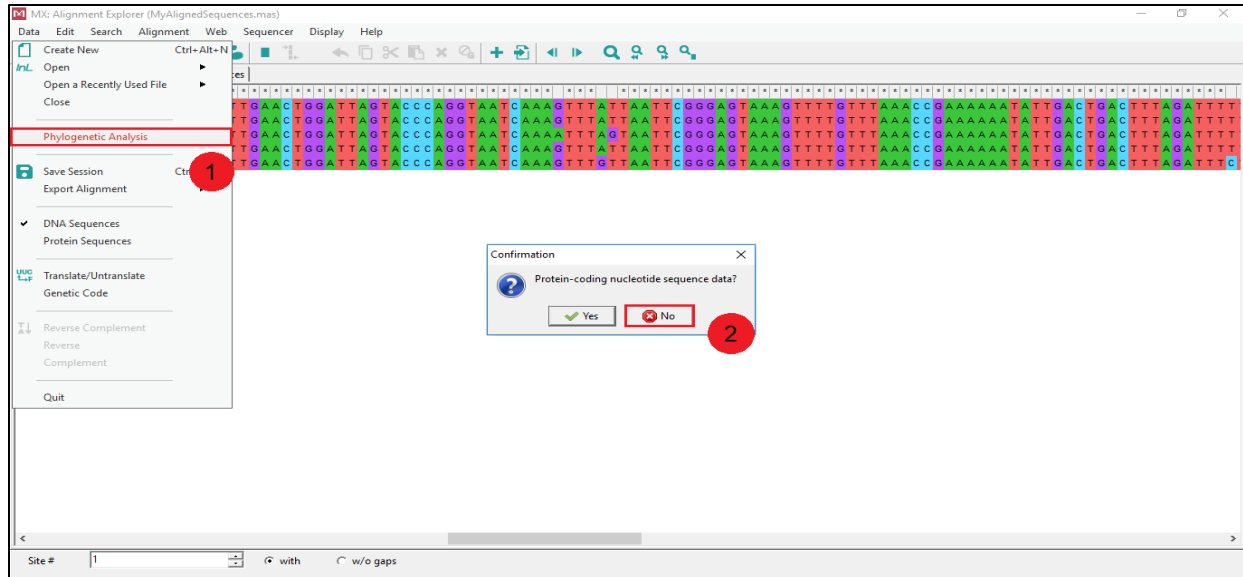
Εικόνα 2.12.: Υλικά και Μέθοδοι - Βήμα 16

ΒΗΜΑ 17: Με επιστροφή στο αρχικό παράθυρο του MEGA-X, απ' όπου και ξεκίνησε η διαδικασία, ακολουθείται η διαδρομή «Align → Edit/Build Alignment». Στη συνέχεια επιλέγεται το «Open a saved alignment session» και το «OK». Εντοπίζεται, το έγγραφο που πριν λίγο αποθηκεύτηκε και ανοίγεται.



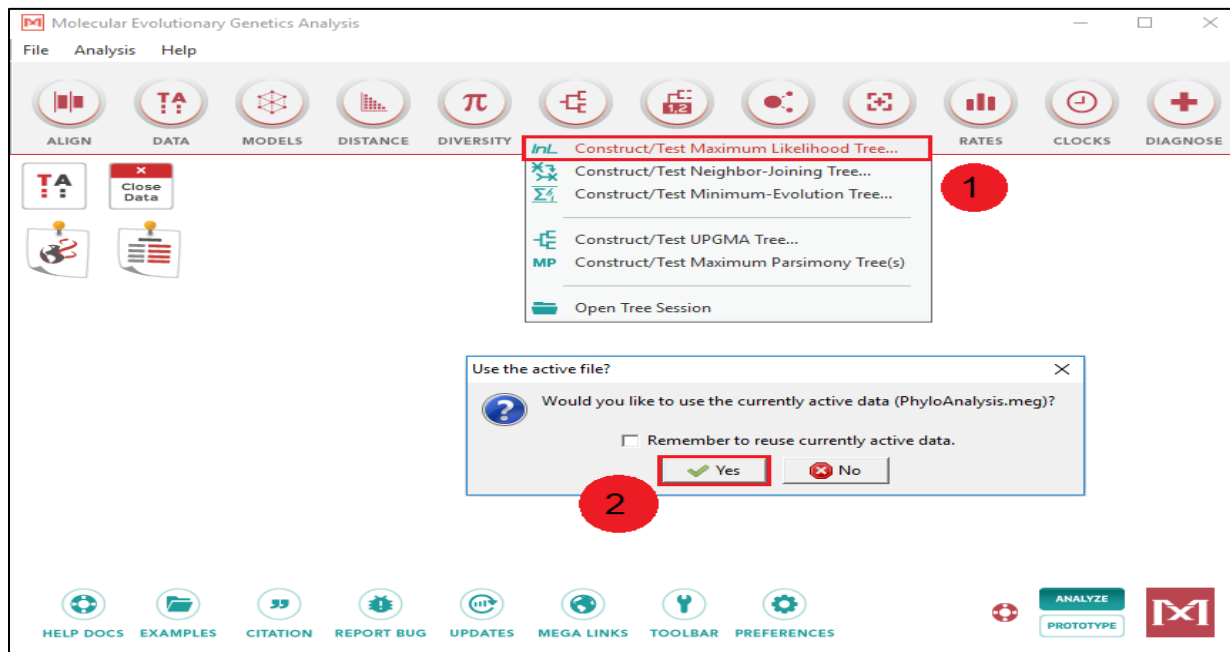
Εικόνα 2.13.: Υλικά και Μέθοδοι - Βήμα 17

ΒΗΜΑ 18: Στη συνέχεια ακολουθείται η διαδρομή «Data → Phylogenetic Analysis» και στην ερώτηση «Protein-coding nucleotide sequence data?» επιλέγεται το «No» για το 12S και 18S rRNA και το «Yes» για το COI.



Εικόνα 2.14.: Υλικά και Μέθοδοι - Βήμα 18

ΒΗΜΑ 19: Γίνεται επιστροφή στο αρχικό παράθυρο του MEGA-X όπου και ακολουθείται η διαδρομή «Phylogeny → Construct/Test Maximum Likelihood Tree...» και στο παράθυρο που εμφανίζεται επιλέγεται το «Yes».



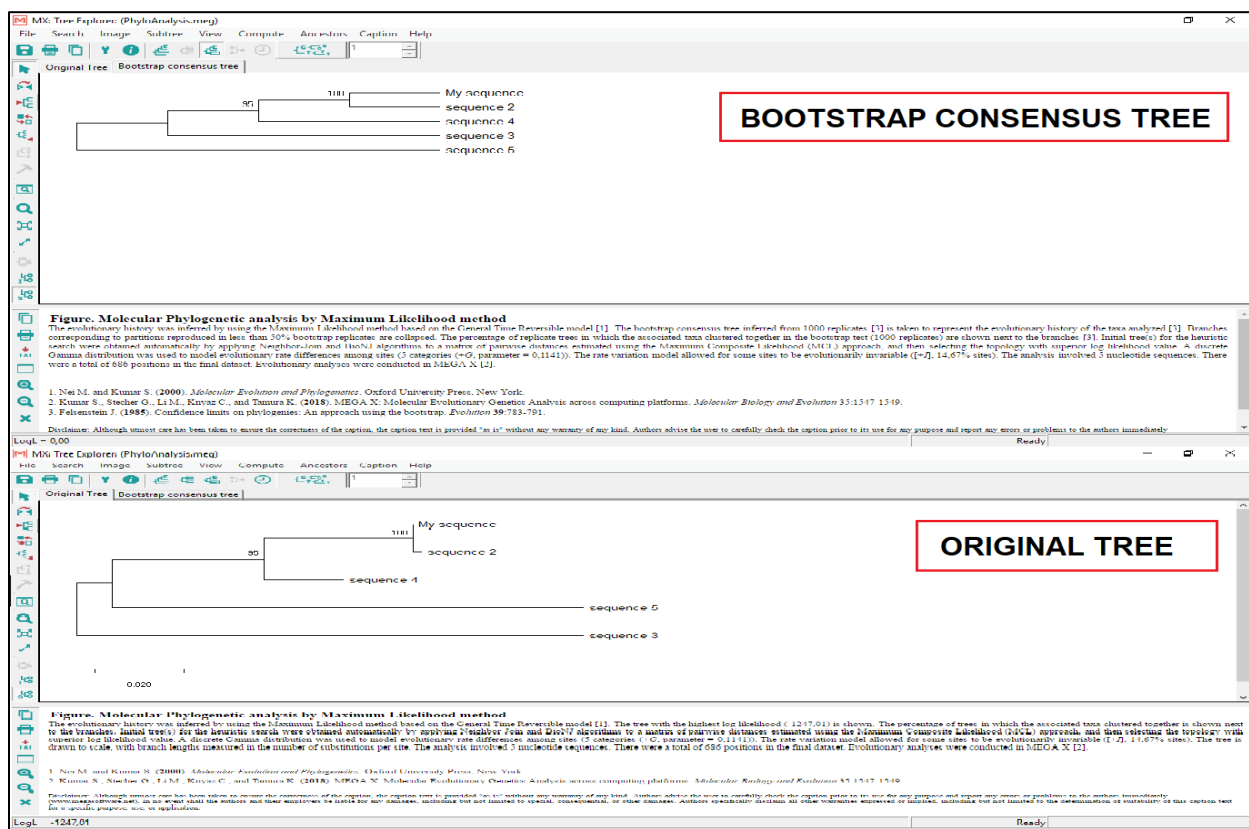
Εικόνα 2.15.: Υλικά και Μέθοδοι - Βήμα 19

ΒΗΜΑ 20: Στο νέο παράθυρο που αναδύεται, στο πλαίσιο που αντιστοιχεί στο «Test of Phylogeny» επιλέγεται το «Bootstrap Method», ενώ στο πλαίσιο που αντιστοιχεί στο «No. of Bootstrap Replications» επιλέγεται το «1000». Αντίστοιχα, στο «Model/Method» επιλέγεται το «General Time Reversible Model», στο «Rates among Sites» επιλέγεται το «Gamma Distributed With Invariant Sites (G+I)» και στο «No of Discrete Gamma Categories» επιλέγεται το «5». Στα υπόλοιπα αφήνονται οι προεπιλογές και επιλέγεται το «OK».

MX: Analysis Preferences	
Phylogeny Reconstruction	
Option	Setting
ANALYSIS	
Statistical Method →	Maximum Likelihood
PHYLOGENY TEST	
Test of Phylogeny →	Bootstrap method
No. of Bootstrap Replications →	1000
SUBSTITUTION MODEL	
Substitutions Type →	Nucleotide
Model/Method →	General Time Reversible model
RATES AND PATTERNS	
Rates among Sites →	Gamma Distributed With Invariant Sites (G+I)
No of Discrete Gamma Categories →	5
DATA SUBSET TO USE	
Gaps/Missing Data Treatment →	Use all sites
Site Coverage Cutoff (%) →	Not Applicable
TREE INFERENCE OPTIONS	
ML Heuristic Method →	Nearest-Neighbor-Interchange (NNI)
Initial Tree for ML →	Make initial tree automatically (Default - NJ/BioNJ)
Initial Tree File →	Not Applicable
Branch Swap Filter →	None
SYSTEM RESOURCE USAGE	
Number of Threads →	3
<input type="button" value="Help"/> <input type="button" value="Cancel"/> <input type="button" value="OK"/>	

Εικόνα 2.16.: Υλικά και Μέθοδοι - Βήμα 20

ΒΗΜΑ 21: Έχουν πλέον δημιουργηθεί τα φυλογενετικά δέντρα «Original Tree» και «Bootstrap consensus tree». Καθένα τους διαθέτει και υπόμνημα. Πλέον, είναι εφικτή η μορφοποίηση των δέντρων, η αποθήκευσή τους, καθώς και η εκτύπωσή τους σε αρχείο pdf. Αναφορικά με τις μορφοποιήσεις των δύο προαναφερθέντων φυλογενετικών δέντρων, από τις πιο συνηθισμένες στο πρώτο είδος είναι η συνδυαστική ακολούθηση της διαδρομής «View → Show/Hide → Branch Lengths» και επιλογή του, έβδομου κατά σειρά από αριστερά προς δεξιά, εικονιδίου «Toggle scaling of the tree», των οποίων και γίνεται εφαρμογή στην παρούσα εργασία. Εξίσου, στο δεύτερο ακολουθείται η διαδρομή «Subtree → Swar» που μετατρέπει τον κέρσορα σε εργαλείο που παρέχει τη δυνατότητα να μεταθέτει τα υποδέντρα που βρίσκονται υπό του κόμβου που επιλέγεται, προκειμένου να καθίσταται το φυλογενετικό δέντρο πιο ευανάγνωστο, χωρίς να χάνεται η κύρια τοπολογία του δέντρου και οι πληροφορίες που αυτή παρέχει.



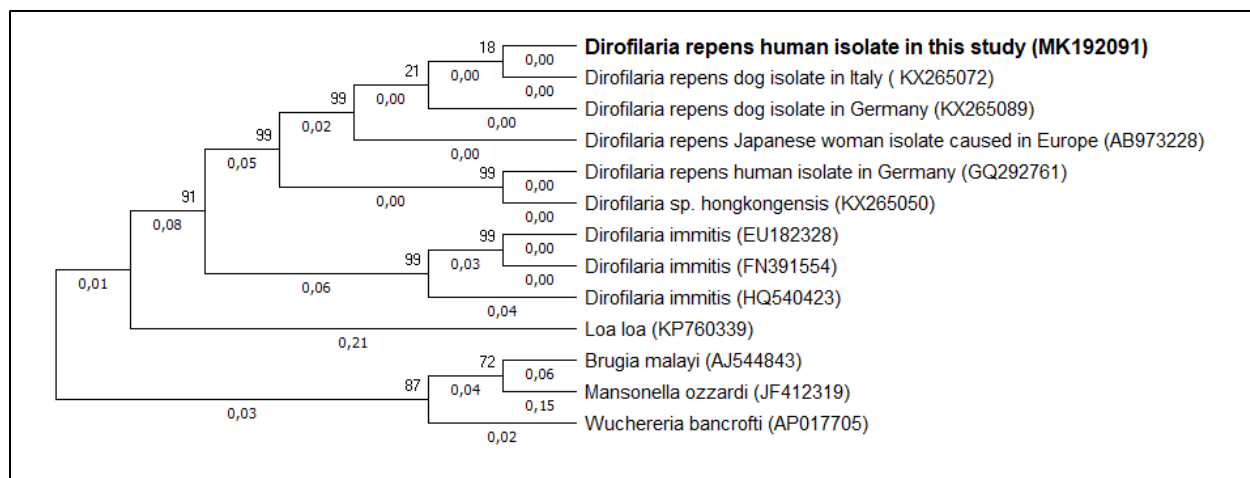
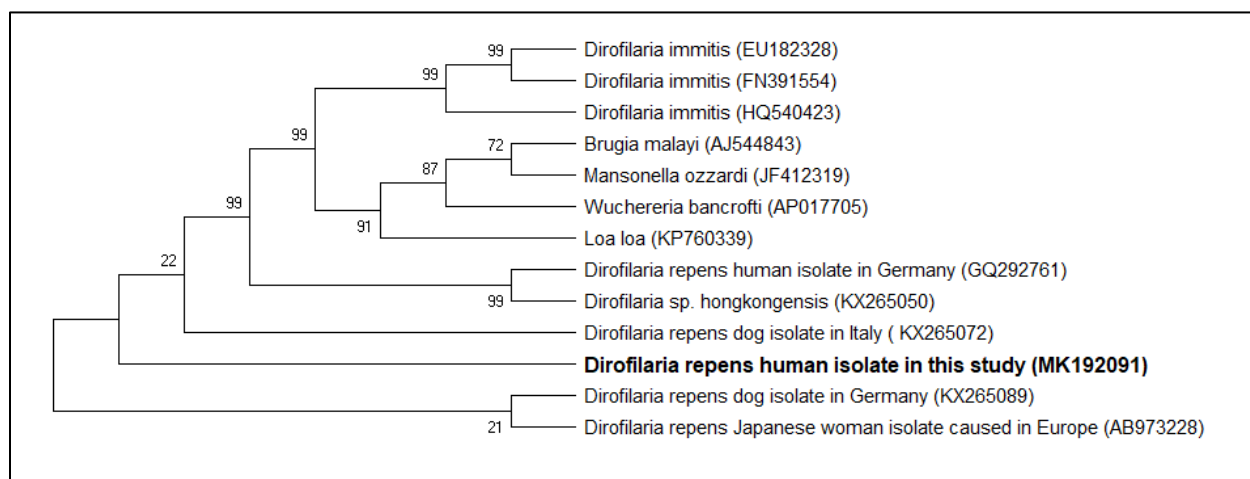
Εικόνα 2.17.: Υλικά και Μέθοδοι - Βήμα 21

3. ΑΠΟΤΕΛΕΣΜΑΤΑ

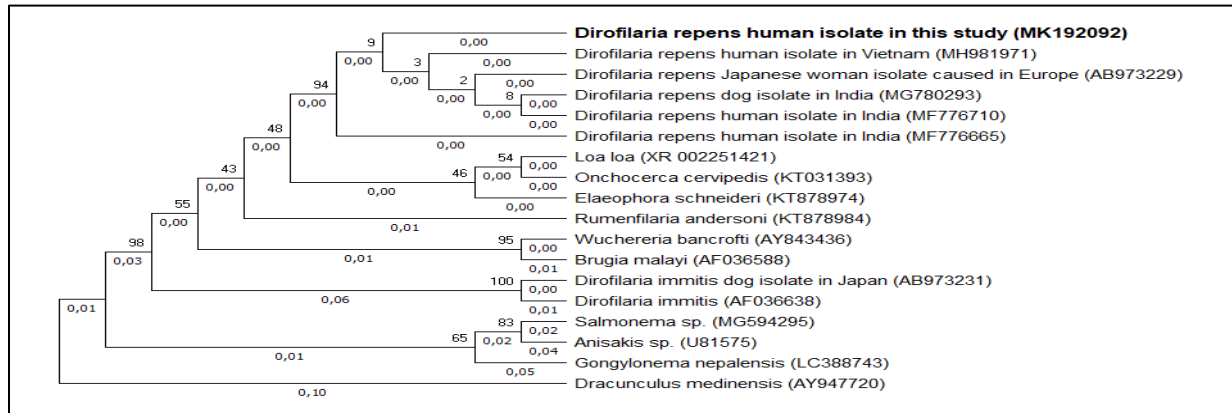
Στα φυλογενετικά δέντρα της ενότητας αυτής αναπαριστώνται οι φυλογενετικές σχέσεις με ανάλυση Μέγιστης Πιθανοφάνειας (ML) μεταξύ αλληλουχιών μιτοχondριακού 12S ριβοσωμικού RNA (A), 18S ριβοσωμικού RNA (B1, B2) και αλληλουχιών της υπομονάδας 1 της μιτοχondριακής οξειδάσης του κυτοχρώματος c ή συντομογραφημένο COI (Γ). Τα δενδρογράμματα αυτά προέκυψαν από ένα γενικό χρονικά αναστρέψιμο μοντέλο (General time reversible model) χρησιμοποιώντας μια διακριτή κατανομή γάμμα (discrete gamma distribution, + G) με πέντε κατηγορίες τιμών και αμετάβλητες θέσεις (rate categories and invariant sites, +I). Η ποσοστιαία πιθανότητα τα σχηματιζόμενα υποδέντρα να εντοπίζονται με την αναπαριστώμενη τοπολογία των εκάστοτε φυλογενετικών δέντρων, αναγράφεται δίπλα από τους βραχίονες (branches), στους εκάστοτε κόμβους (nodes). Η διαφορά μεταξύ των δύο ειδών δέντρων που αναπαριστώνται στην ενότητα αυτή είναι ότι μόνο στα Bootstrap consensus trees υφίσταται η έκφραση των αποτελεσμάτων της ανάλυσης με χρήση της μεθόδου bootstrap, ενώ στα Original trees αναπαριστάται η μία και μοναδική θεωρούμενη ως καλύτερη τοπολογία, βάση της μεθόδου και των παραμέτρων που χρησιμοποιήθηκαν, χωρίς τη συμμετοχή της μεθόδου bootstrap. Οι τιμές που αναγράφονται κατά μήκος των κλαδιών των Original trees αντιπροσωπεύουν τη γενετική απόσταση σε αντικαταστάσεις ενός νουκλεοτιδίου. Οι αριθμοί πρόσβασης της GenBank (NCBI) δίνονται εντός παρενθέσεως μετά την περιγραφή της αλληλουχίας. Οι φυλογενετικές αναλύσεις διεξήχθησαν, όπως έχει προαναφερθεί, με χρήση του πακέτου λογισμικού MEGA X.

Μαζί με τα φυλογενετικά δέντρα, παρατίθεται για κάθε αλληλουχία του παρασίτου ενδιαφέροντος ένας πίνακας, στον οποίο βρίσκονται καταγεγραμμένα τα αποτελέσματα του

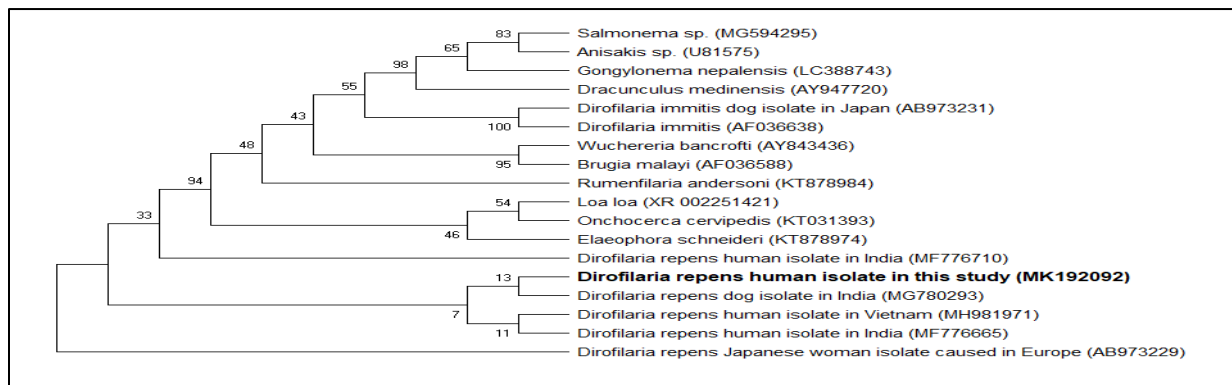
BLAST που αφορούν την ποσοστιαία έκφραση του βαθμού κάλυψης της προαναφερθείσας αλληλουχίας από τις επιμέρους υπόλοιπες που χρησιμοποιήθηκαν για την κατασκευή των εκάστοτε φυλογενετικών δέντρων, καθώς και του βαθμού ομοιότητάς τους ως προς την πρώτη.

(A) 12S rRNA**Εικόνα 3.1.: 12S rRNA - Original tree****Εικόνα 3.2.: 12S rRNA - Bootstrap consensus tree****Πίνακας 3.1.: Αποτελέσματα BLAST για 12S rRNA**

Περιγραφή αλληλουχίας (αριθμός πρόσβασης)	Κάλυψη (%) αλληλουχίας παρασίτου ενδιαφέροντος	Συντελεστής ομοιότητας (%)
<i>D.repens</i> dog isolate in Italy (KX265072)	100	100
<i>D.repens</i> dog isolate in Germany (KX265089)	100	99
<i>D.repens</i> Japanese woman isolate caused in Europe (AB973228)	91	99
<i>Dirofilaria sp. hongkongensis</i> (KX265050)	100	98
<i>D.repens</i> human isolate in Germany (GQ292761)	100	97
<i>D.immitis</i> (HQ540423)	92	91
<i>D.immitis</i> (EU182328)	100	90
<i>D.immitis</i> (FN391554)	95	90
<i>Loa loa</i> (KP760339)	82	89
<i>Wuchereria bancrofti</i> (AP017705)	100	88
<i>Brugia malayi</i> (AJ544843)	91	88
<i>Mansonella ozzardi</i> (JF412319)	96	85

(B1) 18SF rRNA

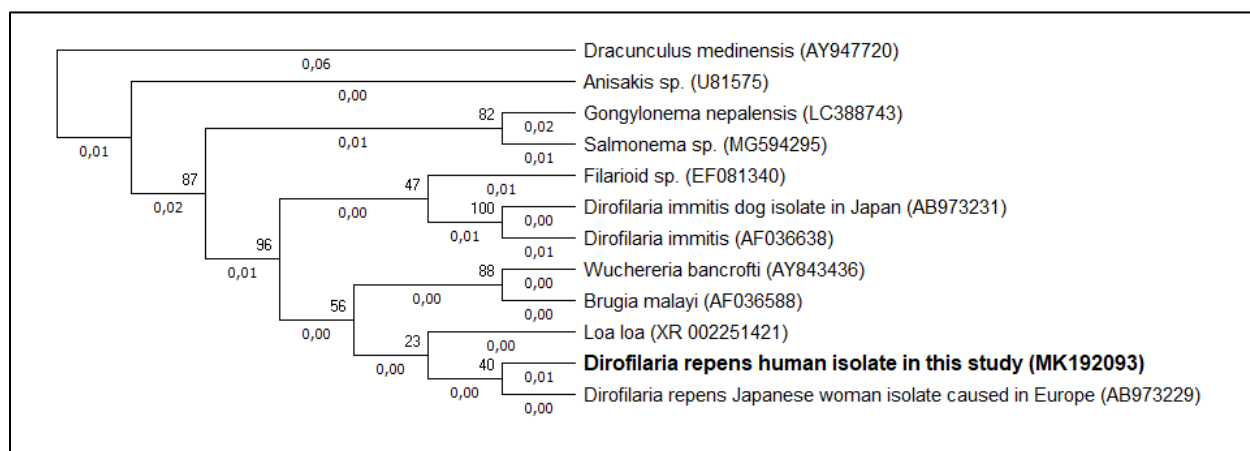
Εικόνα 3.3.: 18SF rRNA - Original tree



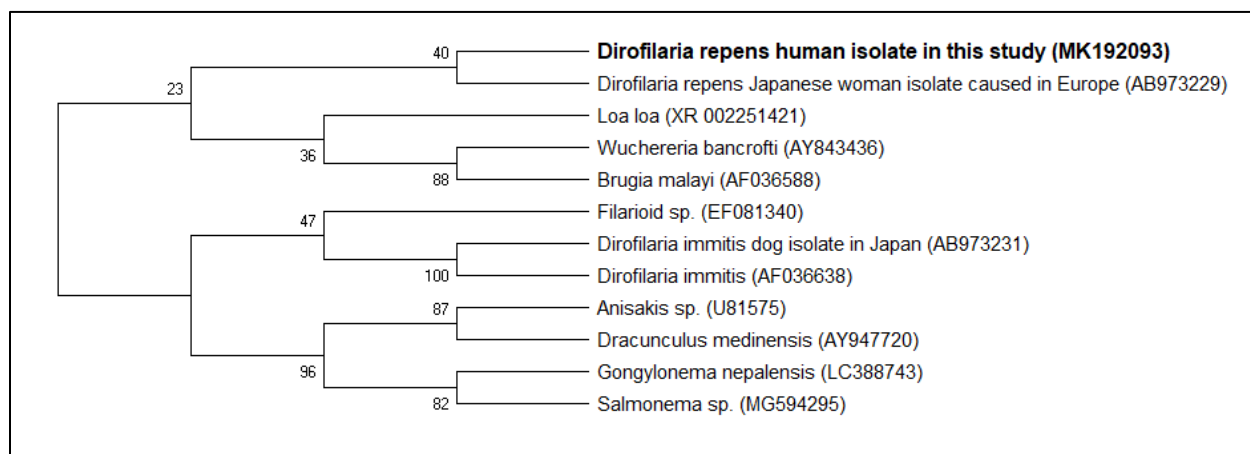
Εικόνα 3.4.: 18SF rRNA - Bootstrap consensus tree

Πίνακας 3.2.: Αποτελέσματα BLAST για 18SF rRNA

Περιγραφή αλληλουχίας (αριθμός πρόσβασης)	Κάλυψη (%) αλληλουχίας παρασίτου ενδιαφέροντος	Συντελεστής ομοιότητας (%)
<i>D.repens</i> human isolate in Vietnam (MH981971)	99	100
<i>D.repens</i> Japanese woman isolate caused in Europe (AB973229)	97	100
<i>D.repens</i> human isolate in India (MF776665)	94	100
<i>D.repens</i> dog isolate in India (MG780293)	93	100
<i>Loa loa</i> (XR 002251421)	100	99
<i>D.repens</i> dog isolate in India (MF776710)	95	99
<i>Onchocerca cervipedis</i> (KT031393)	92	99
<i>Rumenfilaria andersoni</i> (KT878984)	91	99
<i>Elaeophora schneideri</i> (KT878974)	90	99
<i>Brugia malayi</i> (AF036588)	100	98
<i>Wuchereria bancrofti</i> (AY843436)	99	98
<i>Salmonema sp.</i> (MG594295)	99	95
<i>D.immitis</i> dog isolate in Japan (AB973231)	97	95
<i>Gongylnema nepalensis</i> (LC388743)	100	94
<i>Anisakis sp.</i> (U81575)	99	94
<i>D.immitis</i> (AF036638)	100	93
<i>Dracunculus medinensis</i> (AY947720)	100	92

(B2) 18SF2 rRNA

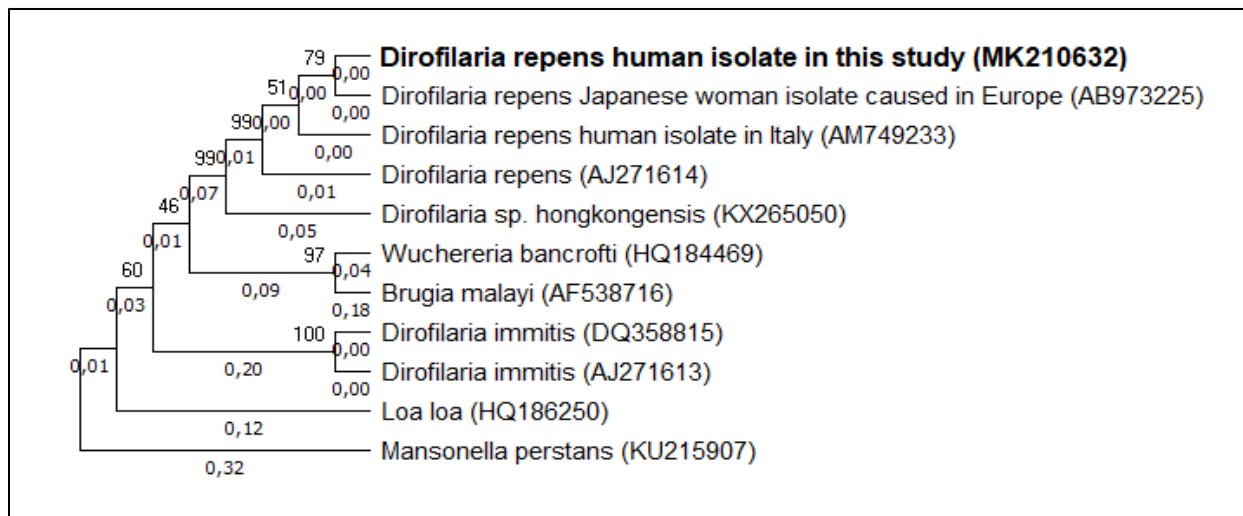
Εικόνα 3.5.: 18SF2 rRNA Original tree



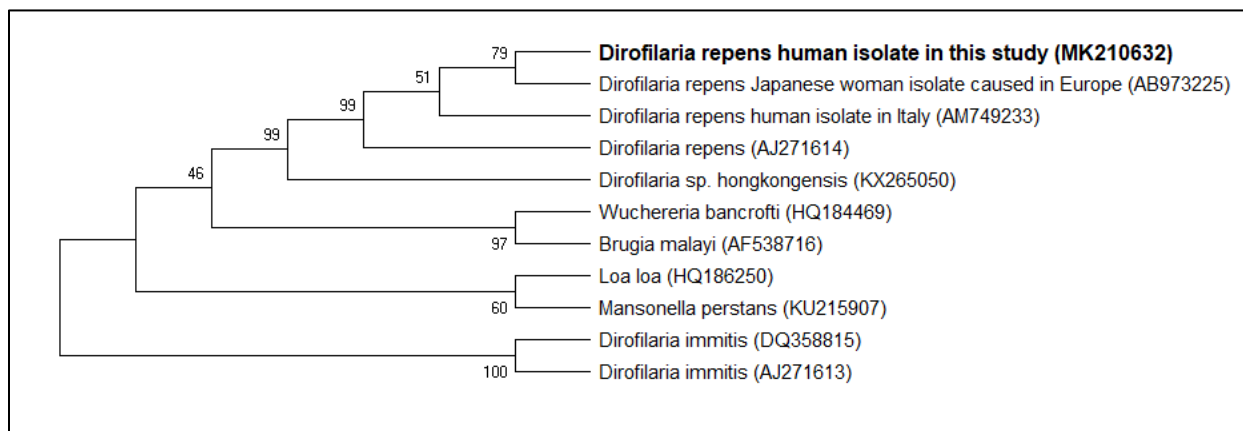
Εικόνα 3.6.: 18SF2 rRNA - Bootstrap consensus tree

Πίνακας 3.3.: Αποτελέσματα BLAST για 18SF2 rRNA

Περιγραφή αλληλουχίας (αριθμός πρόσβασης)	Κάλυψη (%) αλληλουχίας παρασίτου ενδιαφέροντος	Συντελεστής ομοιότητας (%)
<i>D.repens</i> Japanese woman isolate caused in Europe (AB973229)	97	100
<i>Loa loa</i> (XR 002251421)	100	99
<i>Brugia malayi</i> (AF036588)	100	98
<i>Wuchereria bancrofti</i> (AY843436)	99	98
<i>Filaroid sp.</i> (EF081340)	99	98
<i>Salmonema sp.</i> (MG594295)	99	95
<i>D.immitis</i> dog isolate in Japan (AB973231)	97	95
<i>Gongylonema nepalensis</i> (LC388743)	100	94
<i>Anisakis sp.</i> (U81575)	99	94
<i>D.immitis</i> (AF036638)	100	93
<i>Dracunculus medinensis</i> (AY947720)	100	92

(Γ) COI

Εικόνα 3.7.: COI - Original tree



Εικόνα 3.8.: COI - Bootstrap consensus tree

Πίνακας 3.4.: Αποτελέσματα BLAST για COI

Περιγραφή αλληλουχίας (αριθμός πρόσβασης)	Κάλυψη (%) αλληλουχίας παρασίτου ενδιαφέροντος	Συντελεστής ομοιότητας (%)
<i>D.repens</i> Japanese woman isolate caused in Europe (AB973225)	93	99
<i>D.repens</i> (AJ271614)	77	99
<i>D.repens</i> human isolate in Italy (AM749233)	73	99
<i>Dirofilaria sp. hongkongensis</i> (KX265050)	99	96
<i>Wuchereria bancrofti</i> (HQ184469)	99	91
<i>Loa loa</i> (HQ186250)	99	90
<i>D.immitis</i> (DQ358815)	77	90
<i>D.immitis</i> (AJ271613)	77	90
<i>Brugia malayi</i> (AF538716)	99	89
<i>Mansonella perstans</i> (KU215907)	79	87

Λαμβάνοντας, λοιπόν, υπόψη τα αποτελέσματα των παραπάνω φυλογενετικών αναλύσεων, η διαδικασία της ταυτοποίησης του υπό μελέτη παρασίτου ολοκληρώθηκε με την εξαγωγή του συμπεράσματος ότι το προαναφερθέν παράσιτο είναι *Dirofilaria repens*. Το συμπέρασμα αυτό βασίζεται, κυρίως, στην ύπαρξη ταυτόσημων αλληλουχιών 12S και 18S rRNA με άλλα παράσιτα του είδους *Dirofilaria repens*, παρόλο που η αλληλουχία COI ήταν διαφορετική σε σχέση με όλες τις αλληλουχίες που βρίσκονται αναρτημένες στην GenBank (NCBI) για το είδος *Dirofilaria repens*, όπως, επίσης, και σε σχέση με οποιεσδήποτε άλλες αντίστοιχες αλληλουχίες άλλων ειδών, όπως παρατηρήθηκε κατά τη φυλογενετική ανάλυση.

3. ΣΥΖΗΤΗΣΗ

Σε προηγούμενη ενότητα έγινε αναφορά στη γεωγραφική εξάπλωση παρασιτικών ελμίνθων του γένους *Dirofilaria*. Σε κάθε περίπτωση, η γνώση αυτή, γενικότερα για οποιονδήποτε μικροοργανισμό, αποτελεί σημαντικότερη πληροφορία, καθώς επιτρέπει, δεδομένων πάντοτε και των κοινωνικοοικονομικών και πολιτικών συνθηκών της εκάστοτε περιοχής, πρωτίστως τη λήψη μέτρων προφύλαξης απέναντι σε ενδεχόμενες απειλές και δευτερευόντως την αντιμετώπιση προβλημάτων, εφόσον υπάρχουν τα αντίστοιχα στοιχεία που κατευθύνουν τον ειδικό στο ορθό απαιτούμενο συμπέρασμα.

Ωστόσο, αξίζει να σημειωθεί ότι η λανθασμένη αναγνώριση των διαφόρων ειδών και συγκεκριμένα των ειδών *Dirofilaria*, είναι πιθανή όταν αυτή βασίζεται μόνο στην παρατήρηση των μορφολογικών χαρακτηριστικών. Παράγοντες όπως, η ποιότητα των συνθηκών δειγματοληψίας και τα ανώριμα στάδια των παρασίτων καθιστούν την μακροσκοπική παρατήρηση ακόμη πιο δύσκολη. Για το λόγο αυτό αξιοποιούμε τις τεχνικές αλληλούχισης [19].

Συνεχίζοντας, λοιπόν, πρέπει να αναφερθεί ότι η ικανότητα ταυτοποίησης μικροοργανισμών με τεχνικές μοριακής ανάλυσης κρίνεται μέγιστης σημασίας. Αυτό, διότι σε συνδυασμό με άλλες πληροφορίες συχνά προσφέρει στοιχεία σχετικά με την προέλευση του συγκεκριμένου μικροοργανισμού, ενώ είναι αυταπόδεικτο ότι πρόκειται για πολύ πιο αξιόπιστη τεχνική σε σχέση με την απλή παρατήρηση μορφολογικών χαρακτηριστικών.

Πιο συγκεκριμένα, στην παρούσα περίπτωση, η αλληλουχία 12S rRNA (MK192091) του παρασίτου ενδιαφέροντος εμφανίζει υψηλή ομοιότητα με αντίστοιχες αλληλουχίες *Dirofilaria repens* που απομονώθηκαν από σκύλο στην Ιταλία (KX265072 - 100%), από σκύλο, επίσης, στη Γερμανία (KX265089 - 99%) και από άνθρωπο στην Ιαπωνία (AB973228 - 99%), η οποία

τελευταία, ωστόσο, κρίθηκε ως εισαγόμενη από την Ευρώπη. Τα αποτελέσματα για τις δύο αλληλουχίες του 18S rRNA (MK192092, MK192093) του παρασίτου ενδιαφέροντος, που παρατίθενται στην προηγούμενη ενότητα, εξίσου, εμφανίζουν υψηλή ομοιότητα με κάποιες άλλες αντίστοιχες αλληλουχίες. Ανάμεσα σε αυτές με την υψηλή ομοιότητα εμφανίζεται και πάλι εκείνη η περίπτωση της *Dirofilaria repens* που απομονώθηκε από άνθρωπο στην Ιαπωνία (AB973229 – 100% για 18SF και 18SF2). Παρατηρείται το ίδιο φαινόμενο και στα αποτελέσματα της ανάλυσης για COI (MK210632), με την προαναφερθείσα περίπτωση διροφιλαρίωσης (AB973225 – 99%), ενώ επίσης εμφανίζεται υψηλή ομοιότητα και με την αντίστοιχη αλληλουχία μιας *Dirofilaria repens* που απομονώθηκε από άνθρωπο στην Ιταλία (AM749233 – 99%). Εντυφώνοντας λίγο περισσότερο στην περίπτωση της διροφιλαρίωσης του ανθρώπου στην Ιαπωνία, διαπιστώνει κανείς ότι προκλήθηκε από νύγμα εντόμου στην Ιταλία. Έτσι, αρχίζουν και διαμορφώνονται κάποιες υποψίες αναφορικά με την προέλευση της διροφιλαρίωσης της παρούσας εργασίας. Βέβαια, αυτό δε σημαίνει ότι είναι υποχρεωτικό ο πάσχων να είχε ταξιδέψει στη συγκεκριμένη περιοχή, καθώς όντας κοντινή ως προς την Ελλάδα, θα μπορούσε κάλλιστα το συγκεκριμένο αρθρόποδο, το οποίο ευθύνεται για τη διροφιλαρίωση που μελετάται στην παρούσα εργασία, εάν έχει προέλευση από την Ιταλία, να μεταφέρθηκε στην Ελλάδα από την Ιταλία με κάποιο τρόπο (εγκλωβισμένο σε αεροσκάφος, αυτοκίνητο, κιβώτιο κλπ). Επίσης, το γεγονός ότι δεν υπάρχουν αναρτημένες στην αντίστοιχη βάση δεδομένων περιπτώσεις διροφιλαρίωσης στην Ελλάδα δε σημαίνει και ότι αυτές δεν υφίστανται. Επομένως, υπάρχει μεγάλη πιθανότητα να πρόκειται για ένα είδος που απαντάται και στην Ελλάδα και απλώς να είναι η πρώτη φορά που λαμβάνει χώρα η ανάρτηση των αντίστοιχων αλληλουχιών στην αντίστοιχη βάση δεδομένων.

Παρόλα αυτά, αναφορικά με το σκέλος της ταυτοποίησης του παρασίτου ενδιαφέροντος, κάποιος θα μπορούσε να παρατηρήσει ότι οι αλληλουχίες του σε σχέση με εκείνες των *Dirofilaria repens* από την Ιταλία δεν εμφανίζουν παντού 100% ομοιότητα, ενώ για τη φυλογενετική ανάλυση του COI δεν υπάρχει καταγεγραμμένη αλληλουχία κανενός είδους με ομοιότητα 100%.

Είναι όμως αυτά ικανά να δηλώσουν ότι το εξεταζόμενο παράσιτο είναι ένα καινούριο είδος, όπως συνέβη και στην περίπτωση του *Dirofilaria hongkongensis*; Είναι γεγονός ότι πρώτα απ' όλα οι διαφορές στην ομοιότητα είναι πολύ μικρές μεταξύ των εξεταζόμενων αλληλουχιών και πιο συγκεκριμένα διαφέρουν το πολύ έως 1% και αυτό δεν ισχύει για όλες τις αλληλουχίες. Επίσης, οι καταγεγραμμένες αλληλουχίες του παρασίτου ενδιαφέροντος που χρησιμοποιούνται στην παρούσα εργασία δε σημαίνει ότι καλύπτουν όλη την πραγματικά υπάρχουσα αλληλουχία, ούτε αντίστοιχα ισχύει αυτό για το σύνολο των υπολοίπων αλληλουχιών με τις οποίες επιτελούνται οι παρούσες φυλογενετικές αναλύσεις. Σε αντίθετη περίπτωση ενδέχεται τα ποσοστά διαφοροποίησης να ήταν ακόμη μικρότερα. Τέλος, είναι γεγονός ότι παρουσιάζονται πολλά κοινά με αλληλουχίες *Dirofilaria repens* συγκεκριμένης γεωγραφικής προέλευσης και μάλιστα κοντινής ως προς την περιοχή όπου διαμένει ο ασθενής (Ελλάδα), ενώ, επίσης, όπως άλλωστε έχει ήδη αναφερθεί σε προηγούμενη ενότητα, δεν υφίστανται ουσιαστικοί έλεγχοι ποιότητας στις εκάστοτε βάσεις δεδομένων που αναρτώνται οι αντίστοιχες αλληλουχίες, γεγονός που μπορεί να συνεπάγεται λάθη στα οποία και ενδέχεται να οφείλονται μικρές ποσοστιαίες αποκλίσεις ομοιότητας. Επομένως, στην παρούσα περίπτωση αναφερόμαστε ξεκάθαρα σε παράσιτο *Dirofilaria repens* [21].

Επομένως, ανακεφαλαιώνοντας, το παράσιτο ενδιαφέροντος χαρακτηρίστηκε ως προς τις αλληλουχίες των 12S rRNA, 18S rRNA και COI, οι οποίες και έχουν αναρτηθεί στην GenBank

(NCBI). Ωστόσο, αν και υφίστανται κάποιες υποψίες, όπως προαναφέρθηκε, αναφορικά με την περίπτωση διροφιλαρίωσης που αναλύεται στην παρούσα εργασία, αξίζει να σημειωθεί ότι οι λεπτομερείς πληροφορίες ως προς τη γεωγραφική γενετική διαφοροποίηση του *Dirofilaria repens* είναι σε γενικές γραμμές ασαφείς λόγω ποικίλων εξωγενών παραγόντων, οι οποίοι σε τελικό απολογισμό συγκεντρώνονται γύρω από το πρόβλημα του υψηλού κόστους μιας λεπτομερούς γενετικής ανάλυσης. Όπως έχει προαναφερθεί, μπορεί το *Dirofilaria repens* να μην αποτελεί άμεσο κίνδυνο για την επιβίωση του ασθενή, ωστόσο, γενικότερα, το θέμα των παρασιτώσεων συνιστά ένα πρόβλημα μέγιστης ιατροοικονομικής σημασίας. Για το λόγο αυτό θα πρέπει να λαμβάνονται υπόψη, με ιδιαίτερη προσοχή, περιπτώσεις αυξημένης εμφάνισης αντίστοιχων κρουσμάτων, καθώς, επίσης, και παρουσίας κρουσμάτων με παρόντες μη ενδημικούς μικροοργανισμούς. Επομένως, για όλους τους προαναφερθέντες λόγους, κρίνεται απαραίτητη η, κυρίως οικονομική, στήριξη της εκάστοτε πολιτείας, προκειμένου να καταστεί εφικτός ο εμπλουτισμός μιας λεπτομερούς γενετικής βάσης δεδομένων διαφόρων ειδών [56].

Στην εργασία αυτή έγινε μία προσπάθεια εισαγωγής στον τομέα της εξελικτικής βιολογίας που συνδέεται με τη φυλογενετική ανάλυση. Επίσης, επιχειρήθηκε να δοθεί μια συνολική περιγραφή των πιο χαρακτηριστικών μεθοδολογιών και των προβλημάτων της φυλογενετικής ανάλυσης, να καλυφθούν τα διάφορα τμήματα κατασκευής ενός φυλογενετικού δέντρου και να δοθεί η δυνατότητα στον αναγνώστη να μπορέσει, εάν το επιθυμεί, να επαναλάβει τη διαδικασία προσαρμοσμένη στις ανάλογες συνθήκες, καθώς επίσης και να ταυτοποιηθεί μοριακά το παράσιτο ενδιαφέροντος με χρήση των αντίστοιχων συσκευών, προγραμμάτων, βάσεων δεδομένων και εργαλείων που έχουν ήδη αναφερθεί, ενώ ταυτόχρονα επιτεύχθηκε για πρώτη φορά η ανάρτηση αλληλουχιών από το συγκεκριμένου είδους παρασίτου προερχόμενο από τον ελλαδικό χώρο, στην αντίστοιχη βάση δεδομένων του NCBI [35].

Δεδομένης της πολυπλοκότητας των αναλύσεων που περιγράφηκαν στη διαδοχή των πέντε βημάτων θα ήταν καλό σε αυτό το σημείο να συζητηθούν με συγκριτικό τρόπο. Σε ό,τι αφορά τους τρόπους υπολογισμού εξελικτικών αποστάσεων θα πρέπει να ειπωθεί πως τα μοντέλα νουκλεοτιδικών αντικαταστάσεων πρέπει να επιλέγονται με βάση την ανάλυση που κρίνεται ως πιο κατάλληλη να λάβει χώρα. Έτσι είναι προφανές ότι ένα πολυπλοκότερο μοντέλο θα πρέπει να χρησιμοποιηθεί εφόσον μια λεπτομερής ανάλυση είναι απαραίτητη. Από την άλλη πλευρά, όμως, όσο πιο πολύπλοκο είναι ένα μοντέλο, τόσο περισσότερες παραδοχές κάνει και επομένως πολλές φορές ίσως ένα απλούστερο μοντέλο είναι προτιμότερο [13].

Πιο συγκεκριμένα, τώρα, αναφορικά με το πρόβλημα της ανασύστασης των φυλογενετικών δέντρων, είναι αρκετά περίπλοκο, με πολλές και διαφορετικές υπάρχουσες μεθόδους ως προς την επίλυσή του. Ανάλογα με τη μεθοδολογία που επιλέγεται, το δέντρο που προκύπτει μπορεί να είναι διαφορετικό διότι ανάλογα με την πληροφορία που ενδιαφέρει, η επιλεγόμενη μέθοδος και τα επιμέρους χαρακτηριστικά των παραμέτρων της διαφέρουν στην εκάστοτε περίπτωση. Ένα τυπικό παράδειγμα δύο διαφορετικών προσεγγίσεων κοινών προβλημάτων στο χώρο της εξελικτικής βιολογίας προσφέρεται από τους «κλαδιστές» και τους «φαινετιστές», οι οποίοι τελευταίοι είχαν αναφερθεί και στην εισαγωγή. Πρόκειται για δύο αντικρουόμενες σχολές με σχεδόν εντελώς αντίθετες απόψεις. Οι κλαδιστές επικέντρωναν το ενδιαφέρον τους στην πορεία της εξέλιξης και ως πιο σημαντικό σε ένα δέντρο θεωρούσαν την τοπολογία, τη διάρθρωση των κλάδων και τον αριθμό των διακλαδώσεων. Έτσι, προτιμούσαν μεθόδους όπως η MP. Οι φαινετιστές, από την άλλη πλευρά, έκριναν ως σημαντικότερες τις επιμέρους σχέσεις ομοιότητας μεταξύ των ειδών και των αλληλουχιών. Έτσι, προτιμούσαν μεθόδους αποστάσεων όπως η UPGMA για να εξάγουν τα συμπεράσματά τους [13, 32].

Επιπλέον, όσον αφορά την ταχύτητα εκτέλεσης των μεθοδολογιών, οι μέθοδοι απόστασης είναι γενικά γρηγορότερες. Η σχέση ταχύτητας εκτέλεσης είναι ανάλογη με τον αριθμό των υποθέσεων που κάνει η κάθε μέθοδος. Έτσι η UPGMA για παράδειγμα, που έχει ως σημείο έναρξης την απλούστερη υπόθεση για σταθερό ρυθμό αντικαταστάσεων είναι και η ταχύτερη. Προφανώς, η ταχύτητα δε θα πρέπει να είναι το μοναδικό κριτήριο και μια πολύ απλή μέθοδος δεν είναι υποχρεωτικά η καλύτερη. Μέθοδοι που επιχειρούν να λύσουν το πρόβλημα μέσω βελτιστοποίησης όπως η MP και η ML είναι λογικό να απαιτούν περισσότερο χρόνο. Το γεγονός, ωστόσο, ότι αναζητούν μια βέλτιστη λύση συγκρίνοντας ένα μεγάλο αριθμό δέντρων συνιστά την αποφυγή τους όταν τίθενται περιπτώσεις πολύ μεγάλου αριθμού συγκρινόμενων αλληλουχιών. Σε κάθε περίπτωση, η επιλογή της καλύτερης μεθόδου γίνεται με βάση την όσο το δυνατόν καλύτερη γνώση του προβλήματος και έχοντας υπόψη το ακριβές ερώτημα που είναι επιθυμητό να απαντηθεί. Ανάλογα με το αν αυτό είναι, παραδείγματος χάρη, κλαδιστικού ή φαινετικού τύπου, αν τα δεδομένα είναι πολλά ή διαχειρίσιμα, αν προσφέρονται ή όχι για μια ανάλυση αποστάσεων, καθώς και άλλα πολλά. Δυστυχώς, στην περίπτωση αυτή, όπως και σε πολλά υφιστάμενα βιολογικά προβλήματα, δεν υπάρχει μια χρυσή συνταγή [13, 32, 44, 57].

Με βάση τη σχετική βιβλιογραφία και μελέτη των υφιστάμενων μεθοδολογιών αποφασίστηκε να ακολουθηθεί μια οδός που να μπορεί ταυτοχρόνως να προσφέρει ένα ποιοτικό αποτέλεσμα, το οποίο να είναι εύκολο να επαναληφθεί από οποιονδήποτε.

Πρωτίστως, λοιπόν, για τη συγκεκριμένη εργασία έπρεπε να επιλεγεί ένα πρόγραμμα που να παρέχει τη δυνατότητα εύκολης χρήσης και επαρκών επιλογών ως προς τις μεθόδους και τα εργαλεία, προκειμένου να είναι δυνατή η εξαγωγή ενός έγκυρου αποτελέσματος. Τα γνωστότερα από αυτά τα προγράμματα έχουν ήδη αναφερθεί και μάλιστα έμφαση δόθηκε στο MEGA το οποίο και χρησιμοποιήθηκε για τις ανάγκες αυτής της εργασίας.

Σε δεύτερη φάση, χρησιμοποιήθηκε, για τον εντοπισμό των αλληλουχιών που χρησιμοποιήθηκαν στη φυλογενετική ανάλυση, η βάση δεδομένων που προσφέρεται από το NCBI. Ο κύριος λόγος για τον οποίο χρησιμοποιήθηκε ήταν διότι διαθέτει πληθώρα αλληλουχιών χωρίς να είναι εξειδικευμένη μόνο ως προς ένα συγκεκριμένο τύπο αλληλουχίας. Αυτό βέβαια δε συνεπάγεται ότι καθίσταται περισσότερο έγκυρη από άλλες βάσεις δεδομένων. Επιπλέον λόγοι για τους οποίους επιλέχθηκε είναι ότι, α) πρόκειται για βάση δεδομένων υψηλής κινητικότητας, που σημαίνει ότι ενημερώνεται συχνά και εμπλουτίζεται, με πληθώρα πληροφοριών, β) είναι συμβατή με το πακέτο λογισμικού MEGA-X, το οποίο και χρησιμοποιήθηκε, και γ) είναι και η βάση δεδομένων εκείνη στην οποία είχαν κατατεθεί οι αλληλουχίες του παρασίτου ενδιαφέροντος αυτής της εργασίας [26, 30].

Οι προαναφερθείσες, λοιπόν, αλληλουχίες, εντοπίστηκαν κάνοντας το λεγόμενο BLAST. Με αυτό αποκλείεται ο κίνδυνος να τοποθετηθούν στο τέλος, στο ίδιο φυλογενετικό δέντρο, αλληλουχίες οι οποίες δε είναι ομόλογες, γεγονός που διασφαλίζει σ' ένα βαθμό το αποτέλεσμα [13].

Οι αλληλουχίες που προέκυψαν από την παραπάνω διαδικασία και χρησιμοποιήθηκαν στις εκάστοτε αναλύσεις επιλέχθηκαν με βάση δύο κριτήρια: α) Ο συντελεστής ομοιότητας θα έπρεπε να βρίσκεται μεταξύ 85% και 100% β) ο συντελεστής κάλυψης θα έπρεπε να βρίσκεται μεταξύ 80% και 100% για το 12S rRNA και 18S rRNA (βέβαια η πλειοψηφία ήταν μεταξύ 90% και 100%) και μεταξύ 70% και 100% για το COI. Οι δύο αυτές συνθήκες είναι ικανές και αναγκαίες προκειμένου να ληφθεί ένα πιο κατατοπιστικό και ενδεχομένως πιο έγκυρο αποτέλεσμα.

Αναφορικά με το λόγο για τον οποίο χρησιμοποιήθηκαν οι αλληλουχίες 12S, 18S και COI από το παράσιτο ενδιαφέροντος, για τις φυλογενετικές αναλύσεις της παρούσας εργασίας, είναι διότι πρόκειται για τακτική καθιερωμένη σε αντίστοιχες περιπτώσεις, καθώς πρόκειται για ευρέως μελετημένους γενετικούς τύπους, ιδίως οι δύο πρώτοι τύποι αλληλουχίας, σε σύγκριση τουλάχιστον, με αλληλουχίες διαφορετικού τύπου. Αξίζει, βέβαια, να σημειωθεί ότι το COI ως γρηγορότερα εξελισσόμενο (fast evolving) κερδίζει σταδιακά ισχυρή θέση σε φυλογενετικές αναλύσεις που αφορούν σε σύγκριση ειδών. Στις προαναφερθείσες αναλύσεις πολλοί κρίνουν απαραίτητη και τη σύγκριση των 12S rRNA αλληλουχιών. Σαφέστατα, όπως και στα περισσότερα βιολογικά προβλήματα, υπάρχουν πολλές και αντικρουόμενες τοποθετήσεις ως προς την ορθότητα χρήσης τους, με απλούστερο παράδειγμα μια σύντομη αναζήτηση στην ιστοσελίδα «ResearchGate», όπου άνθρωποι με επιστημονικό υπόβαθρο έχουν τη δυνατότητα να αναρτήσουν τα ερωτήματά τους και να ξεκινήσει ένας αντίστοιχος διάλογος [1, 2, 38].

Ως προς το εργαλείο που χρησιμοποιήθηκε για την ευθυγράμμιση οφείλει να αναφερθεί ότι και σε αυτή την περίπτωση υφίσταται διχασμός απόψεων ως προς το ποιο είναι το καλύτερο. Είναι γεγονός ότι η επιστήμη της ανασύστασης των φυλογενετικών δέντρων είναι σε μεγάλο βαθμό θεωρητική και έτσι ως κλασσικό βιολογικό πρόβλημα δε θα μπορούσε παρά να μην προσφέρει με σιγουριά την ιδανική λύση και σε πολλά από τα επιμέρους βήματά της. Αξίζει, βέβαια, να αναφερθεί ότι, εδώ, απορρίφθηκαν εξ αρχής μέθοδοι αποκλειστικής ευθυγράμμισης ζεύγους αλληλουχιών, καθώς ήταν επιθυμητή η ταυτόχρονη σύγκριση περισσότερων των δύο αλληλουχιών προκειμένου να υφίσταται μια πιο ολοκληρωμένη εικόνα της ιστορίας του μικροοργανισμού ενδιαφέροντος. Έγινε χρήση, λοιπόν, μέσω του MEGA-X, του εργαλείου ClustalW το οποίο μπορεί να αναφερθεί ότι διαθέτει μεγαλύτερη μνήμη συγκριτικά με τον μεγάλο του αντίπαλο, το «Muscle», το οποίο θεωρείται γρηγορότερο [13, 25, 34, 36].

Σημαντικό είναι, επίσης, να αναφερθεί, ότι μετά τη διαδικασία της ευθυγράμμισης των αλληλουχιών διαγράφηκαν τα τμήματα εκείνα που επεκτείνονταν πέραν των ορίων της αλληλουχίας του παρασίτου ενδιαφέροντος έτσι ώστε να αποφευχθούν εσφαλμένα αποτελέσματα σύγκρισης. Επομένως, έγινε μια προσπάθεια στο τελικό στάδιο της φυλογενετικής ανάλυσης, εκείνο δηλαδή της κατασκευής του φυλογενετικού δέντρου, να συγκριθούν τα ευθυγραμμισμένα τμήματα των συγκρινόμενων αλληλουχιών που εμφάνιζαν αλληλοεπικάλυψη. Δυστυχώς, ήταν αναπόφευκτο κάποιες αλληλουχίες να εμφανίζουν ελαφρώς μικρότερου βαθμού αλληλοεπικάλυψη με τις αλληλουχίες του παρασίτου ενδιαφέροντος, ωστόσο αυτό περιορίστηκε στο μέγιστο δυνατό βαθμό.

Στις παρούσες φυλογενετικές αναλύσεις αποφασίστηκε να περιοριστεί ο αριθμός των συγκρινόμενων αλληλουχιών σε νούμερο μικρότερο του είκοσι. Αυτό διότι, όπως έχει ήδη αναφερθεί, το φυλογενετικό δέντρο που προκύπτει στο τέλος μιας ανάλυσης είναι απλώς μία υπόθεση. Όσο μεγαλύτερος ο αριθμός των συγκρινόμενων αλληλουχιών, τόσο μεγαλύτερος ο αριθμός των υποθέσεων και επομένως τόσο μεγαλύτερη η πιθανότητα κατασκευής ενός μη ορθού εξελικτικά φυλογενετικού δέντρου. Τυπικό παράδειγμα αποτελεί το γεγονός ότι πολλοί αναλυτές επιλέγουν στα ML φυλογενετικά δέντρα τους να μην εμφανίζουν ποσοστά χαμηλότερα του 50-70%. Προσωπική άποψη είναι ότι από τη στιγμή που έχει λάβει χώρα η χρήση του «bootstrapping», το οποίο στην ουσία αποτελεί ένα στατιστικό έλεγχο, οφείλουν να αναπαριστώνται τα εκάστοτε ποσοστά ακόμη και αν δεν κρίνονται ιδιαίτερα ενθαρρυντικά, όπως στις περιπτώσεις που είναι κάτω του 50% [19, 57].

Μεγάλης σημασίας είναι λοιπόν και η χρήση του «bootstrapping». Αυτό, διότι κάθε έρευνα, προκειμένου να φτάσει σε ένα συμπέρασμα οφείλει, πέραν της αξιοποίησης των δεδομένων που

έχει συλλέξει, να αξιολογεί τη στατιστική σημαντικότητα των αποτελεσμάτων της. Αυτός ακριβώς είναι και ο ρόλος της προαναφερθείσας μεθόδου [13, 36].

Αναφορικά, τώρα, με την ένταξη εξωομάδας στα φυλογενετικά δέντρα της παρούσας διπλωματικής εργασία δεν κρίθηκε απαραίτητη, καθώς αυτό που ενδιέφερε ήταν οι σχέσεις μεταξύ των συγκρινόμενων αλληλουχιών, προκειμένου να ταυτοποιηθεί το παράσιτο ενδιαφέροντος, γεγονός στο οποίο η ύπαρξη ή μη εξωομάδας δε θα διαδραμάτιζε πρωτεύοντα ρόλο [13].

Τέλος, στα αποτελέσματα αποφασίστηκε να παρουσιαστούν τόσο τα «Original trees», όσο και τα «Bootstrap consensus tree». Τα πρώτα ουσιαστικά είναι δέντρα που αναπαριστούν τη μοναδική τοπολογία που θεωρείται ως καλύτερη βάση της μεθόδου που έχει επιλεγεί. Στη βάση της μοναδικής αυτής τοπολογίας είναι δυνατή η αναπαράσταση των αποστάσεων των εκάστοτε βραχιόνων-κλαδιών. Ωστόσο, δε θα πρέπει να υπάρχει σύγχυση με τα «Bootstrap consensus trees» τα οποία με βάση τα όσα ισχύουν τη στιγμή που συγγράφεται η παρούσα εργασία θεωρούνται έγκυρα ως προς τη χρήση τους για την αναπαράσταση των εξελικτικών σχέσεων των συγκρινόμενων αλληλουχιών και τα οποία αποτελούν το ουσιαστικό αποτέλεσμα της παρούσας μελέτης [13, 36].

Στο σημείο αυτό να επισημανθεί ότι η χρήση του όρου «έγκυρο» για τα φυλογενετικά δέντρα της εργασίας αυτής, αντί του όρου «ορθό», γίνεται συνειδητά. Έγκυρα αποκαλούνται από την άποψη της αξιοπιστίας τους ως προς την τήρηση των κανόνων και οδηγιών που υφίστανται στην επιστημονική κοινότητα κατά το χρονικό διάστημα που λαμβάνει χώρα η συγγραφή της παρούσας εργασίας. Ο όρος «ορθό» από την άλλη αποφεύγεται, ακριβώς, διότι δεν είναι εφικτό να γνωρίζουμε γενικότερα στην κατασκευή φυλογενετικών δέντρων εάν το

δέντρο που εν τέλει κατασκευάσαμε είναι όντως μια ρεαλιστική απεικόνιση της εξέλιξης των ειδών. Προς αποφυγή παρεξηγήσεων, τονίζεται ότι ισχύει πως σε φυλογενετικές αναλύσεις, ιδίως μικρού μεγέθους, που θέλουμε να βρούμε την απόσταση μεταξύ διαφόρων αλληλουχιών μπορούμε να το εκφράσουμε ποσοστιαία και τυπικά θα είναι ορθό ως αποτέλεσμα. Ωστόσο, αυτό που επιδιώκεται δια μέσω της προσπάθειας ανασύστασης των φυλογενετικών δέντρων είναι η απεικόνιση της εξέλιξης και εκεί είναι που τίθεται το ερωτηματικό ως προς την ορθότητά τους [13, 36, 57].

Πλέον, η επιστήμη, σε πληθώρα τομέων, τρέχει με πολύ γρήγορους ρυθμούς. Σε κάθε περίπτωση απαραίτητη προϋπόθεση της προόδου της είναι η διεξαγωγή πειραμάτων και οι παρατηρήσεις. Στη μελετούμενη περίπτωση, της φυλογενετικής ανάλυσης, αυτό σημαίνει ότι μελλοντικά, εφόσον ήδη εδώ και κάποια χρόνια υφίσταται η λεπτομερής καταγραφή των ειδών, ενδεχομένως να είναι εφικτή η κατασκευή αντίστοιχων μοντέλων φυλογενετικής ανάλυσης που να προσεγγίζουν ακόμη περισσότερο την πραγματική ιστορία της εξέλιξής τους. Σε μια ιδανική περίπτωση θα είναι δυνατή η ανασύσταση του πραγματικού Δέντρου της Ζωής.

Επίσης, είναι γεγονός ότι στα νεότερα χρόνια, η επιστήμη της βιοπληροφορικής, γενικότερα, έχει προσφέρει τα μέγιστα στον προαναφερθέν κλάδο, τείνοντας μάλιστα να βελτιώνεται συνεχώς, δημιουργώντας, έτσι, τη βάση για την προσφορά βοήθειας στη διαδρομή προς την επίλυση ποικίλων βιολογικών και κατ' επέκταση κοινωνικοοικονομικών προβλημάτων. Όπως και σε κάθε άλλη παρόμοια περίπτωση, έτσι και εδώ, είναι δεδομένο ότι πολλά πράγματα ενδέχεται στο πέρας του χρόνου να τροποποιηθούν. Ο λόγος είναι ότι οι φυλογενετικές αναλύσεις προκειμένου να είναι έγκυρες απαιτούν από τη μια μεριά την εφαρμογή της τρέχουσας βιβλιογραφίας και των σύγχρονων μεθόδων της βιοπληροφορικής και από την άλλη, αντλούν δεδομένα από άκρως ενημερωμένες βάσεις δεδομένων που διαρκώς ανανεώνονται.

5. ΠΕΡΙΛΗΨΗ

Η φυλογενετική, υποκλάδος της επιστήμης της εξελικτικής βιολογίας, με τη βοήθεια της βιοπληροφορικής, αποτελεί το πλέον αξιόπιστο μέσο για την μοριακή ταυτοποίηση των παρασίτων που προκαλούν νόσο στον άνθρωπο. Στην παρούσα μελέτη έγινε μία προσπάθεια εισαγωγής στον τομέα της φυλογενετικής ανάλυσης, μέσω συνολικής περιγραφής των πιο χαρακτηριστικών μεθοδολογιών και προβλημάτων της. Συγκεκριμένα επιχειρήθηκε να περιγραφούν τα διάφορα τμήματα κατασκευής ενός φυλογενετικού δέντρου και να ταυτοποιηθεί μοριακά ένας έλμινθας, ο οποίος είχε αφαιρεθεί χειρουργικά από ανθρώπινο οφθαλμό, μέσω σύγκρισης των 12S rRNA (MK192091), 18S rRNA (MK192092, MK192093) και COI (MK210632) αλληλουχιών του με αντίστοιχες αλληλουχίες άλλων οργανισμών, στην ανάλογη βάση δεδομένων του NCBI, με χρήση του πακέτου λογισμικού MEGA-X. Οι αλληλουχίες 12S και 18S rRNA του υπό μελέτη έλμινθα παρουσίασαν συντελεστή ομοιότητας 100% και η COI 99% με αντίστοιχες αλληλουχίες ελμίνθων του είδους *Dirofilaria repens* με προέλευση από την Ιταλία (KX265072, AB973229 και AB973225 αντίστοιχα), γεγονός που οδήγησε και στην τελική ταυτοποίησή του ως *Dirofilaria repens*.

Λέξεις κλειδιά: *Dirofilaria repens*, παράσιτο, βιοπληροφορική, φυλογενετική ανάλυση, φυλογενετικά δέντρα, μοριακή ταυτοποίηση, MEGA-X, NCBI, 12S rRNA, 18S rRNA, COI

6. ABSTRACT

Phylogenetics, which is a subfield of evolutionary biology, in combination with bioinformatics is currently considered to be the most reliable tool for accurate molecular identification of parasites causing human or animal diseases. An extensive synopsis of the most typical methodologies and problems encountered is presented, together with a phylogenetic tree building process description. Additionally, the molecular identification of a nematode helminth surgically removed from a human eye and microscopically being identified as *Dirofilaria* spp. was performed. The molecular identification was based on the comparison of the 12S rRNA (MK192091), 18S rRNA (MK192092, MK192093) and COI (MK210632) sequences of this helminth with the respective sequences of other organisms using the MEGA-X software package and the proper database provided by NCBI. The 12S rRNA and 18S rRNA parasite of interest sequences proved to be identical (100% similarity), while its COI sequence was 99% similar to *Dirofilaria repens* sequences from Italy (KX265072, AB973229 and AB973225 respectively). These findings resulted to its identification as nematode of the species *Dirofilaria repens*.

Key words: *Dirofilaria repens*, parasite, bioinformatics, phylogenetic analysis, phylogenetic trees, molecular identification, MEGA-X, NCBI, 12S rRNA, 18S rRNA, COI

7. ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Anand Patwardhan, Samit Ray, & Amit Roy. (2014). Molecular markers in phylogenetic studies - a review. *Phylogenetics & Evolutionary Biology*, 2(2).
2. Aneta A. Ptaszyska, Jacek towski, Sebastian Gnat, & Wanda Maek. (23/08/2011). Application of COI sequences in studies of phylogenetic relationships among 40 Apionidae species. *Journal of Insect Science*, 12.
3. Antonio Soares, Ricardo Râbelo, & Alexandre Delbem. (15/07/2017). Optimization based on phylogram analysis. *Expert Systems with Applications*, 78, 32-50.
4. Bioinformatics.org. (n.d.). Retrieved 28/01/2019, from Bioinformatics.org:
<https://www.bioinformatics.org/sms/iupac.html>
5. Brudno M., Malde S., Poliakov A., Do C.B., Couronne O., Dubchak I., & Serafim Batzoglou. (2003). Glocal alignment: finding rearrangements during alignment. *Bioinformatics*, 19 Suppl 1:i54-62, 54–62.
6. Centers for Disease Control and Prevantion. (08/02/2012). *Biology - Life Cycle of D. repens*. Retrieved 23/11/2018, from Centers for Disease Control and Prevantion:
https://www.cdc.gov/parasites/dirofilariasis/biology_d_repens.html
7. Choudhuri, S. (2014). *Bioinformatics for Beginners (1st Edition)*. Academic Press.
8. DDBJ. (n.d.). Retrieved 28/01/2019, from DDBJ: <https://www.ddbj.nig.ac.jp/index-e.html>

9. DEPARTMENT OF BIOLOGICAL SCIENCE AT FLORIDA STATE UNIVERSITY. (n.d.).
Tree Searching Methods. Retrieved 06/02/2019, from DEPARTMENT OF
 BIOLOGICAL SCIENCE AT FLORIDA STATE UNIVERSITY:
<https://www.bio.fsu.edu/~stevet/BSC5936/Swofford.F2003.2.pdf>
10. EMBL-EBI. (n.d.). Retrieved 28/01/2019, from <https://www.ebi.ac.uk/>
11. Gina Cannarozzi. (n.d.). *Phylogeny and Evolution*. Retrieved 24/11/2018, from EMBnet node
 Switzerland: [https://embnet.vital-
 it.ch/CoursEMBnet/Basel06/slides/embnet_basel_day5.pdf](https://embnet.vital-it.ch/CoursEMBnet/Basel06/slides/embnet_basel_day5.pdf)
12. Guilherme Issao Fujjwara, Pete Kruskal, Arkajit Dey, Carlos Pards , Victor Costan, Marten
 van Dijk, Andreea Bodnari, Wes Brown, Sarah Spencer, & Nathaniel Parrish.
 (03/10/2012). *6.047/6.878 Lecture 2: Sequence Alignment and Dynamic*. Retrieved
 24/11/2018, from MIT - Massachusetts Institute of Technology:
[http://web.mit.edu/6.047/scribe_notes/scribe_notes_local/2012-
 bak/Lecture02_DynamicProgramming/Lecture02_DynamicProgramming_standalone.pdf](http://web.mit.edu/6.047/scribe_notes/scribe_notes_local/2012-

 bak/Lecture02_DynamicProgramming/Lecture02_DynamicProgramming_standalone.pdf)
13. Hall, B. (2004). *Tutorial: Create a Tree!* (2nd ed.). MA, USA: Sinauer Associates
 Sunderland.
14. Hesper, B., & Hogeweg , P. (1970). Bioinformatica: een werkconcept. *I*(6), 28–29.
15. Hogeweg , P. (2011). The Roots of Bioinformatics in Theoretical Biology. (D. Searls, Ed.)
7(3).
16. Hogeweg, P. (1978). Simulating the growth of cellular forms. *SAGE*, *31*(3), 90–96.

17. Institute for Systems Biology. (n.d.). *RepeatMasker Web Server*. Retrieved 24/11/2018, from Institute for Systems Biology: <http://www.repeatmasker.org/cgi-bin/WEBRepeatMasker>
18. Joseph Felsenstein. (2004). *Inferring Phylogenies*. Sunderland, Massachusetts: Sinauer Associates.
19. Jun Suzuki, Seiki Kobayashi, Utako Okata, Hitomi Matsuzaki, Mariko Mori, Ko-Ron Chen, & Satoshi Iwata. (27/01/2015). Molecular analysis of *Dirofilaria repens* removed from a subcutaneous nodule in a Japanese woman after a tour to Europ. *Parasite*, 22(2).
20. Kalogeropoulos, C., Stefaniotou, M., Gorgoli, K., Papadopoulou, C., Pappa, C., & Paschidis, C. (2014). Ocular Dirofilariasis: A Case Series of 8 Patients. *Middle East African Journal of Ophthalmology*, 21(4), 312–316.
21. Kelvin K. W. To, Samson S. Y. Wong, Rosana W. S. Poon, Nigel J. Trendell-Smith, Antonio H. Y. Ngan, Jacky W. K. Lam, Tommy H. C. Tang, Ah-Kian AhChong, Joshua Chi-Hang Kan, Kwok-Hung Chan, and Kwok-Yung Yuen (11/2012). A Novel *Dirofilaria* Species Causing Human and Canine Infections in Hong Kong. *Journal of Clinical Microbiology*, 50(11), 3534–3541.
22. King, R. C., Stansfield, W. D., & Mulligan, P. K. (2006). A dictionary of genetics. *Oxford University Press*, 336.
23. Lesk, A. (n.d.). *Bioinformatics*. Retrieved 23/11/2018, from Encyclopaedia Britannica: <https://www.britannica.com/science/bioinformatics>

24. Medical University of Vienna. (n.d.). *Unit 03c: Heuristic methods methods for pairwise pairwise*. Retrieved 24/11/2018, from Medical University of Vienna:
http://www.meduniwien.ac.at/imc/biosim/vo_bioinformatik_05_06/Bi03c_print.pdf
25. MEGA (Molecular Evolutionary Genetics Analysis). (n.d.). *CLUSTALW Options (DNA)*. Retrieved 15/01/2019, from MEGA (Molecular Evolutionary Genetics Analysis):
https://www.megasoftware.net/webhelp/alnbuilder_hc/hc_clustalw_options_dna_dialog.htm
26. MEGA Molecular Evolutionary Genetics Analysis. (n.d.). Retrieved 28/01/2019, from MEGA Molecular Evolutionary Genetics Analysis: <https://www.megasoftware.net/>
27. MEGA (Molecular Evolutionary Genetics Analysis). (n.d.). Retrieved 15/01/2019, from MEGA (Molecular Evolutionary Genetics Analysis):
https://www.megasoftware.net/mega4/WebHelp/part_iv___evolutionary_analysis/constructing_phylogenetic_trees/statistical_tests_of_a_tree_obtained/bootstrap_tests/hc_bootstrap_test_phylogeny.htm
28. Morand, S., Krasnov, B., & Littlewood, D. (Eds.). (2015). *Parasite Diversity and Diversification: Evolutionary Ecology Meets Phylogenetics*. Cambridge, United Kingdom: Cambridge University Press.
29. Mount, D. (2004). *Bioinformatics: Sequence and Genome Analysis* (2nd ed.). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
30. NCBI. (n.d.). Retrieved 28/01/2019, from NCBI: <https://www.ncbi.nlm.nih.gov/>
31. NIH. (n.d.). Retrieved 28/01/2019, from <https://www-bimas.cit.nih.gov/molbio/readseq/>

32. Opperdoes, F. (08/08/1997). *Phenetics versus Cladistics and the pro's and con's of the various phylogeny inference methods*. Retrieved 14/01/2019, from de Duve Institute: <http://www.deduveinstitute.be/~opperd/private/phenetics>
33. Pablo A. Goloboff, & Diego Pol. (07/2007). On Divide-and-Conquer Strategies for Parsimony Analysis of Large Data Sets: Rec-I-DCM3 versus TNT. *Systematic Biology*, 56(3), pp. 485-495.
34. Paulo AS Nuin, Zhouzhi Wang, & Elisabeth RM Tillier. (24/10/2006). The accuracy of several multiple sequence alignment programs for proteins. *BMC Bioinformatics*.
35. Pennisi, E. (21/09/2015). *First comprehensive tree of life shows how related you are to millions of species*. Retrieved 14/01/2019, from Sciencemag: <https://www.sciencemag.org/news/2015/09/first-comprehensive-tree-life-shows-how-related-you-are-millions-species>
36. Philippe Lemey, Marco Salemi, & Anne-Mieke Vandamme (Eds.). (2009). *The Phylogenetic Handbook* (2nd ed.). NY USA: Cambridge University Press.
37. Reeck GR, de Haën C, Teller DC, Doolittle RF, Fitch WM, Dickerson RE, et al. (1987). "Homology" in Proteins and Nucleic Acids: A Terminology. *Cell*, 50(5), 667.
38. ResearchGate. (2013). *Online Discussion in ResearchGate*. Retrieved 14/01/2019, from https://www.researchgate.net/post/How_to_select_genes_for_phylogenetic_analysis_based_on_taxonomic_level
39. S. Gorezis, M. Psilla, I. Asproudis, D. Peschos, C. Papadopoulou, & M.Stefaniotou. (2006). Intravitreal Dirofilariasis: A Rare Ocular Infection. *Orbit*, 25(1), 57-59.

40. Santamaría, R., & Therón, R. (26/05/2009). Treevolution: visual analysis of phylogenetic trees. *Bioinformatics*, 25(15), 1970–1971.
41. Schuh, R. (2000). *Biological Systematics*. Cornell U. Press, 6.
42. Sneath, P., & Sokal, R. (1973). *Numerical taxonomy — The principles and practice of numerical classification*. San Francisco: W. H. Freeman.
43. Soares, A., Râbelo, R., & Delbem, A. (2017). Optimization based on phylogram analysis. In *Expert Systems with Applications* (Vol. 78, pp. 32-50). Elsevier.
44. Sokal R., & Michener C. (1958). A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin*, 38, 1409–1438.
45. Stephen A. Krawetz, & David D. Womble (Eds.). (2003). *Introduction to Bioinformatics: A Theoretical and Practical Approach*. USA: Humana Press.
46. Tandy Warnow. (2017). *Computational Phylogenetics: An introduction to designing methods for phylogeny estimation*.
47. TEXAS TECH UNIVERSITY. (n.d.). Retrieved 28/01/2019, from TEXAS TECH UNIVERSITY:
http://www.faculty.biol.ttu.edu/Strauss/Phylogenetics/LectureNotes/07_UltrametricAdditiveTrees.pdf
48. U.S. National Library of Medicine. (n.d.). *Genetics Home Reference*. Retrieved 28/01/2019, from U.S. National Library of Medicine:
<https://ghr.nlm.nih.gov/primer#mutationsanddisorders>

49. Understanding Evolution. (n.d.). *Understanding Evolution*. Retrieved 24/11/2018, from Trees are hypotheses:
https://evolution.berkeley.edu/evolibrary/article/0_0_0/evotrees_primer_06
50. UniProt. (n.d.). *UniProt Taxonomy*. Retrieved 08/02/2019, from UniProt :
<https://www.uniprot.org/taxonomy/1240989>
51. University of California Museum of Paleontology. (n.d.). *Carl Linnaeus (1707-1778)*. Retrieved 24/11/2018, from University of California Museum of Paleontology:
<http://www.ucmp.berkeley.edu/history/linnaeus.html>
52. Valery O. Polyanovsky, Mikhail A. Roytberg, & Vladimir G. Tumanyan. (2011). Comparative analysis of the quality of a global algorithm and a local algorithm for alignment of two sequences. *Algorithms for Molecular Biology*, 6(1), 25.
53. Xin Wang, I. King Jordan, & Leonard W. Mayer. (2015). A Phylogenetic Perspective on Molecular Epidemiology. In Yi-Wei Tang, *Molecular Medical Microbiology* (2nd ed., pp. 517-536). Academic Press.
54. Yemini, Y. (n.d.). *Chapter 3: Phylogenetics*. Retrieved 28/01/2019, from Computer Science, Columbia University: <http://www.cs.columbia.edu/4761/notes07/chapter3.2-phylo.pdf>
55. Ziemert, N., & Jensen, P. (2012). Natural Product Biosynthesis by Microorganisms and Plants, Part C. In D. Hopwood (Ed.), *Methods in Enzymology* (1st ed., Vol. 517, pp. 161-162). San Diego, California, USA: Academic Press.
56. Παπαδοπούλου, Χ. (1990). *Σημειώσεις Παρασιτολογίας*. Ιωάννινα: Πανεπιστημιακό Τυπογραφείο Ιωαννίνων.

57. Πουλακάκης Νίκος. (n.d.). *Εξελικτική Οικολογία - Διάλεξη 9: Φυλογενετικά Δέντρα*.

Retrieved 14/01/2019, from Πανεπιστήμιο Κρήτης - Τμήμα Βιολογίας:

<http://www.biology.uoc.gr/courses/BIOL471/For%20website/09.%20%CE%9A%CE%B1%CF%84%CE%B1%CF%83%CE%BA%CE%B5%CF%85%CE%AE%20%CF%86%CF%85%CE%BB%CE%BF%CE%B3%CE%B5%CE%BD%CE%B5%CF%84%CE%B9%CE%BA%CF%8E%CE%BD%20%CE%B4%CE%AD%CE%BD%CF%84%CF%81%CF%89%CE%BD%20-%20%CE%91>