



ΓΕΩΠΟΝΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
AGRICULTURAL UNIVERSITY OF ATHENS

ΤΜΗΜΑ ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ

ΕΡΓΑΣΤΗΡΙΟ ΓΕΝΕΤΙΚΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΒΙΟΛΟΓΙΑΣ ΣΥΣΤΗΜΑΤΩΝ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ

*Συγκριτική γονιδιωματική και φυλογενετική
ανάλυση πρωτεϊνών που εμπλέκονται στη δομή
και τη λειτουργία του ενδοπλασματικού δικτύου*

ΑΣΠΑΣΙΑ Ε. ΚΟΝΤΟΥ

Βιολόγος

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ:

ΚΟΥΜΑΝΤΟΥ ΒΑΣΙΛΙΚΗ, Λέκτορας (Επιβλέπων)

ΘΗΡΑΙΟΥ ΤΡΙΑΣ, Επίκουρη Καθηγήτρια

ΧΑΤΖΟΠΟΥΛΟΣ ΠΟΛΥΔΕΥΚΗΣ, Καθηγητής

ΑΘΗΝΑ 2019

ΓΕΩΠΟΝΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΤΜΗΜΑ ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ
ΕΡΓΑΣΤΗΡΙΟ ΓΕΝΕΤΙΚΗΣ
ΜΕΤΑΠΤΥΧΙΑΚΟ ΒΙΟΛΟΓΙΑΣ ΣΥΣΤΗΜΑΤΩΝ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ

**ΣΥΓΚΡΙΤΙΚΗ ΓΟΝΙΔΙΩΜΑΤΙΚΗ ΚΑΙ ΦΥΛΟΓΕΝΕΤΙΚΗ
ΑΝΑΛΥΣΗ ΠΡΩΤΕΪΝΩΝ ΠΟΥ ΕΜΠΛΕΚΟΝΤΑΙ ΣΤΗ
ΔΟΜΗ ΚΑΙ ΤΗ ΛΕΙΤΟΥΡΓΙΑ ΤΟΥ ΕΝΔΟΠΛΑΣΜΑΤΙΚΟΥ
ΔΙΚΤΥΟΥ**

**COMPARATIVE GENOMICS AND PHYLOGENETIC
ANALYSIS OF PROTEINS INVOLVED IN THE
STRUCTURE AND FUNCTIONS OF THE
ENDOPLASMIC RETICULUM**

ΕΠΙΒΛΕΠΟΥΣΑ: ΚΟΥΜΑΝΤΟΥ ΒΑΣΙΛΙΚΗ

ΑΣΠΑΣΙΑ Ε. ΚΟΝΤΟΥ

ΑΘΗΝΑ, 2019

| ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ | | |
|-----------------------------|--|----|
| | ΕΥΧΑΡΙΣΤΙΕΣ | 4 |
| | ΠΕΡΙΛΗΨΗ | 5 |
| | ABSTRACT | 6 |
| 1 | ΕΙΣΑΓΩΓΗ | 7 |
| 1.1 | Συγκριτική γονιδιωματική ανάλυση | 7 |
| 1.2 | Ενδομεμβρανικό σύστημα των ευκαρυωτικών οργανισμών | 10 |
| 1.2.1 | Ενδοπλασματικό δίκτυο | 11 |
| 1.3 | Πρωτεΐνες που εμπλέκονται στη διαμόρφωση του Ενδοπλασματικού Δικτύου | 12 |
| 1.3.1 | Η πρωτεϊνική οικογένεια των reticulons | 16 |
| 1.3.1.1 | Βασικά στοιχεία | 16 |
| 1.3.1.2 | Εξελικτική ιστορία των reticulons | 19 |
| 1.3.2 | Οικογένεια των πρωτεϊνών REEP | 20 |
| 1.3.2.1 | REEP1/SPG31 | 22 |
| 1.3.2.2 | REEP5/Yop1p | 22 |
| 1.3.3 | Atlastin1 (ATL1)/ Sey1p | 22 |
| 1.3.4 | Ufe1p | 23 |
| 1.3.5 | Use1p | 24 |
| 1.3.6 | Dsl1p | 26 |
| 1.3.7 | Arl6IP1/ARMER | 29 |
| 1.3.8 | Protrudin | 29 |
| 1.3.9 | Lunapark/Lnp1p | 30 |
| 1.3.10 | Spastin | 31 |
| 1.3.11 | CLIMP-63 | 32 |
| 1.3.12 | Kinectin | 34 |
| 1.3.13 | Οι πρωτεΐνες TMEM33 | 35 |
| 1.3.14 | STIM1 | 36 |
| 1.4 | Άλλες πρωτεΐνες που αλληλεπιδρούν με τα reticulons, ή εντοπίζονται στο Ενδοπλασματικό δίκτυο | 38 |
| 1.4.1 | TBC1D20 | 38 |
| 1.4.2 | Yip3p | 38 |
| 1.4.3 | Υποδοχέας του RTN4/Nogo A (NogoR) | 39 |
| 1.4.4 | Caspr1 | 40 |
| 1.4.5 | Πρωτεΐνες VAP: VAP-A και VAP-B | 41 |
| 1.5 | Σκοπός | 42 |
| 2 | ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ | 44 |
| 2.1 | Γενική πορεία εργασίας | 44 |
| 2.2 | Συγκριτική γονιδιωματική ανάλυση | 48 |
| 2.2.1 | Αλγόριθμος BLAST | 48 |
| 2.2.2 | Το πρόγραμμα HMMER | 51 |
| 2.3 | Φυλογενετική ανάλυση | 54 |
| 2.3.1 | Πολλαπλή στοίχιση αλληλουχιών | 54 |

| | | |
|------------------|---|-----|
| 2.3.1.1 | Πρόγραμμα στοιχίσεων MUSCLE | 55 |
| 2.3.2 | Προετοιμασία πολλαπλής στοίχισης για την κατασκευή φυλογενετικών δέντρων | 57 |
| 2.3.3 | Μέθοδοι για τη δημιουργία φυλογενετικών δέντρων | 58 |
| 2.3.3.1 | Μέθοδοι αποστάσεων | 58 |
| 2.3.3.2 | Μέθοδοι ελαχιστοποίησης παραμέτρων | 59 |
| 2.3.3.2.1 | Πρόγραμμα RaxML | 60 |
| 2.3.4 | Εκτίμηση φυλογενετικών δέντρων | 61 |
| 3 | ΑΠΟΤΕΛΕΣΜΑΤΑ-ΣΥΖΗΤΗΣΗ | 63 |
| 3.1 | Αποτελέσματα συγκριτικής γονιδιωματικής ανάλυσης | 63 |
| 3.2 | Σύγκριση αποτελεσμάτων με προηγούμενες αναλύσεις συγκριτικής γονιδιωματικής | 68 |
| 3.3 | Αποτελέσματα φυλογενετικής ανάλυσης | 70 |
| 3.3.1 | Reticulons | 70 |
| 3.3.2 | REEPs | 73 |
| 3.3.3 | Atlastin 1/ Sey1p | 75 |
| 3.3.4 | Spastin | 77 |
| 3.3.5 | CLIMP-63 | 78 |
| 3.3.6 | Kinectin | 79 |
| 3.3.7 | TMEM33 | 80 |
| 3.3.8 | Caspr1 | 81 |
| 3.3.9 | NogoR | 83 |
| 3.3.10 | VAPs | 85 |
| 4 | ΣΥΜΠΕΡΑΣΜΑΤΑ | 87 |
| | ΒΙΒΛΙΟΓΡΑΦΙΑ | 95 |
| | ΠΑΡΑΡΤΗΜΑ | 100 |

ΕΥΧΑΡΙΣΤΙΕΣ

Η παρούσα διπλωματική εργασία εκπονήθηκε στα πλαίσια του μεταπτυχιακού προγράμματος με τίτλο «Βιολογία Συστημάτων» του Γεωπονικού Πανεπιστημίου Αθηνών.

Αρχικά θα ήθελα να ευχαριστήσω θερμά τη Λέκτορα κυρία Βασιλική Κουμάντου η οποία ήταν υπεύθυνη για την επίβλεψη της εργασίας μου, για την άριστη συνεργασία μας καθώς και την πολύτιμη καθοδήγηση της και την εμπιστοσύνη που μου έδειξε.

Εν συνεχεία θα ήθελα να εκφράσω τις ευχαριστίες μου στον κ. Mark Field, Καθηγητή κυτταρικής βιολογίας και Παρασιτολογίας στο πανεπιστήμιο του Dundee της Αγγλίας, στον κ. Joel B. Dacks Αναπληρωτή Καθηγητή στο τμήμα Ιατρικής και Οδοντιατρικής του Πανεπιστημίου της Alberta του Καναδά, καθώς και στην κ. Emily Herman μεταδιδακτορική ερευνήτρια του Πανεπιστημίου της Alberta του Καναδά, χωρίς την βοήθεια των οποίων δε θα ήταν δυνατή η εκπόνηση της συγκεκριμένης διπλωματικής εργασίας. Πιο συγκεκριμένα, οι αναλύσεις με τη χρήση των προγραμμάτων BLAST και HMMer έγιναν για εξοικονόμηση χρόνου με αυτοματοποιημένες μεθόδους σε συνεργασία με το εργαστήριο του κ. Joel Dacks στο Πανεπιστήμιο της Alberta του Καναδά.

Τέλος τις ευχαριστίες μου εκφράζω και στους καθηγητές Πολυδεύκη Χατζόπουλο και Τριάς Θηραίου που δέχτηκαν να είναι μέλη της τριμελούς επιτροπής αξιολόγησης της μεταπτυχιακής μου εργασίας.

Περίληψη

Το ενδομεμβρανικό σύστημα των ευκαρυωτικών οργανισμών αποτελεί ένα πολύπλοκο σύστημα μεμβρανών το οποίο καταλαμβάνει σημαντικό τμήμα του ευκαρυωτικού κυττάρου. Πολλές και εξαιρετικής σημασίας λειτουργίες όπως η σύνθεση, τροποποίηση και στόχευση μεμβρανικών πρωτεϊνών καθώς και η κυστιδιακή μεταφορά επιτελούνται από τμήματα του συστήματος αυτού. Πολλαπλοί μηχανισμοί ασθενειών προκύπτουν από ελαττώματα στις λειτουργίες αυτές. Συνεπώς, η μελέτη της εξέλιξης του ενδομεμβρανικού συστήματος, αν λάβουμε υπόψιν και την απουσία κατά κανόνα ενδομεμβρανικών σχηματισμών από τα προκαρυωτικά συστήματα, παρουσιάζει τεράστιο ενδιαφέρον. Παράλληλα όμως, αν και έχουν γίνει αρκετές μελέτες σε αυτήν την κατεύθυνση, υπάρχουν ακόμα πολλά κενά. Μέσα στα πλαίσια της παρούσας διπλωματικής εργασίας πραγματοποιήθηκε μελέτη της εξελικτικής πορείας τμημάτων του ενδομεμβρανικού συστήματος, όπως το ενδοπλασματικό δίκτυο και η ενδομεμβρανική κυκλοφορία. Για τον σκοπό αυτό εφαρμόστηκε ένα πρωτόκολλο βιοπληροφορικής το οποίο περιελάμβανε αναλύσεις συγκριτικής γονιδιωματικής και φυλογένεσης 23 πρωτεϊνών του ενδομεμβρανικού συστήματος. Από τα αποτελέσματα των αναλύσεων προέκυψε μεγάλος αριθμός πρωτεϊνών οι οποίες εμφανίζουν υψηλό βαθμό συντήρησης στο ευρύ φάσμα των ευκαρυωτικών οργανισμών. Τέτοιες πρωτεΐνες είναι τα reticulons, οι REEPs και οι Atlastins που συμμετέχουν στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου και η Ufe1p, η Use1p, η Dsl1p, η TBC1D20, η Yip3p και οι VAPs που εμπλέκονται στην ενδομεμβρανική κυκλοφορία. Συνολικά από όλα τα αποτελέσματα της παρούσας μελέτης συνάγεται η αρχαία προέλευση του ενδομεμβρανικού συστήματος καθώς και ο υψηλός βαθμός πολυπλοκότητας βασικών μηχανισμών στον τελευταίο κοινό πρόγονο των ευκαρυωτικών οργανισμών (LECA).

Επιστημονική περιοχή μελέτης: Εξελικτική Βιολογία - Βιοπληροφορική

Λέξεις κλειδιά: ενδομεμβρανικό σύστημα, ενδοπλασματικό δίκτυο, κυστιδιακή μεταφορά, τελευταίος κοινός πρόγονος των ευκαρυωτικών οργανισμών (LECA), reticulons, εξέλιξη, συγκριτική γονιδιωματική, φυλογενετική ανάλυση.

ABSTRACT

The endomembrane system is a complex membrane system which occupies a significant portion of the eukaryotic cell. Parts of this system perform many functions of great importance like the synthesis, post-translational modifications and targeting of membrane proteins, as well as vesicular transport. Multiple disease mechanisms arise from defects in these functions. Therefore, the study of the endomembrane system is a field of great interest, especially when considering the fact that endomembrane formations are usually absent from most prokaryotic systems. However, there are still many gaps in this field even though a lot of research has been conducted. The goal of this master's thesis was to study the evolution of parts of the endomembrane system such as the endoplasmic reticulum and membrane transport. This was achieved with the development of a bioinformatics protocol, including comparative genomics and phylogenetic analysis of 23 proteins of the endomembrane system. The results revealed a large number of proteins, such as reticulons, REEPs and atlastins (involved in the endoplasmic reticulum formation), Ufe1p, Use1p, Dsl1p, TBC1D20, Yip3p and VAPs (involved in endomembrane trafficking), which are highly conserved in eukaryotes. Moreover, the present study reveals the ancient origin of the endomembrane system and the high degree of complexity of basic mechanisms in the last common ancestor of eukaryotic organisms (LECA).

Title: Comparative genomics and phylogenetic analysis of proteins involved in the structure and functions of the endoplasmic reticulum.

Field of study: Evolutionary Biology - Bioinformatics

Key words: endomembrane system, endoplasmic reticulum, vesicular trafficking, Last Eukaryotic Common Ancestor (LECA), reticulons, evolution, comparative genomics, phylogenetic analysis.

1 Εισαγωγή

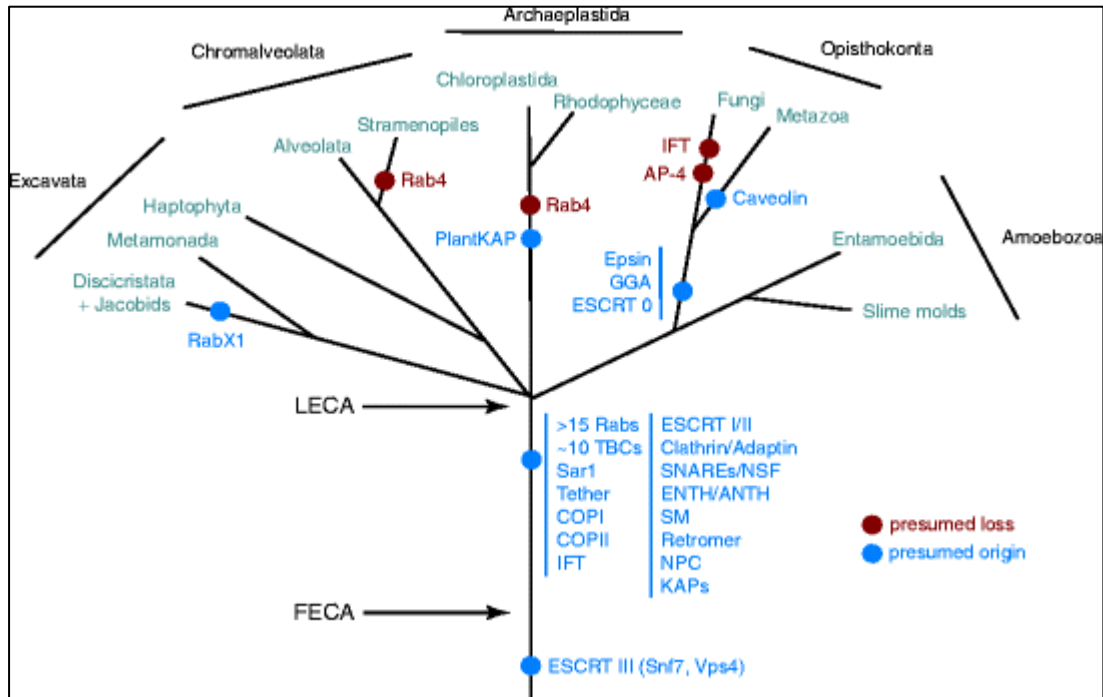
1.1 Συγκριτική γονιδιωματική ανάλυση

Η συγκριτική γονιδιωματική είναι ένας ερευνητικός τομέας της βιολογίας ο οποίος περιλαμβάνει τη σύγκριση γονιδιωματικών χαρακτηριστικών (αλληλουχίες DNA, γονίδια, σειρά γονιδίων, ρυθμιστικές αλληλουχίες και άλλα γονιδιακά στοιχεία) ανάμεσα σε διαφορετικά είδη, με τη χρήση ποικίλων εργαλείων βιοπληροφορικής.

Η συγκριτική γονιδιωματική ανάλυση μέσω του προσδιορισμού των ομοιοτήτων αλλά και διαφορών που μπορεί να υπάρχουν μεταξύ των οργανισμών μπορεί να εξυπηρετήσει μια σειρά από σκοπούς. Αρχικά ο προσδιορισμός του βαθμού συντήρησης γονιδίων ανάμεσα σε διαφορετικούς οργανισμούς επί εκατομμύρια χρόνια μπορεί να οδηγήσει στον εντοπισμό γονιδίων τα οποία είναι απαραίτητα για τη ζωή. Επιπλέον, μέσω της σύνδεσης συγκεκριμένων γονιδίων με διαφορετικά βιολογικά συστήματα παρέχεται η δυνατότητα για την ανάπτυξη καινοτόμων προσεγγίσεων για τη θεραπεία των ανθρώπινων ασθενειών και τη βελτίωση της ανθρώπινης υγείας. Παράλληλα η συγκριτική γονιδιωματική αποτελεί ένα ισχυρότατο εργαλείο για τη μελέτη της εξέλιξης. Πιο συγκεκριμένα, η αξιοποίηση της γνώσης που αποκτάται γύρω από τις εξελικτικές σχέσεις μεταξύ των ειδών μπορεί να οδηγήσει στην κατανόηση των μηχανισμών που οδήγησαν στην εμφάνιση αυτών καθώς των αρχών που καθορίζουν τη συμπεριφορά και τη βιολογία τους. Τέλος στη σημερινή εποχή η μείωση του κόστους και η αύξηση της ισχύος των μεθόδων αλληλούχησης του DNA έχει οδηγήσει λόγω της εξάλειψης των συχνά λεπτών διαφορών μεταξύ των ζωικών ειδών, στην ανάπτυξη εφαρμογών της συγκριτικής γονιδιωματικής στη γεωργία, τη βιοτεχνολογία και τη ζωολογία. Με αυτόν τον τρόπο είναι δυνατή η βελτίωση της υγείας των κατοικίδιων ζώων και η ανάπτυξη νέων στρατηγικών για τη διατήρηση σπάνιων και απειλούμενων ειδών.

Τα εργαλεία τα οποία έχουν αναπτυχθεί σήμερα επιτρέπουν την αυτοματοποίηση των αναλύσεων σε μεγάλο βαθμό, μειώνοντας σημαντικά το χρόνο που καταναλώνεται. Επιπλέον, η ραγδαία αύξηση του ρυθμού αλληλούχησης γονιδιωμάτων έχει ως αποτελέσματα την ύπαρξη βάσεων δεδομένων με πλήρως αλληλουχημένα γονιδιώματα από όλες τις γενεαλογίες των οργανισμών. Τα παραπάνω δεδομένα έχουν επιτρέψει την εφαρμογή μεθόδων συγκριτικής γονιδιωματικής, με σκοπό την ταυτόχρονη μελέτη πολλών διαφορετικών γονιδίων και κατ' επέκταση πρωτεϊνών. Είναι πλέον δυνατή δηλαδή η μελέτη ολόκληρων συστημάτων, όπως το ενδομεμβρανικό σύστημα των ευκαρυωτικών οργανισμών. Χαρακτηριστική λειτουργία που επιτελείται από το σύστημα αυτό είναι και η ενδομεμβρανική κυκλοφορία για την οποία έχουν πραγματοποιηθεί αρκετές μελέτες μέχρι σήμερα. Όπως φαίνεται και στην εικόνα 1, στην οποία συνοψίζονται ορισμένα από τα μέχρι σήμερα αποτελέσματα, μια σειρά από πρωτεΐνες που συμμετέχουν σε αυτή τη λειτουργία εμφανίζουν υψηλό βαθμό συντήρησης. Πιο συγκεκριμένα, φαίνεται ότι στοιχεία της ενδομεμβρανικής κυκλοφορίας υπήρχαν ήδη στον πρώτο κοινό πρόγονο των ευκαρυωτικών οργανισμών (FECA), προτού η λειτουργία αυτή εδραιωθεί στον τελευταίο κοινό πρόγονο των ευκαρυωτικών οργανισμών (LECA). Δηλαδή, η ενδομεμβρανική κυκλοφορία αποτελεί μια «αρχαία»

λειτουργία η οποία εμφάνιζε υψηλό βαθμό πολυπλοκότητας ήδη από τον LECA.



Εικόνα 1. Σχηματική απεικόνιση των φυλογενετικών σχέσεων ευκαρυωτικών οργανισμών στην οποία σημειώνεται η προέλευση πολλών συστατικών της ενδομεμβρανικής κυκλοφορίας. LECA: last eukaryotic common ancestor, FECA: first eukaryotic common ancestor (Koumandou and Field, 2011).

Παρόλη την πρόοδο όμως που έχει πραγματοποιηθεί στην προσπάθεια ανασύστασης της εξέλιξης της ενδομεμβρανικής κυκλοφορίας παραμένουν ακόμα αναπάντητα ερωτήματα για τον ακριβή μηχανισμό αυτής στον LECA, καθότι δεν έχουν ελεγχθεί ακόμα όλες οι σύγχρονες πρωτεΐνες που συμμετέχουν στο μηχανισμό αυτό. Παράλληλα, μεγάλα κενά υπάρχουν και στη μελέτη άλλων τμημάτων του ενδομεμβρανικού συστήματος, όπως το ενδοπλασματικό δίκτυο. Στον πίνακα 1 υπάρχει μια λίστα με πρωτεΐνες για τις οποίες μέχρι στιγμής είτε δεν έχει πραγματοποιηθεί μελέτη της εξελικτικής ιστορίας είτε η μελέτη που έχει πραγματοποιηθεί είναι πολύ μικρότερης έκτασης.

Πίνακας 1. Λειτουργίες στις οποίες συμμετέχουν οι 23 πρωτεΐνες οι οποίες χρησιμοποιήθηκαν στη συγκεκριμένη μελέτη. Οι πρωτεΐνες είναι κατηγοριοποιημένες ανάλογα με το λειτουργικό τους ρόλο στο κύτταρο.

| <i>Λειτουργίες</i> | <i>Πρωτεΐνες</i> | |
|--|-------------------------|----------------|
| <i>Δομή ενδοπλασματικού δικτύου (ΕΔ)</i> | <i>Αγωγοί</i> | Reticulons |
| | | REEP5/Yop1p |
| | | REEP1 |
| | | Lunapark |
| | | Protrudin |
| | | Arl6IP1 |
| | | Spastin |
| | <i>Κόμβοι αγωγών ΕΔ</i> | Atlastin/Sey1p |
| | | Ufe1p |
| | | Use1p |
| | | Dsl1p |
| | <i>Ασκοί</i> | Reticulons |
| | | CLIMP-63 |
| Kinectin | | |
| TMEM33 | | |
| <i>Αλληλεπίδραση ΕΔ-μικροσωληνίσκων</i> | STIM1 | |
| | Spastin | |
| | CLIMP-63 | |
| | REEP1 | |
| <i>Ενδομεμβρανική κυκλοφορία</i> | Reticulons | |
| | Spastin | |
| | Ufe1p/ Use1p/ Dsl1p | |
| | TBC1D20 | |
| | Yip3p | |
| | VAPs | |
| <i>Άλλες λειτουργίες</i> | NogoR | |
| | Caspr | |

Από τον παραπάνω πίνακα είναι εμφανές ότι η πλειοψηφία των πρωτεϊνών συμμετέχει στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου. Το ενδοπλασματικό δίκτυο αποτελεί έναν βασικό σχηματισμό του ενδομεμβρανικού συστήματος, ο οποίος επιτελεί καθοριστικές λειτουργίες όπως η σύνθεση, τροποποίηση και στόχευση μεμβρανικών πρωτεϊνών. Σημαντικός είναι και ο ρόλος της αλληλεπίδρασης του κυτταροσκελετού με τους αγωγούς του ενδοπλασματικού δικτύου, στη διαμόρφωση της δομής του τελευταίου. Για το λόγο αυτό επιλέχθηκαν και πρωτεΐνες οι οποίες συμμετέχουν στη διαδικασία αυτή. Τέλος θεμελιώδης είναι και ο ρόλος της ενδομεμβρανικής κυκλοφορίας στο ευκαρυωτικό κύτταρο το οποίο διαθέτει πολύ μεγαλύτερο μέγεθος συγκριτικά με το προκαρυωτικό κύτταρο. Επτά είναι οι πρωτεΐνες οι οποίες, όπως φαίνεται στον πίνακα 1, συμμετέχουν σε αυτή την κυτταρική λειτουργία. Αξίζει να σημειωθεί ότι οι πρωτεΐνες reticulons εντάσσονται σε δύο διαφορετικές λειτουργικές ομάδες, συμμετέχοντας τόσο στη διαμόρφωση της δομής των αγωγών του ενδοπλασματικού δικτύου όσο και στην κυστιδιακή μεταφορά.

Παράλληλα, από βιβλιογραφικά δεδομένα, βρέθηκε ότι πολλές από τις 23 αυτές πρωτεΐνες αλληλεπιδρούν μεταξύ τους με αποτέλεσμα να σχηματίζεται ένα δίκτυο αλληλεπιδράσεων. Όπως φαίνεται και από τον πίνακα 2 οι αλληλεπιδράσεις αυτές εντοπίζονται κυρίως ανάμεσα στις πρωτεΐνες που συμμετέχουν στη διαμόρφωση της δομής του σωληνοειδούς ενδοπλασματικού δικτύου. Επιπλέον, παρατηρούμε ότι τα reticulons αλληλεπιδρούν με πρωτεΐνες που συμμετέχουν στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου, αλλά και με πρωτεΐνες που σχετίζονται με την κυστιδιακή μεταφορά καθώς και με τις πρωτεΐνες NogoR και Caspr1 του νευρικού συστήματος. Το γεγονός αυτό αντικατοπτρίζει την πληθώρα των λειτουργιών στις οποίες συμμετέχουν τα reticulons.

Πίνακας 2. Πίνακας των αλληλεπιδράσεων ανάμεσα στις πρωτεΐνες οι οποίες χρησιμοποιήθηκαν στη συγκεκριμένη μελέτη.

| | | Αγωγοί ΕΔ | | | | | | | Κόμβοι ΕΔ | | | | Ασκοί ΕΔ | | | ΕΔ-μικροσωληνίσκοι | Κυστιδιακή μεταφορά | | | Άλλη λειτουργία | |
|---------------------|----------------|------------|-------------|-------|----------|-----------|---------|---------|----------------|-------|-------|-------|----------|----------|--------|--------------------|---------------------|-------|------|-----------------|-------|
| | | reticulons | REEP5/Yop1p | REEP1 | lunapark | protrudin | Arl6IP1 | Spastin | Atlastin/Sey1p | Ufe1p | Use1p | Dsl1p | CLIMP-63 | kinectin | TMEM33 | STIM1 | TBC1D20 | Yip3p | VAPs | NogoR | Caspr |
| Αγωγοί ΕΔ | reticulons | + | | + | + | | + | + | | | | | | | + | | + | + | | + | + |
| | REEP5/Yop1p | | | + | + | | | + | | | | | | | | | | | | | |
| | REEP1 | | | | + | | + | + | | | | | | | | | | | | | |
| | lunapark | | | | | | | + | | | | | | | | | | | | | |
| | protrudin | | | | | | | + | + | | | | | | | | | | | | |
| | Arl6IP1 | | | | | | | | + | | | | | | + | | | | | | |
| | Spastin | | | | | | | | + | | | | | | | | | | | | |
| Κόμβοι ΕΔ | Atlastin/Sey1p | | | | | | | | | | | | | | | | | | | | |
| | Ufe1p | | | | | | | | | | | | | | | | | | | | |
| | Use1p | | | | | | | | | | | | | | | | | | | | |
| | Dsl1p | | | | | | | | | | | | | | | | | | | | |
| Ασκοί ΕΔ | CLIMP-63 | | | | | | | | | | | | | | | | | | | | |
| | kinectin | | | | | | | | | | | | | | | | | | | | |
| | TMEM33 | | | | | | | | | | | | | | | | | | | | |
| ΕΔ-μικροσωληνίσκοι | STIM1 | | | | | | | | | | | | | | | | | | | | |
| Κυστιδιακή μεταφορά | TBC1D20 | | | | | | | | | | | | | | | | | | | | |
| | Yip3p | | | | | | | | | | | | | | | | | | | | |
| | VAPs | | | | | | | | | | | | | | | | | | | | |
| Άλλη λειτουργία | NogoR | | | | | | | | | | | | | | | | | | | | |
| | Caspr | | | | | | | | | | | | | | | | | | | | |

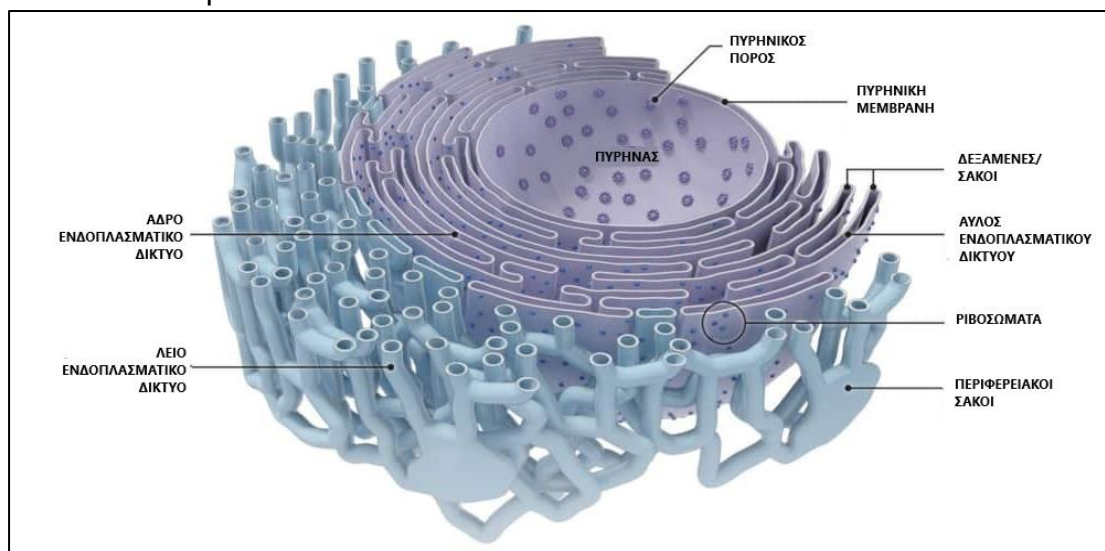
1.2 Ενδομεμβρανικό σύστημα των ευκαρυωτικών οργανισμών

Το ενδομεμβρανικό σύστημα είναι ένα διασυνδεδεμένο δίκτυο μεμβρανών το οποίο εντοπίζεται στους ευκαρυωτικούς οργανισμούς. Απαρτίζεται από την πυρηνική μεμβράνη, το ενδοπλασματικό δίκτυο, το σύστημα Golgi, την πλασματική μεμβράνη και τα κυστίδια μεταφοράς. Το σύστημα αυτό είναι υπεύθυνο για τη βιοσύνθεση και τη μεταφορά πρωτεϊνών και λιπιδίων και παίζει σημαντικό ρόλο στην υποκυτταρική διαμερισματοποίηση. Η σημασία του για το κύτταρο γίνεται εμφανής από τη σχεδόν καθολική παρουσία του στην πλειοψηφία των ευκαρυωτικών οργανισμών. Η προέλευση και η εξέλιξη του δεν είναι προφανείς διότι οι προκαρυωτικοί οργανισμοί κατά κανόνα δεν διαθέτουν οργανίδια ή μεμβρανικούς σχηματισμούς, οι οποίοι θα μπορούσαν να δώσουν γένεση στο

σύστημα αυτό. Επομένως, η δημιουργία του αποτελεί ένα κρίσιμο στάδιο στη μετάβαση από τους προκαρυωτικούς στους ευκαρυωτικούς οργανισμούς. Αυτό συμβαίνει διότι ένα σύστημα μεμβρανών φαίνεται να ήταν απαραίτητο για την αποτελεσματική μεταφορά μακρομορίων εντός του κυτταροπλάσματος των ευκαρυωτικών κυττάρων, τα οποία έχουν εκατονταπλάσιο όγκο από τα προκαρυωτικά. Επιπλέον, η ενδοκύτωση και η στοχευμένη μεταφορά πρωτεϊνών στην κυτταρική μεμβράνη είναι δύο ακόμη σημαντικές δυνατότητες τις οποίες προσέφερε το ενδομεμβρανικό σύστημα των ευκαρυωτικών κυττάρων.

1.2.1 Ενδοπλασματικό δίκτυο

Το ενδοπλασματικό δίκτυο (εικόνα 2) αποτελεί τμήμα του ενδομεμβρανικού συστήματος των ευκαρυωτικών οργανισμών. Πιο συγκεκριμένα, πρόκειται για ένα δίκτυο μεμβρανών οι οποίες εκτείνονται σε όλο το μήκος και το πλάτος του κυττάρου, από την πυρηνική μέχρι την πλασματική μεμβράνη. Αποτελείται από ένα δίκτυο πεπλατυσμένων σάκων και κυλινδρικών αγωγών. Το ενδοπλασματικό δίκτυο παίζει σημαντικό ρόλο στη βιοσύνθεση και μεταφορά μεμβρανικών ή εκκριτικών πρωτεϊνών και ορμονών, στη βιοσύνθεση λιπαρών οξέων και λιπιδίων, στο μεταβολισμό των υδατανθράκων καθώς και στην ομοίωση του ασβεστίου (Schönthal, 2012). Από τα μορφολογικά του χαρακτηριστικά διακρίνεται στο αδρό και το λείο ενδοπλασματικό δίκτυο. Το πρώτο παίρνει το όνομά του από την παρουσία ριβοσωμάτων στην επιφάνεια του, τα οποία του προσδίδουν μια «αδρή» όψη. Το τμήμα αυτό είναι υπεύθυνο για τη σύνθεση και έκκριση των πρωτεϊνών. Το λείο ενδοπλασματικό δίκτυο είναι υπεύθυνο για όλες τις υπόλοιπες λειτουργίες του ενδοπλασματικού δικτύου.



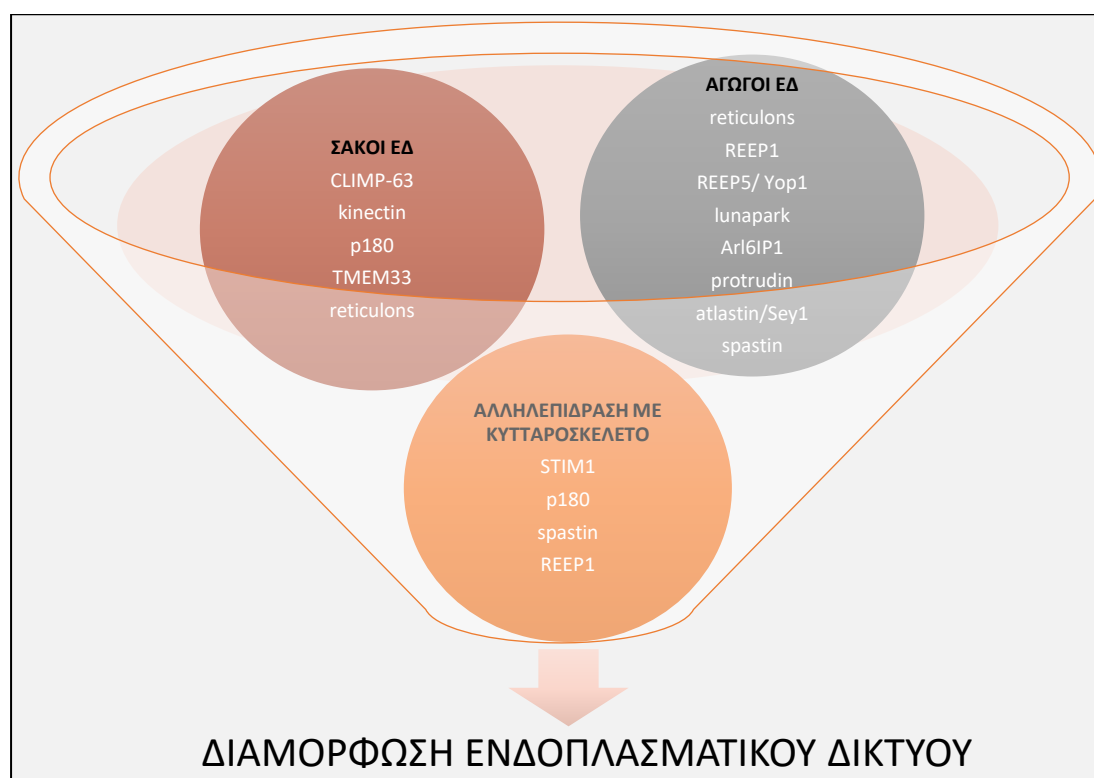
Εικόνα 2. Το ενδοπλασματικό δίκτυο των ευκαρυωτικών κυττάρων (<https://factslegend.org/40-endoplasmic-reticulum-facts-types-structure-functions/>).

Η σημασία του ενδοπλασματικού δικτύου είναι εμφανής και από το ρόλο του σε μια σειρά από διαταραχές. Πιο συγκεκριμένα, δυσλειτουργία του ενδοπλασματικού δικτύου μπορεί να οδηγήσει σε ασθένειες όπως ο διαβήτης, η φλεγμονή, η παχυσαρκία και νευροεκφυλιστικές ασθένειες όπως η νόσος του Parkinson, η νόσος του Alzheimer και η διπολική διαταραχή (Schönthal,

2012; Yoshida, 2007). Σημαντικό ρόλο σε αυτές τις διαταραχές παίζει η αδυναμία απομάκρυνσης από το ενδοπλασματικό δίκτυο πρωτεϊνών που δεν έχουν διπλωθεί σωστά ώστε να αποκτήσουν μια λειτουργική τρισδιάστατη δομή (Yoshida, 2007).

1.3 Πρωτεΐνες που εμπλέκονται στη διαμόρφωση του Ενδοπλασματικού Δικτύου

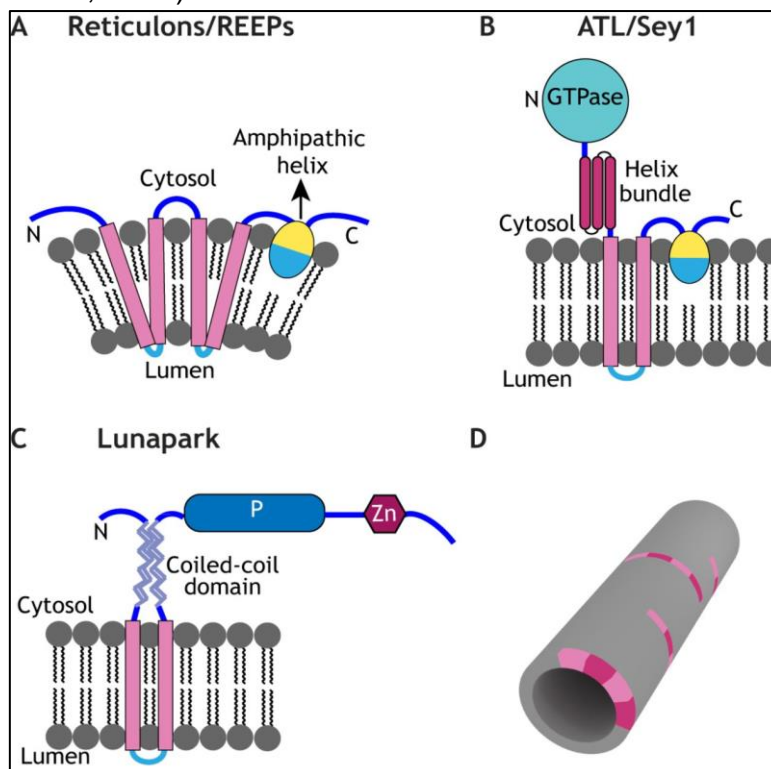
Η διαμόρφωση της δομής του ενδοπλασματικού δικτύου διαμεσολαβείται από μια σειρά από πρωτεΐνες. Πιο συγκεκριμένα, οι πρωτεΐνες reticulons, REEP1, REEP5/DP1/Yop1, atlastin/Sey1p, Ufe1p, Use1p, Dsl1p, Arl6lp1, protrudin, Lunapark και spastin συμμετέχουν στη διαμόρφωση της δομής του δικτύου των αγωγών, οι πρωτεΐνες reticulons, CLIMP-63, kinectin, p180 και TMEM33 συμμετέχουν στη διαμόρφωση των πεπλατυσμένων σάκων και οι πρωτεΐνες STIM1, spastin, CLIMP-63, p180 και REEP1 συμμετέχουν στις αλληλεπιδράσεις του ενδοπλασματικού δικτύου με τον κυτταροσκελετό (εικόνα 3).



Εικόνα 3. Πρωτεΐνες που συμμετέχουν στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου. Στον κόκκινο κύκλο φαίνονται οι πρωτεΐνες οι οποίες συμμετέχουν στη διαμόρφωση των σάκων-ασκών του ενδοπλασματικού δικτύου, στο γκρι κύκλο οι πρωτεΐνες που συμμετέχουν στη διαμόρφωση του δικτύου των αγωγών, ενώ στον πορτοκαλί κύκλο οι πρωτεΐνες οι οποίες διαμεσολαβούν στις αλληλεπιδράσεις του ενδοπλασματικού δικτύου με τον κυτταροσκελετό και συγκεκριμένα με τους μικροσωληνίσκους.

Οι πρώτες πρωτεΐνες οι οποίες ταυτοποιήθηκαν λόγω της συμμετοχής τους στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου είναι η REEP5 και τα reticulons. Πιο συγκεκριμένα, οι πρωτεΐνες αυτές δρουν συνεργιστικά, ώστε να προσδώσουν καμπυλότητα στους αγωγούς του σωληνοειδούς

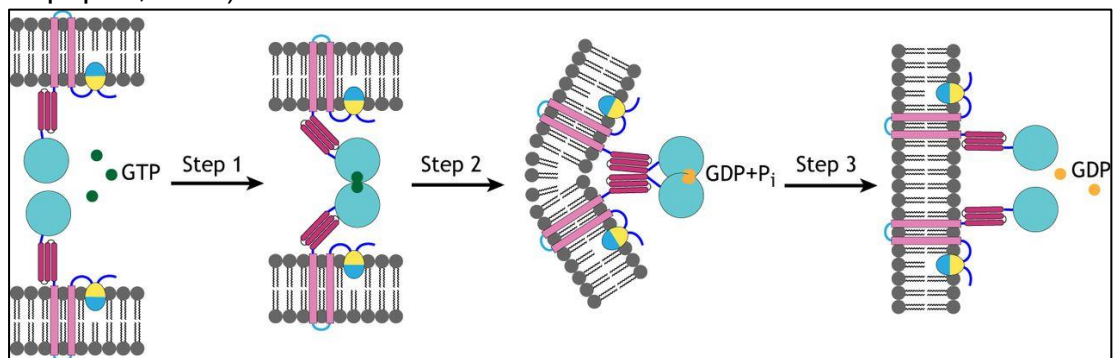
ενδοπλασματικού δικτύου καθώς και στις εξαιρετικά καμπύλες άκρες των ασκών (εικόνα 4A,D) (Voeltz et al., 2006). Σημαντικό ρόλο στην παραπάνω λειτουργία παίζει και η δυνατότητα ολιγομερισμού των reticulons καθώς και οι δομές διπλής φουρκέτας των reticulons και REEP5 οι οποίες εισέρχονται με τη μορφή σφήνας στη μεμβράνη την οποία και λυγίζουν (Friedman and Voeltz, 2011). Την ίδια λειτουργία με τα reticulons και την REEP5 επιτελεί και η πρωτεΐνη Arl6IP1 της οποίας η διαμεμβρανική επικράτεια έχει και αυτή δομή διπλής φουρκέτας παραπλήσια με των πρωτεϊνών αυτών. Σημειώνεται ότι ενώ τα reticulons και η REEP5 είναι συντηρημένες σε όλους τους ευκαρυώτες, η Arl6IP1 εμφανίζει αντιπρόσωπους μόνο στους πολυκύτταρους οργανισμούς (Yamamoto et al., 2014).



Εικόνα 4. Πρωτεΐνες που συμμετέχουν στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου. A) Τοπολογία των πρωτεϊνών reticulons και REEPs στη μεμβράνη. Η δομή διπλής φουρκέτας σε συνδυασμό με την αμφιπαθή έλικα στο καρβοξυτελικό άκρο (με κίτρινο η υδρόφιλη επιφάνεια και με μπλε η υδρόφοβη) θεωρείται ότι προσδίδουν θετική καμπυλότητα στις μεμβράνες του ενδοπλασματικού δικτύου. B) Δομή και τοπολογία της πρωτεΐνης atlastin. Η πρωτεΐνη Sey1p έχει παρόμοια δομή με την atlastin με τη διαφορά ότι η επικράτεια της σπειροειδούς έλικας (helix bundle) εμφανίζει μεγαλύτερο βαθμό πολυπλοκότητας. C) Δομή και τοπολογία της πρωτεΐνης lunapark. P: η επικράτεια η οποία φωσφορυλιώνεται στη διάρκεια της μίτωσης στους ανώτερους ευκαρυωτικούς οργανισμούς. Zn: η επικράτεια δακτύλων ψευδαργύρου η οποία εμπλέκεται στο διαμερισμό της πρωτεΐνης. D) Οι πρωτεΐνες reticulons και REEPs ολιγομερίζονται και σχηματίζουν τοξοειδείς δομές με αποτέλεσμα την κάμψη της μεμβράνης (Wang and Rapoport, 2019).

Οι πρωτεΐνες μέλη της οικογένειας των atlastin/Sey1p παίζουν καθοριστικό ρόλο στην ομοτυπική σύντηξη των μεμβρανών του ενδοπλασματικού δικτύου, μέσω μιας ενεργότητας GTPάσης την οποία διαθέτουν (Εικόνα 4B, 5). Το αποτέλεσμα είναι η δημιουργία ενός διακλαδισμένου δικτύου μεμβρανών. Οι atlastins αλληλεπιδρούν τόσο με τα reticulons, όσο και με τις REEPs (Farhan and Hauri, 2009; Urade et al.,

2014). Σε αντίθεση με τα κύτταρα των θηλαστικών και της *Drosophila melanogaster*, ο *Saccharomyces cerevisiae* διαθέτει έναν εναλλακτικό μηχανισμό μεμβρανικής σύντηξης ανεξάρτητο από την Sey1p. Στον μηχανισμό αυτό συμμετέχουν οι πρωτεΐνες Use1p, Ufe1p, Sec22p (πρωτεΐνες SNARE), καθώς και το σύμπλοκο Dsl1, που αποτελείται από τις πρωτεΐνες Tip20p, Dsl1p και Sec39p/ Dsl3. Παρόλα αυτά δεν είναι ακόμα γνωστό το κατά πόσο οι δύο εναλλακτικοί μηχανισμοί μεμβρανικής σύντηξης στο σακχαρομύκητα λειτουργούν παράλληλα (Anwar et al., 2012; Hu and Rapoport, 2016).



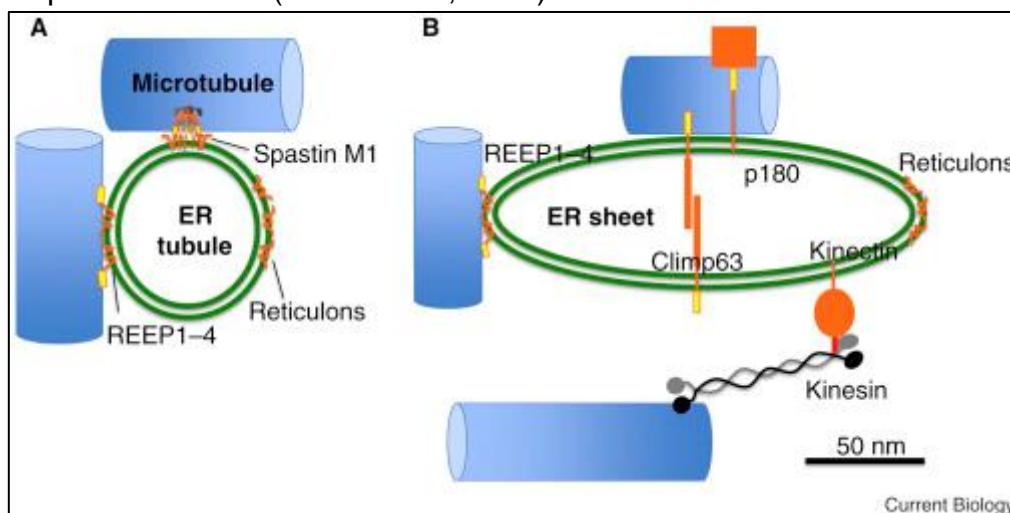
Εικόνα 5. Ομοτυπική σύντηξη μεμβρανών από την πρωτεΐνη atlastin. Οι δύο πρωτεΐνες atlastin που βρίσκονται σε διαφορετικές μεμβράνες συνδέονται αρχικά με GTP και στη συνέχεια μεταξύ τους, σχηματίζοντας διμερές (βήμα 1). Ακολουθεί υδρόλυση των μορίων GTP και αλλαγές στη δομή των πρωτεϊνών atlastin που επιτρέπουν την αλληλεπίδραση των δύο περιοχών σπειροειδούς έλικας. Οι δύο μεμβράνες πλησιάζουν μεταξύ τους και συντήκονται (βήμα 2). Τέλος, Τα μόρια GDP και Pi απελευθερώνονται και οι δύο πρωτεΐνες atlastin αποσυνδέονται (βήμα 3) (Wang and Rapoport, 2019).

Οι πρωτεΐνες protrudin και lunapark (εικόνα 15C) λειτουργούν ανταγωνιστικά με την atlastin ως προς τη μεμβρανική σύντηξη. Επιπλέον, οι πρωτεΐνες αυτές αποτελούν ρυθμιστές της πυκνότητας του σωληνοειδούς ενδοπλασματικού δικτύου καθώς και της αναλογίας των αγωγών και των ασκών του ενδοπλασματικού δικτύου (Chang et al., 2013).

Η sprastin είναι μια πρωτεΐνη γνωστή για τον ρόλο της στη ρύθμιση του μήκους των μικροσωληνίσκων. Το γονίδιο της SPAST στον άνθρωπο κωδικοποιεί για δύο πρωτεϊνικές ισομορφές διαφορετικού μήκους, τις M1 και M87. Η πρωτεΐνη REEP1 ανήκει σε μια μεγάλη πρωτεϊνική οικογένεια που εμφανίζει 6 μέλη στον άνθρωπο. Τόσο η REEP1 όσο και η ισομορφή M1 της sprastin, οι οποίες αλληλεπιδρούν μεταξύ τους, συμμετέχουν στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου, χωρίς να είναι γνωστός ο ακριβής τους ρόλος (Park et al., 2010; Solowska and Baas, 2015).

Οι πρωτεΐνες οι οποίες εντοπίζονται σε μεγαλύτερο βαθμό στους ασκούς του ενδοπλασματικού δικτύου είναι οι CLIMP-63, kinectin και p180 (εικόνα 6B) (Goyal and Blackstone, 2013). Χαρακτηριστικό και των 3 αυτών πρωτεϊνών το οποίο παίζει σημαντικό ρόλο στη λειτουργία τους είναι μια επικράτεια σπειροειδούς έλικας. Η πρωτεΐνη CLIMP-63 λειτουργεί σαν γέφυρα ανάμεσα στις μεμβράνες των ασκών του ενδοπλασματικού δικτύου, διατηρώντας με αυτόν τον τρόπο σταθερό πάχος στους αυλούς του ενδοπλασματικού δικτύου. Το μέγεθος των αυλών είναι ιδιαίτερα σημαντικό

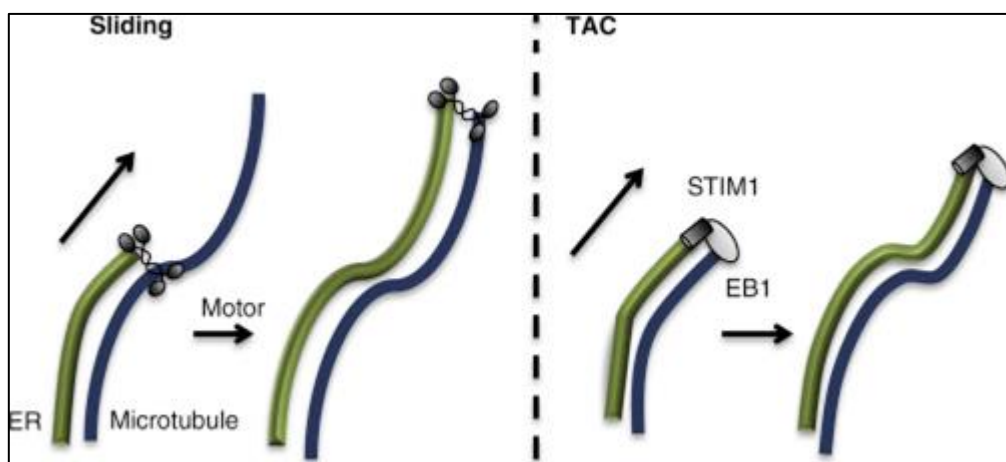
καθώς θα πρέπει από τη μια πλευρά να μπορεί να χωρέσει πρωτεϊνικά μόρια, όπως chaperons (πρωτεΐνες-συνοδούς), και από την άλλη να επιτρέπει στους ασκούς να πακετάρονται καταλαμβάνοντας το μικρότερο δυνατό χώρο (Shibata et al., 2010). Από την άλλη πλευρά οι επικράτειες σπειροειδούς έλικας των kinectin και p180 μπορούν να λειτουργήσουν σαν ράβδοι που σταθεροποιούν την επιπεδότητα των ασκών του ενδοπλασματικού δικτύου (Goyal and Blackstone, 2013). Οι πρωτεΐνες αυτές όμως δεν είναι απαραίτητες για το σχηματισμό *per se* των ασκών του ενδοπλασματικού δικτύου. Αντίθετα τα reticulons και η REEP5/Υοp1p προσδίδουν την απαραίτητη καμπυλότητα στις άκρες των ασκών, χωρίς να είναι ξεκάθαρο κατά πόσο η παρουσία τους αποτελεί αναγκαία συνθήκη για τον σχηματισμό των μεμβρανικών αυτών δομών. Πρόσφατες σχετικά μελέτες έδειξαν ότι σε αυτή τη διαδικασία συμμετέχει η πρωτεΐνη TMEM33, η οποία εντοπίζεται κατά προτίμηση στους ασκούς του ενδοπλασματικού δικτύου όπου περιορίζει τη δράση των reticulons (Urade et al., 2014).



Εικόνα 6. Α) Οι πρωτεΐνες reticulons, spastin M1 και REEPs εντοπίζονται στις μεμβράνες των ασκών του ενδοπλασματικού δικτύου και συμμετέχουν στη διαμόρφωση της δομής αυτών. Μέλη της οικογένειας των REEPs έχουν στο καρβόξυ τελικό τους άκρο μια περιοχή αλληλεπίδρασης με μικροσωληνίσκους. Η spastin M1 διαθέτει μια περιοχή πρόσδεσης στους μικροσωληνίσκους, των οποίων το μήκος περιορίζει. Β) Οι πρωτεΐνες Climp-63, p180, kinectin, reticulons και REEPs συμμετέχουν στη διαμόρφωση της δομής των ασκών του ενδοπλασματικού δικτύου. Η πρωτεΐνη p180 διαθέτει μια επικράτεια πρόσδεσης σε μικροσωληνίσκους. Η kinectin αλληλεπιδρά με την πρωτεΐνη kinesin (Goyal and Blackstone, 2013).

Η κατανομή του ενδοπλασματικού δικτύου και η ισορροπημένη αναλογία αγωγών/ασκών επηρεάζονται από τους μικροσωληνίσκους στα κύτταρα των θηλαστικών. Υπάρχουν δύο μηχανισμοί με τους οποίους πραγματοποιείται η εξαρτώμενη από τους μικροσωληνίσκους κίνηση του ενδοπλασματικού δικτύου: α) το σύμπλεγμα TAC (tip attachment complex) και β) ο μηχανισμός του “sliding” (εικόνα 7). Πιο συγκεκριμένα, το σύμπλεγμα TAC απαρτίζεται από τις πρωτεΐνες STIM1, μια διαμεμβρανική πρωτεΐνη του ενδοπλασματικού δικτύου και την EB1 (+TIP protein end binding 1), μια πρωτεΐνη προσδεσμένη στο + άκρο των μικροσωληνίσκων. Η αλληλεπίδραση των δύο αυτών πρωτεϊνών επιτρέπει στους αγωγούς του ενδοπλασματικού δικτύου να επιμηκύνονται ή να συρρικνώνονται ανάλογα με

τη διαμόρφωση του μήκους των μικροσωληνίσκων. Ο μηχανισμός του sliding αφορά στην μετακίνηση των αγωγών του ενδοπλασματικού δικτύου κατά μήκος ενός μικροσωληνίσκου μέσω της πρωτεΐνης kinesin (Friedman and Voeltz, 2011; Gurel et al., 2014). Άλλες πρωτεΐνες του ενδοπλασματικού δικτύου με ικανότητα πρόσδεσης στους μικροσωληνίσκους του κυτταροσκελετού είναι οι spastin, CLIMP-63, p180 και REEP1. Δεν είναι γνωστό όμως κατά πόσο αυτές οι αλληλεπιδράσεις επηρεάζουν την οργάνωση του ενδοπλασματικού δικτύου (Gurel et al., 2014).



Εικόνα 7. Δύο μηχανισμοί μετατόπισης του ενδοπλασματικού δικτύου με βάση τους μικροσωληνίσκους, οι 'Sliding' και 'TAC' (tip attachment complex). Ο πρώτος περιγράφει την κίνηση των αγωγών του ενδοπλασματικού δικτύου κατά μήκος των ήδη σχηματισμένων μικροσωληνίσκων μέσω κινητικών πρωτεϊνών, όπως η κινεσίνη. Ο δεύτερος αφορά στην αυξομείωση του μήκους των αγωγών του ενδοπλασματικού δικτύου, παράλληλα με τη αντίστοιχη αύξηση ή μείωση του μήκους των μικροσωληνίσκων και στηρίζεται στην αλληλεπίδραση των πρωτεϊνών STIM1 και EBI (Gurel et al., 2014).

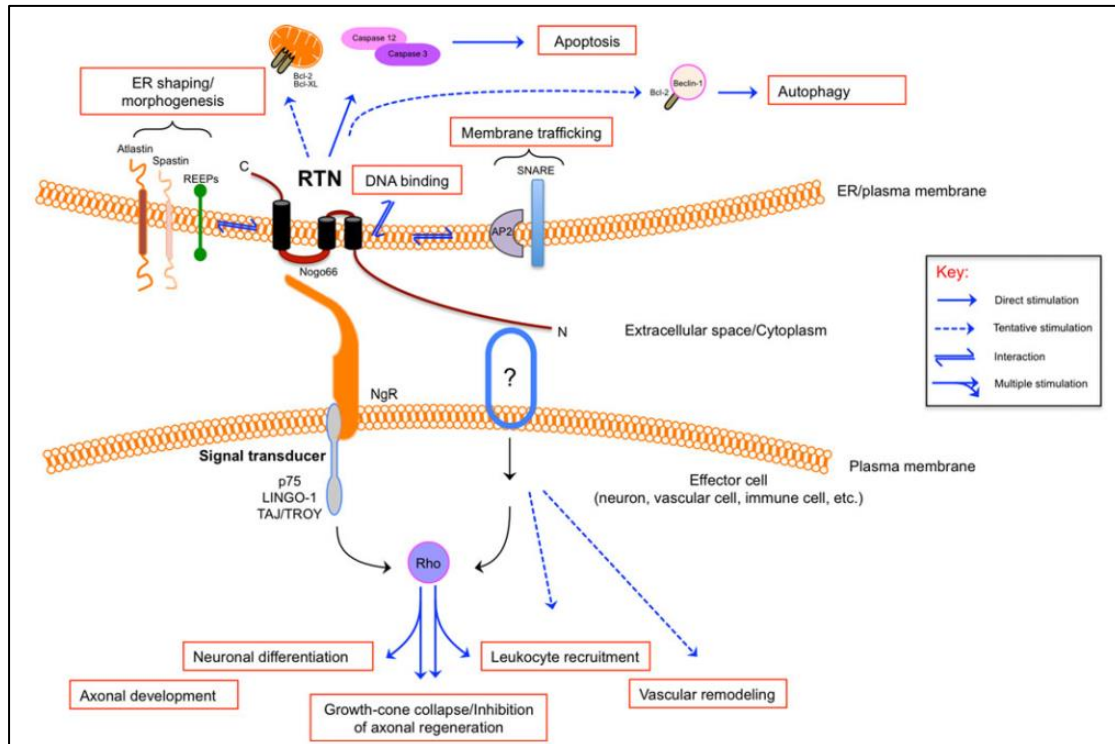
Παρακάτω δίνονται μερικά γενικά στοιχεία για κάθεμια από τις προαναφερθείσες πρωτεΐνες.

1.3.1 Η πρωτεϊνική οικογένεια των reticulons

1.3.1.1 Βασικά στοιχεία

Τα reticulons (RTNs) είναι μια οικογένεια μεμβρανικών πρωτεϊνών οι οποίες είναι συντηρημένες από το σακχαρομήκητα μέχρι τον άνθρωπο (Yang and Strittmatter, 2007). Το πρώτο μέλος της οικογένειας που ανακαλύφθηκε ήταν η RTN1 λόγω της παρουσίας του cDNA της σε νευρικούς ιστούς. Στη συνέχεια, η πρωτεΐνη αυτή χαρακτηρίστηκε ως ένα αντιγόνο ειδικό για νευροενδοκρινικά κύτταρα (Yang and Strittmatter, 2007). Ωστόσο, η οικογένεια των reticulons ήρθε στο προσκήνιο με την ανακάλυψη του RTN4A ή NogoA, που έγινε γνωστό για το ρόλο του ως αναστολέας της αύξησης των νευρώνων καθώς και της αναγέννησης των νευραξόνων στο κεντρικό νευρικό σύστημα (Chiurchiù et al., 2014). Σήμερα, πέρα από την αναστολή της αξονικής ανάπτυξης, είναι γνωστό ότι τα μέλη της οικογένειας των reticulons συμμετέχουν σε μια σειρά άλλων λειτουργιών όπως είναι η διαμόρφωση της δομής του ενδοπλασματικού δικτύου, η κυστιδιακή μεταφορά, και η απόπτωση (εικόνα 8). Τελευταίες μελέτες έχουν δείξει ότι οι πρωτεΐνες αυτές έχουν επιπλέον κάποιους όχι τόσο αναμενόμενους ρόλους όπως η πρόσδεση

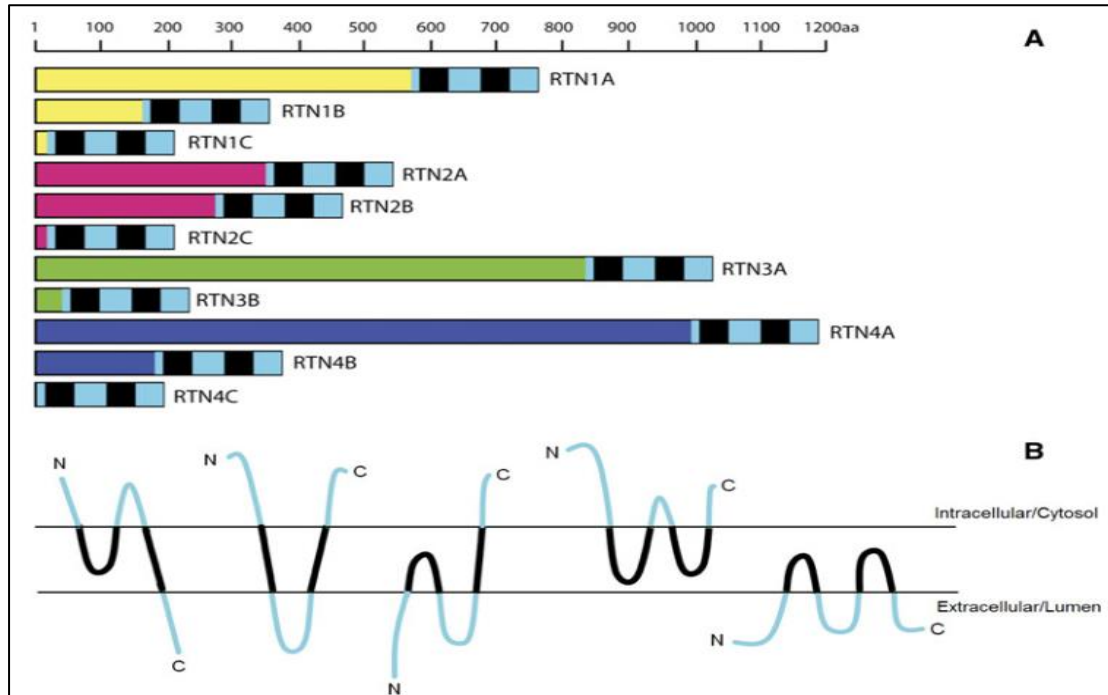
στο DNA, η αυτοφαγία και μερικές λειτουργίες σχετιζόμενες με τη φλεγμονή (εικόνα 8). Για το λόγο αυτό θεωρείται ότι τα reticulons παίζουν ρόλο σε μια σειρά από νευροεκφυλιστικές ασθένειες, όπως η νόσος του Alzheimer, αμυοτροφική πλευρική σκλήρυνση, η σκλήρυνση κατά πλάκας, καθώς και η κληρονομική σπαστική παραπληγία (Chiurchiù et al., 2014). Γενικότερα η ετερογένεια που παρουσιάζουν τα reticulons ως προς τη δομή, την τοπολογία και τον υποκυτταρικό εντοπισμό αντανακλά τη μεγάλη ποικιλία των λειτουργιών στις οποίες συμμετέχουν μέσα στο κύτταρο (Oertle et al., 2003).



Εικόνα 8. Βασικές κυτταρικές λειτουργίες στις οποίες συμμετέχουν τα reticulons. Οι υδρόφοβες περιοχές των reticulons α) μπορούν να αλληλεπιδράσουν με άλλες μεμβρανικές πρωτεΐνες που συμμετέχουν στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου (REEPs, atlastin, spastin) και στην κυστιδιακή μεταφορά (SNARE, AP2), β) μπορούν να προσδένονται άμεσα στο DNA ή γ) να ρυθμίζουν την απόπτωση και την αυτοφαγία μέσα από την ενεργοποίηση των κασπασών και τον έλεγχο των πρωτεϊνών της οικογένειας Bcl-2. Οι περισσότερες όμως από τις λειτουργίες στις οποίες συμμετέχουν τα reticulons διαμεσολαβούνται είτε από την αλληλεπίδραση της επικράτειας Nogo-66 με τον υποδοχέα NogoR, είτε μέσω της αλληλεπίδραση του αμινοτελικού άκρου των πρωτεϊνών αυτών με κάποιον υποδοχέα. Στην περίπτωση του RTN4, οι αλληλεπιδράσεις αυτές οδηγούν στην ενεργοποίηση των πρωτεϊνών Rho και των σηματοδοτικών μονοπατιών στα οποία αυτές συμμετέχουν (Chiurchiù et al., 2014).

Τα μέλη της πρωτεϊνικής οικογένειας τα οποία εμφανίζονται στα χορδωτά χαρακτηρίζονται ως reticulons, ενώ οι ομόλογες πρωτεΐνες που υπάρχουν σε όλους τους υπόλοιπους ευκαρυώτες χαρακτηρίζονται ως reticulon-like πρωτεΐνες (Oertle et al., 2003). Στα θηλαστικά υπάρχουν 4 γονίδια που κωδικοποιούν reticulons (RTN1-4), από τα οποία προκύπτουν 11 εναλλακτικά μετάγραφα (προϊόντα εναλλακτικού ματίσματος) και ακολούθως 11 πρωτεϊνικές ισομορφές, όπως φαίνεται και στην εικόνα 9 (RTN1A, RTN1B, RTN1C, RTN2A, RTN2B, RTN2C, RTN3A, RTN3B, RTN4A-NogoA, RTN4B-NogoB, RTN4C-NogoC) (Oertle et al., 2003). Στον σακχαρομήκητα

Saccharomyces cerevisiae υπάρχουν δύο γονίδια (RTN1, RTN2), και 21 και 17 γονίδια στα φυτά *Arabidopsis thaliana* και *Oryza sativa* αντίστοιχα (Chiurchiù et al., 2014).



Εικόνα 9. Α. Σχηματική απεικόνιση των reticulons των θηλαστικών. Με τα μαύρα και μπλε πλαίσια απεικονίζεται το RHD (reticulon homology domain) σε κάθε πρωτεΐνη, όπου τα μαύρα πλαίσια περιλαμβάνουν υδρόφοβες περιοχές. Τα αμινοτελικά άκρα, τα οποία διαφέρουν ανάμεσα στις πρωτεϊνικές ισομορφές απεικονίζονται με διαφορετικά χρώματα. Η κλίμακα αναφέρεται στον αριθμό των αμινοξέων. Β. Σχηματική απεικόνιση των πιθανών διαμορφώσεων των reticulons στις μεμβράνες (ενδοπλασματικού δικτύου και πλασματική μεμβράνη). Οι διαμορφώσεις αυτές προκύπτουν από προβλέψεις από βιοχημικά πειράματα ή από *in silico* αναλύσεις (Chiurchiù et al., 2014).

Το μέγεθος των reticulons κυμαίνεται από 200 έως 1200 αμινοξικά κατάλοιπα (Oertle et al., 2003; Yang and Strittmatter, 2007). Βασικό δομικό χαρακτηριστικό της οικογένειας αυτής είναι η ύπαρξη μιας συντηρημένης επικράτειας (reticulon homology domain-RHD) προς το καρβόξυ-τελικό άκρο, η οποία έχει μήκος 150-200 αμινοξικά κατάλοιπα και είναι πολύ σημαντική για τον υποκυτταρικό εντοπισμό και τη λειτουργία των reticulons, καθώς και για την αλληλεπίδραση τους με άλλες πρωτεΐνες. Η επικράτεια αυτή αποτελείται από δυο υδρόφοβα διαμεμβρανικά τμήματα μεγάλου μήκους (περίπου 30-35 αμινοξικά κατάλοιπα το καθένα), τα οποία περιβάλλουν έναν υδρόφιλο βρόγχο, μήκους 66 αμινοξικών καταλοίπων (Nogo66). Στη συνέχεια, υπάρχει ένα καρβόξυ-τελικό άκρο μικρού μήκους (Chiurchiù et al., 2014; Oertle et al., 2003). Παρόλο που η πρωτοταγής αλληλουχία της RHD επικράτειας εμφανίζει μικρό βαθμό συντήρησης, η συνολική δομή της έχει διατηρηθεί από τα φυτά μέχρι το σακχαρομύκητα και τον άνθρωπο. Το γεγονός αυτό καταδεικνύει τη σημασία της τρισδιάστατης δομής της επικράτειας στη λειτουργικότητα των πρωτεϊνών (Yang and Strittmatter, 2007). Η πρωτεΐνη NogoA μέσω της επικράτειας RHD μπορεί να προσδεθεί στον μεμβρανικό υποδοχέα Nogo-R, προκαλώντας κατάρρευση του αυξητικού κώνου των νευρώνων.

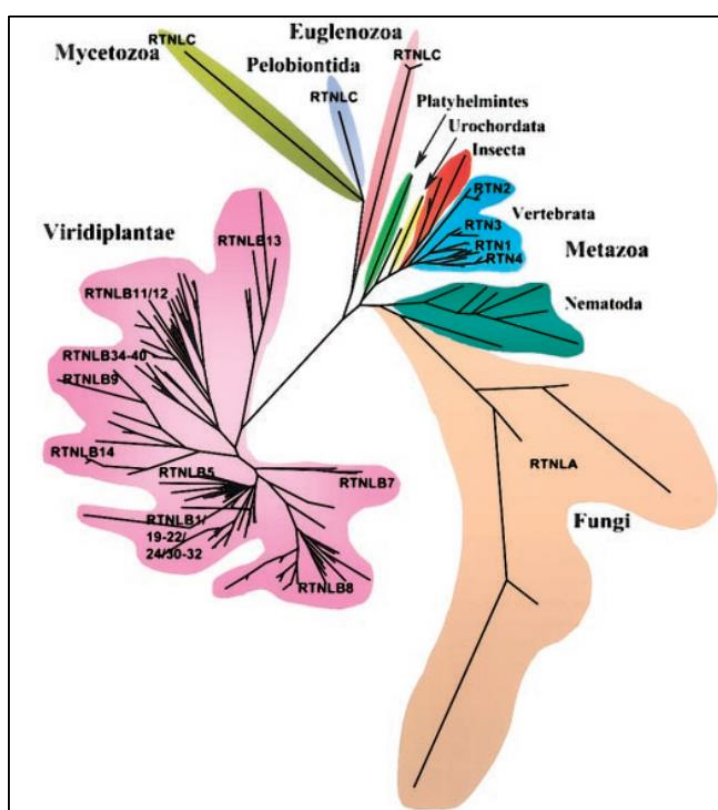
Σε αντίθεση με το καρβόξυ-τελικό άκρο, το άμινο-τελικό άκρο των reticulons εμφανίζει τεράστιες διάφορες ανάμεσα στα μέλη της οικογένειας ως προς το μήκος και την ακολουθία των αμινοξέων (Chiurchiù et al., 2014; Yang and Strittmatter, 2007). Πιο συγκεκριμένα, σε κάποια RTNs, όπως τα RTN1C, RTN2C και RTN4C το άμινο-τελικό άκρο έχει πολύ μικρό μήκος, ενώ αντίθετα σε κάποια RTNs όπως τα RTN1A/B, RTN2A/B, RTN3A και RTN4A/B η περιοχή αυτή είναι μεγάλου μήκους και πιθανά συνεισφέρει σε κάποια επιπλέον βιολογική λειτουργία. Παρόλα αυτά η λειτουργικότητα των περιοχών αυτών, στις οποίες δεν έχει βρεθεί μέχρι σήμερα κάποια δομική επικράτεια, δεν είναι ακόμα γνωστή (Chiurchiù et al., 2014).

Τα reticulons εντοπίζονται κυρίως στις μεμβράνες του ενδοπλασματικού δικτύου. Ωστόσο μέλη της πρωτεϊνικής οικογένειας έχουν βρεθεί τόσο στο σύμπλεγμα Golgi, όσο και στην πλασματική μεμβράνη (Chiurchiù et al., 2014). Παράδειγμα αποτελεί η πρωτεΐνη NogoA η οποία εντοπίζεται στη μεμβράνη των ολιγοδενδροκυττάρων των θηλαστικών. Η ακριβής θέση και διαμόρφωση των reticulons στις μεμβράνες του κυττάρου δεν είναι πλήρως καθορισμένη μέχρι σήμερα. Παρόλα αυτά με βάση τη δομή της RHD επικράτειας σε συνδυασμό με βιοχημικές και *in silico* αναλύσεις έχουν προκύψει πολλές διαφορετικές διαμορφώσεις τις οποίες είναι δυνατό δυνητικά να λαμβάνουν τα reticulons (εικόνα 9).

1.3.1.2 Εξελικτική ιστορία των reticulons

Έχουν γίνει προσπάθειες με σκοπό την ανασύσταση της εξελικτικής πορείας των reticulons. Πιο συγκεκριμένα, μέσω μιας εκτενούς μελέτης που πραγματοποιήθηκε το 2003 (Oertle et al., 2003) αποδείχτηκε η ευρεία κατανομή των reticulons σε ολόκληρο το φάσμα των ευκαρυωτικών οργανισμών και παράλληλα η πλήρης απουσία αυτών από τους προκαρυώτες και τα αρχαία (εικόνα 10). Μέχρι σήμερα έχουν ταυτοποιηθεί πάνω από 250 γονίδια που εντάσσονται στην οικογένεια των reticulons και τα οποία εντοπίζονται τόσο στα ζώα και τα φυτά, όσο και στους μύκητες και σε άλλους κατώτερους ευκαρυωτικούς οργανισμούς (Oertle et al., 2003). Ειδικότερα, έχει διαπιστωθεί η παρουσία πολλαπλών ορθόλογων γονιδίων στους φυτικούς οργανισμούς, τα οποία προέκυψαν μέσα από διπλασιασμούς. Λόγω της απουσίας των reticulons από τους προκαρυώτες, η έναρξη της εξέλιξης των reticulons τοποθετείται στα αρχικά στάδια της εξέλιξης του ευκαρυωτικού κυττάρου και πιθανών συνέβη παράλληλα με την εξέλιξη του ενδομεμβρανικού συστήματος των ευκαρυωτικών οργανισμών 1.7 δισεκατομμύρια χρόνια πριν. Επιπλέον, τα reticulons θεωρείται ότι προέκυψαν από έναν πρωτεϊνικό πρόγονο πλούσιο σε ιντρόνια. Απώλειες αλλά και σε ορισμένες περιπτώσεις προσθήκες ιντρονίων, οδήγησαν στη δημιουργία των γονιδίων που κωδικοποιούν τα reticulons στα διαφορετικά φύλα (Oertle et al., 2003). Οι ανακατατάξεις αυτές είχαν ως αποτέλεσμα τα νέα μέλη της πρωτεϊνικής οικογένειας που προέκυψαν να έχουν διαφορετικές αλληλουχίες εξωνίων στο άμινο-τελικό άκρο, οι οποίες τους προσέδιδαν καινούργιες λειτουργίες. Με αυτόν τον τρόπο προέκυψε η μεγάλη ποικιλομορφία στο αμινοτελικό άκρο των reticulons, η οποία παρατηρείται σήμερα.

Την περίοδο πραγματοποίησης της αρχικής μελέτης με σκοπό την ανασύσταση της εξελικτικής ιστορίας της πρωτεϊνικής οικογένειας των reticulons (Oertle et al., 2003), ο αριθμός των αλληλουχημένων γονιδιωμάτων από το ευρύ φάσμα των ευκαρυωτικών οργανισμών ήταν μικρός. Αυτό είχε ως αποτέλεσμα την υπερεκπροσώπηση των ανώτερων ευκαρυωτικών οργανισμών στη μελέτη αυτή. Αντίθετα, σήμερα υπάρχουν διαθέσιμα τα γονιδιώματα και κατ' επέκταση τα πρωτεώματα από περισσότερους από 7.000 ευκαρυωτικούς οργανισμούς. Παράλληλα τα τελευταία χρόνια έχουν υπάρξει σημαντικές αλλαγές στην ταξινόμηση των ευκαρυωτικών οργανισμών, λόγω της εκτενούς χρήσης μοριακών δεικτών. Τέλος, η ανάπτυξη διαφορετικών αλγορίθμων για την πραγματοποίηση γονιδιακών μελετών έχει επεκτείνει τη δυνατότητα αναζήτησης περισσότερων ορθόλογων γονιδίων σε απομακρυσμένους εξελικτικά οργανισμούς.

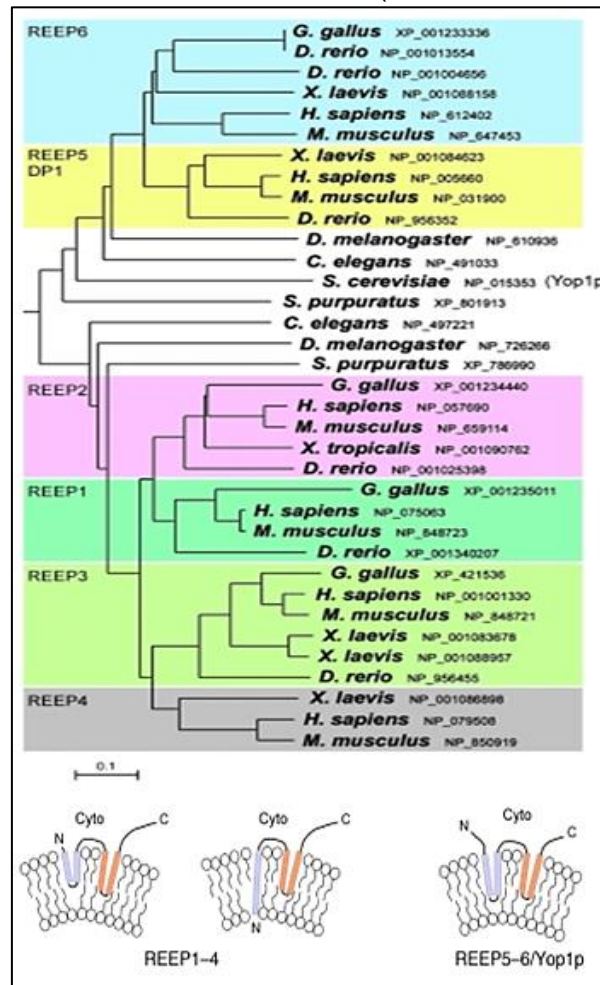


Εικόνα 10. Δενδρόγραμμα στο οποίο είναι εμφανής η κατανομή των γονιδίων των reticulons στα διαφορετικά βασίλεια των ευκαρυωτικών οργανισμών (με βάση παλαιότερη ταξινόμηση των ευκαρυωτικών οργανισμών). Η αναζήτηση των παράλογων γονιδίων πραγματοποιήθηκε με κριτήριο την ύπαρξη και την στοίχιση της δομικής επικράτειας reticulon homology domain (RHD). Παρατηρούμε ότι αντιπρόσωποι των reticulons υπάρχουν στα περισσότερα ευκαρυωτικά βασίλεια από τα πρώιμα μέχρι τα ζώα (Oertle et al., 2003).

1.3.2 Οικογένεια των πρωτεϊνών REEP

Οι REEPs (Receptor Expression Enhancing Proteins) είναι μια οικογένεια μεμβρανικών πρωτεϊνών που είναι συντηρημένες από το σακχαρομύκητα μέχρι τον άνθρωπο και ανήκουν στην οικογένεια Yip (Yip-interacting protein) (Lim et al., 2015). Στον άνθρωπο και γενικότερα στα θηλαστικά υπάρχουν 6 πρωτεΐνες REEP 1-6, ενώ στο σακχαρομύκητα μόνο μια, η Yop1p (YIP 1 partner) που είναι ορθόλογη της ανθρώπινης REEP5. Οι

πρωτεΐνες REEP των θηλαστικών με βάση την ομολογία στην αλληλουχία και τη δομή μπορούν να χωριστούν σε δύο ομάδες, τις REEP1-4 και REEP 5-6, τα μέλη των οποίων εμφανίζουν μεγαλύτερη συγγένεια μεταξύ τους (εικόνα 11). Εξελικτικές αναλύσεις έχουν δείξει ότι η ομάδα REEP1-4 προέκυψε από δύο διπλασιασμούς ολόκληρου του γονιδιώματος, που οδήγησαν αρχικά στη δημιουργία των υποοικογενειών REEP 1-2 και REEP 3-4, και τελικά στην οικογένεια REEP 1-4 (Hurt et al., 2014). Εκτός από τα θηλαστικά και το σακχαρομύκητα οι υπόλοιποι ευκαρυώτες διαθέτουν από έναν αντιπρόσωπο της οικογένειας REEP, που εμφανίζει μεγαλύτερη συγγένεια προς μια από τις δύο αυτές ομάδες των REEP1-4 και REEP 5-6 (Park et al., 2010).



Εικόνα 11. Στο πάνω μέρος της εικόνας δίνεται ένα δέντρο των πρωτεϊνών της οικογένειας REEP. Στο κάτω μέρος της εικόνας εμφανίζεται η τοπολογία των πρωτεϊνών REEP εντός των μεμβρανών, η οποία διαφέρει ανάμεσα στις REEP 1-4 (ύπαρξη δύο πιθανών μοντέλων) και REEP 5-6 (Park et al., 2010).

Το χαρακτηριστικό στοιχείο της οικογένειας των πρωτεϊνών REEP αποτελεί μια επικράτεια μήκους περίπου 200 αμινοξικών καταλοίπων. Η επικράτεια αυτή αποτελείται από δύο μακριά υδρόφοβα τμήματα καθένα από τα οποία αποκτά μια μικρή δομή φουρκέτας εντός της λιπιδικής διπλοστοιβάδας (Yamamoto et al., 2014). Η δομή αυτή όμως διαφέρει ελαφρά ανάμεσα στις δύο ομάδες των REEP που προαναφέρθηκαν (εικόνα 11). Πιο συγκεκριμένα, το πρώτο διαμεμβρανικό τμήμα των πρωτεϊνών της ομάδας

REEP1-4 είναι μικρότερο σε μήκος από το αντίστοιχο των REEP 5-6. Επιπλέον, οι πρωτεΐνες REEP1-4 δεν διαθέτουν αμινοτελικό άκρο εντός του κυτταροπλάσματος, ενώ το καρβοξυτελικό τους άκρο είναι μεγαλύτερο σε μήκος (Park et al., 2010). Εκτός από τη διαμεμβρανική επικράτεια που προαναφέρθηκε βρέθηκε από *in silico* μελέτες ότι οι πρωτεΐνες REEP περιέχουν μια ακόμα συντηρημένη επικράτεια την TB2_DP1_HVA22.

Χαρακτηριστική λειτουργία των πρωτεϊνών REEP αποτελεί η συμμετοχή τους στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου. Πιο συγκεκριμένα μέσω των διαμεμβρανικών τους τμημάτων σχηματίζουν φουρκέτα η οποία εισέρχεται στις μεμβράνες του ενδοπλασματικού δικτύου, προσδίδοντας τους καμπυλότητα (Lim et al., 2015).

1.3.2.1 REEP1/SPG31

Το γονίδιο REEP1 αποτελείται από 7 εξώνια και κωδικοποιεί μια πρωτεΐνη μήκους 200 αμινοξέων. Η 3' αμετάφραστη περιοχή του αποτελεί στόχο για το miR-140. Η πρωτεΐνη REEP1 εκφράζεται σε μια σειρά από νευρικούς και μη, ανθρώπινους ιστούς, συμπεριλαμβανομένου του νωτιαίου μυελού (Züchner et al., 2006). Μεταλλάξεις στο γονίδιο της REEP1 σχετίζονται με ασθένειες όπως η αυτοσωμική επικρατής οικογενής σπαστική παραπληγία και η περιφερική κληρονομική κινητική νευροπάθεια τύπου V (Hurt et al., 2014; Lim et al., 2015).

Η πρωτεΐνη αυτή ανακαλύφθηκε λόγω της ικανότητας της να ενισχύει την ετερόλογη έκφραση γονιδίων που κωδικοποιούν οσφρητικούς υποδοχείς και άλλους υποδοχείς συζευγμένους με G πρωτεΐνες (Hurt et al., 2014). Η REEP1 εντοπίζεται τόσο στο ενδοπλασματικό δίκτυο, όσο και στα μιτοχόνδρια και διαμεσολαβεί στις αλληλεπιδράσεις ανάμεσα στα δύο αυτά οργανίδια (Lim et al., 2015).

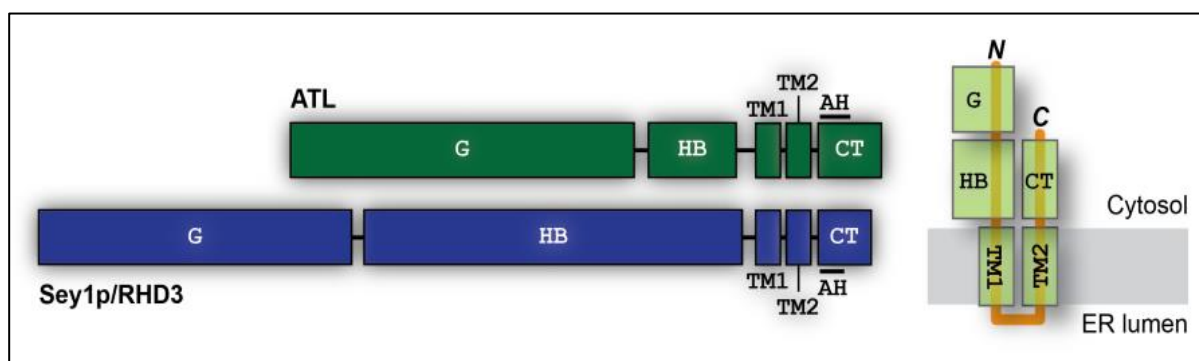
1.3.2.2 REEP5/Yop1p

Η πρωτεΐνη αυτή ήταν το πρώτο μέλος της οικογένειας των REEPs το οποίο ανακαλύφθηκε σε μία έλλειψη χρωμοσωμικού τμήματος μήκους 100 kb από το χρωμόσωμα 5, σε ασθενή με αδενωματώδη πολυποδίαση. Για το λόγο αυτό αρχικά ονομάστηκε DP1 (deleted in polyposis locus 1) (Grodin et al., 1991). Η REEP5 είναι ορθόλογη της Yop1p του *Saccharomyces cerevisiae*. Βασική λειτουργία των πρωτεϊνών αυτών είναι η διαμόρφωση της δομής των αγωγών του ενδοπλασματικού δικτύου μαζί με τις πρωτεΐνες reticulons και atlastins (Wang and Rapoport, 2019).

1.3.3 Atlastin1 (ATL1)/ Sey1p

Οι atlastins (ATLs) και οι πρωτεΐνες Sey1p-like (synthetic enhancement of YOP1 1) είναι μεμβρανικές GTPάσες οι οποίες εντοπίζονται στο ενδοπλασματικό δίκτυο και ανήκουν στην οικογένεια των δυναμινών. Τόσο οι atlastins όσο και οι Sey1ps δεν εμφανίζουν καμία ομοιότητα ως προς την αλληλουχία, αλλά θεωρούνται λειτουργικά ομόλογες. Και οι δύο πρωτεΐνες περιλαμβάνουν μια κυτταροπλασματική επικράτεια GTPάσης στο αμινοτελικό άκρο, μια επικράτεια HB (helical bundle domain), δύο διαμεμβρανικά τμήματα και μια κυτταροπλασματική καρβοξυτελική ουρά η οποία περιέχει μια

αμφιπαθή έλικα (εικόνα 12). Η επικράτεια HB είναι κατά πολύ μεγαλύτερη στις πρωτεΐνες οι οποίες ομοιάζουν στην Sey1p (Hu and Rapoport, 2016).



Εικόνα 12. Στα αριστερά παρουσιάζονται οι επικράτειες που υπάρχουν στις πρωτεΐνες atlastin και Sey1p. G: επικράτεια GTPάσης, HB: helix bundle, TM1, TM2: διαμεμβρανικές περιοχές, CT: καρβοξυτελικό άκρο με αμφιπαθή έλικα (AH). Στα δεξιά παρουσιάζεται η τοπολογία των παραπάνω πρωτεϊνών μέσα στις μεμβράνες του ενδοπλασματικού δικτύου (Hu and Rapoport, 2016).

Οι πρωτεΐνες atlastin και Sey1p-like συμμετέχουν στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου. Πιο συγκεκριμένα, η βασική τους λειτουργία αφορά στη συνένωση των αγωγών του ενδοπλασματικού δικτύου με αποτέλεσμα την επέκτασή του.

Μέλη της πρωτεϊνικής οικογένειας των atlastin/Sey1p εμφανίζονται σε όλους τους ευκαρυωτικούς οργανισμούς. Κανένας, όμως, ευκαρυώτης δεν διαθέτει και atlastin και Sey1p. Στα θηλαστικά υπάρχουν τρεις atlastins, με την ATL1 να εκφράζεται στα νευρικά κύτταρα. Μεταλλαγές στο γονίδιο που κωδικοποιεί αυτή την πρωτεΐνη είναι δυνατό να προκαλέσουν κληρονομική σπαστική παραπληγία, μια νευροεκφυλιστική ασθένεια η οποία χαρακτηρίζεται από μικρότερο μήκος των αξόνων των μεγαλύτερων κινητικών νευρώνων του κορτικοσφαιρικού σωλήνα του σώματος. Πιο συγκεκριμένα, η μεταλλαγμένη atlastin δεν επιτρέπει την συνένωση των αγωγών του ενδοπλασματικού δικτύου και την επέκτασή του δικτύου αυτού σε ολόκληρο το μήκος των νευραξόνων. Αυτό οδηγεί σταδιακά στην αποδυνάμωση και αταξία της κίνησης των κάτω άκρων. Άλλες διαταραχές του νευρικού συστήματος που σχετίζονται με τις πρωτεΐνες atlastin έχουν αναφερθεί εκτός από τον άνθρωπο (atlastin3) στις μύγες και το zebrafish. Παράλληλα, απουσία της Sey1p από την *Candida albicans* μειώνει την μολυσματικότητα του οργανισμού (Hu and Rapoport, 2016).

1.3.4 Ufe1p

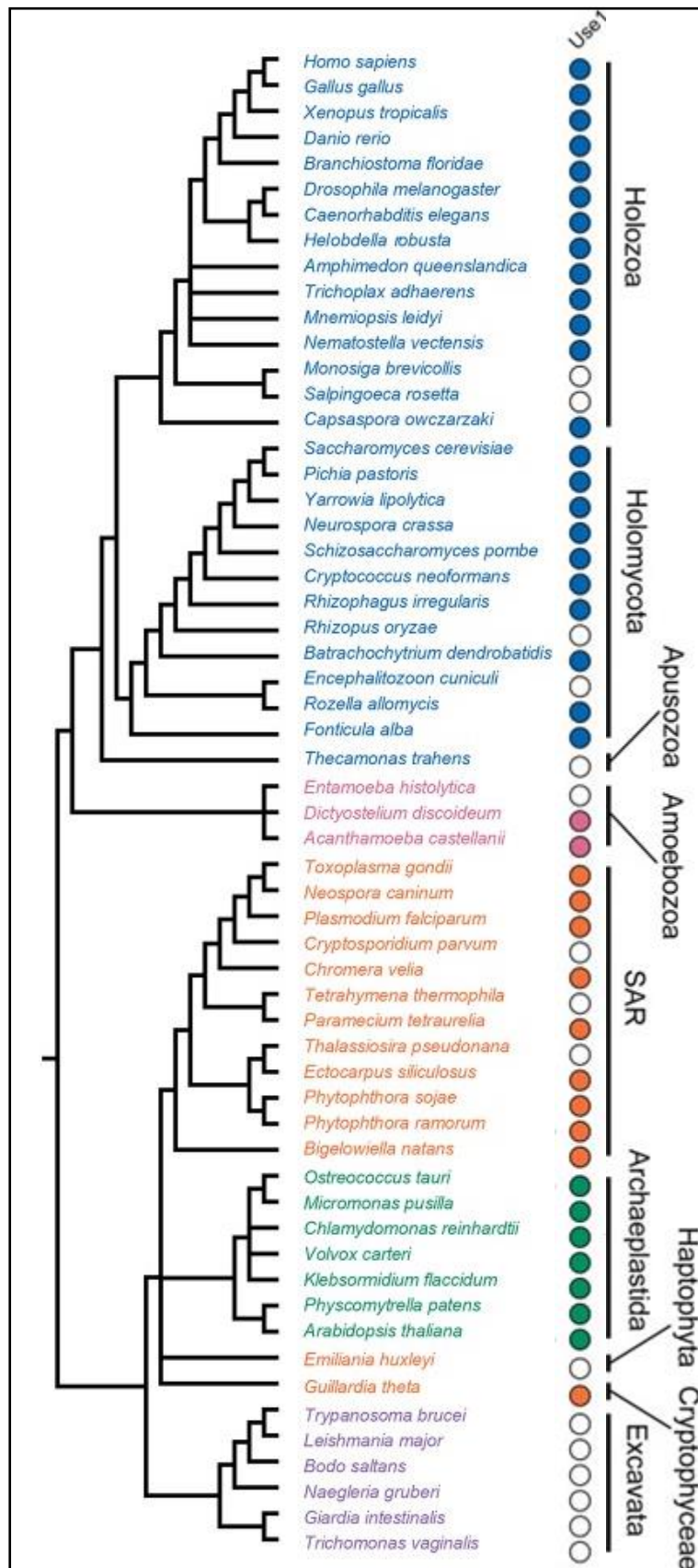
Η Ufe1p είναι μια t-SNARE πρωτεΐνη η οποία βρέθηκε στον *Saccharomyces cerevisiae* και είναι απαραίτητη για την κυστιδιακή μεταφορά από το σύμπλεγμα Golgi προς το ενδοπλασματικό δίκτυο. Η πρωτεΐνη αυτή εντοπίζεται στις μεμβράνες του ενδοπλασματικού δικτύου και μαζί με άλλες πρωτεΐνες t-SNARE (Sec20p, Use1p) δημιουργούν ένα σύμπλοκο που αλληλεπιδρά με την v-SNARE Sec22p, η οποία εντοπίζεται πάνω σε COPI κυστίδια προερχόμενα από το cis-Golgi, ώστε να προκαλέσουν σύντηξη των δύο μεμβρανών (Patel et al., 1998; Rogers et al., 2013). Παράλληλα, η πρωτεΐνη Ufe1p συμμετέχει στην ομοτυπική σύντηξη μεμβρανών (Patel et al.,

1998). Πιο συγκεκριμένα, μαζί με τις πρωτεΐνες Use1p και Dsl1p διαμεσολαβεί στη σύντηξη μεμβρανών του ενδοπλασματικού δικτύου. Επιπλέον, χαρακτηριστική είναι και η συμμετοχή της Ufe1p στη σύντηξη μεμβρανών απλοειδικών πυρήνων μυκήτων, που οδηγεί στη δημιουργία ενός διπλοειδούς πυρήνα κατά τη σύζευξη (Rogers et al., 2013). Τέλος, η Ufe1p συμμετέχει στην αυτοφαγία στο σακχαρομύκητα καθώς μεταφέρεται σε ειδικά COPII κυστίδια από το ενδοπλασματικό δίκτυο προς τα σημεία οργάνωσης του αυτοφαγосώματος (Lemus et al., 2016).

1.3.5 Use1p

Η Use1p (unconventional SNARE in the ER) είναι μια SNARE πρωτεΐνη του *Saccharomyces cerevisiae*, η οποία εντοπίζεται στο ενδοπλασματικό δίκτυο. Η πρωτεΐνη αυτή εμφανίζει το μοτίβο SNARE της αντίστοιχης πρωτεϊνικής οικογένειας, αλλά έχει μικρό βαθμό ομολογίας αμινοξικής ακολουθίας με τα υπόλοιπα μέλη της οικογένειας (Dilcher et al., 2003). Επιπλέον, η πρωτεΐνη Use1p όπως και η Ufe1p συμμετέχει στην κυστιδιακή μεταφορά από το σύμπλεγμα Golgi στο ενδοπλασματικό δίκτυο, στην ομοτυπική σύντηξη μεμβρανών του ενδοπλασματικού δικτύου καθώς και στη μεμβρανική σύντηξη απλοειδικών πυρήνων κατά τη σύζευξη στον οργανισμό *Sacharomyces cerevisiae* (Rogers et al., 2013).

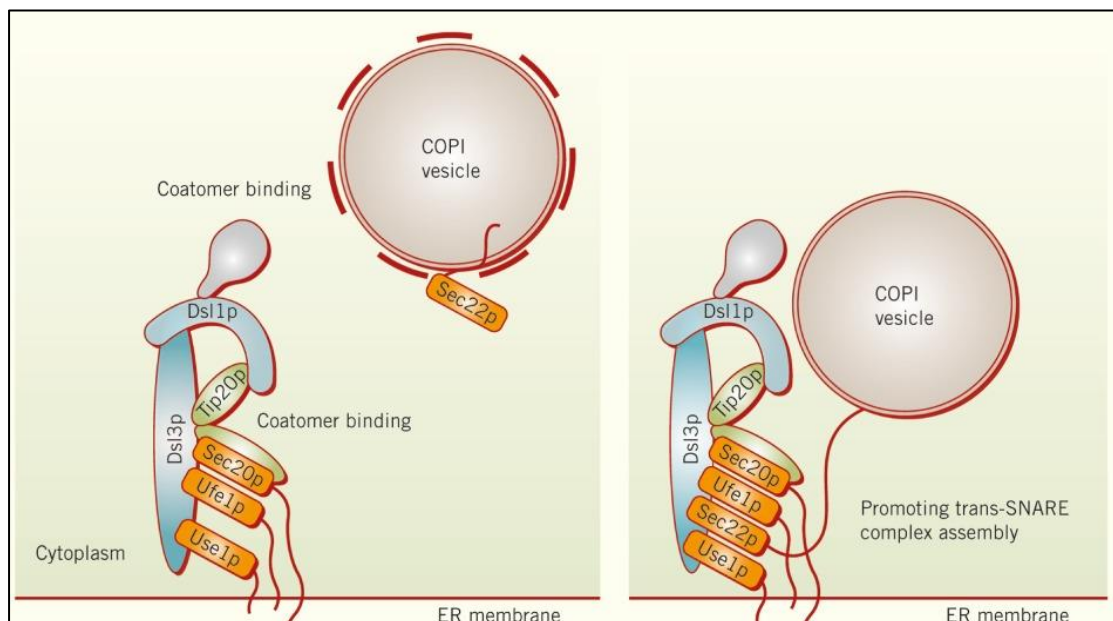
Έχει πραγματοποιηθεί ήδη μια μελέτη συγκριτικής γονιδιωματικής ανάλυσης για την πρωτεΐνη Use1p (Field et al., 2017). Τα αποτελέσματα αυτής συνοψίζονται στην εικόνα 13. Σύμφωνα με αυτά παρατηρούμε ότι ορθόλογες πρωτεΐνες της Use1p εμφανίζονται στα οπισθόκοντα, τα φυτά και την ομάδα των SAR. Παράλληλα η πρωτεΐνη αυτή φαίνεται να απουσιάζει από τα excavata.



Εικόνα 13. Αποτελέσματα συγκριτικής γονιδιωματικής μελέτης για την πρωτεΐνη Use1p στους ευκαρυωτικούς οργανισμούς (Field et al., 2017).

1.3.6 Dsl1p

Η πρωτεΐνη Dsl1p (depends on SLY1-20) είναι μια περιφερειακή μεμβρανική πρωτεΐνη του ενδοπλασματικού δικτύου, μέλος του συμπλόκου Dsl1 στον *Sacharomyces cerevisiae*. Το σύμπλοκο αυτό το οποίο αποτελείται από τις πρωτεΐνες Dsl1p, Dsl3p/Sec39p και Tip20p παίζει βασικό ρόλο στην κυστιδιακή μεταφορά από το σύμπλεγμα Golgi στο ενδοπλασματικό δίκτυο (Andag and Schmitt, 2003; Schmitt, 2010; Spang, 2012). Πιο συγκεκριμένα, φαίνεται ότι το σύμπλοκο Dsl1 μέσω μιας επικράτειας της πρωτεΐνης Dsl1p αναγνωρίζει τα COPI κυστίδια και προωθεί τη σύντηξη τους με τη μεμβράνη του ενδοπλασματικού δικτύου. Η σύντηξη πραγματοποιείται μέσω των πρωτεϊνών SNARE Use1p, Ufe1p και Sec20p, οι οποίες αλληλεπιδρούν με το σύμπλοκο Dsl1 επιτρέποντας τον εντοπισμό του στη μεμβράνη του ενδοπλασματικού δικτύου (εικόνα 14) (Andag and Schmitt, 2003; Rogers et al., 2014; Schmitt, 2010). Η ίδια διαδικασία στα θηλαστικά διαμεσολαβείται από το σύμπλοκο της συνταξίνης 18, στο οποίο συμμετέχει η πρωτεΐνη Zw10 ομόλογη της Dsl1p στα θηλαστικά. Το σύμπλοκο αυτό παίζει σημαντικό ρόλο στη μείωση στα θηλαστικά, όπου η πρωτεΐνη Zw10 παρακολουθεί τη σωστή πρόσδεση των μικροσωληνίσκων στον κινητοχώρο (Schmitt, 2010). Μια επιπλέον, λειτουργία στην οποία συμμετέχει η πρωτεΐνη Dsl1p, μέσω του συμπλόκου Dsl1, μαζί με τις πρωτεΐνες Use1p και Ufe1p αφορά στην ομοτυπική σύντηξη ανάμεσα στις μεμβράνες του ενδοπλασματικού δικτύου (Rogers et al., 2014).

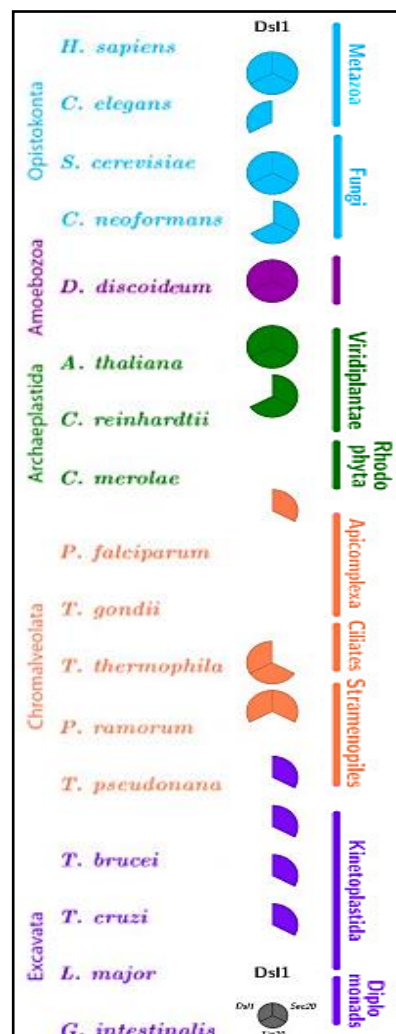


Εικόνα 14. Ο ρόλος της πρωτεΐνης Dsl1p στην κυστιδιακή μεταφορά από το σύμπλεγμα Golgi προς το ενδοπλασματικό δίκτυο (Andag and Schmitt, 2003).

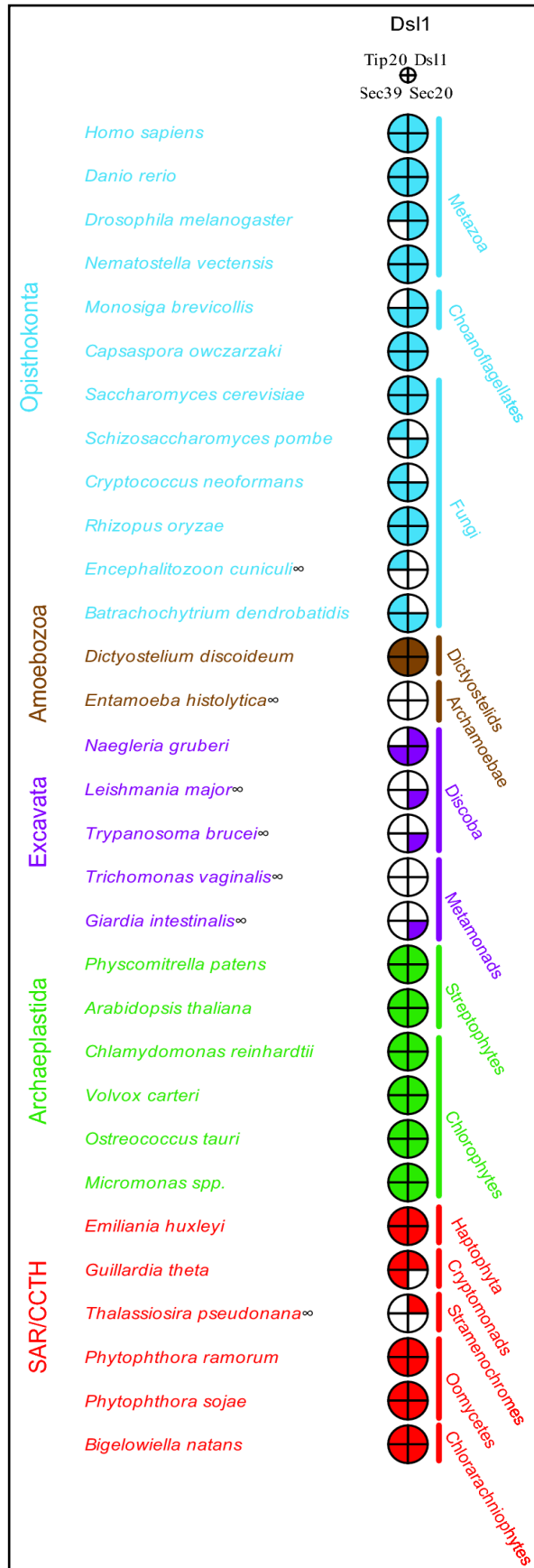
Έχουν αναγνωριστεί τρεις επικράτειες στην πρωτεΐνη Dsl1p: μια επικράτεια πρόσδεσης στην πρωτεΐνη Tip20p στο αμινοτελικό άκρο, μια «όξινη» περιοχή στο κέντρο της πρωτεΐνης τα κατάλοιπα τρυπτοφάνης της οποίας είναι απαραίτητα για την αλληλεπίδραση με την πρωτεΐνη delta-COP του καλύμματος των κυστιδίων COPI, καθώς και μια εξελικτικά συντηρημένη

επικράτεια στο καρβοξυτελικό άκρο της πρωτεΐνης (Andag and Schmitt, 2003).

Η εξελικτική ιστορία της πρωτεΐνης Dsl1p μαζί με τις υπόλοιπες πρωτεΐνες μέλη του συμπλόκου Dsl1 έχει ήδη μελετηθεί και σε προηγούμενες λιγότερο εκτενείς μελέτες (Klinger et al., 2013; Koumandou et al., 2007). Τα αποτελέσματα αυτών φαίνονται στις εικόνες 15 και 16. Από τις εικόνες αυτές παρατηρούμε ότι η πρωτεΐνη Dsl1p εμφανίζει ευρεία εξάπλωση στους ευκαρυωτικούς οργανισμούς. Παράλληλα είναι εμφανές ότι ο αριθμός των οργανισμών που χρησιμοποιήθηκαν σε κάθε μια από αυτές τις μελέτες είναι σχετικά μικρός. Το γεγονός αυτό δικαιολογείται πιθανά από τον μειωμένο αριθμό πλήρως αλληλουχημένων γονιδιωμάτων κατά το χρονικό διάστημα πραγματοποίησης της συγκεκριμένης μελέτης.



Εικόνα 15. Παρουσία των πρωτεϊνών που απαρτίζουν το σύμπλοκο Dsl (Dsl1p, Tip20p, Sec20p) στους ευκαρυωτικούς οργανισμούς (Koumandou et al., 2007).



Εικόνα 16. Αποτελέσματα συγκριτικής γονιδιωματικής μελέτης για τις πρωτεΐνες μέλη του συμπλόκου Dsl (Dsl1p, Tip20p, Sec20p) στους ευκαρυωτικούς οργανισμούς (Klinger et al., 2013).

1.3.7 Arl6IP1/ARMER

Η Arl6IP1 (ADP-ribosylation factor-like 6 interacting protein 1, γνωστή και ως ARMER) είναι μια μεμβρανική πρωτεΐνη που διαπερνά τη μεμβράνη 4 φορές μέσα από 4 ισχυρά υδρόφοβες δομές φουρκέτας, δομή που προσομοιάζει την επικράτεια RHD (reticulon homology domain) των reticulons. Η πρωτεΐνη αυτή ανιχνεύτηκε για πρώτη φορά στη διάρκεια πειραμάτων για τον προσδιορισμό πρωτεϊνών που αλληλεπιδρούν με την πρωτεΐνη Arl6 (ADP-ribosylation factor-like 6), με τεχνικές yeast two-hybrid (Ingle et al., 1999; Yamamoto et al., 2014). Η Arl6IP1 εντοπίζεται μόνο σε πολυκύτταρους ευκαρυωτικούς οργανισμούς και συγκεκριμένα στα θηλαστικά, στα ψάρια, στους νηματώδεις και στα αμφίβια, σε αντίθεση με οργανισμούς όπως οι μύκητες (Yamamoto et al., 2014).

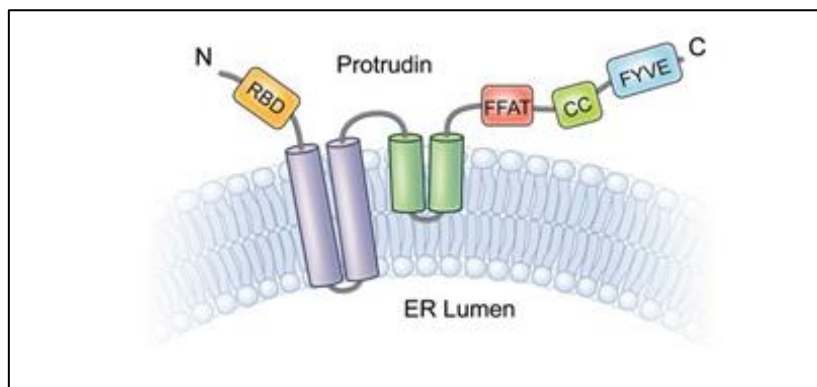
Στην πρωτεΐνη Arl6IP1 έχουν αποδοθεί μια σειρά από διαφορετικοί ρόλοι. Πιο συγκεκριμένα, μια καλά χαρακτηρισμένη λειτουργία της αφορά στη ρύθμιση της κυτταρικής επιβίωσης. Μελέτες έχουν δείξει ότι αποπτωτικά ερεθίσματα οδηγούν στον περιορισμό της έκφρασης της Arl6IP1, ενώ αντίθετα υπερέκφραση της πρωτεΐνης αυτής εξωγενώς σε κύτταρα οδηγεί στην αναστολή της έκφρασης της κασπάσης-9, προστατεύοντας τα από την απόπτωση (Yamamoto et al., 2014). Μια ακόμη λειτουργία την οποία επιτελεί η Arl6IP1 αφορά στη ρύθμιση της μεταφοράς του γλουταμικού στους νευρώνες, μέσα από την πρόσδεση της στην πρωτεΐνη GTRAP3-18/addictin (γνωστή και ως ARL6IP5), πρωτεΐνη ρυθμιστή του μεταφορέα του γλουταμικού. Παράλληλα, στο zebrafish η πρωτεΐνη Arl6IP1 εμπλέκεται στην ανάπτυξη της νευρικής ακρολοφίας στη διάρκεια της εμβρυογένεσης καθώς και στην πρώιμη ανάπτυξη της καρδιάς, του αμφιβληστροειδούς και του πτερυγίου. Σημαντικός είναι και ο ρόλος της πρωτεΐνης αυτής στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου και συγκεκριμένα στο σχηματισμό των αγωγών αυτού, λειτουργία την οποία επιτελούν και οι πρωτεΐνες reticulons και REEPs/DP1/Yop1. Καθοριστικής σημασίας για τη λειτουργία αυτή είναι η δομή της Arl6IP1 η οποία προσομοιάζει αυτή των reticulons, καθώς και η αλληλεπίδραση της με την atlastin (Yamamoto et al., 2014). Τέλος, μεταλλαγές στην αλληλουχία της πρωτεΐνης Arl6IP1 σχετίζονται με την νευροεκφυλιστική ασθένεια σπαστική παραπληγία, όπως συμβαίνει και με αρκετές άλλες μεμβρανικές πρωτεΐνες του ενδοπλασματικού δικτύου (Novarino et al., 2014).

1.3.8 Protrudin

Η πρωτεΐνη protrudin είναι μια μεμβρανική πρωτεΐνη που εντοπίζεται κυρίως στους αγωγούς του ενδοπλασματικού δικτύου και ρυθμίζει την πολωμένη κυστιδιακή μεταφορά στους νευρώνες (Hashimoto et al., 2014; Ohnishi et al., 2014). Επιπλέον, η πρωτεΐνη protrudin συμμετέχει στη δημιουργία και στη σταθεροποίηση των αγωγών του ενδοπλασματικού δικτύου (Ohnishi et al., 2014).

Η δομή της φαίνεται στην εικόνα 17. Πιο συγκεκριμένα, η πρωτεΐνη protrudin περιέχει μια επικράτεια πρόσδεσης της πρωτεΐνης Rab11 (RBD), 2 διαμεμβρανικά τμήματα, ένα μοτίβο FFAT (two phenylalanines in an acidic tract) και μια επικράτεια FYVE (Fab-1, YGL023, Vps27, and EEA1). Η

πρωτεΐνη αυτή είναι γνωστή και ως SPG33 (Spastic paraplegia 33), διότι μεταλλαγές στην αλληλουχία της σχετίζονται με την ασθένεια οικογενή σπαστική παραπληγία (Hashimoto et al., 2014; Urade et al., 2014).



Εικόνα 17. Τοπολογία της πρωτεΐνης protrudin στη μεμβράνη του ενδοπλασματικού δικτύου. CC, coiled-coil; RBD, Rab-binding domain, FFAT: two phenylalanines in an acidic tract, FYVE: Fab-1, YGL023, Vps27, and EEA1 (Chang et al., 2013).

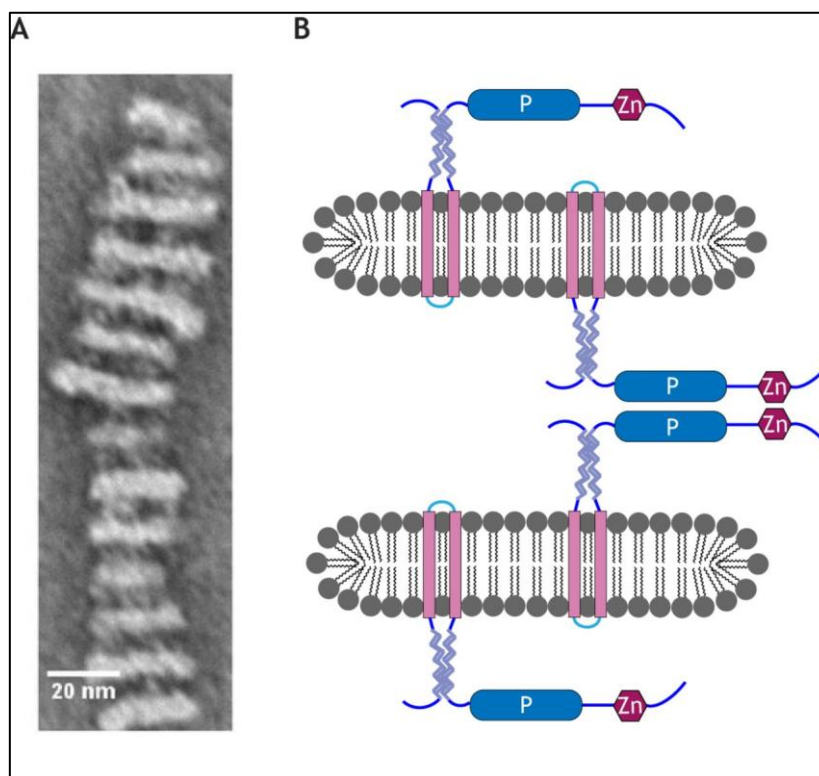
Η protrudin μέσω της επικράτειας FYVE αλληλεπιδρά με την πρωτεΐνη VAP η οποία είναι σημαντικός ρυθμιστής του υποκυτταρικού εντοπισμού της στο ενδοπλασματικό δίκτυο. Επιπλέον, μέσω της ίδιας επικράτειας η protrudin αλληλεπιδρά με την πρωτεΐνη Kif5A (SPG10), η οποία ρυθμίζει την κυστιδιακή μεταφορά προς του άξονες των νευρώνων (Hashimoto et al., 2014). Η protrudin αλληλεπιδρά με τις πρωτεΐνες των οικογενειών atlastin και REEP (Urade et al., 2014).

1.3.9 Lunapark/Lnp1p

Οι πρωτεΐνες lunapark ανήκουν σε μια οικογένεια πρωτεϊνών συντηρημένη από τον άνθρωπο μέχρι το σακχαρομύκητα. Οι πρωτεΐνες μέλη της οικογένειας περιέχουν ένα μοτίβο N-μυριστοϋλίωσης, δύο διαμεμβρανικές επικράτειες και μία επικράτεια δακτύλων ψευδαργύρου. Στα θηλαστικά, δίπλα στην επικράτεια ψευδαργύρου υπάρχει η αλληλουχία LNPARK, από την οποία πήρε το όνομα της η πρωτεΐνη (Chen et al., 2012).

Η πρωτεΐνη Lnp1p (Lunapark family member 1) του σακχαρομύκητα εντοπίζεται στους κόμβους του σωληνοειδούς ενδοπλασματικού δικτύου, του οποίου τη δομή διαμορφώνει λειτουργώντας συνεργιστικά με τις πρωτεΐνες reticulons, Yor1p και ανταγωνιστικά με την Sey1p (Urade et al., 2014; Yamamoto et al., 2014). Πιο συγκεκριμένα, η Lnp1p και η Sey1p δρουν ανταγωνιστικά ως προς την αλληλεπίδραση με τα reticulons καθώς και τη διαμόρφωση του πολυγωνικού δικτύου των αγωγών του ενδοπλασματικού δικτύου (Chen et al., 2012; Urade et al., 2014). Παράλληλα η Sey1p ρυθμίζει τον εντοπισμό της Lnp1p στο ενδοπλασματικό δίκτυο (Chen et al., 2012). Επιπλέον, η Lnp1p αλληλεπιδρά με την πρωτεΐνη gp78, μια λιγάση ουβικιτίνης που συμμετέχει στον έλεγχο ποιότητας του ενδοπλασματικού δικτύου. Παρόλο που η Lnp1p διαθέτει μια ενδογενή δράση λιγάσης της ουβικιτίνης, μέχρι σήμερα δεν έχει συσχετιστεί η δράση της με την αποικοδόμηση ελαττωματικών πρωτεϊνών του ενδοπλασματικού δικτύου (Zhao et al., 2016). Τέλος, απομόνωση της πρωτεΐνης Lnp από κύτταρα τόσο

του ανθρώπου όσο και του *Xenopus laevis* και ανάμειξη με λιπίδια, οδήγησε στη δημιουργία πολλών μεμβρανικών μικυλλίων τοποθετημένων το ένα πάνω στο άλλο (εικόνα 18). Πιο συγκεκριμένα, η δομή αυτή διαμορφώθηκε λόγω της παρουσίας πρωτεϊνών Lnp και στις δύο μη καμπύλες επιφάνειες κάθε πεπλατυσμένου δίσκου καθώς και λόγω των αλληλεπιδράσεων πρωτεϊνών Lnp από διαδοχικούς δίσκους (Wang and Rapoport, 2019).



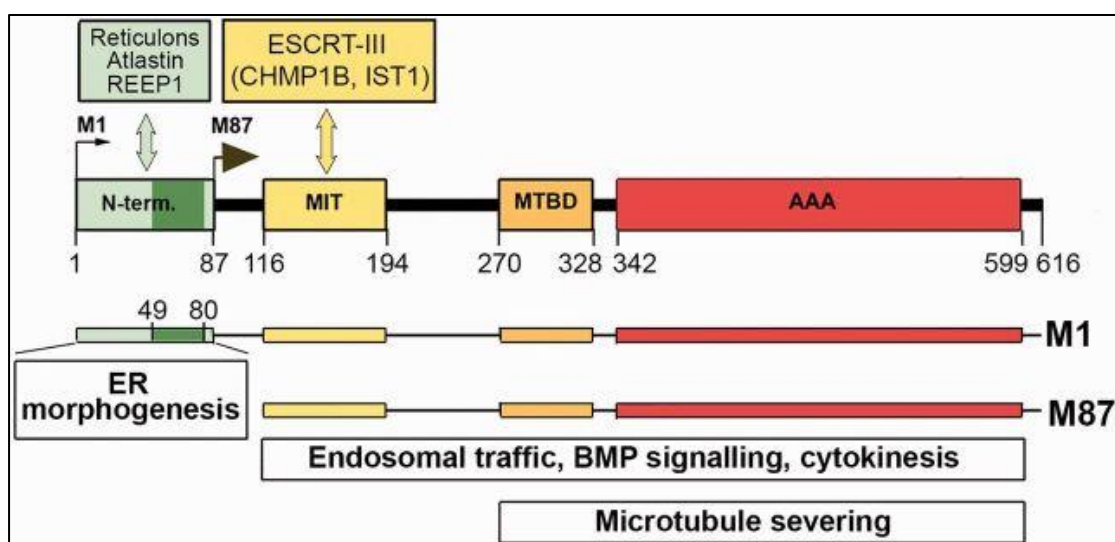
Εικόνα 18. Δημιουργία μικυλλίων από την πρωτεΐνη luncapark. A) Η απομονωμένη πρωτεΐνη luncapark από τον *Xenopus laevis* αναμείχτηκε με λιπίδια και οι δομές που σχηματίστηκαν φαίνονται μέσω αρνητικής χρώσης σε ηλεκτρονικό μικροσκόπιο. B) Η πρωτεΐνη luncapark εντοπίζεται και στις δύο επιφάνειες μεμβρανικών μικυλλίων. Οι επικράτειες φωσφορυλίωσης και δακτύλων ψευδαργύρου πρωτεϊνών από διαδοχικούς δίσκους αλληλεπιδρούν, με αποτέλεσμα την τοποθέτηση του ενός δίσκου πάνω στον άλλο (Wang and Rapoport, 2019).

1.3.10 Spastin

Η spastin ανήκει σε μια ευρεία οικογένεια ΑΤΡασών που συμμετέχουν σε ποικίλες κυτταρικές λειτουργίες και ονομάζονται AAA-ATPases (ATPases Associated with diverse cellular Activities). Η πρωτεΐνη αυτή παίζει σημαντικό ρόλο στη λειτουργία των μικροσωληνίσκων του κυτταροσκελετού τους οποίους τέμνει σε μικρότερα τμήματα, στη μεταφορά κυτταρικών συστατικών και στην κυτταρική διαίρεση. Το γονίδιο που κωδικοποιεί την πρωτεΐνη αυτή είναι το πιο συχνά μεταλλάσσόμενο γονίδιο στη νόσο της κληρονομικής σπαστικής παραπληγίας. Μεταλλάξεις στο γονίδιο αυτό μπορεί να εμφανίζονται στο 15-40% των ασθενών, ανάλογα με το εθνικό τους υπόβαθρο (Solowska and Baas, 2015).

Το γονίδιο SPAST έχει μέγεθος περίπου 90kb, περιλαμβάνει 17 εξώνια και κωδικοποιεί για δύο ισομορφές της πρωτεΐνης spastin στον άνθρωπο, τις M1 και M87, οι οποίες προκύπτουν από δύο εναλλακτικά κωδικόνια έναρξης

(εικόνα 19). Η ισομορφή M1 περιέχει μια επιπλέον υδρόφοβη επικράτεια μήκους 86 αμινοξικών καταλοίπων στο αμινοτελικό της άκρο, η οποία της επιτρέπει να εντοπίζεται και να συμμετέχει στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου. Αυτό συμβαίνει μέσω μιας δομής φουρκέτας την οποία αποκτούν τα αμινοξικά κατάλοιπα 49-80 και η οποία μπορεί να εισέρχεται μερικώς στις μεμβράνες του ενδοπλασματικού δικτύου. Παράλληλα, οι επικράτειες MTBD (Microtubule-binding domain) και AAA είναι απαραίτητες για τον εξαμερισμό της spastin και τη διαμόρφωση του μήκους των μικροσωληνίσκων καθώς και για την κυστιδιακή μεταφορά πρωτεϊνών και την κυτοκίνηση. Αντίθετα μέσω της επικράτειας MIT (microtubule interacting and trafficking domain) πραγματοποιείται η αλληλεπίδραση με τις πρωτεΐνες ESCRT-III (endosomal sorting complex required for transport III), CHMP1B (charged multivesicular body protein 1B) και IST1 (increased sodium tolerance), οι οποίες συμμετέχουν στην κυστιδιακή μεταφορά πρωτεϊνών και την κυτοκίνηση (Solowska and Baas, 2015).



Εικόνα 19. Από το γονίδιο SPAST στον άνθρωπο προκύπτουν δύο πρωτεΐνες, οι M1 και M87, οι οποίες έχουν διαφορετικό μήκος. Η M1 spastin περιέχει στο αμινοτελικό άκρο μια επιπλέον επικράτεια, η οποία θεωρείται ότι της επιτρέπει να εντοπίζεται στη μεμβράνη του ενδοπλασματικού δικτύου και να συμμετέχει στη διαμόρφωση της δομής αυτού μέσα από την αλληλεπίδραση της με τις πρωτεΐνες reticulons, atlastin και REEP1. Αλληλεπίδραση της επικράτειας MIT με την πρωτεΐνη ESCRT-III ρυθμίζει τη μεταφορά κυστιδίων και την κυτοκίνηση.

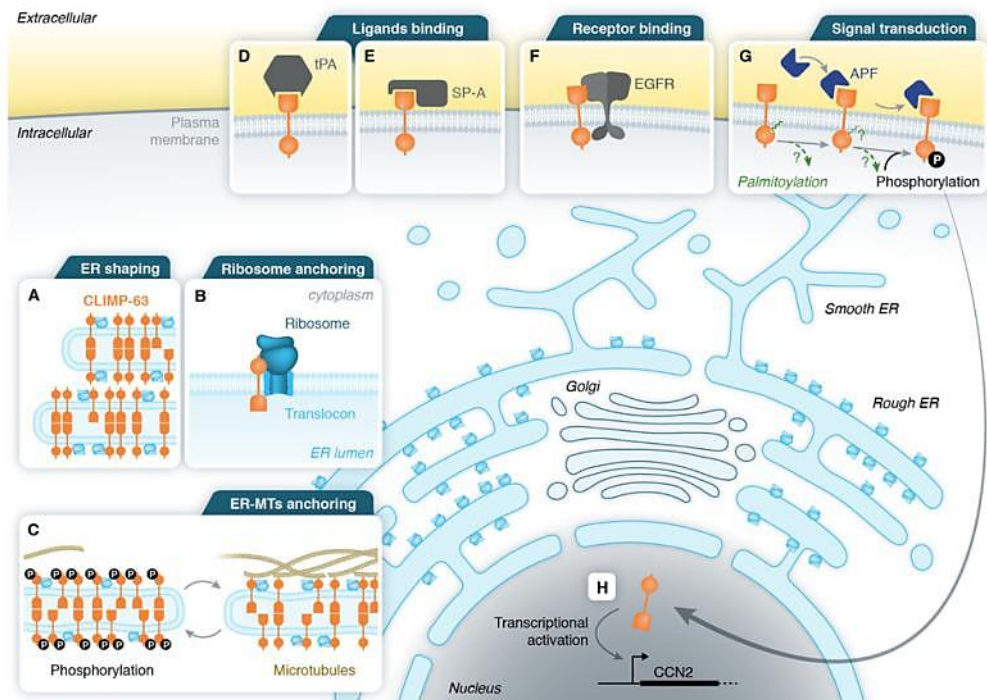
Στα τρωκτικά η ισομορφή M87 είναι η πλέον επικρατής σε όλους τους ιστούς και σε όλα τα αναπτυξιακά στάδια, ενώ η M1 εντοπίζεται κυρίως στο νωτιαίο μυελό σε ενήλικα ζώα. Αντίστοιχα σε μελέτες στο ανθρώπινο κεντρικό νευρικό σύστημα η M87 βρέθηκε τόσο στο νωτιαίο μυελό όσο και στο φλοιό του εγκεφάλου, ενώ η M1 εντοπίστηκε σε σημαντικά επίπεδα μόνο στο νωτιαίο μυελό. Τα παραπάνω τονίζουν ότι η ισομορφή M87 είναι η επικρατής στο κεντρικό νευρικό σύστημα (Solowska and Baas, 2015).

1.3.11 CLIMP-63

Η πρωτεΐνη CLIMP-63 (cytoskeleton-linking membrane protein 63) ή αλλιώς CKPA4 (cytoskeleton-associated protein 4), αρχικά είχε λάβει το όνομα p63. Ανακαλύφθηκε ταυτόχρονα από δύο ανεξάρτητες ερευνητικές

ομάδες ως η πιο παλμιτούλιωμένη πρωτεΐνη στη διάρκεια της μίτωσης, καθώς και ως αντιγόνο μονοκλωνικών αντισωμάτων που δημιουργήθηκαν έναντι κυτταρικού κλάσματος πλούσιου σε μεμβράνες του συστήματος Golgi. Η μοναδική πρωτεϊνική ισομορφή της CLIMP-63 εμφανίζεται στους ανώτερους ευκαρυωτικούς οργανισμούς, ενώ απουσιάζει από τα βακτήρια, τους μύκητες, τους σκώληκες και τα έντομα. Στον άνθρωπο αποτελείται από μια κυτταροπλασματική επικράτεια μήκους 106 αμινοξικών καταλοίπων, μια διαμεμβρανική περιοχή και μια μεγάλη εξωκυτταροπλασματική περιοχή μήκους 474 καταλοίπων. Πρόκειται για μια διαμεμβρανική πρωτεΐνη τύπου II, που εντοπίζεται κυρίως στο αδρό ενδοπλασματικό δίκτυο, αλλά έχει βρεθεί και στην πλασματική μεμβράνη (Sandoz and van der Goot, 2015).

Η πρωτεΐνη CLIMP-63 επιτελεί πολλαπλούς ρόλους σε διαφορετικές περιοχές του κυττάρου. Η επικράτεια της που εντοπίζεται στο κυτταρόπλασμα έχει την ικανότητα να προσδένεται στους μικροσωληνίσκους παρέχοντας ένα μέσο σύνδεσης με τον κυτταροσκελετό (εικόνα 20C). Αντίθετα, η επικράτεια που εντοπίζεται στον αυλό του ενδοπλασματικού δικτύου μπορεί να δημιουργήσει όμο-ολιγομερή με άλλα μόρια που βρίσκονται είτε στην ίδια είτε στην αντίθετη (*in trans*) πλευρά του αυλού του ενδοπλασματικού δικτύου. Με αυτόν τον τρόπο η πρωτεΐνη CLIMP-63 καθορίζει το πλάτος των ασκών του ενδοπλασματικού δικτύου (εικόνα 20A). Παράλληλα, η CLIMP-63 συμμετέχει στα αρχικά στάδια του σχηματισμού του συμπλόκου ριβοσώματος-translocase (εικόνα 20B), ενώ έχει δειχτεί ότι αλληλεπιδρά και με την πρωτεΐνη Dicer. Αντίθετα, στην πλασματική μεμβράνη ο ρόλος της πρωτεΐνης είναι διαφορετικός, καθώς λειτουργεί σαν υποδοχέας διαφορετικών μορίων κάποια από τα οποία φαίνονται στην εικόνα 20D,E,F,G (του tPA - tissue plasminogen activator - που αυξάνει την ενεργοποίηση του πλασμινογόνου στα λεία μυϊκά κύτταρα των αγγείων του ανθρώπου, της SP-A 9 (surfactant protein A) σε πνευμονοκύτταρα αρουραίου τύπου II και σε ανθρώπινα κύτταρα A549, του υποδοχέα EGF-R, παρεμποδίζοντας την ανάπτυξη και τις μεταστάσεις ηπατοκυτταρικού καρκινώματος, του παράγοντα APF). Τέλος, η CLIMP-63 φαίνεται να έχει και τη δυνατότητα μετακίνησης στον πυρήνα όπου δεσμεύεται στον εγγύτερο υποκινητή του γονιδίου *CCN2*, που δρα στον πολλαπλασιασμό, την προσκόλληση, τη μετανάστευση, τη διαφοροποίηση και την επιβίωση των κυττάρων (εικόνα 20H) (Sandoz and van der Goot, 2015).



Εικόνα 20. Λειτουργίες στις οποίες συμμετέχει η CLIMP-63 στο ευκαρυωτικό κύτταρο καθώς και πρωτεΐνες με τις οποίες αλληλεπιδρά. Α) Η CLIMP-63 μέσω αλληλεπιδράσεων δημιουργεί ολιγομερή τα οποία συμμετέχουν στη διατήρηση του σχήματος (σταθερού πλάτους) των ασκών του ενδοπλασματικού δικτύου. Β) Μέσω αλληλεπιδράσεων με τα ριβοσώματα συμβάλλει στον εντοπισμό τους στο αδρό ενδοπλασματικό δίκτυο. Γ) Η μη-φωσφορυλιωμένη μορφή της CLIMP-63 αλληλεπιδρά με τους μικροσωληνίσκους. Στην πλασματική μεμβράνη η CLIMP-63 είναι υποδοχέας Δ) της tPA, Ε) της SP-A στα πνευμονοκύτταρα τύπου II και F) αλληλεπιδρά με τον υποδοχέα EGFR, όταν αυτός δεν είναι προσδεμένος με τον προσδέτη EGF. Γ) Επιπλέον, στην πλασματική μεμβράνη λειτουργεί ως υποδοχέας για τον παράγοντα APF, μετά την πρόσδεση του οποίου μπορεί να υποστεί αποπαλμιτοϋλίωση, φωσφορυλίωση και μεταφορά στον πυρήνα. Η) Στον πυρήνα η CLIMP-63 λειτουργεί ως μεταγραφικός παράγοντας του γονιδίου *CCN2* (Sandoz and van der Goot, 2015).

1.3.12 Kinectin

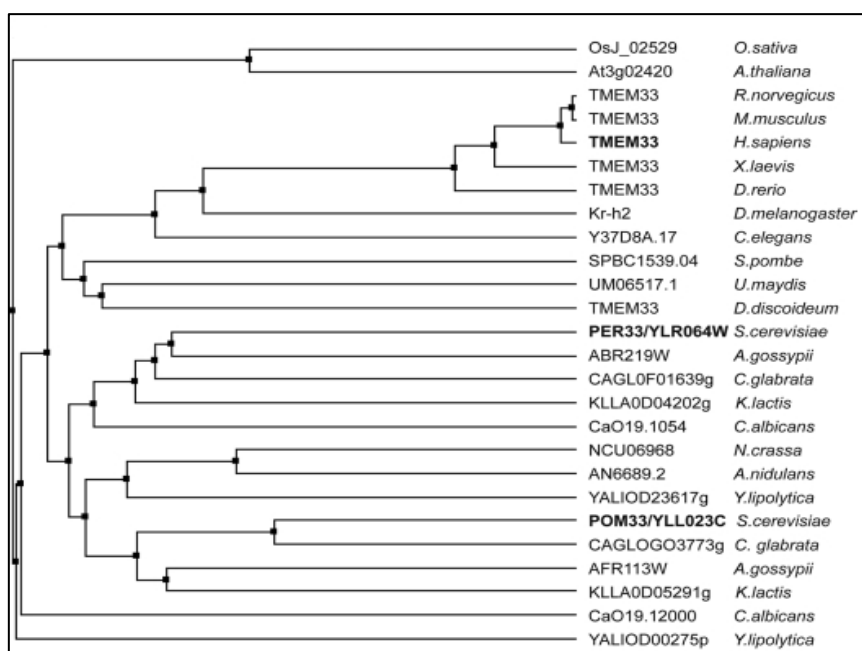
Η kinectin είναι μια μεμβρανική πρωτεΐνη η οποία ανακαλύφθηκε αρχικά ως υποδοχέας της πρωτεΐνης κινεσίνη-1 των μικροσωληνίσκων (Toyoshima et al., 1992). Έχει βρεθεί ότι διακοπή αυτής της αλληλεπιδράσης είναι δυνατό να περιορίσει τη δυνατότητα μετανάστευσης των κυττάρων (Ng et al., 2016).

Η kinectin είναι μια πρωτεΐνη μήκους 1357 αμινοξέων η οποία αλληλεπιδρά με την GTPάση Rho και διαμεσολαβεί στην πρόσδεση του συμπλόκου της μετάφρασης EF1 (elongation factor-1) στο RNA. Η kinectin εντοπίζεται στις μεμβράνες του ενδοπλασματικού δικτύου και συγκεκριμένα στους πεπλατυσμένους ασκούς αυτού μαζί με την CLIMP-63 και την p180. Μέχρι σήμερα έχουν ταυτοποιηθεί τουλάχιστον 15 ισομορφές της πρωτεΐνης αυτής, οι οποίες παρουσιάζουν διαφορές ως προς τον κυτταρικό τύπο και το αναπτυξιακό στάδιο στο οποίο εκφράζονται. Μια από αυτές τις ισομορφές, μεγέθους 120kDa, σχετίζεται με το μιτοχόνδριο και μέσω της αλληλεπιδράσης της με την κινεσίνη επηρεάζει την κινητική των οργανιδίων αυτών (Schönthal, 2012).

Η αμινοξική ακολουθία της kinectin αποτελείται από ένα βραχύ καρβοξυτελικό άκρο, το οποίο εντοπίζεται στον αυλό του ενδοπλασματικού δικτύου, μια διαμεμβρανική περιοχή και ένα κυτταροπλασματικό αμινοτελικό άκρο, το οποίο έχει δομή σπειροειδούς έλικας. Η πρωτεΐνη αυτή θεωρείται ότι σχηματίζει διμερή μέσω αλληλεπιδράσεων στο αμινοτελικό της άκρο (Schönthal, 2012).

1.3.13 Οι πρωτεΐνες TMEM33

Η οικογένεια των πρωτεϊνών TMEM33 είναι συντηρημένη από τους μύκητες μέχρι τα θηλαστικά όπως φαίνεται και στην εικόνα 21. Χαρακτηριστικό των πρωτεϊνών αυτών αποτελούν οι διαμεμβρανικές επικράτειες τις οποίες διαθέτουν. Στον μύκητα *Saccharomyces cerevisiae*, υπάρχουν δύο μέλη της οικογένειας TMEM33, οι πρωτεΐνες Pom33 and Per33, ενώ στον *Schizosaccharomyces pombe* υπάρχει μόνο ένα, η πρωτεΐνη Tts1p (tetra-spanning protein 1). Η πρωτεΐνη Pom33p εντοπίζεται στους πυρηνικούς πόρους και ρυθμίζει την κατανομή των συμπλόκων των πυρηνικών πόρων στην πυρηνική μεμβράνη. Επιπλέον, η πρωτεΐνη αυτή αλληλεπιδρά με την Rtn1p, που ανήκει στην πρωτεϊνική οικογένεια των reticulons. Αντίθετα, η Per33 εντοπίζεται στη μεμβράνη του ενδοπλασματικού δικτύου αλλά η λειτουργία της δεν είναι γνωστή. Η Tts1p εντοπίζεται τόσο στην πυρηνική μεμβράνη, την οποία και αναδιαμορφώνει στη διάρκεια της μίτωσης, όσο και στη μεμβράνη του περιφερειακού ενδοπλασματικού δικτύου. Η παρουσία της πρωτεΐνης αυτής στο ενδοπλασματικό δίκτυο σε συνδυασμό με τις πρωτεΐνες Rtn1p και Yor1p συμβάλλει στην διαμόρφωση των αγωγών του ενδοπλασματικού δικτύου καθώς και στον καθορισμό του επιπέδου της κυτταρικής διαίρεσης. Σημαντικό ρόλο σε όλες τις λειτουργίες της Tts1p διαδραματίζει μια αμφιπαθής έλικα η οποία εντοπίζεται στο καρβοξυτελικό άκρο της πρωτεΐνης (Urade et al., 2014; Zhang and Oliferenko, 2014).



Εικόνα 21. Φυλογενετικό δέντρο στο οποίο φαίνεται η παρουσία της TMEM33 σε θηλαστικά και μύκητες. Παρατηρούμε ότι σε αντίθεση με τον κλάδο των *Saccharomycotina*, σε όλα τα υπόλοιπα είδη υπάρχει από ένα ορθόλογο γονίδιο (Chadrin et al., 2010).

Μια πρόσφατη μελέτη απέδειξε ότι η ανθρώπινη TMEM33 αλληλεπιδρά με τις πρωτεΐνες RTN1A,-2B,-3C,4C καθώς και την Arl6IP1. Η πρωτεΐνη αυτή εντοπίζεται τόσο στην πυρηνική μεμβράνη όσο και στη μεμβράνη του ενδοπλασματικού δικτύου, όταν εκφράζεται εξωγενώς σε ανθρώπινα καρκινικά κύτταρα. Έχει προταθεί ότι η ανθρώπινη TMEM33 μπορεί να καταστείλει τη λειτουργία των reticulons στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου (Urade et al., 2014). Επιπλέον, το ανθρώπινο γονίδιο που κωδικοποιεί την TMEM33 είναι ένα από τα γονίδια τα οποία πιθανά συμμετέχουν στη διαμόρφωση του χρώματος του δέρματος, σε πληθυσμούς που προέρχονται από την ανατολική Ασία. Η ανθρώπινη πρωτεΐνη TMEM33 συμμετέχει σε διάφορα αναπτυξιακά στάδια της μελανογένεσης και αποτελεί καθοριστικό παράγοντα στο στρες του ενδοπλασματικού δικτύου σε καρκινικά κύτταρα (Sakabe et al., 2015; Urade et al., 2014). Παρόλα αυτά οι βιολογικές λειτουργίες των πρωτεϊνών αυτών δεν έχουν προσδιοριστεί πλήρως ακόμα.

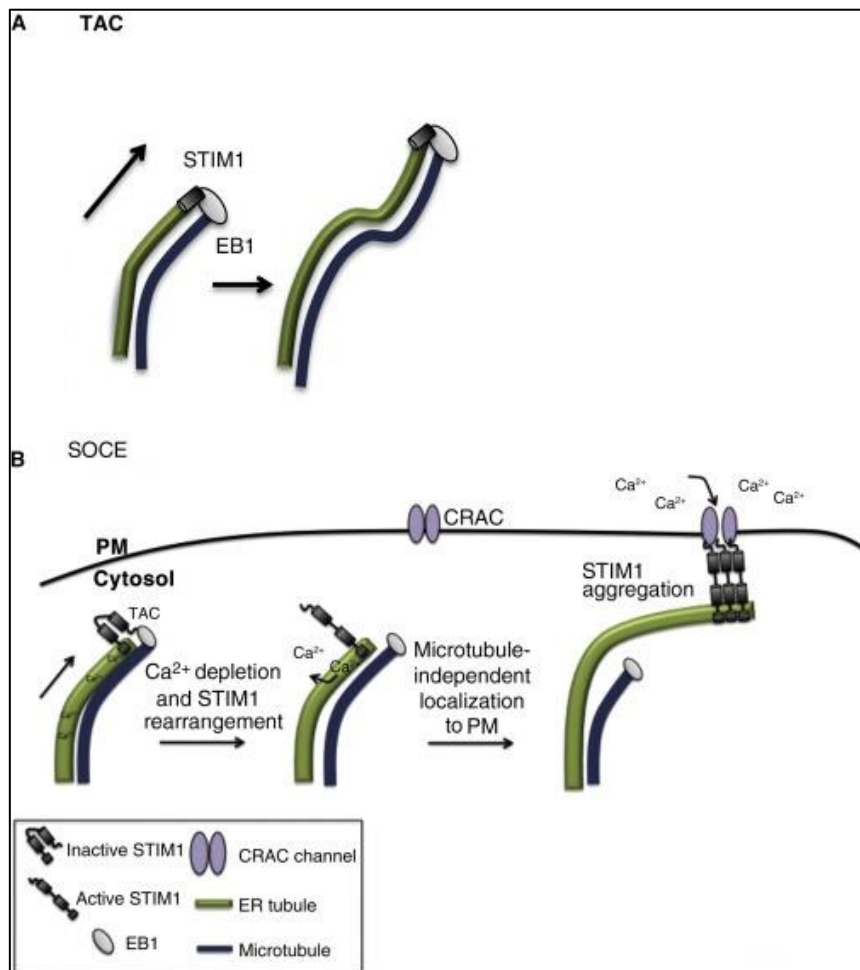
Από όλα τα παραπάνω παρατηρούμε ότι όλα τα μέλη της πρωτεϊνικής οικογένειας της TMEM33 τα οποία έχουν μελετηθεί, εμφανίζουν κοινή τοπολογία στο κύτταρο. Πιο συγκεκριμένα, τόσο η ανθρώπινη TMEM33, όσο και οι πρωτεΐνες Pom33p, Per33 και Tts1p είναι μεμβρανικές και εντοπίζονται τόσο στο ενδοπλασματικό δίκτυο, όσο και στον πυρηνικό φάκελο. Το γεγονός αυτό υποδηλώνει πιθανά ότι αυτές οι πρωτεΐνες πραγματοποιούν κοινές λειτουργίες και συμμετέχουν στη διαμόρφωση της δομής των μεμβρανών αυτών, χωρίς αυτές οι λειτουργίες να έχουν διευκρινιστεί και περιγραφεί πλήρως.

Επιπλέον, ένα ακόμη ζήτημα το οποίο χρήζει περαιτέρω διερεύνησης αφορά στην κατανομή των πρωτεϊνών της οικογένειας TMEM33 στους ευκαρυωτικούς οργανισμούς. Όπως προαναφέρθηκε έχει ήδη πραγματοποιηθεί μια φυλογενετική μελέτη σε αυτήν την κατεύθυνση τα αποτελέσματα της οποίας φαίνονται στην εικόνα 13. Στη μελέτη αυτή όμως ελέγχθηκε η ύπαρξη ορθόλογων γονιδίων μόνο στους μύκητες και κάποια ζώα.

1.3.14 STIM1

Η STIM1 (Stromal interaction molecule 1) είναι μια διαμεμβρανική πρωτεΐνη τύπου II που εντοπίζεται κατά κύριο λόγο στο ενδοπλασματικό δίκτυο και σε ένα μικρότερο ποσοστό της τάξης του 20% στην πλασματική μεμβράνη (Pascual-Caro et al., 2018). Η πρωτεΐνη αυτή περιέχει στο αμινοτελικό της άκρο ένα μοτίβο EF-hand, που εντοπίζεται στον αυλό του ενδοπλασματικού δικτύου για την πρόσδεση ιόντων Ca^{2+} . Επιπλέον περιλαμβάνει μια διαμεμβρανική περιοχή και μια περιοχή πρόσδεσης στους μικροσωληνίσκους στο καρβοξυτελικό της άκρο (MT-binding domain). Η πρωτεΐνη STIM1 συμμετέχει τόσο στη σηματοδότηση μέσω Ca^{2+} στο ενδοπλασματικό δίκτυο, όσο και στη μετακίνηση του σωληνοειδούς ενδοπλασματικού δικτύου μέσω του μηχανισμού TAC (Tip Attachment

Complex) (English et al., 2009) (εικόνα 22). Πιο συγκεκριμένα, απουσία ή μείωση του Ca^{2+} από το ενδοπλασματικό δίκτυο οδηγεί στην αποσύνδεση των ιόντων αυτών από το μοτίβο EF-hand της STIM1, με αποτέλεσμα την αλλαγή της δομής και τον ολιγομερισμό της πρωτεΐνης αυτής και τη συσσώρευση της στα σημεία επαφής του ενδοπλασματικού δικτύου με την πλασματική μεμβράνη όπου ενεργοποιεί τα κανάλια εισαγωγής ασβεστίου στο κύτταρο (SOC- store-operated Ca^{2+} channels) (Kurosaki and Baba, 2010). Ανάμεσα σε αυτά τα κανάλια συμπεριλαμβάνονται και τα κανάλια Orai, τα οποία μαζί με την STIM1 ρυθμίζουν μια σειρά από κυτταρικές λειτουργίες. Επιπλέον, η STIM1 συμμετέχει και στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου, αλληλεπιδρώντας με την πρωτεΐνη EB1 (end binding 1) που βρίσκεται στο + άκρο των μικροσωληνίσκων. Φωσφορυλίωση της STIM1 σε 3 περιοχές-στόχους των πρωτεϊνών ERK1/2 (Ser575, Ser608, and Ser621) οδηγεί στην αποσύνδεση της από την EB1 στη διάρκεια της μίτωσης και συνεπώς στον αποκλεισμό των ασκών του ενδοπλασματικού δικτύου από την μιτωτική άτρακτο (Goyal and Blackstone, 2013; Pascual-Caro et al., 2018).



Εικόνα 22. Α) Η πρωτεΐνη STIM-1 συμμετέχει στον μηχανισμό TAC (Tip Attachment Complex). Πιο συγκεκριμένα, αλληλεπιδρά με την πρωτεΐνη EB1, που βρίσκεται προσδεσμένη στο + άκρο των μικροσωληνίσκων, με αποτέλεσμα την μετακίνηση του ενδοπλασματικού δικτύου παράλληλα με την αύξηση του μήκους των μικροσωληνίσκων. Β) Η πρωτεΐνη STIM-1 αποτελεί αισθητήρα ασβεστίου στο ενδοπλασματικό δίκτυο, μέσω των μοτίβων EF hand που βρίσκονται στον αυλό του

ΕΔ. Η απουσία ή μείωση της ποσότητας του ασβεστίου στο ενδοπλασματικό δίκτυο οδηγεί στην αλλαγή της διαμόρφωσης της STIM-1, η οποία δημιουργεί συσσωματώματα και μέσω του μηχανισμού TAC έρχεται σε επαφή με κανάλια ασβεστίου στην πλασματική μεμβράνη και προωθεί την εισαγωγή ασβεστίου στο κύτταρο, μέσω ενός μηχανισμού που ονομάζεται SOCE (Store Operated Calcium Entry) (Gurel et al., 2014).

1.4 Άλλες πρωτεΐνες που αλληλεπιδρούν με τα reticulons, ή εντοπίζονται στο ενδοπλασματικό δίκτυο

1.4.1 TBC1D20

Η πρωτεΐνη TBC1D20 (Rab1 GAP Tbc domain family member 20) είναι μια μεμβρανική πρωτεΐνη τύπου II του ενδοπλασματικού δικτύου, της οποίας η επικράτεια TBC εντοπίζεται στο κυτταρόπλασμα. Η πρωτεΐνη αυτή ανήκει στη μεγάλη οικογένεια των πρωτεϊνών GAP (GTPase- activating protein), οι οποίες περιέχουν μια εξελικτικά πολύ συντηρημένη επικράτεια την TBC (Tre2, Bub2, and Cdc16). Οι πρωτεΐνες GAP ενισχύουν την ενδογενή ικανότητα των GTPασών Rab να υδρολύουν το GTP, όταν βρίσκονται προσδεσμένες σε αυτό. Η υδρόλυση αυτή έχει σαν αποτέλεσμα την απενεργοποίηση των πρωτεϊνών Rab (Rab-GDP) (Sidjanin et al., 2016). Βιοχημικοί έλεγχοι έχουν δείξει ότι η TBC1D20 λειτουργεί σαν GAP πρωτεΐνη και διαμεσολαβεί την υδρόλυση του GTP από τις πρωτεΐνες Rab1a and Rab2b, με αποτέλεσμα να τις απενεργοποιεί (Haas et al., 2007; Sidjanin et al., 2016). Οι πρωτεΐνες Rab εμπλέκονται στην κυστιδιακή μεταφορά και τη διαμόρφωση του συμπλέγματος Golgi. Επιπλέον, η πρωτεΐνη TBC1D20 έχει βρεθεί να αλληλεπιδρά με την RTN1 στο ενδοπλασματικό δίκτυο (Haas et al., 2007).

Η πρωτεΐνη TBC1D20 παίζει σημαντικό ρόλο στην κυστιδιακή μεταφορά, και όταν υπερεκφράζεται μπορεί να οδηγήσει στη διάσπαση του συμπλέγματος Golgi. Επιπλέον, η πρωτεΐνη αυτή παίζει ουσιαστικό ρόλο στην ωρίμανση των αυτοφαγοσωμάτων. Στον άνθρωπο, η απώλεια της λειτουργίας της πρωτεΐνης TBC1D20 προκαλεί το σύνδρομο Warburg Micro 4 (WARBM4), μια αυτοσωμική υπολειπόμενη διαταραχή που χαρακτηρίζεται από ανωμαλίες στα μάτια, τον εγκέφαλο και τα γεννητικά όργανα (Sidjanin et al., 2016). Τέλος, έχει βρεθεί ότι η πρωτεΐνη TBC1D20 είναι απαραίτητη για την αποτελεσματική αντιγραφή του ιού της ηπατίτιδας C στα ανθρώπινα κύτταρα, μέσω της αλληλεπίδρασης της με την πρωτεΐνη NS5A του ιού (Sklan et al., 2007).

1.4.2 Yip3p

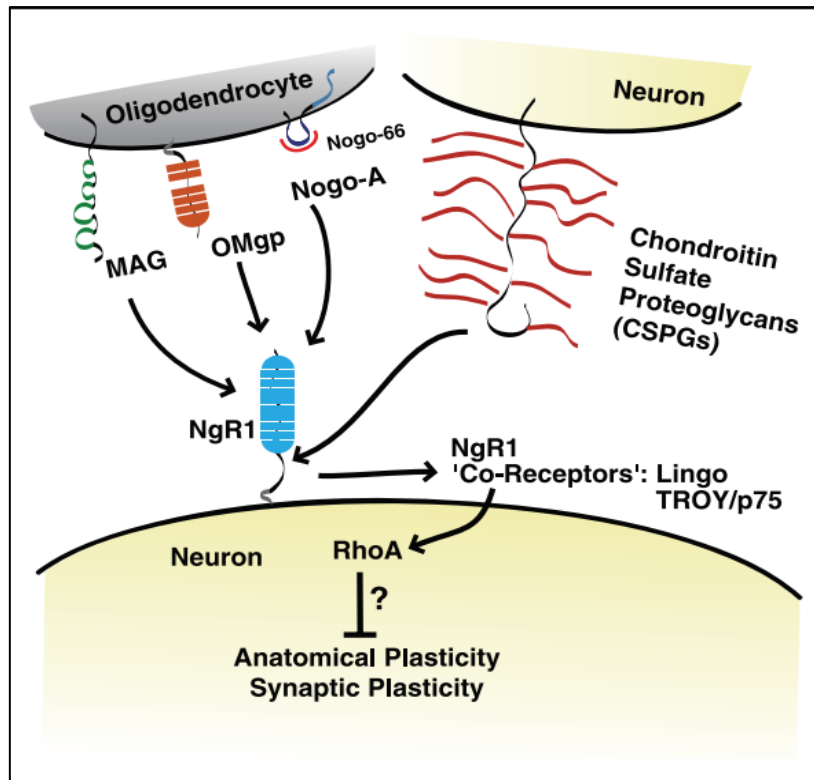
Η Yip3p είναι μια πρωτεΐνη του σακχαρομύκητα *Saccharomyces cerevisiae*, που περιέχει δύο εκτενείς υδρόφοβες επικράτειες και ένα καρβοξυτελικό άκρο με όξινα αμινοξικά κατάλοιπα. Τα χαρακτηριστικά αυτά είναι κοινά με τον παράγοντα PRA1, ορθόλογη πρωτεΐνη της Yip3p στα θηλαστικά (Martincic et al., 1997; Yang and Strittmatter, 2007). Ο PRA1 είναι μια πρωτεΐνη GDF (Rab-GDI displacement factor), που σημαίνει ότι συμβάλλει στο σωστό υποκυτταρικό εντοπισμό των πρωτεϊνών Rab, συμμετέχοντας στη διάσπαση του συμπλόκου των πρωτεϊνών Rab με τους παράγοντες GDI (Rab GDP dissociation inhibitor) (Geng et al., 2005). Η πρωτεΐνη Yip3p εντοπίζεται κυρίως στις μεμβράνες του συμπλέγματος Golgi και δευτερευόντως στις

μεμβράνες του ενδοπλασματικού δικτύου, αλλά έχει βρεθεί και στα κυστίδια COPII (Geng et al., 2005; Otte et al., 2001). Βασικός της ρόλος είναι η συμμετοχή της στην κυστιδιακή μεταφορά ανάμεσα στο ενδοπλασματικό δίκτυο και το σύμπλεγμα Golgi. Επιπλέον, αλληλεπιδρά με πρωτεΐνες που συμμετέχουν στην κυστιδιακή μεταφορά όπως η πρωτεΐνη Ypt1p, μέλος της οικογένειας των πρωτεϊνών Rab. Τέλος, έχει βρεθεί ότι η πρωτεΐνη Yip3p αλληλεπιδρά με την πρωτεΐνη Rtn1p (ορθόλογη των reticulons στο σακχαρομύκητα), ενισχύοντας το ρόλο των reticulons στην κυστιδιακή μεταφορά (Geng et al., 2005).

1.4.3 Υποδοχέας του RTN4A/Nogo A (NogoR)

Η πρωτεΐνη NogoR (Nogo receptor) αναγνωρίστηκε αρχικά ως υποδοχέας της πρωτεΐνης NogoA (ή αλλιώς RTN4A), ο οποίος εντοπίζεται στην πλασματική μεμβράνη των νευρώνων (Yang and Strittmatter, 2007). Σήμερα είναι γνωστό ότι ο υποδοχέας αυτός διαθέτει και άλλους προσδέτες, όπως οι πρωτεΐνες Nogo-B, Nogo-C καθώς και οι MAG (Myelin-associated glycoprotein) και OMgp (oligodendrocyte-myelin glycoprotein) (Seiler et al., 2015). Πρόκειται για μια πρωτεΐνη που εμφανίζει εξειδικευμένη έκφραση στον εγκέφαλο και είναι πλούσια σε επαναλήψεις λευκίνης (Otte et al., 2001). Ο υποδοχέας NogoR δεν διαθέτει διαμεμβρανικό τμήμα, αλλά προσδένεται στη μεμβράνη των νευρώνων μέσω μιας γέφυρας GPI (γλυκοζυλοφωσφατιδυλινοσιτόλη).

Όπως προαναφέρθηκε (στην ενότητα 1.3.1) η πρόσδεση της πρωτεΐνης RTN4A στον υποδοχέα NogoR μέσω της επικράτειας Nogo-66, μπορεί να οδηγήσει στην αναστολή της αύξησης των νευρώνων, μετά από τραυματισμό στο κεντρικό νευρικό σύστημα (κατάρρευση του αυξητικού κώνου των νευρώνων) (Yang and Strittmatter, 2007). Ο μηχανισμός δράσης του υποδοχέα NogoR που οδηγεί στην αναστολή της νευρωνικής αύξησης είναι γνωστός. Όπως φαίνεται και στην εικόνα 23, η πρωτεΐνη NogoA, μέσω της επικράτειας Nogo-66 προσδένεται στον υποδοχέα NogoR, ο οποίος αλληλεπιδρά με μια σειρά από άλλους υποδοχείς καθοδικά από αυτόν. Πιο συγκεκριμένα, ο NogoR σχηματίζει ένα σύμπλοκο με τις πρωτεΐνες: LINGO-1 (Leucine-rich repeat and immunoglobulin domain-containing 1), TROY (Tumor necrosis factor receptor superfamily member 19), και τον υποδοχέα p75. Το σύμπλοκο αυτό ενεργοποιεί το μονοπάτι Rho / ROCK που οδηγεί στην κατάρρευση του αυξητικού κώνου (Seiler et al., 2015; Xu et al., 2015).



Εικόνα 23. Ο υποδοχέας NogoR και οι πρωτεΐνες με τις οποίες αλληλεπιδρά. Οι πρωτεΐνες MAG (Myelin-associated glycoprotein), OMgp (oligodendrocyte-myelin glycoprotein), και η πρωτεΐνη Nogo-A μέσω της επικράτειας της Nogo-66 αποτελούν προσδέτες του υποδοχέα NogoR (NgR1). Καθεμία από αυτές τις πρωτεΐνες προσδέεται σε μια επικράτεια πλούσια σε λευκίνες που διαθέτει ο υποδοχέας NogoR. Ο υποδοχέας αυτός με τη σειρά του αλληλεπιδρά με άλλους υποδοχείς, όπως οι Lingo, TROY και p75, ώστε να ενεργοποιήσει τη μικρή GTPάση RhoA. Επιπλέον, ορισμένα μέλη της οικογένειας των πρωτεογλυκανών CSPGs (chondroitin sulfate proteoglycans) διαθέτουν την ικανότητα πρόσδεσης στον υποδοχέα NogoR (Stephany et al., 2016).

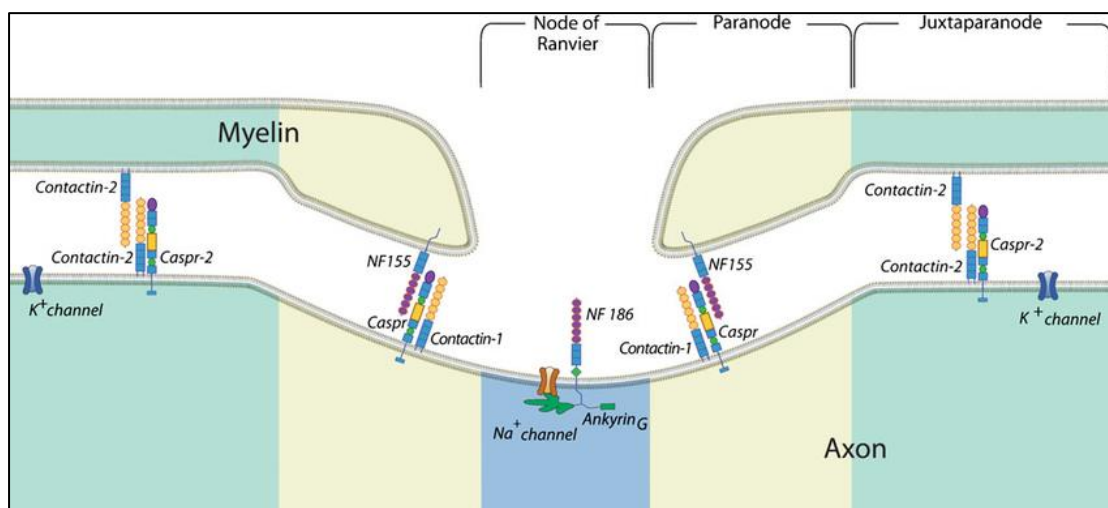
Ο υποδοχέας NogoR, μέσα από την αλληλεπίδραση του με την πρωτεΐνη NogoA έχει συσχετιστεί με μια σειρά από ασθένειες και παθολογικές καταστάσεις του νευρικού συστήματος, όπως η νόσος του Alzheimer, η νόσος του Parkinson η σχιζοφρένεια, η σκλήρυνση κατά πλάκας, η αμυοτροφική πλευρική σκλήρυνση, το γλαύκωμα, το γλοίωμα, το ισχαιμικό εγκεφαλικό επεισόδιο καθώς και η μείωση της μνήμης με την ηλικία. Επιπλέον, άλλες λειτουργίες στις οποίες θεωρείται ότι συμμετέχει ο υποδοχέας NogoR είναι η ρύθμιση των βλαστοκυττάρων του νευρικού συστήματος, η ρύθμιση της δραστηριότητας της μικρογλοίας, η αναστολή της αγγειογένεσης και η παρεμβολή στο σχηματισμό της μνήμης (Seiler et al., 2015).

1.4.4 Caspr1

Η πρωτεΐνη Caspr1 (contactin-associated protein) ανήκει σε μια οικογένεια διαμεμβρανικών πρωτεϊνών που ενώ έχουν παρόμοια δομή, καθεμία εμφανίζει εξειδικευμένες λειτουργίες. Στον άνθρωπο οι πρωτεΐνες Caspr1–5 παίζουν σημαντικό ρόλο στη διαμόρφωση και σταθερότητα των εμμύλων νευραξόνων. Η Caspr1 είναι το πρώτο μέλος της οικογένειας το οποίο ανακαλύφθηκε λόγω της αλληλεπίδρασης της με το πρωτεϊνικό σύμπλοκο contactin-PTPRZ1. Η πρωτεΐνη αυτή εντοπίζεται στους νευράξονες

και συγκεκριμένα στις άκρες των κόμβων του Ranvier (paranodes). Ως μέλος της υπερικογένειας των πρωτεϊνών neurexin, η Caspr1 αποτελείται από μια μεγάλη εξωμεμβρανική επικράτεια, μια απλή διαμεμβρανική περιοχή και μια μικρή ενδοκυτταρική επικράτεια (Zou et al., 2017).

Η πρωτεΐνη Caspr1 σχηματίζει σύμπλοκο με την contactin-1 και την NF-155 (neurofascin-155) (εικόνα 24), διαχωρίζοντας τους κόμβους του Ranvier από τις εμμέλες περιοχές των νευραξόνων ενώ παράλληλα συμμετέχει στη μετάδοση του δυναμικού δράσης καθώς και στη μεταφορά του σήματος. Η αλληλεπίδραση της με την πρωτεΐνη NogoA (RTN4A) είναι σημαντική για τον εντοπισμό των καναλιών ιόντων K^+ στις άκρες των κόμβων του Ranvier (paranodes) (Yang and Strittmatter, 2007). Επιπλέον, η Caspr1 συνεισφέρει στη διατήρηση της σταθερότητας της μυελίνης που περιβάλλει τους νευράξονες και παίζει καθοριστικό ρόλο στην ανάπτυξη του φλοιού του εγκεφάλου. Πρόσφατες μελέτες έχουν δείξει ότι η Caspr1 μπορεί να παίζει ρόλο στον καθορισμό του χρονισμού της διαφοροποίησης των νευρικών προγονικών κυττάρων σε νευρώνες και αστροκύτταρα. Επιπλέον, η πρωτεΐνη Caspr1 έχει συσχετιστεί με νευροεκφυλιστικές ασθένειες όπως η νόσος του Alzheimer και η σκλήρυνση κατά πλάκας (Zou et al., 2017).



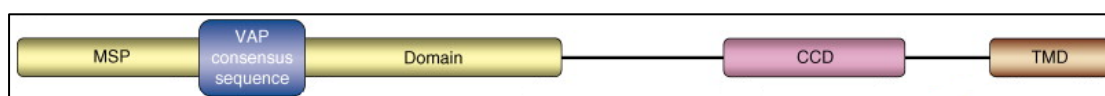
Εικόνα 24. Δομή περιοχών γύρω από τον κόμβο του Ranvier. Στους κόμβους του Ranvier, χάρη στην πρωτεΐνη neurofascin 186 (NF186), συγκεντρώνονται κανάλια ιόντων Na^+ τα οποία είναι απαραίτητα για την ιοντική αγωγιμότητα των νευραξόνων. Η αλληλεπίδραση ανάμεσα στα κανάλια Na^+ και την NF186 μεσολαβείται από την πρωτεΐνη αγκυρίνη G. Στο τμήμα του νευράξονα που ονομάζεται paranode, η μυελίνη συνδέεται στενά με τον νευράξονα σύνδεση που μεσολαβείται από το πρωτεϊνικό σύμπλοκο των Contactin-1, Caspr, NF155. Στο τμήμα του νευράξονα που ονομάζεται juxtaparanode εντοπίζονται κανάλια K^+ , στενά συνδεδεμένα με την πρωτεΐνη Caspr2. Σημαντικό ρόλο στη θέση αυτών των πρωτεϊνών παίζει η πρωτεΐνη contactin-2. Πρωτεϊνικά μόρια contactin-2 εντοπίζονται τόσο στον νευράξονα, όσο και στην πλευρά της μυελίνης και αλληλεπιδρούν μεταξύ τους (Meinl et al., 2010).

1.4.5 Πρωτεΐνες VAP: VAP-A και VAP-B

Οι VAPs (VAMP-associated proteins) είναι διαμεμβρανικές πρωτεΐνες τύπου II του ενδοπλασματικού δικτύου, μικρού μήκους, οι οποίες εκφράζονται σε όλους τους ευκαρυώτες (Lev et al., 2008; Murphy and Levine, 2016). Πήραν το όνομα τους από την αλληλεπίδραση τους με την πρωτεΐνη VAMP, μια SNARE απαραίτητη για τη σύντηξη των συναπτικών κυστιδίων στη

μετασυναπτική μεμβράνη. Οι πρωτεΐνες VAP αλληλεπιδρούν με ένα μεγάλο αριθμό πρωτεϊνών (μέχρι σήμερα έχουν προσδιοριστεί 100 πρωτεΐνες (Murphy and Levine, 2016)), που μπορούν να χωριστούν σε 3 μεγάλες κατηγορίες: α) SNAREs, β) ιικές πρωτεΐνες και γ) και πρωτεΐνες που περιέχουν το μοτίβο FFAT (double phenylalanine in an acidic tract). Επιπλέον, οι πρωτεΐνες VAP συμμετέχουν σε μια σειρά από κυτταρικές λειτουργίες, όπως η κυστιδιακή μεταφορά, η μεταφορά και ο μεταβολισμός λιπιδίων, το μονοπάτι UPR (unfolded protein response) του ενδοπλασματικού δικτύου και η οργάνωση των μικροσωληνίσκων. Εκφράζονται σε διαφορετικούς ιστούς και όργανα και εκτός από το ενδοπλασματικό δίκτυο έχουν βρεθεί και σε άλλες ενδοκυτταρικές μεμβράνες, όπως το σύμπλεγμα Golgi, το ERIC (ER–Golgi intermediate compartment), τα ενδοσώματα, τις νευρομυϊκές συνάψεις και την πλασματική μεμβράνη (Lev et al., 2008).

Στα θηλαστικά έχουν βρεθεί δύο πρωτεΐνες VAP, οι VAP-A και VAP-B, οι οποίες εμφανίζουν μεταξύ τους μεγάλο βαθμό ομοιότητας της αλληλουχίας (σε ποσοστό 63%). Επιπλέον, οι πρωτεΐνες αυτές αποτελούνται από 3 συντηρημένες επικράτειες: α) μια β-πτυχωτή επιφάνεια στο αμινοτελικό άκρο, η οποία μοιάζει σε ποσοστό 22% με την επικράτεια MSP (major sperm protein) των νηματωδών, β) μια κεντρική ακολουθία, η οποία σε κάποιους οργανισμούς εμφανίζει μια σπειροειδή έλικα που μοιάζει με την επικράτεια CCD, που εμφανίζεται τόσο στις πρωτεΐνες VAMP όσο και σε άλλες SNARE και γ) μια διαμεμβρανική επικράτεια α-έλικας (TMD) στο καρβοξυτελικό άκρο που περιλαμβάνει το μοτίβο GxxxG διμερισμού (εικόνα 25) (Lev et al., 2008; Murphy and Levine, 2016). Οι VAP-A και VAP-B υφίστανται ομο- και ετεροδιμερισμό, πιθανά μέσω των περιοχών TMD τους, πράγμα που δεν συμβαίνει με την ορθόλογη πρωτεΐνη Scs2p του σακχαρομύκητα, στην οποία δεν υπάρχει η επικράτεια διμερισμού και η CCD. Παρόλες τις διαφορές όμως που υπάρχουν στην δομή των πρωτεϊνών όλες οι VAPs περιλαμβάνουν μια ιδιαίτερα συντηρημένη ακολουθία 16 αμινοξικών καταλοίπων, την VAP συναινετική ακολουθία, στην επικράτεια MSP του αμινοτελικού άκρου (Lev et al., 2008).



Εικόνα 25. Οι επικράτειες των πρωτεϊνών της οικογένειας Var. Όλα τα μέλη της οικογένειας αυτής αποτελούνται από μια διαμεμβρανική επικράτεια στο καρβοξυτελικό άκρο (TMD) τους, η οποία προεκβάλλει στο κυτταρόπλασμα, και μια επικράτεια MSP (major sperm protein domain) στο αμινοτελικό άκρο, στην οποία περιέχεται μια VAP συναινετική ακολουθία 16 αμινοξέων. Στα μεταξύ υπάρχει μια επικράτεια σπειροειδούς έλικας, δίπλα από το διαμεμβρανικό τμήμα, η οποία μοιάζει με την επικράτεια CCD, που εμφανίζεται σε πολλές πρωτεΐνες SNARE (Lev et al., 2008).

1.5 Σκοπός

Σκοπός της παρούσης διπλωματικής εργασίας είναι η ανασύσταση της εξελικτικής ιστορίας των παραπάνω πρωτεϊνών, οι οποίες εμπλέκονται στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου, για τις οποίες είτε δεν έχουν πραγματοποιηθεί αντίστοιχες μελέτες, είτε οι ήδη υπάρχουσες

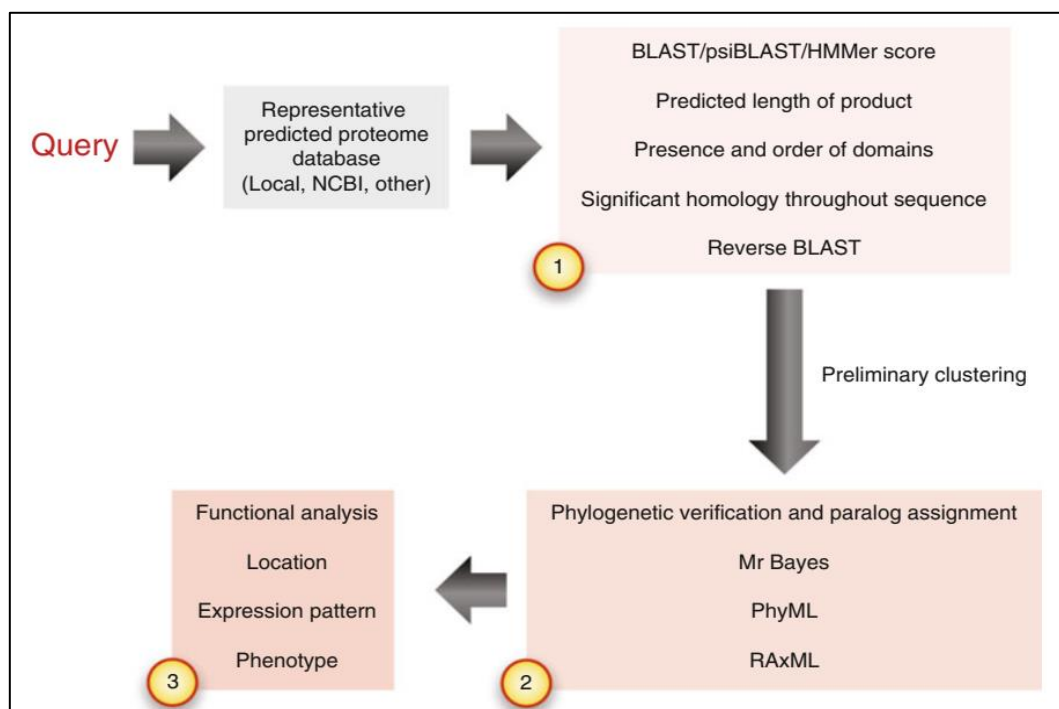
δημοσιευμένες μελέτες είναι παλιές και λιγότερο εκτενείς. Μέσα από την παρούσα μελέτη πρόκειται να προκύψουν κάποια επιπλέον δεδομένα τα οποία θα συμβάλλουν στην ανασύσταση της πολυπλοκότητας του τελευταίου κοινού προγόνου των ευκαρυωτικών οργανισμών (Last Eukaryotic Common Ancestor-LECA).

2 Υλικά και Μέθοδοι

2.1 Γενική πορεία εργασίας

Τα τελευταία χρόνια ο διαρκώς αυξανόμενος αριθμός γονιδιωμάτων τα οποία έχουν πλήρως αλληλουχηθεί από ολόκληρο το φάσμα των ευκαρυωτικών οργανισμών, έχει συμβάλει καθοριστικά στην πραγματοποίηση εκτενών φυλογενετικών αναλύσεων καθώς και στη μελέτη της εξέλιξης των ευκαρυωτικών οργανισμών. Μελέτες που έχουν ήδη πραγματοποιηθεί σε αυτήν την κατεύθυνση υποδεικνύουν την ύπαρξη ενός υψηλού βαθμού πολυπλοκότητας στον τελευταίο κοινό πρόγονο των ευκαρυωτικών οργανισμών (LECA), πολυπλοκότητα η οποία είναι ορατή και σε όλους τους ευκαρυωτικούς οργανισμούς σήμερα.

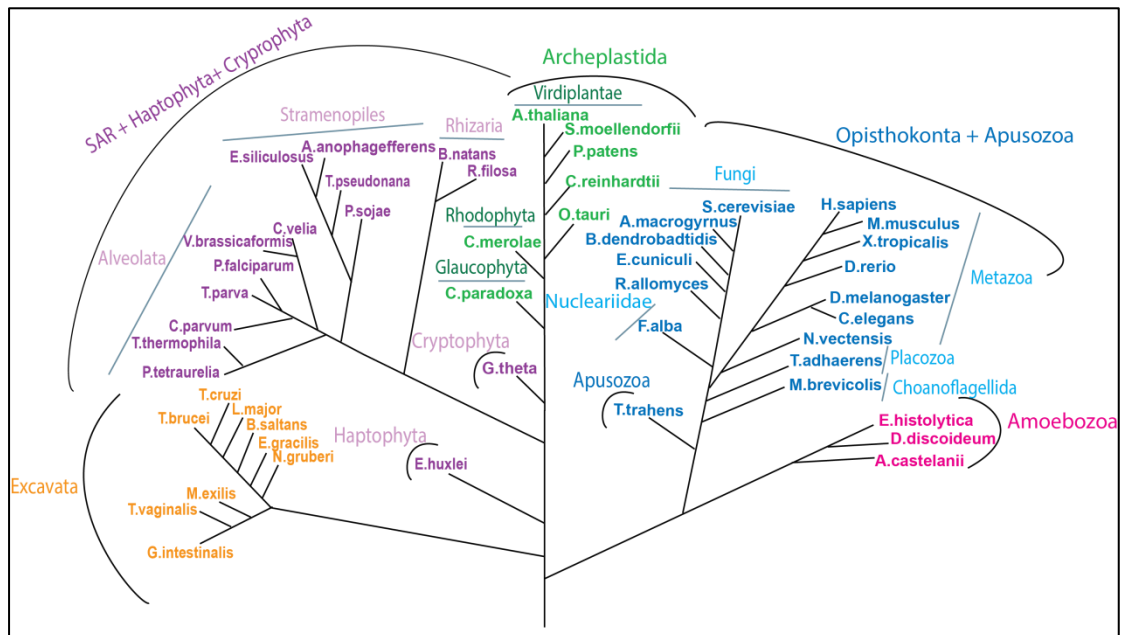
Στην προσπάθεια μελέτης του βαθμού συντήρησης των μηχανισμών που χρησιμοποιούνται για τη διαμόρφωση του ενδοπλασματικού δικτύου στους ευκαρυωτικούς οργανισμούς, πραγματοποιήθηκε γονιδιωματική και φυλογενετική ανάλυση των 23 πρωτεϊνών, οι οποίες περιγράφηκαν στην εισαγωγή, που συμμετέχουν σε αυτές τις λειτουργίες. Η πορεία εργασίας η οποία ακολουθήθηκε περιλαμβάνει μια σειρά από στάδια τα οποία εξασφαλίζουν τη μεγαλύτερη δυνατή ορθότητα και ακρίβεια των αποτελεσμάτων. Η συγκεκριμένη μεθοδολογία έχει χρησιμοποιηθεί και σε προηγούμενες δημοσιευμένες μελέτες, οι οποίες αφορούν την εξελικτική πορεία πρωτεϊνών όπως οι Rabs, πρωτεΐνες κυστιδίων μεταφοράς, ESCRTs και SNAREs (Koumandou and Field, 2011). Το γεγονός αυτό τονίζει την ισχύ της μεθόδου. Τα βασικά βήματα αυτής συνοψίζονται στην εικόνα 26.



Εικόνα 26. Πορεία εργασίας η οποία ακολουθήθηκε για την πραγματοποίηση της συγκριτικής γονιδιωματικής και φυλογενετικής ανάλυσης πρωτεϊνών που συμμετέχουν στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου (Koumandou and Field, 2011).

Στη συγκεκριμένη μελέτη χρησιμοποιήθηκαν συνολικά 23 πρωτεΐνες και 48 επιλεγμένοι οργανισμοί (εικόνα 27), αντιπροσωπευτικοί όλων των υπερομάδων ευκαρυωτικών οργανισμών (Οπισθόκοντα: περιλαμβάνει τους οργανισμούς από τους μύκητες μέχρι και τα ζώα, Amoebozoa, Archaeplastida, Excavata, κλάδος των SAR: περιλαμβάνει Stramenopiles, Alveolates και Rhizaria). Όπου ήταν δυνατό συμπεριλήφθηκαν δύο ή περισσότεροι οργανισμοί από κάθε υπερομάδα οργανισμών έτσι ώστε να είναι δυνατό να διαπιστωθεί κατά πόσο έχει συμβεί δευτερογενής απώλεια κάποιου γονιδίου από ένα μόνο είδος ή πλήρης απουσία γονιδίου από ολόκληρη την υπερομάδα. Με αυτόν τον τρόπο περιορίστηκαν και τα λάθη λόγω ελλείψεων που μπορεί να υπάρχουν στις βάσεις δεδομένων καθώς και λόγω ενδοειδικής ποικιλομορφίας (Koumandou and Field, 2011). Βασικό κριτήριο επιλογής των οργανισμών ήταν να έχουν πλήρως αλληλουχημένο γονιδίωμα. Ως οργανισμοί αναφοράς επιλέχθηκαν οι *Homo sapiens* και *Saccharomyces cerevisiae*, δεδομένου ότι οι περισσότερες μελέτες ως προς τη λειτουργία αυτών των πρωτεϊνών έχουν γίνει σε αυτά τα δύο είδη.

Βάσεις δεδομένων που χρησιμοποιήθηκαν για πρόσβαση στα γονιδιώματα για τις αναλύσεις συγκριτικής γονιδιωματικής είναι αυτές των: NCBI RefSeq (<https://www.ncbi.nlm.nih.gov/refseq/>), Joint Genome Institute (JGI) (http://genome.jgi-psf.org/euk_cur1.html), Broad Institute (<http://www.broadinstitute.org/>), Sanger Institute (<http://www.genedb.org/>), EuPathdb (<http://eupathdb.org/eupathdb/>), καθώς και BLAST servers εξειδικευμένοι για συγκεκριμένους οργανισμούς όπως για τον *C. merolae* (<http://merolae.biol.s.utokyo.ac.jp/blast/blast.html>), τον *T. gondii* (<http://www.toxodb.org/>), τον *C. parvum* (<http://www.cryptodb.org/>), και τη *Giardia intestinalis* (<http://www.giardiadb.org/giardiadb/>).



Εικόνα 27. Δενδρόγραμμα των 50 συνολικά οργανισμών που χρησιμοποιήθηκαν στη συγκεκριμένη μελέτη. Παρατηρούμε ότι αντιπροσωπεύονται όλες οι υπερομάδες των ευκαρυωτών.

Όλες οι αναλύσεις πραγματοποιήθηκαν με τη χρήση πρωτεϊνικών αλληλουχιών, διότι εμφανίζουν μεγαλύτερο βαθμό συντήρησης σε σχέση με το DNA. Αρχικά συγκεντρώθηκε από τη βάση δεδομένων του NCBI από μια πρωτεϊνική ισομορφή για κάθε μια από τις 23 πρωτεΐνες της συγκεκριμένης μελέτης, προερχόμενη από τον έναν ή και από τους δύο οργανισμούς αναφοράς (πίνακας 3). Αυτές οι πρωτεΐνες χρησιμοποιήθηκαν για την πραγματοποίηση εκτενούς αναζήτησης για την ύπαρξη ορθολόγων σε καθέναν από τους υπόλοιπους 48 οργανισμούς. Για το σκοπό αυτό εφαρμόστηκε αρχικά ο αλγόριθμος BLASTp (Altschul et al., 1990), με τον οποίο μπορεί να γίνει σύγκριση μιας ή περισσότερων πρωτεϊνικών ακολουθιών με μία πρωτεϊνική βάση δεδομένων. Η αναζήτηση πραγματοποιήθηκε στις βάσεις δεδομένων Reference Proteins (refseq_protein) ή Non-redundant protein sequences (nr) για κάθε οργανισμό ξεχωριστά με τις εξής παραμέτρους: E-value μικρότερο από 0.1, μέγεθος “λέξεων” (word size) ίσο με 3 και πίνακα αντικατάστασης τον BLOSUM62. Τα αποτελέσματα που προέκυπταν ελέγχθηκαν ως προς τη στατιστική σημαντικότητα (αποτελέσματα με $E\text{-value} \leq e^{-3}$ θεωρούνται σημαντικά), το μέγεθος (έπρεπε να είναι παρόμοιο με το μέγεθος της αρχικής πρωτεΐνης) και την ύπαρξη επικρατειών (θα έπρεπε να υπάρχουν οι ίδιες επικράτειες που εμφανίζονταν και στην αρχική πρωτεΐνη). Αποτελέσματα που πληρούσαν τα παραπάνω κριτήρια γίνονταν αποδεκτά ως υποψήφιες ορθόλογες αλληλουχίες. Επιπλέον, μια ακόμη ένδειξη ότι ένα αποτέλεσμα αποτελεί πραγματικά ορθόλογη πρωτεΐνη αποτελεί η στοίχιση της σε ολόκληρο το μήκος της αρχικής πρωτεΐνης και όχι μόνο σε κάποια επικράτεια. Αυτό συμβαίνει διότι υπάρχουν επικράτειες οι οποίες είναι πολύ κοινές και εμφανίζουν μεγάλο βαθμό συντήρησης ακόμα και μεταξύ μη ορθόλογων πρωτεϊνών. Εάν από την αρχική αναζήτηση προέκυπτε ένα μόνο ξεκάθαρο αποτέλεσμα, όπως για παράδειγμα η περίπτωση όπου το πρώτο hit από το BLASTp είχε πολύ μικρότερο E-value από όλα τα υπόλοιπα, τότε μόνο αυτό εξεταζόταν περαιτέρω. Στην περίπτωση όμως που προέκυπταν πολλαπλά αποτελέσματα με παρόμοιο E-value, τότε πραγματοποιείτο περαιτέρω έλεγχος για όλα ώστε να διαπιστωθεί κατά πόσο αποτελούν παράλογες αλληλουχίες. Σε κάθε περίπτωση ο έλεγχος περιελάμβανε αναλύσεις «αντίστροφου» BLAST έναντι του αρχικού οργανισμού αναφοράς (είτε *Homo sapiens* είτε *Saccharomyces cerevisiae*). Αποδεκτό θεωρείτο ένα αποτέλεσμα από το αρχικό BLAST μόνο εάν έδινε στο «αντίστροφο» BLAST την αρχική πρωτεΐνη στις 5 πρώτες θέσεις των αποτελεσμάτων με $E\text{-value} \leq e^{-3}$.

Πίνακας 3. Οι πρωτεΐνες που χρησιμοποιήθηκαν από τους οργανισμούς αναφοράς *Homo sapiens* και *Saccharomyces cerevisiae*.

| | <i>Homo sapiens</i> | <i>Saccharomyces cerevisiae</i> |
|------------------|---------------------|---------------------------------|
| RTN1 | ✓ | ✓ |
| RTN2 | ✓ | ✓ |
| RTN3 | ✓ | |
| RTN4 | ✓ | |
| REEP5 / Yop1p | ✓ | ✓ |
| REEP1 | ✓ | |
| Atlastin / Sey1p | ✓ | ✓ |
| Ufe1p | | ✓ |
| Use1p | | ✓ |
| Dsl1p | | ✓ |
| Arl6IP1 | ✓ | |
| Protrudin | ✓ | |
| Lunapark | | ✓ |
| Spastin | ✓ | |
| CLIMP-63 | ✓ | |
| Kinectin | ✓ | |
| TMEM33 | ✓ | |
| STIM1 | ✓ | |
| TBC1D20 | ✓ | |
| Yip3p | | ✓ |
| NogoR | ✓ | |
| Caspr | ✓ | |
| VAP-A | ✓ | |
| VAP-B | ✓ | |

Στις περιπτώσεις που οι παραπάνω αναλύσεις δεν έδωσαν αποδεκτό αποτέλεσμα πραγματοποιήθηκε ένας επιπλέον έλεγχος για την ύπαρξη ορθόλογων αλληλουχιών, με τη χρήση του προγράμματος Hmmer (Eddy, 1998), το οποίο εμφανίζει μεγαλύτερη ευαισθησία από το BLAST. Το βασικό μειονέκτημα του προγράμματος αυτού είναι ότι μπορεί να εντοπίζει πρωτεΐνες με πολύ μικρό βαθμό ομοιότητας με την αρχική πρωτεΐνη οι οποίες δεν είναι στην πραγματικότητα συγγενικές εξελικτικά με αυτήν. Στο πρόγραμμα αυτό για κάθε μία από τις πρωτεΐνες της μελέτης, εισήχθησαν όλες οι ομόλογες αλληλουχίες που εντοπίστηκαν μέσω του BLAST, ώστε να κατασκευαστεί ένα πρωτεϊνικό προφίλ, με βάση το οποίο πραγματοποιείται περεταίρω αναζήτηση των πρωτεωμάτων. Όπως και στην περίπτωση του BLAST τα αποτελέσματα ελέγχονταν ως προς το E-value και το μέγεθος, με εκείνα τα οποία περνούσαν τον αρχικό έλεγχο να υποβάλλονται και πάλι σε αναλύσεις «αντίστροφου» BLAST. Με αυτόν τον τρόπο κατέστη δυνατή η ταυτοποίηση κάποιων επιπλέον ορθόλογων αλληλουχιών.

Εδώ θα πρέπει να σημειωθεί ότι όλες οι αναλύσεις με τη χρήση των προγραμμάτων BLAST και HMMer έγιναν με αυτοματοποιημένες μεθόδους σε συνεργασία με το εργαστήριο του κ. Joel Dacks στο Πανεπιστήμιο της Alberta του Καναδά, για εξοικονόμηση χρόνου. Στα πλαίσια της συγκεκριμένης διπλωματικής εργασίας πραγματοποιήθηκε ο έλεγχος αποτελεσμάτων που

προέκυψαν για την εξαγωγή των τελικών συμπερασμάτων ως προς την παρουσία ή όχι ορθόλογων πρωτεϊνών.

Στις περιπτώσεις που δεν υπήρξε θετικό αποτέλεσμα από τις παραπάνω διαδικασίες, θα έπρεπε να πραγματοποιηθούν περαιτέρω έλεγχοι με τη χρήση του αλγορίθμου tBLASTn (Koumandou and Field, 2011). Κάτι τέτοιο όμως δεν πραγματοποιήθηκε στα πλαίσια της συγκεκριμένης διπλωματικής εργασίας.

Μετά από τις αναλύσεις συγκριτικής γονιδιωματικής ακολούθησε φυλογενετική ανάλυση επιλεγμένων πρωτεϊνών. Πιο συγκεκριμένα, για κάθε πρωτεΐνη συγκεντρώθηκαν όλες οι πρωτεϊνικές ακολουθίες των πρωτεϊνών που εντοπίστηκαν μέσω BLAST και Hmmer. Οι αλληλουχίες αυτές στοιχίστηκαν μέσω του προγράμματος πολλαπλών στοιχίσεων MUSCLE, το οποίο είναι διαθέσιμο διαδικτυακά στην ιστοσελίδα του EBI (<https://www.ebi.ac.uk/Tools/msa/muscle/>). Η επεξεργασία των αποτελεσμάτων των στοιχίσεων έγινε μέσω του προγράμματος Mesquite (Maddison and Maddison, 2018) για την αφαίρεση των «κενών» (trimming), που θα μπορούσαν να προκαλέσουν προβλήματα κατά την κατασκευή φυλογενετικών δέντρων. Στη συνέχεια, χρησιμοποιήθηκε το πρόγραμμα ProtTest (Ingley et al., 1999) για τον προσδιορισμό του βέλτιστου μοντέλου νουκλεοτιδικής αντικατάστασης για κάθε πολλαπλή στοιχίση ακολουθιών. Πιο συγκεκριμένα, ο έλεγχος πραγματοποιήθηκε για τα εξής μοντέλα: DAYHOFF, DCMUT, JTT, MTREV, WAG, RTREV, CPREV, VT, BLOSUM62, MTMAM, LG και GTR σε συνδυασμό με την κατανομή GAMMA, τον αριθμό των αμετάβλητων θέσεων και τη συχνότητα εμφάνισης κάθε αμινοξέως. Τα παραπάνω μοντέλα βοηθούν στην εκτίμηση του ρυθμού των μεταλλαγών που έχουν συμβεί στην πραγματικότητα στις αλληλουχίες DNA. Με βάση το μοντέλο που προέκυψε σε κάθε περίπτωση πραγματοποιήθηκε κατασκευή φυλογενετικού δέντρου μέσω του προγράμματος μέγιστης πιθανοφάνειας RaxML (Stamatakis, 2014). Λόγω της μεγάλης υπολογιστικής ισχύος που απαιτείται για την εκτέλεση του προγράμματος αυτού χρησιμοποιήθηκε η διαδικτυακή πλατφόρμα cypress (<https://www.phylo.org/portal2/home.action>) για την πραγματοποίηση των αναλύσεων. Από τα αποτελέσματα της, και μέσω των προγραμμάτων Phylip 3.695 (Baum, 1989) και FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>), είναι δυνατή η οπτικοποίηση ενός τελικού φυλογενετικού δέντρου, του οποίου οι κλάδοι στηρίζονται από το μεγαλύτερο δυνατό αριθμό bootstrap.

2.2 Συγκριτική γονιδιωματική ανάλυση

2.2.1 Αλγόριθμος BLAST

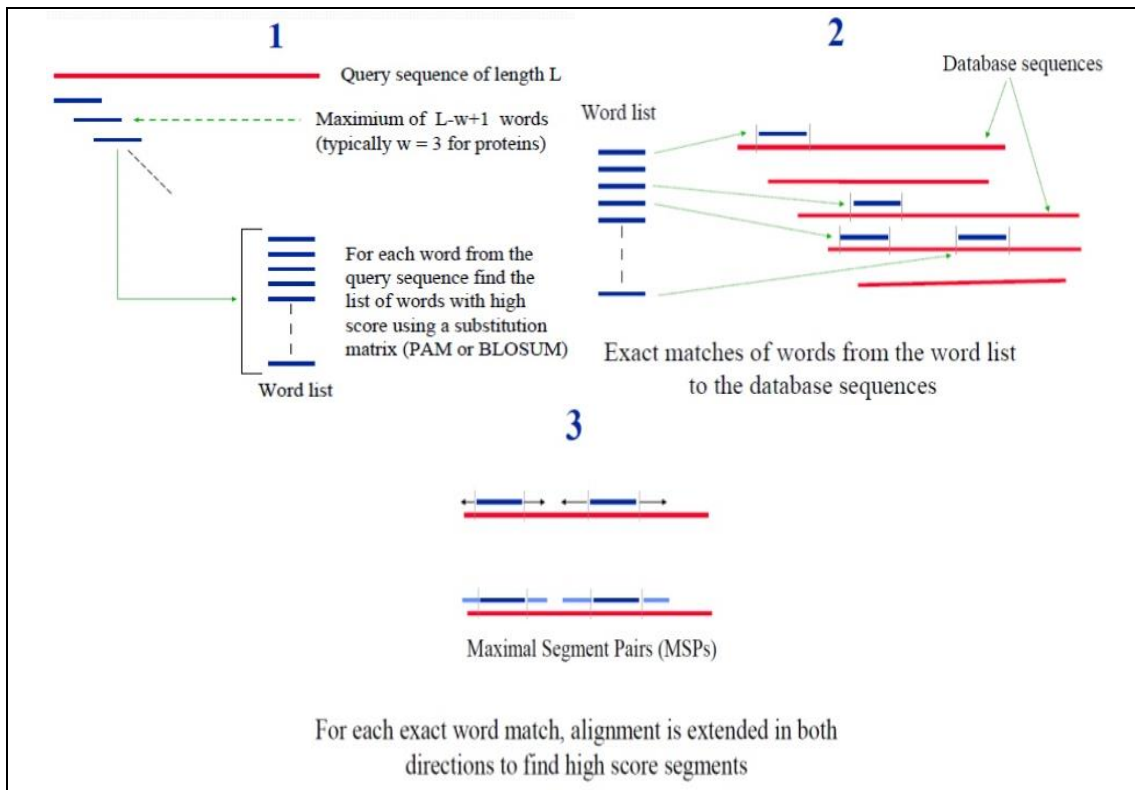
Ένας ευρέως χρησιμοποιούμενος αλγόριθμος στον τομέα της βιοπληροφορικής είναι ο BLAST (Basic Local Alignment Search Tool). Δημιουργήθηκε το 1990 από τους Altschul και συνεργάτες και είναι ελεύθερα προσβάσιμος στο διαδίκτυο μέσω της διεύθυνσης www.ncbi.nlm.nih.gov/BLAST/. Το BLAST αποτελεί έναν ευριστικό αλγόριθμο τοπικών στοιχίσεων για την αναζήτηση ομοιότητας μεταξύ βιολογικών αλληλουχιών. Πιο συγκεκριμένα, μέσω αυτού πραγματοποιείται με μεγάλη

ταχύτητα στοίχιση μιας αλληλουχίας (είτε νουκλεοτιδικής, είτε πρωτεϊνικής) με μία διεθνή βάση δεδομένων (όπως οι EMBL, Genbank, DDBJ κλπ.) καθώς και υπολογισμός της στατιστικής σημαντικότητας αυτής της στοίχισης. Παρόλα αυτά ο αλγόριθμος BLAST δεν εγγυάται τη βέλτιστη λύση, όπως οι αλγόριθμοι δυναμικού προγραμματισμού.

Ο αλγόριθμος BLAST είναι χρήσιμος σε πολλές περιπτώσεις όπως: α) η ταυτοποίηση μιας αλληλουχίας και κατ' επέκταση η ταυτοποίηση ενός είδους ή ομολόγων μεταξύ τους ειδών, β) ο εντοπισμός ομολόγων γονιδίων σε διαφορετικούς οργανισμούς, γ) η ταυτοποίηση μελών γονιδιακών οικογενειών, δ) η χαρτογράφηση γονιδίων μέσω σύγκρισης μιας γονιδιακής περιοχής ενδιαφέροντος με μια βάση δεδομένων, ε) ο εντοπισμός δομικών επικρατειών που είναι δυνατό να υπάρχουν σε μια αλληλουχία.

Μέσω του αλγορίθμου BLAST πραγματοποιείται σύγκριση της πρωτοταγούς δομής μιας νουκλεοτιδικής ή και πρωτεϊνικής ακολουθίας, με τις κατατεθειμένες αλληλουχίες της εκάστοτε βάσης δεδομένων. Πιο συγκεκριμένα, η αρχική αλληλουχία διασπάται σε μικρότερα τμήματα (words) καθορισμένου μήκους, τα οποία συγκρίνονται με αντίστοιχα τμήματα από όλες τις αλληλουχίες που βρίσκονται κατατεθειμένες σε μια ή περισσότερες βάσεις δεδομένων (εικόνα 28). Οι στοιχίσεις μεταξύ των τμημάτων βαθμολογούνται με βάση πίνακες αντικατάστασης. Όταν η βαθμολογία μιας στοίχισης τμημάτων ξεπερνάει κάποιο κατώφλι, τότε η στοίχιση επεκτείνεται και προς τις δύο κατευθύνσεις της αρχικής αλληλουχίας (εικόνα 28). Η παραπάνω διαδικασία έχει ως αποτέλεσμα την πραγματοποίηση τοπικών στοιχίσεων οι οποίες βαθμολογούνται. Μόνο οι στοιχίσεις οι οποίες έχουν την υψηλότερη βαθμολογία συσχέτισης (HSPs - high scoring sequence pairs) λαμβάνονται υπόψιν και παρουσιάζονται στα αποτελέσματα. Συνεπώς, με αυτόν τον τρόπο εντοπίζονται οι αλληλουχίες εκείνες οι οποίες εμφανίζουν το μεγαλύτερο ποσοστό ομοιότητας και κατ' επέκταση ομολογίας με την αρχική αλληλουχία. Τα αποτελέσματα παρουσιάζονται και ταξινομούνται κατά σειρά μειούμενης συγγένειας.

Για τις στοιχίσεις με την υψηλότερη βαθμολογία συσχέτισης υπολογίζεται ένα μέτρο στατιστικής σημαντικότητας το E-value (ή ποσοστό ψευδώς θετικών αποτελεσμάτων), το οποίο δείχνει πόσο πιθανό είναι μια στοίχιση με δεδομένο βαθμό ομοιότητας (ή μεγαλύτερο) να έχει προκύψει τυχαία. Η τιμή του E-value δίνεται από τον τύπο $E\text{-value} = Kmne^{(-\lambda S)}$, όπου m και n τα μήκη των συγκρινόμενων αλληλουχιών, το k σχετίζεται με το μέγεθος της βάσης δεδομένων και S η βαθμολογία (score) της στοίχισης. Όσο μικρότερη είναι η τιμή του E-value τόσο αυξάνεται η αξιοπιστία ενός αποτελέσματος. Συνεπώς, εφόσον τα S και E-value είναι ποσά αντιστρόφως ανάλογα, όσο μεγαλύτερη είναι η βαθμολογία μιας στοίχισης τόσο λιγότερο πιθανό είναι αυτή να έχει προκύψει τυχαία. Επιπλέον, παρατηρούμε ότι το E-value εξαρτάται τόσο από το μέγεθος της βάσης δεδομένων όσο και από το μέγεθος της αρχικής αλληλουχίας.



Εικόνα 28. Διαδοχικά στάδια στην εκτέλεση του αλγορίθμου BLAST. Βήμα 1: Αρχικά πραγματοποιείται διάσπαση της αρχικής αλληλουχίας σε μικρότερα τμήματα καθορισμένου μήκους (words). Βήμα 2: Τα τμήματα που προέκυψαν συγκρίνονται με αντίστοιχα τμήματα του ίδιου μεγέθους προερχόμενα από τις αλληλουχίες που περιέχονται στις βάσεις δεδομένων. Βήμα 3: Οι στοιχίσεις αυτές βαθμολογούνται με βάση καθορισμένους πίνακες αντικατάστασης και αυτές οι οποίες ξεπερνούν ένα συγκεκριμένο κατώφλι επεκτείνονται και προς τις δύο κατευθύνσεις. Οι τελικές στοιχίσεις με την μεγαλύτερη βαθμολογία συσχέτισης (HSPs - high scoring sequence pairs) λαμβάνονται υπόψιν (<https://www.slideshare.net/zahid6/presentation-for-blast-algorithm-bioinformaticce>).

Σημειώνεται ότι υπάρχουν πολλές παραλλαγές του BLAST έτσι ώστε να είναι δυνατή η σύγκριση όλων των πιθανών συνδυασμών ανάμεσα σε μια αρχική πρωτεϊνική ή νουκλεοτιδική αλληλουχία έναντι μιας πρωτεϊνικής ή νουκλεοτιδικής βάσης δεδομένων (Πίνακας 4). Από τα προγράμματα αυτά εκείνα που χρησιμοποιούνται πιο συχνά είναι τα BLASTn και BLASTp, γιατί χρησιμοποιούν απευθείας συγκρίσεις χωρίς να είναι απαραίτητη η πραγματοποίηση κάποιας μετάφρασης αλληλουχίας. Παρόλα αυτά εφόσον οι πρωτεϊνικές ακολουθίες είναι πιο συντηρημένες εξελικτικά από τις νουκλεοτιδικές, το tBLASTn, το tBLASTx και το BLASTx παρέχουν πιο αξιόπιστα και ακριβή αποτελέσματα όσον αφορά το κωδικό DNA. Αντίθετα το PSI-BLAST χρησιμοποιείται για την αναζήτηση απομακρυσμένων εξελικτικά συγγενικών πρωτεϊνών καθώς εμφανίζει πολύ μεγαλύτερη ευαισθησία από το απλό BLASTp.

Πίνακας 4. Στον παρακάτω πίνακα περιγράφεται η λειτουργία καθενός από τα 5 διαφορετικά προγράμματα BLAST τα οποία υπάρχουν σήμερα.

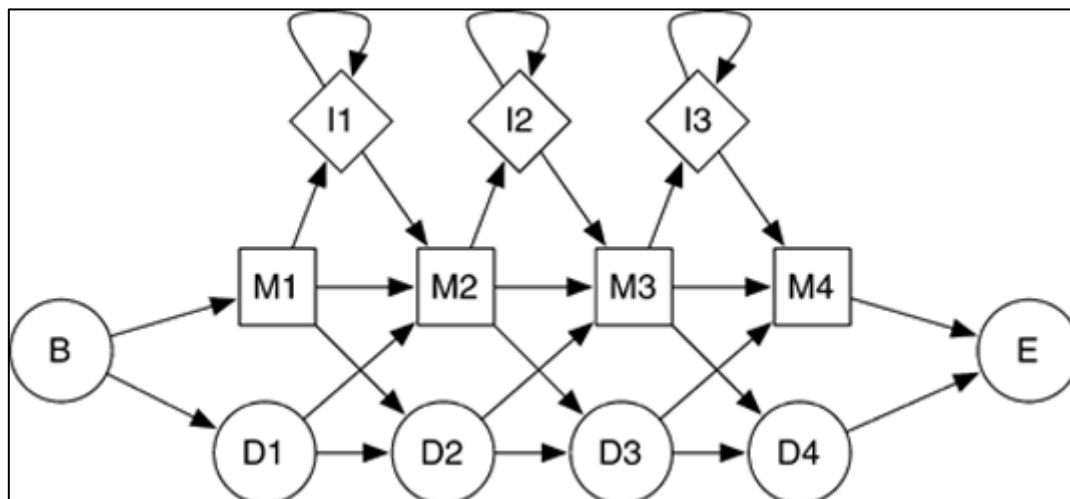
| Πρόγραμμα | Περιγραφή λειτουργίας |
|------------------|--|
| blastn | Σύγκριση μιας νουκλεοτιδικής αλληλουχίας με μια νουκλεοτιδική βάση δεδομένων. |
| blastp | Σύγκριση μιας πρωτεϊνικής αλληλουχίας με μια πρωτεϊνική βάση δεδομένων. |
| blastx | Μετάφραση μιας αρχικής νουκλεοτιδικής αλληλουχίας σε όλα τα ανοιχτά πλαίσια ανάγνωσης και σύγκριση με μια πρωτεϊνική βάση δεδομένων. |
| tblastn | Σύγκριση μιας πρωτεϊνικής αλληλουχίας με μια νουκλεοτιδική βάση δεδομένων, μεταφρασμένη σε όλα τα ανοιχτά πλαίσια ανάγνωσης. |
| tblastx | Μετάφραση μιας αρχικής νουκλεοτιδικής αλληλουχίας σε όλα τα ανοιχτά πλαίσια ανάγνωσης και σύγκριση με μια νουκλεοτιδική βάση δεδομένων, μεταφρασμένη σε όλα τα ανοιχτά πλαίσια ανάγνωσης. |
| psi-blast | Αρχικά πραγματοποίηση blastp από το οποίο προκύπτει ένας πίνακας PSSM (Position-Specific-Scoring-Matrix). Ο πίνακας αυτός χρησιμοποιείται στη συνέχεια ως προφίλ για την πραγματοποίηση αναζήτησης σε βάση πρωτεϊνικών δεδομένων με αποτέλεσμα να προκύπτει καινούργιος πίνακας. Η διαδικασία επαναλαμβάνεται για το νέο πίνακα. |
| megablast | Κάνει συγκρίσεις όπως και το blastn για μεγάλο αριθμό αρχικών αλληλουχιών. |

2.2.2 Το πρόγραμμα Hmmer

Το Hmmer είναι ένα πρόγραμμα για την αναζήτηση σε βάσεις δεδομένων αλληλουχιών, καθώς και για την πραγματοποίηση στοιχίσεων. Η αναζήτηση αυτή μπορεί να πραγματοποιηθεί τόσο για μια ξεχωριστή αλληλουχία όσο και για ένα σύνολο στοιχισμένων αλληλουχιών από τις οποίες κατασκευάζεται ένα προφίλ. Το πρόγραμμα αυτό έχει πολύ μεγάλη ευαισθησία και μπορεί να εντοπίσει ομολογία ανάμεσα σε αλληλουχίες, με βάση τον βαθμό ομοιότητάς τους. Η μεγάλη όμως αυτή ευαισθησία είναι πιθανό να οδηγήσει και σε ψευδώς θετικά αποτελέσματα. Επιπλέον, παρέχει τη δυνατότητα για αυτοματοποιημένο σχολιασμό πρωτεϊνικών επικρατειών μέσα από διάφορες μεγάλες βάσεις δεδομένων με στοιχίσεις που έχουν περιγραφεί και σχολιαστεί επαρκώς και μοντέλα Hmmer γνωστών επικρατειών, συμπεριλαμβανομένης της Pfam και άλλων.

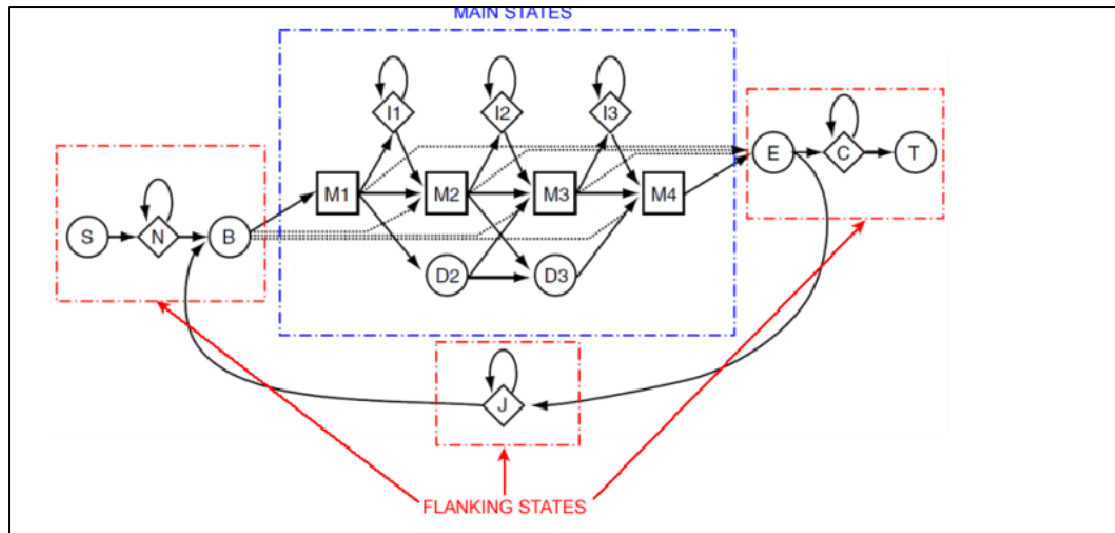
Η λειτουργία του Hmmer βασίζεται στη χρήση πιθανολογικών μοντέλων, που ονομάζονται profile hidden Markov models (profile HMMs - εικόνα 29). Πιο συγκεκριμένα, με αφετηρία μια αρχική στοιχίση αλληλουχιών κατασκευάζεται ένα profile HMM, με βάση το οποίο προκύπτει ένα σύστημα βαθμολόγησης. Το profile είναι ένα μοντέλο βαθμολόγησης συγκεκριμένης θέσης (position-specific scoring model) που περιγράφει ποια σύμβολα είναι πιθανό να εμφανιστούν σε κάθε θέση καθώς και πόσο συχνά συμβαίνουν ενθέσεις ή ελλείψεις σε κάθε θέση, σε μια πολλαπλή στοιχίση αλληλουχιών. Δηλαδή, το Hmmer χρησιμοποιεί αλγόριθμους συναρτήσεων που εξετάζουν όλες τις πιθανές στοιχίσεις, οι οποίες σταθμίζονται από τη σχετική τους πιθανότητα. Αυτός είναι ένας λόγος για τον οποίο το Hmmer αποκτά περισσότερη ισχύ από τα εργαλεία που εξαρτώνται από τη μοναδική βέλτιστη ευθυγράμμιση, όπως το BLAST. Στη συνέχεια, μπορεί να πραγματοποιηθεί σύγκριση του profile που έχει προκύψει με μια ή περισσότερες βάσεις

δεδομένων. Αλληλουχίες οι οποίες μετά από σύγκριση με το profile αποκτούν καλύτερη βαθμολόγηση από μια μηδενική υπόθεση θεωρούνται πραγματικά ομόλογες με τις αρχικές αλληλουχίες. Συχνά το πρόγραμμα Hmmer χρησιμοποιείται μαζί με κάποια βάση δεδομένων που περιέχει profiles αλληλουχιών, όπως η Pfam ή πολλές άλλες βάσεις δεδομένων που συμμετέχουν στην Interpro (Eddy, 2018).



Εικόνα 29. Στην εικόνα φαίνεται ένα profile hidden Markov model, μέσω του οποίου ορίζονται μια σειρά από πιθανές διαδοχικές καταστάσεις (M, I, D). Επιπλέον, το μοντέλο ορίζει και τις πιθανότητες ταύτισης σε μια κατάσταση (M), καθώς και εμφάνισης μια ένθεσης (I) ή έλλειψης (D) μετά από κάθε κατάσταση M. Η πιθανότητα εμφάνισης μιας κατάστασης ονομάζεται πιθανότητα γενέσεως. Η πιθανότητα μετάβασης από μια κατάσταση σε μια άλλη ονομάζεται πιθανότητα μετάβασης (<https://bibiserv.cebitec.uni-bielefeld.de/sadr2/databasesearch/hmmer/index.html>).

Οι εκδόσεις HMMER2 και HMMER3 χρησιμοποιούν μια δομή για την κατασκευή των profile HMMs που ονομάζεται «Plan 7 architecture» λόγω των 7 καταστάσεων που περιλαμβάνονται στο μοντέλο (εικόνα 30). Πιο συγκεκριμένα, εκτός από τις βασικές καταστάσεις (M, I and D), το μοντέλο περιλαμβάνει άλλες 6 για την ανίχνευση μη-ομόλογων πλευρικών αλληλουχιών σε μια στοίχιση (Eddy, 1998).



Εικόνα 30. Η δομή του Plan 7 πάνω στο οποίο βασίζεται η λειτουργία των HMMER2 και HMMER3. Η διαφορά με τα απλά HMMs είναι ότι εδώ υπάρχουν ενσωματωμένες και άλλες 6 παράμετροι για την αναγνώριση μη-ομόλογων πλευρικών αλληλουχιών σε μια στοίχιση (Giraldo, 2011).

Το στοιχείο που διαφοροποιεί το Hmmer από άλλους αλγορίθμους στοίχισης όπως οι BLAST και FASTA και του προσδίδει μεγαλύτερη ευαισθησία είναι ότι οι τελευταίοι χρησιμοποιούν πίνακες αντικατάστασης, όπως οι BLOSUM και PAM, που δεν λαμβάνουν υπόψιν τη θέση των αμινοξέων στην ακολουθία, πράγμα που συμβαίνει με τα profiles. Επιπλέον, τα profiles έχουν πολλές παραμέτρους, πολύ περισσότερες από έναν πίνακα αντικατάστασης. Πιο συγκεκριμένα, οι πίνακες αντικατάστασης αμινοξικών καταλοίπων BLOSUM και PAM έχουν μόνο 210 ελεύθερες παραμέτρους, οι οποίες έχουν προκύψει ως μέσοι όροι από μια μεγάλη συλλογή από πολλές διαφορετικές γνωστές στοίχισεις αλληλουχιών. Αντίθετα ένα προφίλ έχει τουλάχιστον 22 παραμέτρους για κάθε μια από τις περίπου 200 συναινετικές θέσεις που εμφανίζονται σε μια τυπική πρωτεϊνική επικράτεια. Οι παράμετροι αυτοί αντιστοιχούν σε μια μόνο στοίχιση οικογένειας πρωτεϊνικών ακολουθιών. Τα HMMs επιτρέπουν τη χρήση της θεωρίας των πιθανοτήτων για να ορίσει και να ερμηνεύσει αυτόν το μεγάλο αριθμό των ελεύθερων παραμέτρων ενός προφίλ (Eddy, 2018).

Σημειώνεται ότι υπάρχουν πολλές παραλλαγές του Hmmer οι οποίες παρέχουν πολλές δυνατότητες επεξεργασίας πρωτεϊνικών αλληλουχιών (πίνακας 5). Τα προγράμματα rhmmer και jackhmmer πραγματοποιούν αναζήτηση μιας πρωτεϊνικής ακολουθίας σε μια πρωτεϊνική βάση δεδομένων, όπως τα BLASTp and PSI-BLAST αντίστοιχα (η διαφορά είναι ότι κατασκευάζουν ένα profile από την αλληλουχία με βάση το οποίο πραγματοποιείται στη συνέχεια αναζήτηση). Το προγράμματα nhmmer και nhmmscan μπορούν να πραγματοποιήσουν αναζήτηση για αλληλουχίες DNA και είναι ισοδύναμα με τα hmsearch και rhmmer το πρώτο και hmmscan το δεύτερο, με τη διαφορά ότι μπορούν να χρησιμοποιήσουν αλληλουχίες DNA μεγέθους χρωμοσωμάτων (Eddy, 2018).

Πίνακας 5. Στον παρακάτω πίνακα περιγράφεται η λειτουργία καθενός από τα 18 διαφορετικά προγράμματα Hmmer τα οποία υπάρχουν σήμερα.

| Πρόγραμμα | Λειτουργία |
|-------------------|---|
| hmmbuild | Κατασκευή ενός profile από μια αρχική αλληλουχία. |
| hmmalign | Πολλαπλή στοίχιση αλληλουχιών με τη χρήση ενός profile. |
| hmmsearch | Αναζήτηση ενός profile σε μια βάση δεδομένων αλληλουχιών. |
| hmmscan | Αναζήτηση μιας αλληλουχίας σε μια βάση δεδομένων με profiles αλληλουχιών. |
| hmmpress | Προετοιμασία μιας βάσης δεδομένων, με profiles αλληλουχιών για το hmmscan. |
| phmmer | Αναζήτηση μιας αλληλουχίας σε μια βάση δεδομένων αλληλουχιών. |
| jackhmmmer | Επαναλαμβανόμενη αναζήτηση μιας αλληλουχίας σε μια βάση δεδομένων. |
| nhmmer | Αναζήτηση μιας αλληλουχίας DNA σε μια νουκλεοτιδική βάση δεδομένων. |
| nhmmscan | Αναζήτηση μιας αλληλουχίας DNA σε μια νουκλεοτιδική βάση με profiles αλληλουχιών. |
| hmmfetch | Ανάκτηση profile(s) από ένα αρχείο με profiles |
| hmmstat | Παρουσίαση περιληπτικών στατιστικών στοιχείων για ένα προφίλ. |
| hmmemit | Δημιουργία (δείγματος) ακολουθιών από ένα προφίλ. |
| hmmlogo | Δημιουργία γραφικού λογότυπου συναινετικής ακολουθίας από ένα προφίλ. |
| hmmconvert | Μετατροπή μεταξύ διαφορετικών τύπων αρχείων προφίλ. |
| hmmpgmd | Αναζήτηση στο πρόγραμμα daemon για την ιστοσελίδα hmmer.org. |
| makehmmdb | Προετοιμασία μιας binary βάσης δεδομένων nhmmer. |
| hmmsim | Συλλογή της κατανομής των score σε τυχαίες αλληλουχίες. |
| alimask | Προσθήκη μιας στήλης mask σε μια πολλαπλή στοίχιση αλληλουχιών. |

2.3 Φυλογενετική ανάλυση

Η φυλογενετική ανάλυση αποτελεί τη μελέτη των εξελικτικών σχέσεων μεταξύ οργανισμών με τη χρήση δεδομένων μοριακής ακολουθίας. Αποτελεί μια ιδιαίτερα χρήσιμη διαδικασία μέσω της οποίας προκύπτει μια οπτική απεικόνιση των φυλογενετικών σχέσεων. Επιπλέον, μέσω της φυλογενετικής ανάλυσης είναι δυνατή η διαλεύκανση των φυλογενετικών σχέσεων και της εξελικτικής πορείας ενός γονιδίου ή μιας πρωτεΐνης ή και ενός οργανισμού, η επιβεβαίωση της συγγένειας μεταξύ αλληλουχιών καθώς και η ακριβής περιγραφή της ταξινομικής ταυτότητας ενός νέου είδους.

Η πραγματοποίηση της φυλογενετικής ανάλυσης περιλαμβάνει 5 βασικά στάδια: α) την συλλογή ομόλογων αλληλουχιών, β) τη στοίχιση αλληλουχιών και τον υπολογισμό βαθμών ομοιότητας, γ) τον υπολογισμό εξελικτικών αποστάσεων μεταξύ αλληλουχιών, δ) την κατασκευή φυλογενετικών δέντρων και ε) τη στατιστική εκτίμηση των φυλογενετικών δέντρων (Νικολάου, 2015). Στη συνέχεια θα αναπτυχθούν τα στάδια της στοίχισης των αλληλουχιών και της κατασκευής και εκτίμησης φυλογενετικών δέντρων.

2.3.1 Πολλαπλή στοίχιση αλληλουχιών

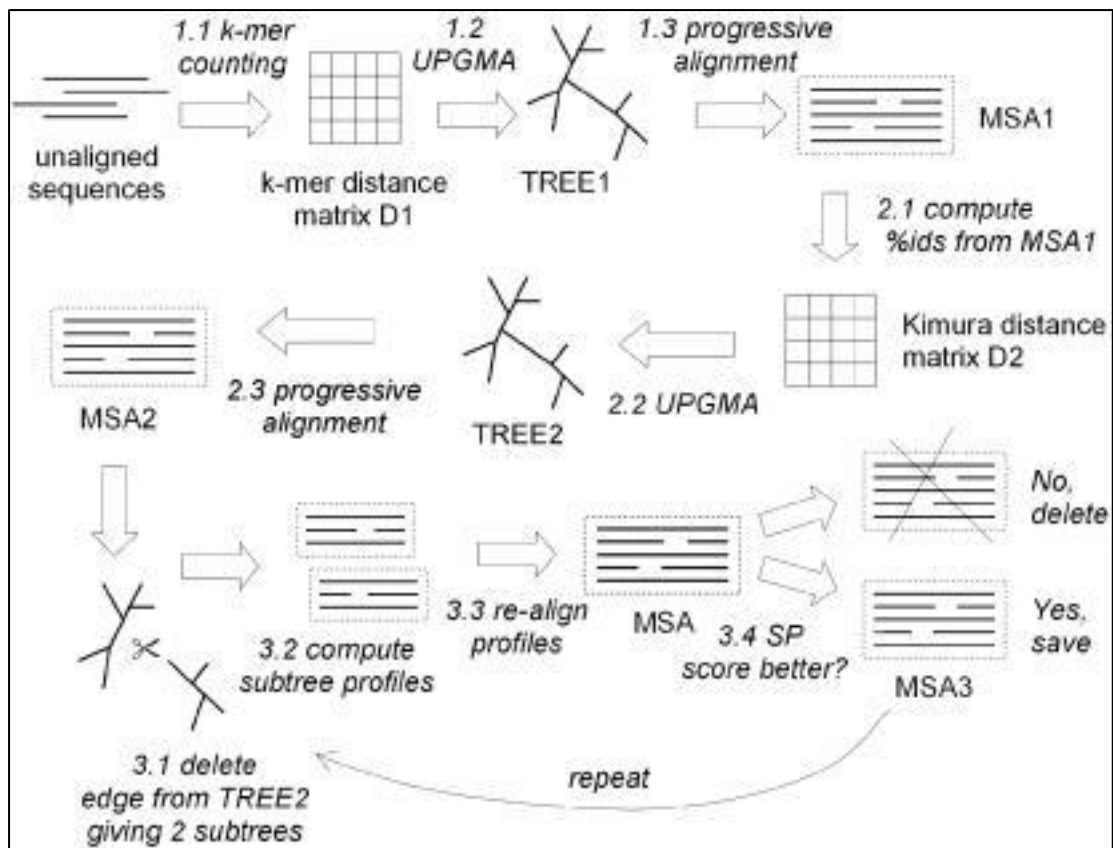
Για τη μελέτη της βαθιάς φυλογένεσης χρειάζονται αρχαίες, καθολικές, ορθόλογες αλληλουχίες οι οποίες όμως μπορεί να είναι πολύ διαφορετικές με

αποτέλεσμα να μην στοιχίζονται σωστά. Στις περιπτώσεις αυτές επιλέγονται οι πρωτεϊνικές αλληλουχίες, οι οποίες δεν επηρεάζονται τόσο από τις μεταλλαγές που συμβαίνουν στο γενετικό υλικό (Ajawatanawong, 2017).

Η στοιχίση πολλαπλών πρωτεϊνικών αλληλουχιών είναι χρήσιμη σε πολλές περιπτώσεις, όπως η κατασκευή φυλογενετικών δέντρων, η πρόβλεψη της δομής πρωτεϊνικών ή και νουκλεοτιδικών αλληλουχιών καθώς και ο εντοπισμός συντηρημένων επικρατειών και η ταυτοποίηση συγκεκριμένων καταλοίπων σε καθορισμένες θέσεις αλληλουχιών (R. C. Edgar, 2004). Οι στρατηγικές μοριακής φυλογένεσης στηρίζονται σε μεγάλο βαθμό στην ακρίβεια μιας στοιχίσης. Μια αρκετά διαδεδομένη στρατηγική αποτελεί αυτή της προοδευτικής στοιχίσης (progressive method) κατά την οποία κατασκευάζεται προσεγγιστικά ένα «δέντρο οδηγός» και στη συνέχεια, με βάση αυτό, πραγματοποιείται μια σειρά διαδοχικών στοιχίσεων αλληλουχιών κατά ζεύγη. Πιο συγκεκριμένα, αρχικά στοιχίζονται οι πιο συγγενικές μεταξύ τους αλληλουχίες (εντοπίζονται στα φύλλα του δέντρου), ενώ σταδιακά προστίθενται όλο και πιο απομακρυσμένες αλληλουχίες. Έτσι είναι δυνατό να πραγματοποιηθούν στοιχίσεις μεταξύ μεμονωμένων αλληλουχιών, μεταξύ προφίλ πολλαπλής στοιχίσης και αλληλουχίας αλλά και μεταξύ δύο ανεξάρτητων προφίλ στοιχίσης. Ο αλγόριθμος της προοδευτικής στοιχίσης εφαρμόζεται σε αρκετά προγράμματα, όπως το MUSCLE, MAFFT και το Clustal Omega (Ajawatanawong, 2017; R. C. Edgar, 2004). Στη συγκεκριμένη μελέτη χρησιμοποιήθηκε το πρόγραμμα MUSCLE το οποίο περιγράφεται στη συνέχεια.

2.3.1.1 Πρόγραμμα στοιχίσεων MUSCLE

Το MUSCLE (**MU**ltiple **S**equences **C**omparison by **L**og-**E**xpectation) αποτελεί ένα ευρέως χρησιμοποιούμενο πρόγραμμα για την πολλαπλή στοιχίση νουκλεοτιδικών και πρωτεϊνικών αλληλουχιών. Μετά από ελέγχους και δοκιμές αποδείχτηκε ότι το πρόγραμμα αυτό πετυχαίνει καλύτερες ταχύτητες και μεγαλύτερη ακρίβεια αποτελεσμάτων σε σύγκριση με τα Clustal Omega και T-Coffee, ανάλογα με τις παραμέτρους που εφαρμόζονται σε κάθε περίπτωση (R. C. Edgar, 2004; Robert C. Edgar, 2004). Το πρόγραμμα βασίζεται στην αρχή της προοδευτικής στοιχίσης, την οποία και βελτιστοποιεί. Πιο συγκεκριμένα, ένα κοινό πρόβλημα της προοδευτικής στοιχίσης είναι τα λάθη που πραγματοποιούνται στη στοιχίση κατά ζεύγη τα οποία μπορεί να οφείλονται στη χαμηλή πιστότητα του «δέντρου-οδηγού», στις παραμέτρους βαθμολόγησης που δεν είναι οι βέλτιστες για κάθε σετ αλληλουχιών ή και σε άλλες ακατάλληλες συνθήκες. Για το λόγο αυτό μέσω του προγράμματος MUSCLE εφαρμόζονται επιπλέον βήματα βελτιστοποίησης του αρχικού αλγορίθμου (R. C. Edgar, 2004). Ο αλγόριθμος πάνω στον οποίο στηρίζεται η λειτουργία του προγράμματος φαίνεται στην εικόνα 31.



Εικόνα 31. Αλγόριθμος πάνω στον οποίο στηρίζεται η λειτουργία του προγράμματος MUSCLE (Novarino et al., 2014).

Το MUSCLE περιλαμβάνει 3 στάδια: α) της “πρόχειρης” προοδευτικής στοίχισης (*Draft progressive*), β) της βελτιωμένης προοδευτικής στοίχισης (*Improved progressive*) και γ) ένα στάδιο βελτιώσεων (*Refinement*). Στο τέλος κάθε σταδίου προκύπτει μια πολλαπλή στοίχιση και ο αλγόριθμος είναι δυνατό να τερματιστεί (Robert C. Edgar, 2004). Στο πρώτο στάδιο ο αλγόριθμος κατασκευάζει μία “πρόχειρη” πολλαπλή στοίχιση, δίνοντας έμφαση στην ταχύτητα παρά στην ακρίβεια. Στο επόμενο στάδιο πραγματοποιείται διόρθωση με τη μέθοδο Kimura και επαναπροσδιορισμός του δυαδικού δέντρου το οποίο προέκυψε από το στάδιο 1, με αποτέλεσμα να κατασκευάζεται μια πιο ακριβής πολλαπλή στοίχιση. Στο τελικό στάδιο βελτιώνεται περαιτέρω η στοίχιση που προέκυψε στο στάδιο 2 (Robert C. Edgar, 2004). Πιο συγκεκριμένα, οι ακριβείς αναλύσεις που πραγματοποιούνται σε κάθε στάδιο είναι οι εξής:

- Στάδιο 1: Πραγματοποιείται καταμέτρηση των μικρών αλληλουχιών μήκους k (k mers ή k -tuples) που έχουν κοινές μεταξύ τους οι αλληλουχίες κατά ζεύγη, χωρίς όμως την πραγματοποίηση πολλαπλής στοίχισης (εικόνα 31 βήμα 1.1). Αυτή η διαδικασία βασίζεται στο γεγονός ότι δύο συγγενικές αλληλουχίες αναμένεται να έχουν περισσότερα κοινά k mers από όσα θα προέκυπταν κατά τύχη. Η διαδικασία αυτή επειδή δεν απαιτεί την πραγματοποίηση στοίχισης προσδίδει μεγάλη ταχύτητα στον αλγόριθμο και είναι 3,000 φορές γρηγορότερη από την αντίστοιχη του CLUSTALW, αλλά μπορεί να είναι λιγότερο ακριβής. Με βάση τον πίνακα αποστάσεων

που προκύπτει κατασκευάζεται δυαδικό δέντρο με τη μέθοδο UPGMA (εικόνα 31 βήμα 1.2). Το δέντρο αυτό χρησιμοποιείται ως «δέντρο οδηγός» για την κατασκευή προοδευτικής στοίχισης (εικόνα 31 βήμα 1.3), όπως περιγράφηκε παραπάνω (Πεδίο 2.3.1 Πολλαπλή στοίχιση αλληλουχιών). Η πολλαπλή στοίχιση που προκύπτει σε αυτό το στάδιο δεν είναι η βέλτιστη λόγω κυρίως του τρόπου κατασκευής του πίνακα των *kmers* και θα βελτιωθεί στη συνέχεια από τα επόμενα 2 στάδια του αλγορίθμου (R. C. Edgar, 2004).

- Στάδιο 2: Από την πολλαπλή στοίχιση που προέκυψε από το στάδιο 1 είναι δυνατός ο υπολογισμός της ομοιότητας των αλληλουχιών κατά ζεύγη και η εκτίμηση της μεταξύ τους απόστασης. Στις αποστάσεις αυτές εφαρμόζεται η διόρθωση Kimura για πολλαπλές αντικαταστάσεις σε μία μόνο θέση (εικόνα 31 βήμα 2.1). Η μέθοδος αυτή είναι πιο ακριβής από τη μέθοδο των *kmers*. Από τον πίνακα αποστάσεων που προκύπτει κατασκευάζεται ένα νέο δυαδικό δέντρο με τη μέθοδο UPGMA (εικόνα 31 βήμα 2.2). Τα δέντρα από τα στάδια 1 και 2 συγκρίνονται και για τα κλαδιά τα οποία διαφέρουν πραγματοποιείται εκ νέου προοδευτική στοίχιση με βάση το καινούργιο δέντρο (εικόνα 31 βήμα 2.3). Εάν τα δύο δέντρα είναι πανομοιότυπα το βήμα αυτό παραλείπεται. Εάν τα δύο δέντρα δεν έχουν καμία ομοιότητα μεταξύ τους (αρκετά απίθανο) τότε η προοδευτική στοίχιση πραγματοποιείται εξ' αρχής. Το στάδιο 2 επαναλαμβάνεται μέχρι έναν μέγιστο αριθμό επαναλήψεων ή έως ότου σταθεροποιηθεί το δέντρο. Το στάδιο 2 δηλαδή αποτελεί ένα πρώτο βήμα βελτιστοποίησης του αρχικού δέντρου αλλά και της αρχικής στοίχισης (R. C. Edgar, 2004).
- Στάδιο 3: Στο τελευταίο στάδιο του αλγορίθμου επιλέγεται ένας κλάδος από το τελικό δέντρο του σταδίου 2 και το δέντρο αυτό διαιρείται σε δύο επιμέρους δέντρα (εικόνα 31 βήμα 3.1). Ένα profile κατασκευάζεται για κάθε ένα από τα δύο δέντρα (εικόνα 31 βήμα 3.2) και τα δύο profiles που προκύπτουν στοιχίζονται και πάλι μεταξύ τους με βάση την αρχή της στοίχισης κατά ζεύγη, που χρησιμοποιείται κατά τη διάρκεια της προοδευτικής στοίχισης (εικόνα 31 βήμα 3.3). Εάν η νέα πολλαπλή στοίχιση που προκύπτει έχει καλύτερο score (sum-of pairs score (SP): μέσος όρος βαθμολογίας κάθε στοίχισης κατά ζεύγη), το οποίο αντικατοπτρίζει την ποιότητα της στοίχισης, τότε αυτή διατηρείται. Σε αντίθετη περίπτωση απορρίπτεται (εικόνα 31 βήμα 3.4). Το στάδιο 3 επαναλαμβάνεται μέχρι έναν μέγιστο αριθμό επαναλήψεων ή έως ότου προκύψει το βέλτιστο score (R. C. Edgar, 2004).

2.3.2 Προετοιμασία πολλαπλής στοίχισης για την κατασκευή φυλογενετικών δέντρων

Για την κατασκευή φυλογενετικών δέντρων από τα αποτελέσματα της πολλαπλής στοίχισης απαιτούνται μια σειρά από ενδιάμεσα βήματα. Πιο συγκεκριμένα η παρουσία ενός ή περισσότερων κενών σε μια πολλαπλή στοίχιση είναι δυνατό να προκαλέσει προβλήματα κατά την κατασκευή των φυλογενετικών δέντρων. Για το λόγο αυτό τα κενά αυτά θα πρέπει να αφαιρούνται μέσω μιας διαδικασίας που ονομάζεται "masking and trimming". Πιο συγκεκριμένα πραγματοποιείται έλεγχος των στοιχίσεων και αποκοπή

των περιοχών διαφορούμενου χαρακτήρα. Η επεξεργασία αυτή μπορεί να πραγματοποιηθεί μέσω του προγράμματος Mesquite (<http://www.mesquiteproject.org/>) το οποίο είναι ένα αρθρωτό, επεκτάσιμο λογισμικό που χρησιμοποιείται στον τομέα της εξελικτικής βιολογίας. Σχεδιάστηκε για να εξυπηρετεί την οργάνωση και ανάλυση συγκριτικών δεδομένων για τους οργανισμούς. Η έμφαση δίνεται στη φυλογενετική ανάλυση, αλλά ορισμένες από τις ενότητες του προγράμματος αφορούν τη γενετική πληθυσμών, ενώ άλλες κάνουν μη φυλογενετική πολυπαραγοντική ανάλυση.

Το επόμενο βήμα πριν την κατασκευή φυλογενετικών δέντρων με το πρόγραμμα RaxML, αφορά στην εύρεση του βέλτιστου μοντέλου εξέλιξης αλληλουχιών. Για το σκοπό αυτό χρησιμοποιείται το πρόγραμμα Prottest (Ingley et al., 1999). Πρόκειται για ένα βιοπληροφορικό εργαλείο για την επιλογή των καλύτερων μοντέλων αντικατάστασης αμινοξέων για τα υπάρχοντα σε κάθε περίπτωση δεδομένα. Πιο συγκεκριμένα, το πρόγραμμα αυτό υπολογίζει τα κριτήρια Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) score και Decision Theory Criterion (DT) για καθένα από τα υποψήφια μοντέλα. Το μοντέλο με τη χαμηλότερη τιμή για ένα ή περισσότερα από τα παραπάνω κριτήρια επιλέγεται.

2.3.3 Μέθοδοι για τη δημιουργία φυλογενετικών δέντρων

Για την κατασκευή φυλογενετικών έχουν αναπτυχθεί αρκετοί αλγόριθμοι οι οποίοι βασίζονται σε διαφορετικές αρχές και διαφέρουν ως προς την ταχύτητα λειτουργίας, καθώς και την αξιοπιστία των αποτελεσμάτων. Οι αλγόριθμοι αυτοί χωρίζονται σε δύο βασικές κατηγορίες: α) αυτούς που χρησιμοποιούν κάποιο πίνακα αποστάσεων για την κατασκευή μιας πολλαπλής στοίχισης (χαρακτηριστικοί είναι οι UPGMA-Unweighted Pair Group Method with Arithmetic Averages και Neighbour Joining-NJ) και β) αυτούς που βασίζονται στην ελαχιστοποίηση παραμέτρων (χαρακτηριστικοί είναι οι μέθοδοι Μέγιστης Φειδωλότητας και Μέγιστης Πιθανοφάνειας) (Νικολάου, 2015).

2.3.3.1 Μέθοδοι αποστάσεων

Στις μεθόδους αποστάσεων η κατασκευή του δέντρου στηρίζεται σε έναν πίνακα ο οποίος περιέχει τις ανά ζεύγη αποστάσεις μεταξύ των αλληλουχιών. Η κατασκευή του πίνακα πραγματοποιείται με βάση μια πολλαπλή στοίχιση αλληλουχιών σε συνδυασμό με κάποιο μοντέλο αντικατάστασης. Στόχος των μεθόδων αυτών είναι να προκύψει ένα δέντρο στο οποίο η τοπολογία και τα μήκη των κλάδων να αντιστοιχούν όσο το δυνατό περισσότερο στις αποστάσεις που περιλαμβάνονται στο πίνακα. Οι μέθοδοι αποστάσεων είναι πολύ γρήγορες στη λειτουργία τους και χρήσιμες για τη δημιουργία μεγάλων δέντρων από ένα μεγάλο σύνολο δεδομένων (>1000 αλληλουχίες), με υψηλό βαθμό ομοιότητας. Παρόλα αυτά έχουν το μειονέκτημα ότι όλες οι πληροφορίες γύρω από τις αλληλουχίες χάνονται, λόγω της δημιουργίας του πίνακα αποστάσεων (Ajawatanawong, 2017; Νικολάου, 2015). Δύο βασικές μέθοδοι αποστάσεων είναι οι UPGMA και Neighbour Joining.

Όσον αφορά τη μέθοδο UPGMA το φυλογενετικό δέντρο κατασκευάζεται σταδιακά με βάση έναν πίνακα αποστάσεων. Πιο συγκεκριμένα, οι αλληλουχίες προστίθενται μια στο δέντρο, με βάση τις μεταξύ τους αποστάσεις από τον πίνακα αποστάσεων. Αρχικά προστίθενται οι αλληλουχίες με τις μικρότερες και στη συνέχεια αυτές με τις μεγαλύτερες μεταξύ τους αποστάσεις. Η UPGMA αποτελεί την απλούστερη μέθοδο δημιουργίας δέντρων αποστάσεων, παράγοντας ικανοποιητικά δέντρα ακόμα και για μεγάλο αριθμό αλληλουχιών. Η μέθοδος αυτή στηρίζεται σε δύο βασικές παραδοχές. Η πρώτη αφορά στην ισχύ του μοριακού ρολογιού, σύμφωνα με το οποίο ο ρυθμός των αντικαταστάσεων εντός ενός γονιδίου είναι σταθερός για όλους τους κλάδους του δέντρου. Η δεύτερη παραδοχή είναι ότι τα αθροίσματα των μηκών των κλάδων που οδηγούν από μία προγονική αλληλουχία προς καθεμιά από τις θυγατρικές αλληλουχίες, είναι ίσα (Νικολάου, 2015).

Η μέθοδος Neighbour Joining αποτελεί μια μέθοδο ελαχιστοποίησης του συνολικού μήκους των κλάδων ενός δέντρου. Με τη μέθοδο Neighbour Joining η κατασκευή ενός δέντρου πραγματοποιείται σταδιακά μέσω της σύνδεσης των οργανισμών ή αλληλουχιών με τη μικρότερη γειτνίαση. Πιο συγκεκριμένα, με βάση έναν αρχικό πίνακα αποστάσεων κατασκευάζεται σε κάθε στάδιο ένας νέος πίνακας που αντικατοπτρίζει τον βαθμό γειτνίασης των οργανισμών ή αλληλουχιών. Η μέθοδος αυτή κατασκευάζει δέντρα που χαρακτηρίζονται από την ιδιότητα της προσθετικότητας. Δηλαδή, οι αποστάσεις ανάμεσα σε δύο φύλλα του δέντρου ισούνται με τις αντίστοιχες αποστάσεις ανάμεσα στις αλληλουχίες που αντιστοιχούν στα φύλλα αυτά, που προκύπτουν από τον αρχικό πίνακα αποστάσεων. Επιπλέον, η μέθοδος NJ παράγει με μεγάλη ταχύτητα δέντρα χωρίς ρίζα και δεν βασίζεται στην αρχή του μοριακού ρολογιού, όπως η UPGMA. Το αρνητικό της μεθόδου είναι ότι σε πολλές περιπτώσεις προσδίδει αρνητικά μήκη στα κλαδιά του δέντρου που παράγει, ενώ παρέχει πολύ χαμηλότερο βαθμό ακρίβειας σε σύγκριση με τις μεθόδους ελαχιστοποίησης παραμέτρων (Νικολάου, 2015).

2.3.3.2 Μέθοδοι ελαχιστοποίησης παραμέτρων

Οι μέθοδοι ελαχιστοποίησης παραμέτρων είναι πιο χρονοβόρες και υπολογιστικά απαιτητικές από τις μεθόδους αποστάσεων αλλά χρησιμοποιούν απευθείας όλη τη βιολογική πληροφορία που προέρχεται από τις αλληλουχίες για την κατασκευή και αξιολόγηση του βέλτιστου φυλογενετικού δέντρου, χωρίς να πραγματοποιούνται μετατροπές. Για το λόγο αυτό οι μέθοδοι ελαχιστοποίησης παραμέτρων θεωρούνται πολύ ακριβείς. Σκοπός των μεθόδων ελαχιστοποίησης παραμέτρων είναι να ελαχιστοποιήσουν μια ποσότητα.

Ο αλγόριθμος στον οποίο βασίζονται περιλαμβάνει ορισμένα βασικά στάδια. Αρχικά κατασκευάζονται όλα τα πιθανά φυλογενετικά δέντρα για n αριθμό αλληλουχιών. Το βέλτιστο δέντρο είναι αυτό το οποίο εμφανίζει την υψηλότερη βαθμολόγηση. Είναι όμως πρακτικά αδύνατο να βαθμολογηθούν όλα τα δέντρα όταν ο αριθμός των taxa ξεπερνά τα 20, με τη χρήση μιας μεθόδου που αναζητά όλα τα δυνατά δέντρα. Για το λόγο αυτό υπάρχουν ορισμένοι υπολογιστικοί αλγόριθμοι (όπως οι 'branch-and-bound' και

'heuristic') που επιτρέπουν την επιλογή και βαθμολόγηση όλων των πιθανών δέντρων ταυτόχρονα. Παράλληλα, περιορίζουν τον αριθμό των πιθανών δέντρων παραλείποντας τις τοπολογίες που είναι θεωρητικά απίθανες (Ajawatanawong, 2017).

Η μέθοδος Μέγιστης Φειδωλότητας αποτελεί την παλαιότερη μέθοδο για την κατασκευή φυλογενετικών δέντρων και περιγράφηκε αρχικά για μορφολογικά χαρακτηριστικά, για τα οποία είναι δύσκολο να πραγματοποιηθεί μέτρηση του ρυθμού των εξελικτικών αλλαγών. Βασίζεται στην αρχή ότι το πιο ακριβές δέντρο είναι εκείνο που εξηγείται από τον ελάχιστο αριθμό αντικαταστάσεων (κριτήριο φειδωλότητας). Σε μια αρχική πολλαπλή στοίχιση κάθε στήλη αντιμετωπίζεται ως ένας ξεχωριστός χαρακτήρας. Πληροφοριακοί θεωρούνται μόνο οι χαρακτήρες οι οποίοι εμφανίζουν τουλάχιστον δύο καταστάσεις (δηλαδή για τους οποίους παρατηρείται τουλάχιστον μια αντικατάσταση νουκλεοτιδίου ή αμινοξέος). Μόνο αυτοί χρησιμοποιούνται για την κατασκευή του δέντρου (Ajawatanawong, 2017; Νικολάου, 2015).

Η μέθοδος Μέγιστης Φειδωλότητας μπορεί να δώσει αξιόπιστα αποτελέσματα για ομάδες αλληλουχιών στις οποίες η ομοπλασία παρατηρείται σπάνια ή και καθόλου. Επιπλέον, μπορεί να χρησιμοποιηθεί για περιπτώσεις όπου παρατηρούνται αρκετές ενθέσεις και ελλείψεις. Ωστόσο, οι περισσότερες αλληλουχίες δεν εξελίσσονται με μικρό ρυθμό και ως αποτέλεσμα μπορεί να είναι δύσκολο να στοιχηθούν. Το γεγονός αυτό καθιστά τη μέθοδο Μέγιστης Φειδωλότητας λιγότερο αποτελεσματική. Παράλληλα η μέθοδος αυτή είναι χρονοβόρα και δεν συνιστάται για συνεχείς αλληλουχίες πολλαπλών γονιδίων ή για αλληλουχίες με υψηλά επίπεδα ποικιλομορφίας (Ajawatanawong, 2017).

Η μέθοδος της Μέγιστης Πιθανοφάνειας αποτελεί τη δεύτερη πιο δημοφιλή μέθοδο κατασκευής φυλογενετικών δέντρων. Μέσω της μεθόδου αυτής πραγματοποιείται σταδιακά η δημιουργία ενός φυλογενετικού δέντρου με βάση ένα μοντέλο αντικατάστασης και μια στοίχιση αλληλουχιών, υπολογίζοντας την πιθανοφάνεια για κάθε κόμβο του δέντρου δεδομένων των θυγατρικών του φύλλων/κόμβων. Πιο συγκεκριμένα, υπολογίζεται η πιθανότητα καθένα από τα δυνατά φυλογενετικά δέντρα και μοντέλα εξέλιξης να ταιριάζει και να προσαρμόζεται στο υπό μελέτη σύνολο αλληλουχιών. Βέλτιστο δέντρο θεωρείται εκείνο το οποίο εμφανίζει τη μέγιστη πιθανοφάνεια.

Η μέθοδος της Μέγιστης Πιθανοφάνειας θεωρείται εξαιρετικά ακριβής καθότι η ανάλυση στηρίζεται κυρίως στα μοντέλα εξέλιξης. Επιπλέον, σε αντίθεση με τη μέθοδο Μέγιστης Φειδωλότητας, αξιοποιεί όλη τη δυνατή πληροφορία προερχόμενη από τις αλληλουχίες ώστε να υπολογίσει την μέγιστη τιμή πιθανοφάνειας (Ajawatanawong, 2017). Τέλος, το γεγονός ότι απαιτεί μεγάλη υπολογιστική ισχύ και χρόνο για την πραγματοποίηση της ανάλυσης είναι δυνατό να ξεπεραστεί με τη χρήση διαδικτυακών server μέσω των οποίων οι αναλύσεις πραγματοποιούνται εξ' αποστάσεως σε υπολογιστές με μεγάλη ισχύ.

2.3.3.2.1 Πρόγραμμα RaxML

Για την κατασκευή φυλογενετικών δέντρων στην παρούσα μελέτη επιλέχθηκε το πρόγραμμα RaxML το οποίο βασίζεται στη μέθοδο της

Μέγιστης Πιθανοφάνειας. Η επιλογή αυτή βασίστηκε τόσο στην αξιοπιστία της μεθόδου όσο και στην ευκολία χρήσης του προγράμματος. Επειδή όμως το πρόγραμμα αυτό απαιτεί μεγάλη υπολογιστική ισχύ οι αναλύσεις πραγματοποιήθηκαν με βάση την πλατφόρμα cipress (<http://www.phylo.org/>). Πρόκειται για μια ελεύθερα διαθέσιμη διαδικτυακή πλατφόρμα για την πραγματοποίηση αναλύσεων φυλογένεσης μέσω υπερυπολογιστών και άλλων πόρων της πλατφόρμας XSEDE του NSF.

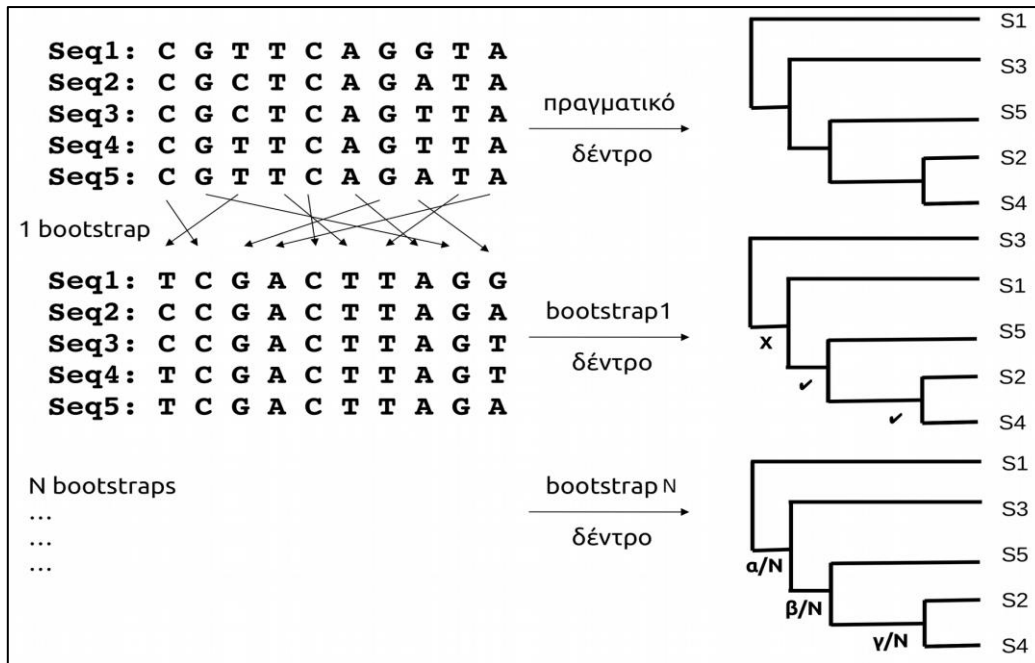
Το πρόγραμμα RAxML (Randomized Axelerated Maximum Likelihood) αναπτύχθηκε για την κατασκευή φυλογενετικών δέντρων για εξαιρετικά μεγάλα σύνολα δεδομένων, τόσο ως προς τον αριθμό των αλληλουχιών όσο και ως προς το μέγεθος αυτών. Επιπλέον, μπορεί να χρησιμοποιηθεί για μέτα-ανάλυση ομάδας φυλογενετικών δέντρων, ανάλυση στοιχίσεων και προσδιορισμό της εξέλιξης αλληλουχημένων ακολουθιών μικρού μεγέθους.

2.3.4 Εκτίμηση φυλογενετικών δέντρων

Σημαντικό στάδιο στην πραγματοποίηση μιας φυλογενετικής μελέτης αποτελεί η αξιολόγηση των φυλογενετικών δέντρων που προκύπτουν. Μια από τις ευρύτερα διαδεδομένες μεθόδους ελέγχου της αξιοπιστίας δεδομένων, που χρησιμοποιείται για αυτό το σκοπό αποτελεί η ανάλυση «bootstrap». Η μέθοδος αυτή βασίζεται σε μια ισχυρή στατιστικά τεχνική επαναλαμβανόμενης δειγματοληψίας (resampling) και εμφανίζει ευρύ πεδίο εφαρμογών. Πιο συγκεκριμένα, σκοπός της είναι ο καθορισμός του στατιστικού λάθους σε καταστάσεις στις οποίες η βασική κατανομή των δειγμάτων είναι είτε άγνωστη είτε δύσκολο να προκύψει με αναλυτικό τρόπο.

Τα βασικά στάδια στην πραγματοποίηση του ελέγχου bootstrap με βάση μια πολλαπλή στοίχιση αλληλουχιών και ένα αρχικό φυλογενετικό δέντρο είναι τα εξής (εικόνα 32):

- a) Αρχικά πραγματοποιείται τυχαία ανακατάταξη των νουκλεοτιδικών ή αμινοξικών θέσεων στην πολλαπλή στοίχιση αλληλουχιών.
- b) Για κάθε ανακαταταγμένη στοίχιση πραγματοποιείται κατασκευή φυλογενετικού δέντρου με την ίδια μέθοδο που χρησιμοποιήθηκε για την κατασκευή του αρχικού δέντρου.
- c) Σε καθένα από τα δέντρα που προκύπτουν καταμετρώνται οι κόμβοι που βρίσκονται στην ίδια θέση με το αρχικό δέντρο, για τους οποίους η τιμή bootstrap αυξάνεται κατά 1.
- d) Η διαδικασία αυτή επαναλαμβάνεται για έναν καθορισμένο αριθμό επαναλήψεων, συνήθως πάνω από 100 και στο τέλος αποδίδεται σε κάθε κόμβο η συνολική τιμή bootstrap.



Εικόνα 32. Βασικά στάδια αξιολόγησης ενός φυλογενετικού δέντρου με τη χρήση της μεθόδου bootstrap (Νικολάου, 2015).

Σε ένα φυλογενετικό δέντρο αξιόπιστοι κλάδοι θεωρούνται εκείνοι οι οποίοι υποστηρίζονται από υψηλή τιμή bootstrap μεγαλύτερη του 70-75%. Σε αυτήν την περίπτωση η συναγόμενη τοπολογία στη συγκεκριμένη θέση θεωρείται βιολογικά σημαντική. Εάν όλοι οι κλάδοι ενός δέντρου στηρίζονται από μεγάλες τιμές bootstrap τότε αυτό θεωρείται γενικότερα αξιόπιστο. Η μέθοδος bootstrap ως στατιστική διαδικασία αποδίδει πιο αξιόπιστα αποτελέσματα όσο μεγαλώνει το αρχικό δείγμα των αλληλουχιών. Βασικό πλεονέκτημα της είναι ότι μπορεί να χρησιμοποιηθεί σε όλες τις μεθόδους κατασκευής φυλογενετικών δέντρων (Νικολάου, 2015).

3 Αποτελέσματα-Συζήτηση

Οι αναλύσεις πραγματοποιήθηκαν με βάση την πορεία εργασίας που περιγράφηκε στο πεδίο 2.1 “Γενική πορεία εργασίας” της ενότητας Υλικά και Μέθοδοι. Τα αποτελέσματα αυτών των αναλύσεων παρατίθενται στις παρακάτω ενότητες.

3.1 Αποτελέσματα συγκριτικής γονιδωματικής ανάλυσης

Οι αρχικές αναλύσεις που πραγματοποιήθηκαν βασίστηκαν στη χρήση του αλγορίθμου BLAST. Τα αποτελέσματα παρατίθενται στον πίνακα 6. Στη συνέχεια, για τους οργανισμούς και τις πρωτεΐνες για τις οποίες δεν βρέθηκαν ορθόλογες πρωτεΐνες πραγματοποιήθηκε περαιτέρω έλεγχος με τη χρήση του προγράμματος Hmmer. Τα συγκεντρωτικά αποτελέσματα που προέκυψαν από τις αναλύσεις των BLAST και Hmmer παρατίθενται στον πίνακα 7.

Στο κάτω μέρος καθενός από τους πίνακες 6 και 7 βρίσκονται τα ονόματα των οργανισμών μελέτης, ενώ στα αριστερά βρίσκεται αρχικά η επεξήγηση του χρωματικού κώδικα που χρησιμοποιείται και στη συνέχεια τα ονόματα των πρωτεϊνών της μελέτης. Επιπλέον, με τα σύμβολα + και - δηλώνεται η παρουσία ή απουσία αντίστοιχα ορθόλογης πρωτεΐνης και οι αριθμοί αντιπροσωπεύουν τον αριθμό των πρωτεϊνών αυτών όταν αυτός υπερβαίνει τη μονάδα.

Από τους παραπάνω πίνακες παρατηρούμε τα εξής:

- Τα αποτελέσματα των αναλύσεων για τις 4 πρωτεΐνες της οικογένειας των reticulons, για τις REEP5-Yop1p και για τις atlastin-sey1p εμφανίζονται συγκεντρωτικά στους παραπάνω πίνακες, αν και οι αναλύσεις πραγματοποιήθηκαν ξεχωριστά για κάθε μία πρωτεΐνη. Αυτό συμβαίνει για τις πρωτεΐνες REEP5 / Yop1p και atlastin / sey1p διότι είναι μεταξύ τους ομόλογες ως προς την αλληλουχία και τη λειτουργία αντίστοιχα, οπότε όπως αναμέναμε εμφάνισαν κοινά αποτελέσματα. Επίσης οι πρωτεΐνες μέλη της οικογένειας των reticulons εμφάνισαν κοινά αποτελέσματα σε μεγάλο αριθμό κατωτέρων κυρίως οργανισμών.
- Μέσω του προγράμματος Hmmer ταυτοποιήθηκαν αρκετές ορθόλογες πρωτεΐνες οι οποίες δεν βρέθηκαν με τον αλγόριθμο BLAST, κυρίως για τις πρωτεΐνες Ufe1p και Dsl1p (μαύροι κύκλοι στον πίνακα 4).
- Οι πρωτεΐνες reticulons, REEP5/Yop1p και atlastin/sey1p που δρουν συνεργιστικά με σκοπό τη διαμόρφωση της δομής του ενδοπλασματικού δικτύου, εμφανίζουν μεγάλο βαθμό συντήρησης με παρουσία αντιπροσώπων τους σε όλες ή σχεδόν όλες τις υπερομάδες των ευκαρυωτικών οργανισμών. Το ενδομεμβρανικό σύστημα θεωρείται βασικό στοιχείο των ευκαρυωτικών οργανισμών με την εξέλιξη του να τοποθετείται παράλληλα με την εξέλιξη του ευκαρυωτικού κυττάρου. Επιπλέον, μέσα από προηγούμενη φυλογενετική μελέτη (Oertle et al., 2003) αποδείχτηκε ότι η έναρξη της εξέλιξης των reticulons τοποθετείται στα αρχικά στάδια της εξέλιξης του ευκαρυωτικού κυττάρου και πιθανών συνέβη παράλληλα με την εξέλιξη του ενδομεμβρανικού συστήματος των ευκαρυωτικών οργανισμών. Κάτι τέτοιο επιβεβαιώνεται και από την παρούσα μελέτη, με την οποία ταυτόχρονα ταυτοποιούνται και δύο ακόμα πρωτεΐνες με κοινή εξελικτική πορεία με τα reticulons, οι REEP5/ Yop1p και atlastin/sey1p.
- Η πρωτεΐνη Arl6IP1 η οποία παίζει καθοριστικό ρόλο στην απόπτωση και συμμετέχει στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου, εμφανίζει αντιπροσώπους μόνο στα ανώτερα μετάζωα. Το γεγονός αυτό έρχεται σε συμφωνία με στοιχεία από τη βιβλιογραφία τα οποία αναφέρουν ότι η λειτουργία της πρωτεΐνης αυτής στη διαμόρφωση των αγωγών του σωληνοειδούς ενδοπλασματικού δικτύου αποτελεί χαρακτηριστικό των πολυκύτταρων οργανισμών (Yamamoto et al., 2014). Επειδή η λειτουργία αυτή επιτελείται και από τις πρωτεΐνες reticulons και REEP5 είναι πιθανό η Arl6IP1 να λειτουργεί σαν «δικλείδα ασφαλείας» σε περίπτωση μεταλλαγών και δυσλειτουργίας των 2 αυτών πρωτεϊνών στα μετάζωα.
- Οι πρωτεΐνες Ufe1p, Use1p και Dsl1p οι οποίες συμμετέχουν στην κυστιδιακή μεταφορά και την ομοτυπική σύντηξη μεμβρανών του ενδοπλασματικού δικτύου στον σακχαρομύκητα, λειτουργίες με κεντρικό ρόλο στο ευκαρυωτικό κύτταρο, εμφανίζουν μεγάλο βαθμό συντήρησης. Το γεγονός αυτό υποδηλώνει ότι τόσο η κυστιδιακή μεταφορά όσο και η ομοτυπική μεμβρανική σύντηξη είναι λειτουργίες πολύ συντηρημένες, η εμφάνιση των οποίων τοποθετείται στα πρώτα στάδια της εξέλιξης του ευκαρυωτικού κυττάρου.

- Οι πρωτεΐνες TBC1D20 και Yip3p οι οποίες συμμετέχουν στην ενδομεμβρανική κυκλοφορία, εμφανίζουν αντιπροσώπους σε όλους τους ευκαρυωτικούς οργανισμούς. Το γεγονός αυτό αποτελεί μια ακόμα ένδειξη για την «αρχαία» προέλευση της κυστιδιακής μεταφοράς.
- Η οικογένεια των πρωτεϊνών Vars εμφανίζει επίσης μεγάλο βαθμό συντήρησης. Οι πρωτεΐνες αυτές συμμετέχουν σε πολλές διαφορετικές λειτουργίες αλληλεπιδρώντας με πολλές διαφορετικές πρωτεΐνες εντός του ευκαρυωτικού κυττάρου. Δεν ήταν δυνατός ο διαχωρισμός των πρωτεϊνών VAP-A και VAP-B, στους οργανισμούς εκτός των μεταζώων.
- Οι Climp-63 και kinectin, που συμμετέχουν στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου και συγκεκριμένα των ασκών αυτού, εντοπίζονται κυρίως σε ορισμένα μετόζωα. Η λειτουργία την οποία επιτελούν είναι θεμελιώδης εντός του ευκαρυωτικού κυττάρου. Οι πρωτεΐνες αυτές όμως δεν είναι απαραίτητες για το σχηματισμό *per se* των ασκών του ενδοπλασματικού δικτύου. Αντίθετα τα reticulons και η REEP5/Yop1p εντοπίζονται στις άκρες των ασκών, προσδίδοντας τους την απαραίτητη καμπυλότητα. Συνεπώς, είναι πιθανό το σχήμα των ασκών του ενδοπλασματικού δικτύου να διαμορφώνεται μόνο από τα reticulons και την REEP5/Yop1p στους οργανισμούς εκτός των μεταζώων. Παράλληλα, η παρουσία των πρωτεϊνών Climp-63 και kinectin στα μετόζωα πιθανά να λειτουργεί ως «δικλείδα ασφαλείας». Για τις πρωτεΐνες αυτές βρέθηκαν και κάποια αμφίβολα αποτελέσματα σε οργανισμούς εκτός μετόζωων (*Trichoplax adhaerens*, *Selaginella moellendorffii* και *Euglena gracilis* για την Climp-63 και *Euglena gracilis* για την kinectin), τα οποία ελέγχθηκαν στη συνέχεια και μέσω πολλαπλής στοίχισης.
- Η πρωτεΐνη TMEM33 είναι συντηρημένη στα μετόζωα, τους μύκητες και τις αμοιβάδες και σε κάποιους κατώτερους ευκαρυωτικούς οργανισμούς. Η πρωτεΐνη αυτή θεωρείται ότι συμμετέχει στη διαμόρφωση της δομής των ασκών του ενδοπλασματικού δικτύου. Συνεπώς, με βάση και τα παραπάνω λειτουργεί πιθανά συμπληρωματικά με όλες τις υπόλοιπες πρωτεΐνες που προαναφέρθηκαν.
- Οι πρωτεΐνες lunapark και protrudin αποτελούν ρυθμιστές της πυκνότητας του σωληνοειδούς ενδοπλασματικού δικτύου καθώς και της αναλογίας των αγωγών και των ασκών του ενδοπλασματικού δικτύου. Η protrudin φαίνεται να είναι συντηρημένη στα μετόζωα και τα πλακόζωα, ενώ αντιπρόσωποι της πρωτεϊνικής οικογένειας της lunapark εμφανίζονται από τα μετόζωα και τα πλακόζωα μέχρι τις αμοιβάδες, τους μύκητες και τα φυτά.
- Πολλά αμφίβολα αποτελέσματα υπήρξαν για την πρωτεΐνη spastin. Πιο συγκεκριμένα, δεν ήταν δυνατός ο διαχωρισμός της πρωτεΐνης spastin από την πρωτεΐνη fidjetin-like protein 1 στους οργανισμούς εκτός των μετόζωων. Η τελευταία ανήκει στην οικογένεια των AAA-ATPases και είναι πολύ συγγενική με τη spastin.
- Για την πρωτεΐνη NogoR υπήρξαν πολλά αμφίβολα αποτελέσματα στους οργανισμούς εκτός μετόζωων. Το γεγονός αυτό πιθανά οφείλεται στην ύπαρξη επαναλήψεων λευκίνης στην αλληλουχία του ανθρώπινου

υποδοχέα του RTN4A, μοτίβο το οποίο εμφανίζεται σε ένα μεγάλο αριθμό πρωτεϊνών. Πιο συγκεκριμένα, από τις αναλύσεις reverse BLAST για όλες αυτές τις πρωτεΐνες προέκυψε ένας μεγάλος αριθμός πρωτεϊνών με κοινό χαρακτηριστικό την παρουσία επαναλήψεων LRR (Leucine Rich Repeat), οι οποίες έφεραν και παρόμοια E-value.

- Η πρωτεΐνη REEP1 με τη σειρά της ανήκει σε μια πρωτεϊνική οικογένεια η οποία εμφανίζει 6 μέλη στον άνθρωπο. Οι 6 αυτές πρωτεΐνες παρουσιάζουν μεγάλο βαθμό ομοιότητας μεταξύ τους με αποτέλεσμα ο διαχωρισμός τους να καθίσταται ιδιαίτερα δύσκολος όσο απομακρυνόμαστε εξελικτικά από τον άνθρωπο. Αυτό είχε ως αποτέλεσμα την εμφάνιση πολλών αμφίβολων αποτελεσμάτων στην προσπάθεια εύρεσης ομόλογων πρωτεϊνών της REEP1 στους ευκαρυώτες. Πιο συγκεκριμένα, σε πολλούς οργανισμούς αν και βρέθηκαν πρωτεΐνες οι οποίες φαίνεται να ανήκουν στην οικογένεια των REEPs, δεν ήταν δυνατό να προσδιοριστεί κατά πόσο εμφανίζουν ξεκάθαρη ομολογία με μία από τις REEP 1-6. Εξαίρεση αποτελούν οι οργανισμοί *M.musculus*, *D.rerio* και *X.tropicalis*.
- Για την πρωτεΐνη Caspr1 υπήρξαν αμφίβολα αποτελέσματα στον οργανισμό *Dictyostelium discoideum*, τα οποία δεν ήταν ξεκάθαρο κατά πόσο αποτελούν πραγματικά ορθόλογες αλληλουχίες της ανθρώπινης Caspr1. Επιπλέον, σε αρκετούς οργανισμούς εντός των μεταζώων δεν κατέστη δυνατός ο διαχωρισμός της Caspr1 από τα υπόλοιπα 4 μέλη που εμφανίζει η πρωτεϊνική αυτή οικογένεια στον άνθρωπο.
- Η πρωτεΐνη STIM1 που διαμεσολαβεί στην αλληλεπίδραση του ενδοπλασματικού δικτύου με τον κυτταροσκελετό είναι συντηρημένη κυρίως στα μετάρζωα. Επιπλέον, είναι γνωστό ότι οι πρωτεΐνες spastin, CLIMP-63 και REEP1 αλληλεπιδρούν με τον κυτταροσκελετό χωρίς να είναι γνωστό κατά πόσο αυτές οι αλληλεπιδράσεις επηρεάζουν την οργάνωση του ενδοπλασματικού δικτύου. Από τα αποτελέσματα μέχρι στιγμής της συγκριτικής γονιδιωματικής δεν είναι ξεκάθαρο κατά πόσο κάποια από αυτές τις πρωτεΐνες θα μπορούσε να επιτελεί τη λειτουργία της STIM1 σε όλους τους υπόλοιπους ευκαρυώτες εκτός των μεταζώων.

3.2 Σύγκριση αποτελεσμάτων με προηγούμενες αναλύσεις συγκριτικής γονιδιωματικής

Η πλειοψηφία των πρωτεϊνών που χρησιμοποιήθηκαν στην παρούσα διπλωματική εργασία είναι πρωτεΐνες για τις οποίες δεν είχαν πραγματοποιηθεί στο παρελθόν αναλύσεις συγκριτικής γονιδιωματικής και φυλογένεσης. Εξαίρεση αποτελούν οι πρωτεΐνες reticulons (Oertle et al., 2003), Use1p (Field et al., 2017) και Dsl1p (Klinger et al., 2013). Πιο συγκεκριμένα, για τις πρωτεΐνες Use1p και Dsl1p η αναζήτηση ορθόλογων πραγματοποιήθηκε με τη χρήση των προγραμμάτων BLAST και Hmmer, όπως στη συγκεκριμένη διπλωματική εργασία. Τα αποτελέσματα των αναλύσεων φαίνονται στον παρακάτω πίνακα (πίνακας 8). Για τις πρωτεΐνες reticulons η σύγκριση των αποτελεσμάτων θα πραγματοποιηθεί στην επόμενη ενότητα «Αποτελέσματα φυλογενετικής ανάλυσης».

Πίνακας 8. Σύγκριση των αποτελεσμάτων της συγκριτικής γονιδιωματικής της συγκεκριμένης διπλωματικής εργασίας με αποτελέσματα προηγούμενων μελετών για τις πρωτεΐνες Use1p και Dsl1p. Στα κόκκινα πλαίσια εντοπίζονται οι οργανισμοί για τους οποίους παρατηρούνται διαφωνίες ανάμεσα στις μελέτες. Στα πορτοκαλί πλαίσια εντοπίζονται οι νέες ορθόλογες πρωτεΐνες που προέκυψαν από την παρούσα διπλωματική εργασία.

| Υπαρξη πρωτεΐνης | | | | |
|-------------------------------------|-------------|--------------|-------------|----------------|
| Απουσία πρωτεΐνης | | | | |
| Οργανισμοί | Use1p | | Dsl1p | |
| | Διπλωματική | Vankatesh et | Διπλωματική | Klinger et al. |
| <i>Saccharomyces cerevisiae</i> | + | | + | |
| <i>Homo sapiens</i> | + | + | + | + |
| <i>Mus musculus</i> | + | | + | |
| <i>Xenopus tropicalis</i> | + | + | + | |
| <i>Danio rerio</i> | + | +2 | + | + |
| <i>Drosophila melanogaster</i> | + | + | + | + |
| <i>Caenorhabditis elegans</i> | + | | + | |
| <i>Nematostella vectensis</i> | + | + | + | + |
| <i>Trichoplax adhaerens</i> | + | + | + | |
| <i>Monosiga brevicollis</i> | - | - | + | + |
| <i>Fonticula alba</i> | - | | - | |
| <i>Allomyces macrogynus</i> | + | | + | |
| <i>Batrachomyxium dendrobatidis</i> | + | + | + | - |
| <i>Rozella allomycis</i> | + | + | - | |
| <i>Encephalitozoon cuniculi</i> | - | - | - | - |
| <i>Thecamonas trahens</i> | - | - | + | |
| <i>Acanthamoeba castellanii</i> | + | + | + | |
| <i>Dictyostelium discoideum</i> | + | + | + | + |
| <i>Entamoeba histolytica</i> | - | - | - | |
| <i>Arabidopsis thaliana</i> | +2 | +2 | + | + |
| <i>Physcomitrella patens</i> | + | + | + | + |
| <i>Selaginella moellendorffii</i> | + | | + | |
| <i>Chlamydomonas reinhardtii</i> | + | + | + | + |
| <i>Ostreococcus tauri</i> | + | + | + | + |
| <i>Cyanidioschyzon merolae</i> | - | | - | |
| <i>Cyanophora paradoxa</i> | - | | - | |
| <i>Aureococcus anophagefferens</i> | - | | - | |
| <i>Ectocarpus siliculosus</i> | + | + | + | |
| <i>Phytophthora sojae</i> | - | | + | + |
| <i>Thalassiosira pseudonana</i> | - | - | + | + |
| <i>Cryptosporidium parvum</i> | - | - | + | + |
| <i>Paramecium tetraurelia</i> | + | + | - | |
| <i>Plasmodium falciparum</i> | - | + | - | |
| <i>Tetrahymena thermophila</i> | - | - | - | |
| <i>Theileria parva</i> | - | | - | |
| <i>Chromera velia</i> | - | + | + | |
| <i>Vitrella brassicaformis</i> | + | | + | |
| <i>Bigelowiella natans</i> | + | + | - | + |
| <i>Reticulomyxa filosa</i> | - | | + | |
| <i>Emiliania huxleyi</i> | - | - | - | + |
| <i>Guillardia theta</i> | + | | + | + |
| <i>Bodo saltans</i> | - | - | - | |
| <i>Giardia intestinalis</i> | - | - | - | |
| <i>Leishmania major</i> | - | - | - | |
| <i>Trypanosoma brucei</i> | - | - | - | |
| <i>Trypanosoma cruzi</i> | - | | - | |
| <i>Euglena gracilis</i> | - | | + | |
| <i>Naegleria gruberi</i> | - | - | + | + |
| <i>Trichomonas vaginalis</i> | + | + | - | |
| <i>Monocercomonoides exilis</i> | - | | - | |

Πιο συγκεκριμένα, από τον πίνακα 6 παρατηρούμε αρχικά ότι στα πλαίσια της συγκεκριμένης ανάλυσης πραγματοποιήθηκε έλεγχος σε μεγαλύτερο αριθμό οργανισμών με αποτέλεσμα να προκύψουν κάποιες επιπλέον ορθόλογες πρωτεΐνες (πορτοκαλί πλαίσια). Παράλληλα είναι εμφανές ότι υπάρχει υψηλός βαθμός συμφωνίας των αποτελεσμάτων ανάμεσα στις μελέτες. Παρατηρούνται διαφορές μόνο στους οργανισμούς *Plasmodium falciparum* και *Chromera velia* για την Use1p και

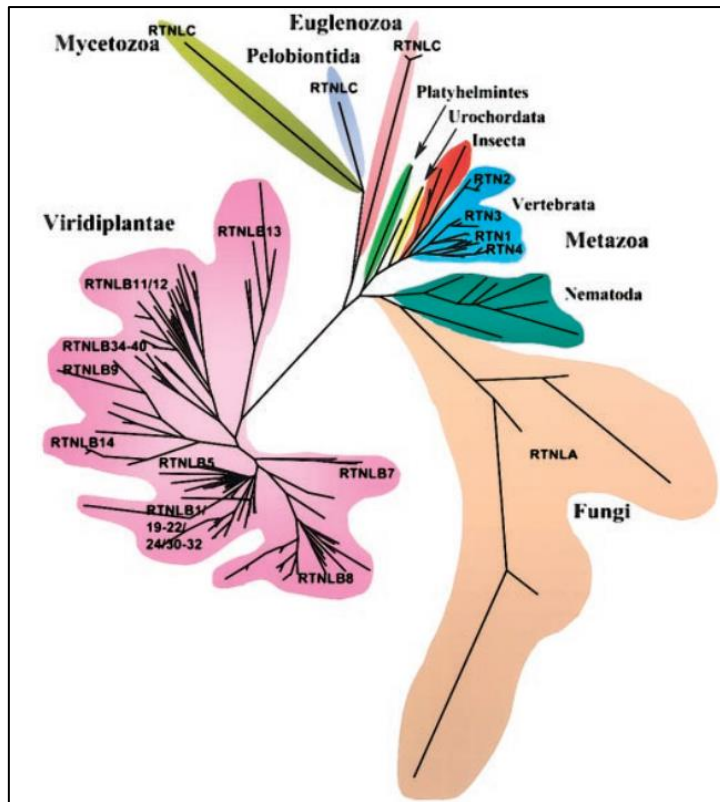
Batrachochytrium dendrobatidis, *Bigelowiella natans* και *Emiliana huxleyi* για την Dsl1p. Δεδομένου ότι η μεθοδολογία που χρησιμοποιήθηκε ήταν κοινή για όλες τις μελέτες συμπεραίνουμε ότι οι διαφορές αυτές οφείλονται στο βασικό στατιστικό κριτήριο αξιολόγησης των αποτελεσμάτων που είναι το E-value. Η τιμή του E-value εξαρτάται από το μέγεθος της βάσης δεδομένων που χρησιμοποιείται. Συνεπώς, η χρησιμοποίηση διαφορετικών βάσεων δεδομένων καθώς και ο διαφορετικός χρόνος πραγματοποίησης των αναλύσεων οδήγησαν πιθανά στις διαφορές που παρατηρούνται. Σε κάθε περίπτωση όμως από όλες τις αναλύσεις προκύπτει ένα κοινό συμπέρασμα που αφορά στον υψηλό βαθμό συντήρησης των πρωτεϊνών Use1p και Dsl1p στους ευκαρυωτικούς οργανισμούς.

3.3 Αποτελέσματα φυλογενετικής ανάλυσης

Για κάποιες πρωτεΐνες οι οποίες εμφανίζουν μεγάλο βαθμό συντήρησης σε ολόκληρο το φάσμα των ευκαρυωτικών οργανισμών, όπως τα μέλη της οικογένειας των reticulons, οι πρωτεΐνες REEP 1 και REEP 5, η atlastin/Sey1p, η TMEM33 και οι Vaps, κατασκευάστηκαν φυλογενετικά δέντρα. Επιπλέον, η κατασκευή φυλογενετικών δέντρων κρίθηκε απαραίτητη και για τις πρωτεΐνες NogoR και spastin, οι οποίες εμφάνιζαν ένα ή περισσότερα αμφίβολα αποτελέσματα. Η διαλεύκανση των αμφίβολων αποτελεσμάτων για τις Climp-63, kinectin και Caspr1 πραγματοποιήθηκε μέσω πολλαπλής στοίχισης. Τα αποτελέσματα των αναλύσεων παρατίθενται στη συνέχεια.

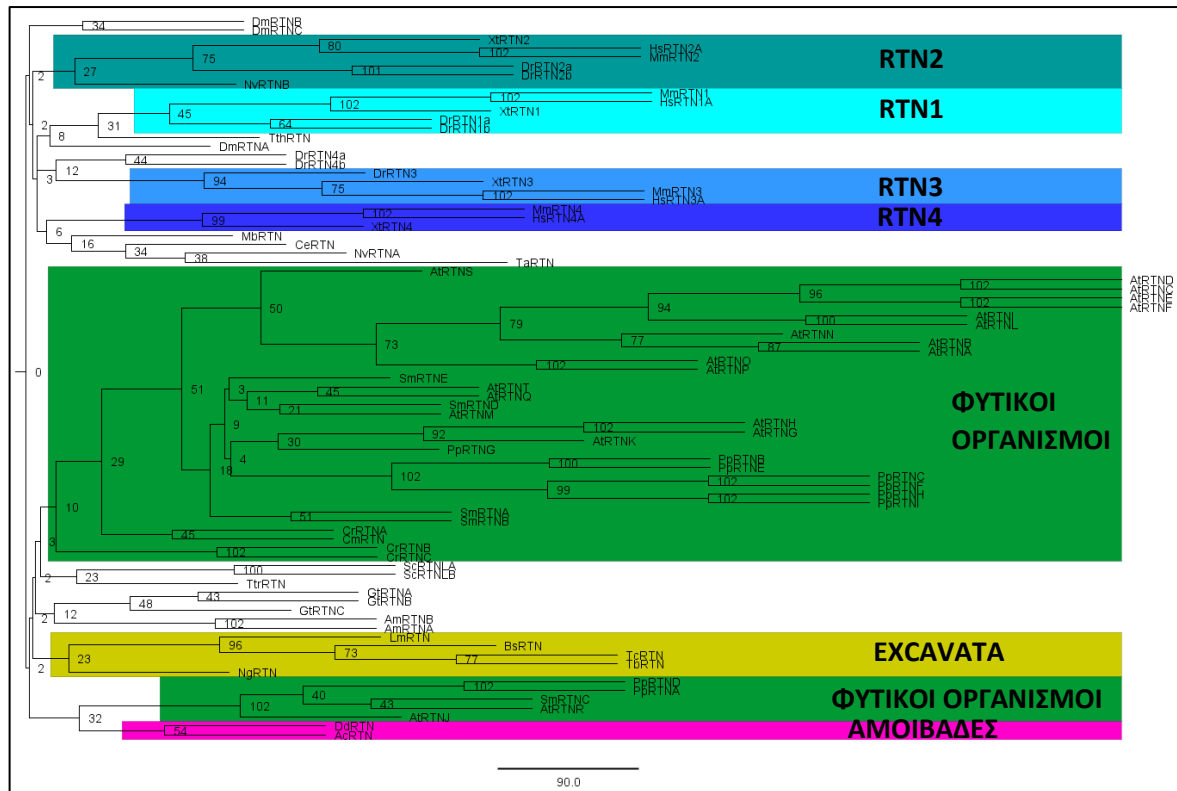
3.3.1 Reticulons

Φυλογενετική ανάλυση πραγματοποιήθηκε για την οικογένεια των reticulons, ώστε να προσδιοριστούν οι ακριβείς σχέσεις μεταξύ των πρωτεϊνών-μελών της. Τα αποτελέσματα αυτής της ανάλυσης συγκρίθηκαν με μια παρόμοια παλαιότερη φυλογενετική μελέτη η οποία πραγματοποιήθηκε το 2003 (Oertle et al., 2003). Σε αυτήν την μελέτη όπως φαίνεται από την εικόνα 33, παρατηρούμε ότι υπάρχει υπερεκπροσώπηση των ανώτερων ευκαρυωτικών οργανισμών και συγκεκριμένα των ζώων και των φυτών. Ο λόγος που αυτό συνέβη είναι ο μικρός αριθμός των αλληλουχημένων γονιδιωμάτων που υπήρχαν την περίοδο πραγματοποίησης της συγκεκριμένης ανάλυσης. Αυτό έχει ως αποτέλεσμα το φυλογενετικό δέντρο το οποίο είχε κατασκευαστεί να μην αντιπροσωπευτικό των πρωτεϊνών-μελών της οικογένειας των reticulons σήμερα.



Εικόνα 33. Δενδρόγραμμα στο οποίο είναι εμφανής η κατανομή των γονιδίων των reticulons στα διαφορετικά βασίλεια των ευκαρυωτικών οργανισμών (με βάση παλαιότερη ταξινόμηση των ευκαρυωτικών οργανισμών). Η αναζήτηση των παράλογων γονιδίων πραγματοποιήθηκε με κριτήριο την ύπαρξη και την στοίχιση της δομικής επικράτειας reticulon homology domain (RHD). Παρατηρούμε ότι αντιπρόσωποι των reticulons υπάρχουν στα περισσότερα ευκαρυωτικά βασίλεια από τα πρώιμα μέχρι τα ζώα (Oertle et al., 2003).

Αντίθετα στην εικόνα 34 παρατίθεται το φυλογενετικό δέντρο το οποίο κατασκευάστηκε στα πλαίσια της συγκεκριμένης διπλωματικής εργασίας.



Εικόνα 34. Φυλογενετικό δέντρο των reticulons. Οι αριθμοί αντιπροσωπεύουν την τιμή των bootstraps που στηρίζουν κάθε κλάδο. Στο όνομα κάθε πρωτεΐνης τα 2 ή 3 πρώτα γράμματα περιγράφουν τον οργανισμό από τον οποίο προέρχεται, σύμφωνα με τον πίνακα 15 του Παραρτήματος, ενώ οι επόμενοι 3 χαρακτήρες περιγράφουν την πρωτεΐνη RTN: reticulon.

Από το φυλογενετικό δέντρο της εικόνας 34 παρατηρούμε ότι έχουν σχηματιστεί 8 διακριτές ομάδες πρωτεϊνών. Πιο συγκεκριμένα:

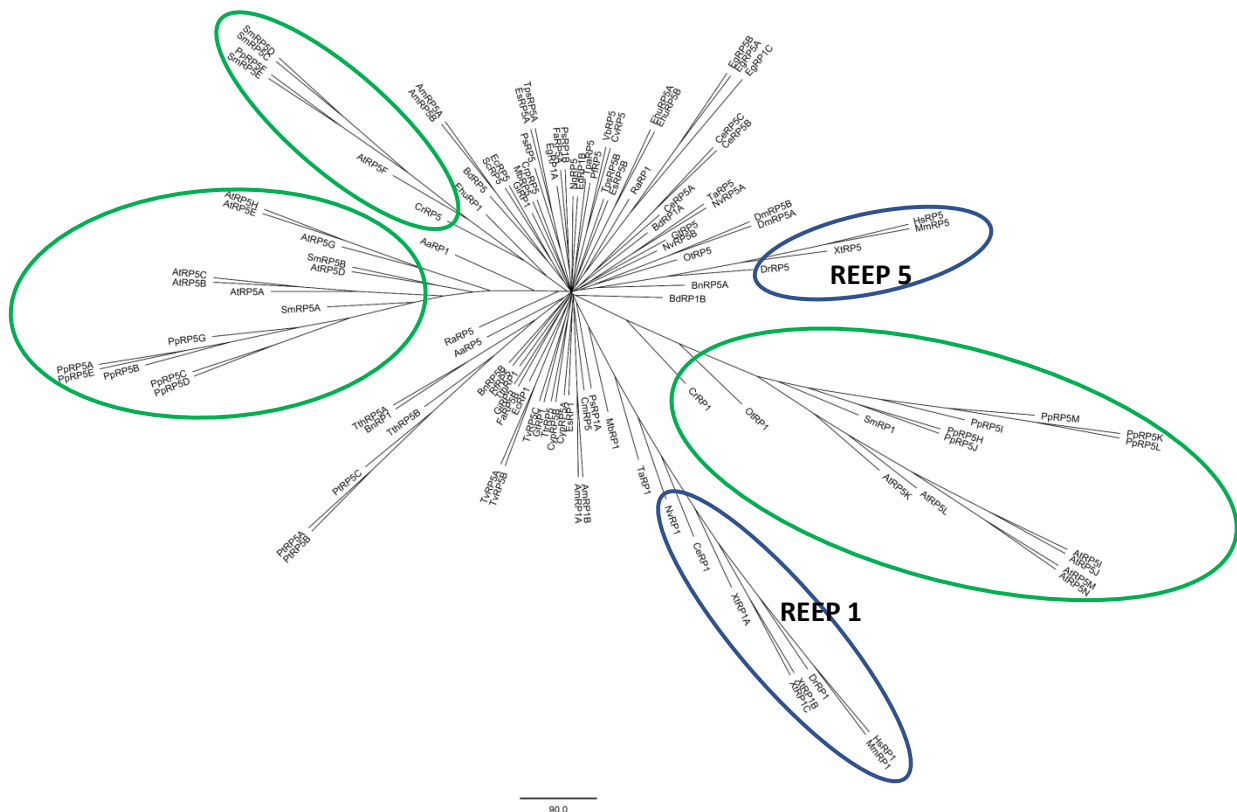
- Στους φυτικούς οργανισμούς οι οποίοι έχουν διαχωριστεί από όλους τους υπόλοιπους και καταλαμβάνουν δύο ξεχωριστούς κλάδους, φαίνεται να έχουν πραγματοποιηθεί πολλοί εσωτερικοί διπλασιασμοί με αποτέλεσμα κάθε οργανισμός να έχει αποκτήσει τελικά ένα μεγάλο αριθμό πρωτεϊνών reticulons.
- Οι πρωτεΐνες reticulons που υπάρχουν τόσο στους μύκητες όσο και στις αμοιβάδες έχουν χωριστεί σε δύο διακριτούς κλάδους.
- Τέσσερις πρωτεΐνες RTN 1-4 είναι παρούσες στους οργανισμούς *H.sapiens*, *M.musculus*, *X.tropicalis* και *D.rerio*. Δηλαδή, ο διαχωρισμός των πρωτεϊνών αυτών πρέπει να πραγματοποιήθηκε μετά από διπλασιασμούς στον κοινό πρόγονο των 4 αυτών οργανισμών.
- Στους οργανισμούς *A.macrogynus*, *S.cerevisiae* και *Guillardia theta* έχουν πραγματοποιηθεί εσωτερικοί διπλασιασμοί με αποτέλεσμα να προκύψουν 2 ή 3 πρωτεΐνες reticulon ανά είδος.
- Σε όλους τους υπόλοιπους οργανισμούς που παρατίθενται στο φυλογενετικό δέντρο υπάρχει από μια πρωτεΐνη reticulon

Η απουσία αντιπροσώπου των reticulons από τους προκαρυωτικούς οργανισμούς έχει οδηγήσει στο συμπέρασμα ότι η εξέλιξη των reticulons πραγματοποιήθηκε παράλληλα με την εξέλιξη των ευκαρυωτικών

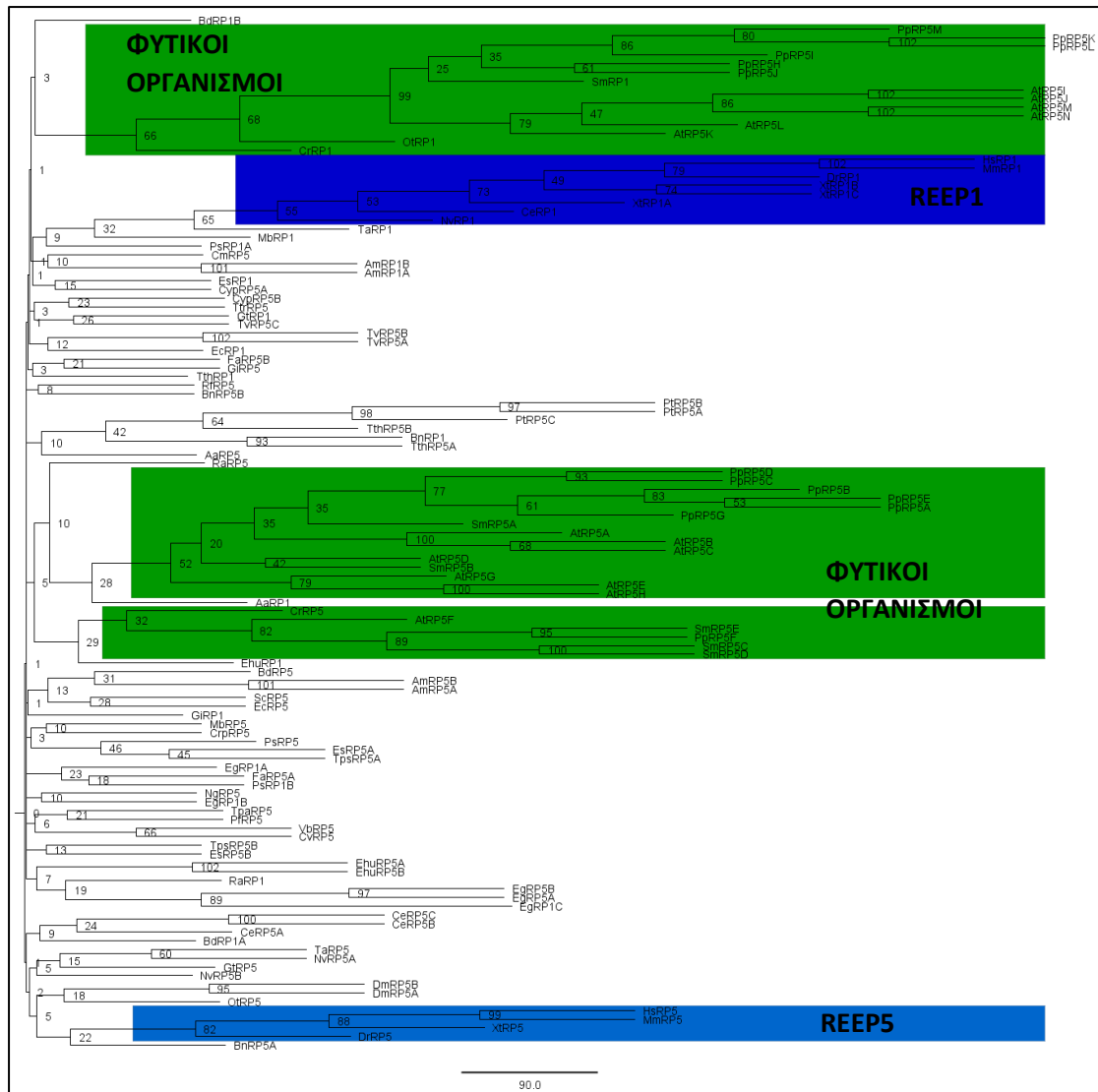
οργανισμών. Πιο συγκεκριμένα, ο βασικός δομικός ρόλος των reticulons στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου σε συνδυασμό με τον υψηλό βαθμό συντήρησης τους στους ευκαρυωτικούς οργανισμούς έχουν τοποθετήσει την εξέλιξη των πρωτεϊνών αυτών παράλληλα με την εξέλιξη του ενδοπλασματικού δικτύου των ευκαρυωτικών οργανισμών. Συμπληρωματικά, σε αυτά τα δεδομένα μπορούμε να προσθέσουμε με βάση τη φυλογενετική ανάλυση που πραγματοποιήθηκε στα πλαίσια της συγκεκριμένης μελέτης ότι στον κοινό πρόγονο των ευκαρυωτικών οργανισμών υπήρχε ένας αντιπρόσωπος της οικογένειας των reticulons. Στη διάρκεια της εξέλιξης φαίνεται να υπήρξαν εσωτερικοί διπλασιασμοί του αρχικού γονιδίου σε αρκετούς οργανισμούς, με το φαινόμενο αυτό να είναι εμφανώς πιο έντονο στα φυτά.

3.3.2 REEPs

Στη συγκεκριμένη μελέτη έγινε έλεγχος της ύπαρξης ορθόλογων γονιδίων για δύο πρωτεΐνες από την οικογένεια των πρωτεϊνών REEP. Σε πολλούς οργανισμούς παρουσιάστηκε αδυναμία προσδιορισμού ξεκάθαρης ομολογίας με μία μόνο από τις ανθρώπινες REEP 1-6. Για το λόγο αυτό κατασκευάστηκε φυλογενετικό δέντρο (εικόνες 35,36).



Εικόνα 35. Φυλογενετικό δέντρων των πρωτεϊνών της οικογένειας των REEPs. Διακρίνονται 2 με 3 ομάδες πρωτεϊνών REEPs στους φυτικούς οργανισμούς (πράσινα πλαίσια). Ο διαχωρισμός ανάμεσα στις πρωτεΐνες REEP1 και REEP5 είναι δυνατός μόνο εντός των μεταζώων (μπλε πλαίσια). Στο όνομα κάθε πρωτεΐνης τα 2 ή 3 πρώτα γράμματα περιγράφουν τον οργανισμό από τον οποίο προέρχεται, σύμφωνα με τον πίνακα 15 του Παραρτήματος, ενώ οι επόμενοι 5 χαρακτήρες περιγράφουν την πρωτεΐνη REEP1 ή REEP5.



Εικόνα 36. Φυλογενετικό δέντρο των πρωτεϊνών της οικογένειας των REEPs. Διακρίνονται 2 με 3 ομάδες πρωτεϊνών REEPs στους φυτικούς οργανισμούς (πράσινα πλαίσια). Ο διαχωρισμός ανάμεσα στις πρωτεΐνες REEP1 και REEP5 είναι δυνατός μόνο εντός των μεταζώων (μπλε πλαίσια). Τα νούμερα αντιπροσωπεύουν τον αριθμό των bootstrap που στηρίζουν κάθε κλάδο. Στο όνομα κάθε πρωτεΐνης τα 2 ή 3 πρώτα γράμματα περιγράφουν τον οργανισμό από τον οποίο προέρχεται, σύμφωνα με τον πίνακα 15 του Παραρτήματος, ενώ οι επόμενοι 5 χαρακτήρες περιγράφουν την πρωτεΐνη REEP1 ή REEP5.

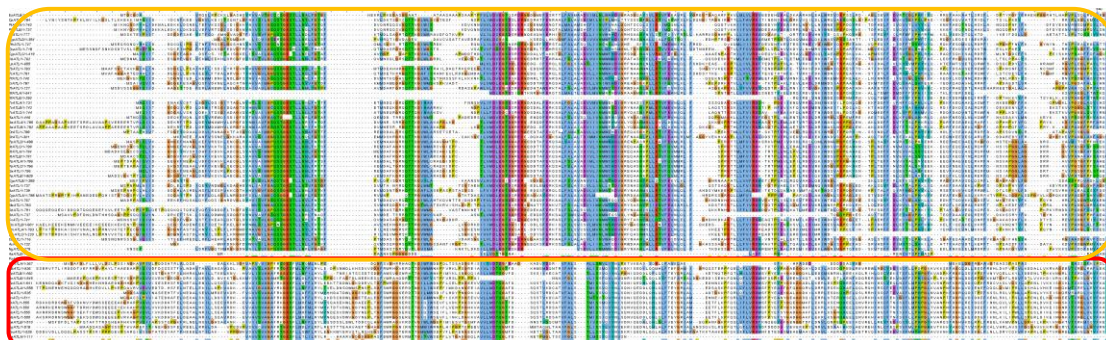
Από τα αποτελέσματα (εικόνες 35, 36) αρχικά παρατηρούμε ότι παρόλο που επιβεβαιώνεται ότι όλες οι πρωτεΐνες που βρέθηκαν μέσω της συγκριτικής γονιδιωματικής ανάλυσης είναι μέλη της οικογένειας των REEPs, δεν υπάρχει καλή στήριξη των φυλογενετικών σχέσεων στο δέντρο (μικρός αριθμός των bootstrap στα περισσότερα κλαδιά). Αυτό συμβαίνει πιθανά λόγω του υψηλού βαθμού ομοιότητας μεταξύ των πρωτεϊνών μελών αυτής της οικογένειας ο οποίος καθιστά δύσκολο το διαχωρισμό τους, όπως είχε παρατηρηθεί ήδη από τις αναλύσεις BLAST και Hmmer που προηγήθηκαν. Ο διαχωρισμός των πρωτεϊνών REEP 1 και 5 είναι δυνατός μόνο εντός των μεταζώων (μπλε πλαίσια στις εικόνες 35 και 36). Επιπλέον, όπως και στα reticulons παρατηρούμε ότι υπάρχει μεγάλος αριθμός εσωτερικών

διπλασιασμών στους φυτικούς οργανισμούς, με αποτέλεσμα να δημιουργούνται 2-3 ομάδες πρωτεϊνών REEP. Για όλες τις υπόλοιπες πρωτεΐνες όμως δεν μπορούν να εξαχθούν ασφαλή συμπεράσματα. Για το σκοπό αυτό θα ήταν απαραίτητη η εφαρμογή μεθόδων συγκριτικής γονιδιωματικής για όλα τα μέλη της οικογένειας των REEPs.

3.3.3 Atlastin 1/ Sey1p

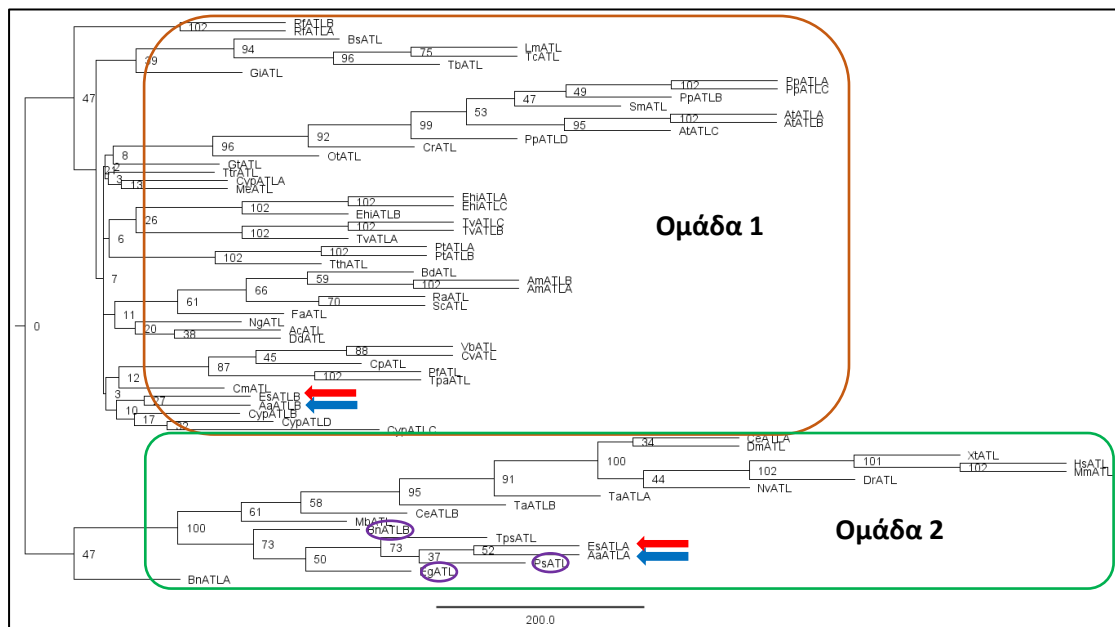
Από τα αποτελέσματα των αναλύσεων που προηγήθηκαν φαίνεται ότι η οικογένεια των πρωτεϊνών atlastin 1 / Sey1p είναι συντηρημένη στους ευκαρυωτικούς οργανισμούς. Από τη βιβλιογραφία είναι γνωστό ότι οι πρωτεΐνες atlastin 1 και Sey1p δεν εμφανίζουν μεγάλη ομοιότητα ως προς την αλληλουχία, αλλά θεωρούνται λειτουργικά ομόλογες, με την atlastin 1 να εμφανίζεται κυρίως στα μετάζωα και την Sey1p να υπάρχει σε όλους τους υπόλοιπους ευκαρυώτες. Τα αποτελέσματα των αναλύσεων συγκριτικής γονιδιωματικής και φυλογένεσης που πραγματοποιήθηκαν επιβεβαιώνουν την παραπάνω διαπίστωση. Πιο συγκεκριμένα, τόσο από τα αποτελέσματα της στοίχισης των αλληλουχιών, όσο και από το φυλογενετικό δέντρο παρατηρούμε ότι οι πρωτεΐνες διακρίνονται σε δύο κατηγορίες ανάλογα με το βαθμό ομολογίας τους ως προς τις atlastin 1 και Sey1p (εικόνες 37, 38). Στο φυλογενετικό δέντρο (εικόνα 38) έχουν δημιουργηθεί δύο διακριτοί κλάδοι πρωτεϊνών με μεγάλη απόσταση μεταξύ τους. Παράλληλα, παρατηρούμε ότι κατά κανόνα σε κάθε οργανισμό το μέλος της πρωτεϊνικής οικογένειας το οποίο εμφανίζεται είναι ομόλογο είτε προς την atlastin 1 είτε προς την Sey1p. Εξαιρέση αποτελούν οι οργανισμοί *Aureococcus anophagefferens*, *Ectocarpus siliculosus*. Για καθέναν από αυτούς παρατηρούμε ότι οι δύο πρωτεΐνες οι οποίες βρέθηκαν μέσω της γονιδιωματικής ανάλυσης, εντοπίζονται τόσο στον κλάδο των ομολόγων της atlastin 1, όσο και στον κλάδο των ομολόγων της Sey1p.

Από τη συγκριτική γονιδιωματική ανάλυση προέκυψαν κάποιες πρωτεϊνικές ακολουθίες για τις οποίες δεν ήταν ξεκάθαρο κατά πόσο είναι ομόλογες με την atlastin 1 ή κάποια άλλη από τις ανθρώπινες atlastins (atlastin 2 ή 3) (εικόνα 38 μωβ κύκλοι). Κάτι τέτοιο δεν ήταν δυνατό να διευκρινιστεί ούτε από τη φυλογενετική ανάλυση. Σημειώνεται ότι τόσο η atlastin 1 όσο και η atlastin 3 συμμετέχουν στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου. Συνεπώς, συμπεραίνουμε ότι οι οργανισμοί *Phytophthora sojae*, *Thalassiosira pseudonana*, *Bigelowiella natans* και *Euglena gracilis* διαθέτουν μέλος της πρωτεϊνικής οικογένειας των atlastins, το οποίο παίζει ρόλο στη διαμόρφωση της δομής του ενδοπλασματικού δικτύου.





Εικόνα 37. Στοίχιση των πρωτεϊνών της οικογένειας atlastin/Sey1p, μετά από περικοπή πολλαπλών κενών. Παρατηρούμε ότι μέσα στο κόκκινο πλαίσιο βρίσκονται πρωτεΐνες που είναι ομόλογες με την ανθρώπινη atlastin 1, ενώ στο πορτοκαλί πρωτεΐνες ομόλογες με την Sey1p. Όπως είναι ορατό οι πρωτεΐνες που είναι ομόλογες προς την Sey1p είναι μεγαλύτερες σε μέγεθος. Επιπλέον, η στοίχιση είναι καλύτερη στο αμινοτελικό άκρο όλων των πρωτεϊνών. Αυτό οφείλεται στην ύπαρξη μιας επικράτειας P-loop NTPase τόσο στην ανθρώπινη atlastin 1, όσο και στην Sey1p του *S.cerevisiae*.



Εικόνα 38. Φυλογενετικό δέντρο πρωτεϊνικής οικογένειας atlastin/sey1p. Στο δέντρο διακρίνονται δύο ομάδες πρωτεϊνών με την ομάδα 1 να απαρτίζεται από πρωτεΐνες ομόλογες προς την Sey1p του σακχαρομύκητα και την ομάδα 2 από πρωτεΐνες ομόλογες προς την ανθρώπινη atlastin 1. Με τα κόκκινα και μπλε βέλη σημειώνονται οι πρωτεΐνες atlastin και sey1p των οργανισμών *E.siliculosus* και *A.anophagefferens* αντίστοιχα. Μέσα στους μωβ κύκλους βρίσκονται οι πρωτεΐνες οι οποίες δεν εμφάνισαν ξεκάθαρη ομολογία για μια από τις 3 ανθρώπινες atlastins. Στο όνομα κάθε πρωτεΐνης τα 2 ή 3 πρώτα γράμματα περιγράφουν τον οργανισμό από τον οποίο

προέρχεται, σύμφωνα με τον πίνακα 15 του Παραρτήματος, ενώ οι επόμενοι 3 χαρακτήρες περιγράφουν την πρωτεΐνη ATL: atlastin.

3.3.4 Spastin

Η πρωτεΐνη spastin ανήκει σε μια μεγάλη οικογένεια πρωτεϊνών που ονομάζονται AAA ATPases, χαρακτηριστικό της οποίας είναι η ύπαρξη μίας ή περισσότερων επικρατειών ATPάσης. Από τις αναλύσεις που πραγματοποιήθηκαν με τα προγράμματα BLAST και Hmmer δεν ήταν δυνατή η εξαγωγή ασφαλών συμπερασμάτων για την παρουσία της πρωτεΐνης spastin, στους περισσότερους ευκαρυωτικούς οργανισμούς. Αυτό συμβαίνει λόγω του μεγάλου βαθμού ομοιότητας της πρωτεΐνης αυτής με άλλα μέλη της οικογένειας AAA και ιδίως με την πρωτεΐνη fidgetin-like protein 1, με την οποία μοιράζεται μεγάλο ποσοστό ομοιότητας στον άνθρωπο της τάξης του 44%. Για το λόγο αυτό κατασκευάστηκε φυλογενετικό δέντρο για όλα τα αποτελέσματα τα οποία συγκεντρώθηκαν. Για την κατασκευή του δέντρου χρησιμοποιήθηκε ως εξωομάδα η ανθρώπινη πρωτεΐνη fidgetin-like protein 1.



Εικόνα 39. Φυλογενετικό δέντρο πρωτεΐνης spastin. Για την κατασκευή του δέντρου χρησιμοποιήθηκε ως εξωομάδα η ανθρώπινη fidgetin-like protein 1 (κόκκινο βέλος). Στο μπλε πλαίσιο βρίσκονται οι πρωτεΐνες που είναι ομόλογες με την ανθρώπινη spastin (μετάζωα και πλακόζωα) και στο πράσινο πλαίσιο οι πρωτεΐνες που είναι ομόλογες με την ανθρώπινη fidgetin-like protein 1. Για τις υπόλοιπες πρωτεΐνες δεν μπορούν να εξαχθούν ασφαλή συμπεράσματα. Στο όνομα κάθε πρωτεΐνης τα 2 ή 3 πρώτα γράμματα περιγράφουν τον οργανισμό από τον οποίο προέρχεται, σύμφωνα με τον πίνακα 15 του Παραρτήματος, ενώ οι επόμενοι 2 ή 4 χαρακτήρες περιγράφουν την πρωτεΐνη SP: spastin ή FLP1: fidgetin-like protein 1.

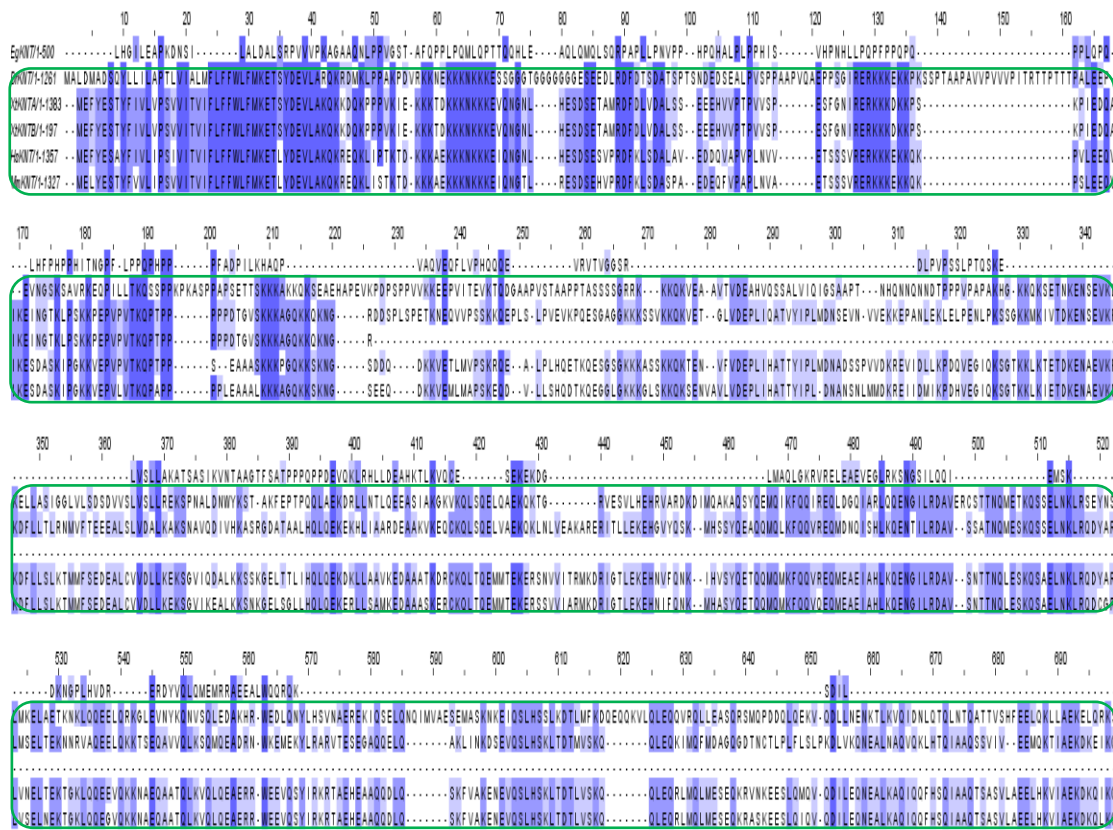
Από τα αποτελέσματα παρατηρούμε ότι και πάλι δεν είναι δυνατό να προσδιοριστεί με βεβαιότητα ποιες πρωτεΐνες ανήκουν στην οικογένεια των πρωτεϊνών spastin (εικόνα 39). Εξαιρεση αποτελούν τα μετάζωα και τα πλακόζωα. Επιπλέον, από την εικόνα 39 διακρίνεται και ένας κλάδος με πρωτεΐνες ομόλογες με την ανθρώπινη fidgetin-like protein 1 (πράσινο πλαίσιο). Με βάση αυτά τα αποτελέσματα δεν μπορούμε να συνάγουμε κατά

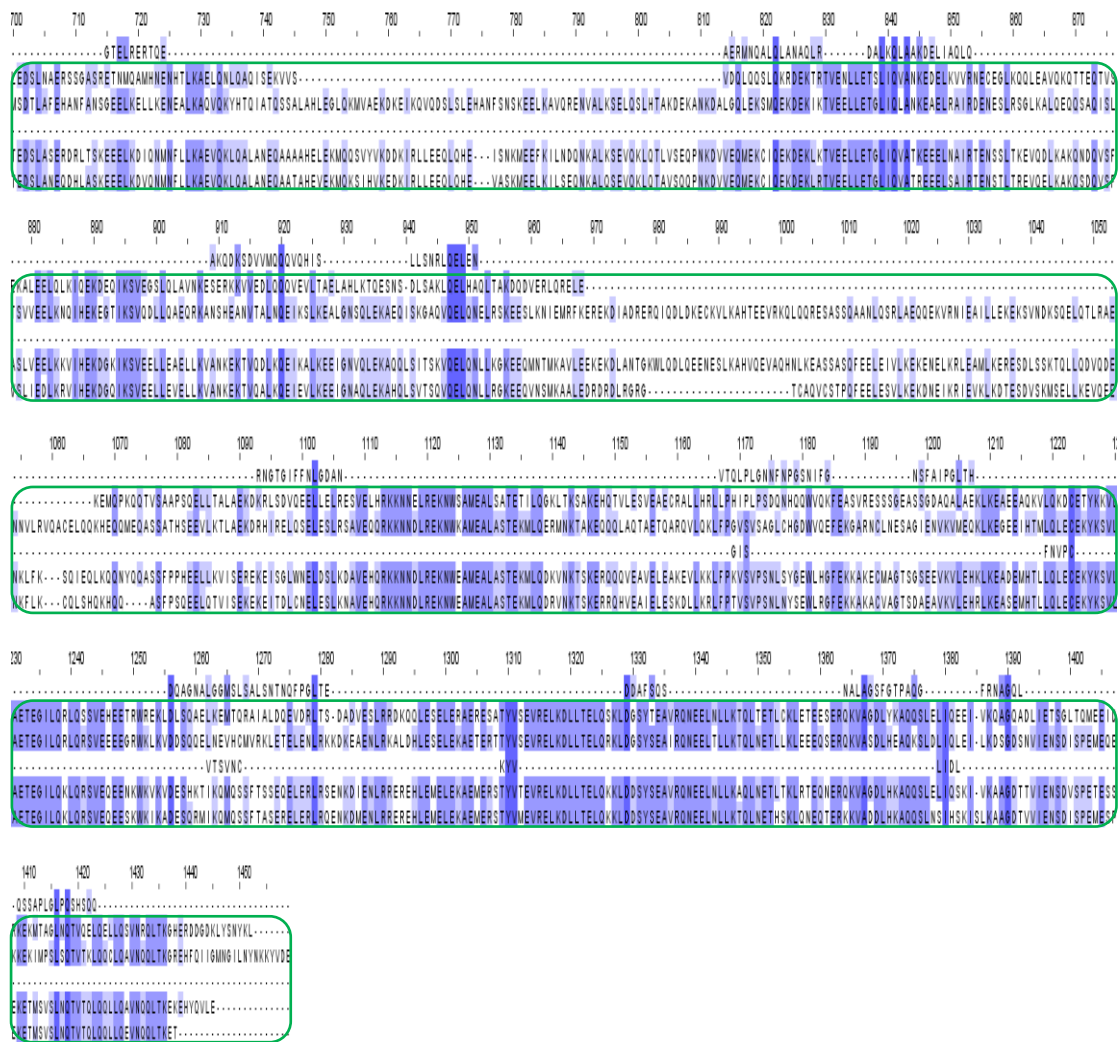
υποδηλώνει το % ποσοστό ομοιότητας μεταξύ των αλληλουχιών, με την αυξανόμενη ένταση του χρώματος να αντιστοιχεί: σε άσπρο <= 40% ομοιότητα, ανοιχτό γκρι > 40% ομοιότητα, ανοιχτό μπλε >60% ομοιότητα, σκούρο μπλε >80% ομοιότητα. Στο όνομα κάθε πρωτεΐνης τα 2 ή 3 πρώτα γράμματα περιγράφουν τον οργανισμό από τον οποίο προέρχεται, σύμφωνα με τον πίνακα 15 του Παραρτήματος, ενώ οι επόμενοι 5 χαρακτήρες περιγράφουν την πρωτεΐνη CLIMP: CLIMP-63.

Από την εικόνα 40 είναι φανερό ότι το ποσοστό ομοιότητας είναι υψηλό μόνο ανάμεσα στις 4 πρωτεΐνες των μεταζώων (πράσινο πλαίσιο). Η διαπίστωση αυτή σε συνδυασμό με τις προηγούμενες παρατηρήσεις οδηγεί στο συμπέρασμα ότι οι οργανισμοί *Trichoplax adhaerens*, *Selaginella moellendorffii* και *Euglena gracilis* δε διαθέτουν πραγματικά ομόλογες αλληλουχίες της ανθρώπινης CLIMP-63.

3.3.6 Kinectin

Από τα αποτελέσματα των προηγούμενων αναλύσεων βρέθηκαν 5 ομόλογες πρωτεΐνες της kinectin στους ευκαρυωτικούς οργανισμούς. Πιο συγκεκριμένα, βρέθηκαν 5 στα μετόζωα και άλλη 1 στον οργανισμό *Euglena gracilis*. Το τελευταίο αποτελέσματα όμως, όπως και στην περίπτωση της CLIMP-63, ήταν αρκετά αμφίβολο, διότι προέκυψε μόνο από την αναζήτηση με τον αλγόριθμο BLAST, χωρίς να στηρίζεται από ισχυρό E-value. Για το λόγο αυτό πραγματοποιήθηκε στοίχιση όλων των πρωτεϊνικών αλληλουχιών που βρέθηκαν και τα αποτελέσματα παρατίθενται στην εικόνα 41.





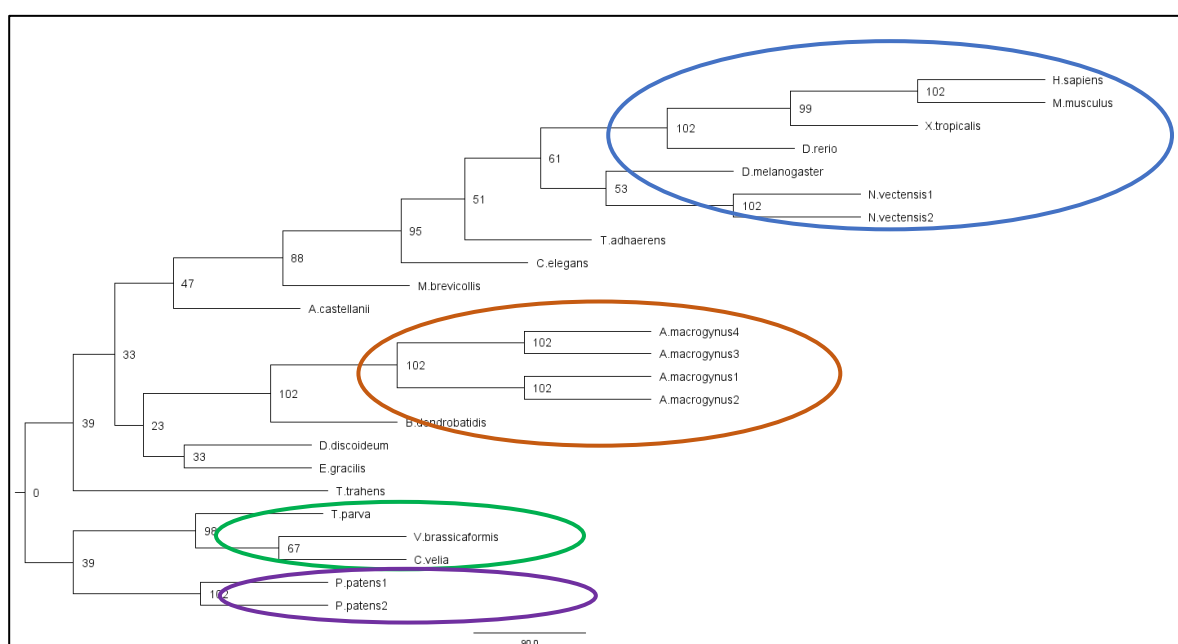
Εικόνα 41. Πολλαπλή στοίχιση πρωτεϊνικών αλληλουχιών που προέκυψαν από την αναζήτηση ορθολόγων με τη χρήση του BLAST και του Hmmer. Μέσα στο πράσινο πλαίσιο βρίσκονται οι πρωτεΐνες που εντοπίστηκαν στα μεταξύα. Για την οπτικοποίηση της στοίχισης χρησιμοποιήθηκε το πρόγραμμα Jalview (<http://www.jalview.org/>). Το χρώμα στα αμινοξικά κατάλοιπα υποδηλώνει το % ποσοστό ομοιότητας μεταξύ των αλληλουχιών, με την αυξανόμενη ένταση του χρώματος να αντιστοιχεί: σε άσπρο $\leq 40\%$ ομοιότητα, ανοιχτό γκρι $> 40\%$ ομοιότητα, ανοιχτό μπλε $>60\%$ ομοιότητα, σκούρο μπλε $>80\%$ ομοιότητα. Στο όνομα κάθε πρωτεΐνης τα 2 ή 3 πρώτα γράμματα περιγράφουν τον οργανισμό από τον οποίο προέρχεται, σύμφωνα με τον πίνακα 15 του Παραρτήματος, ενώ οι επόμενοι 3 χαρακτήρες περιγράφουν την πρωτεΐνη KNT: kinectin.

Από τα αποτελέσματα της στοίχισης παρατηρούμε ότι το ποσοστό ομοιότητας των αμινοξικών καταλοίπων μεταξύ των πρωτεϊνικών ακολουθιών των μεταξύων είναι πολύ μεγάλο, πράγμα που δεν συμβαίνει με την πρωτεϊνική ακολουθία από την *E. Gracilis*. Συνεπώς, η πρωτεΐνη αυτή δεν είναι ομόλογη με την ανθρώπινη kinectin.

3.3.7 TMEM33

Η πρωτεΐνη TMEM33 εμφανίζει αντιπροσώπους σε πολλές ομάδες ευκαρυωτικών οργανισμών. Από το φυλογενετικό δέντρο που κατασκευάστηκε για την πρωτεΐνη αυτή (εικόνα 42) παρατηρούμε ότι οι οργανισμοί διαχωρίζονται σε δύο βασικούς κλάδους. Ο πρώτος περιλαμβάνει τα οπισθόκοντα, τις αμοιβάδες, τα aruzozoa και τα excavata, ενώ ο δεύτερος

τα φυτά και τα alveolata. Επιπλέον, όλοι οι οργανισμοί κατά κανόνα περιέχουν μια πρωτεΐνη μέλος της οικογένειας των TMEM33. Εξαιρέση αποτελούν η *Nematostella vectensis* με 2, η *Allomyces macrogynus* με 4 και η *Physcomitrella patens* με 2. Παράλληλα παρατηρούμε η TMEM33 εμφανίζει ευρεία εξάπλωση στα οπισθόκοντα, τους μύκητες και τις αμοιβάδες. Κάτι τέτοιο όμως δε συμβαίνει με τους κατώτερους ευκαρυωτικούς οργανισμούς που ανήκουν στις ομάδες των SAR και excavata, από τους οποίους μόνο 4 συνολικά οργανισμοί εμφανίζουν ορθόλογες ως προς την ανθρώπινη TMEM33 πρωτεΐνες. Το γεγονός αυτό υποδηλώνει ότι η εμφάνιση του γονιδίου που κωδικοποιεί την πρωτεΐνη TMEM33 εντοπίζεται σε αρκετά πρώιμο στάδιο της εξέλιξης των ευκαρυωτικών οργανισμών και συνοδεύεται από μαζικές απώλειες αυτού σε κατώτερους κυρίως ευκαρυωτικούς οργανισμούς.



Εικόνα 42. Φυλογενετικό δέντρο για την πρωτεΐνη TMEM33. Στο μπλε πλαίσιο βρίσκονται τα μετόζωα, στο καφέ πλαίσιο οι μύκητες, στο πράσινο πλαίσιο τα alveolata και στο μωβ πλαίσιο τα φυτά.

3.3.8 Caspr1

Η πρωτεΐνη Caspr1 από τα αποτελέσματα των αναλύσεων που πραγματοποιήθηκαν με τα προγράμματα BLAST και Hmmer φαίνεται ότι εμφανίζει ορθόλογες πρωτεΐνες κυρίως στα μετόζωα και τα πλακόζωα. Παρατηρούνται όμως και 7 πρωτεϊνικές ακολουθίες προερχόμενες από τον οργανισμό *Dictyostelium discoideum*, οι οποίες αν και είναι μικρές σε μήκος περιέχουν την επικράτεια FA58C η οποία είναι παρούσα στο αμινοτελικό άκρο της ανθρώπινης Caspr1. Τα αποτελέσματα αυτά όμως ήταν αρκετά αμφίβολα διότι δεν στηρίζονταν από αρκετά μικρό E-value, ενώ παράλληλα δεν έδιναν στο reverse BLAST ως πρώτη πρωτεΐνη την ανθρώπινη Caspr1. Για το λόγο αυτό πραγματοποιήθηκε πολλαπλή στοίχιση όλων των πρωτεϊνικών αλληλουχιών που εντοπίστηκαν από τις αναλύσεις BLAST και Hmmer και τα αποτελέσματα παρατίθενται στην εικόνα 43.


```

1410      1420      1430      1440      1450      1460      1470      1480      1490      1500      1510      1520      1530      1540      1550
SVRFYNTIGTALTSQD...OSITISIAQVVFENVIIVNLEWVYVLRKAKKGIELTHMKEIERTDQDPIVSTSDSHNSKROSDITKIPRNSQENVDNDLQVNLCHNLEEHLMDSDTLEKHQHDSDVGRHSND
TSQSI SPRVAQIVDPR...VIVAVIVIVVLLILOAIGTFIFKYSRSHNSSEHNIIDVDSTKFSMEYID...ITRNSDRHOQSTNOYHLADHKQVSSNSGSSAKKKEIHL...
LSRTRILATLPSVQAD...ANVTGQIVGIIIFLLVLOIITALLVYRKSQDQAKYITEEAKAAEH...QATVEFKAIQPEVDTINQKKELVV...
IKTSTEDIKIFVSQNHQDLVSKAIGGQILALQEFILMSQSLIYMRSRPEVNHKINTEQE...NCSPRSEELVHNTTNNNNKPTVASKKEFFC...
PPFLVDEEKLRLKAYNE...VDSVLLACLVIIFLLILMFFIQRVYLRHNSQDVLTHEDQD...ADQADPRDDAVLHSTTQOVRKTEIF...
RWF TGLFVYVHODVYFRWELIVTVSMLERFELVY...LYLYLVYKSDVTHNEPQ...LESPSAR...PLDTLRKDRKALPEIDESPKE...
--RWYMGVTVRHHVDD...GLAQI...IAFTVIFELIATAALFLVYHYHRYKQSYLTNERKA...IDAPSASK...POSTRKEQLPQIEEEERKE...
--RWYLPDPFVYHDE...QWVAILLQFLVAFELQLQVLMVLFVYLNHRYKSDVYHNERKAAHEVYHSGKPLPTSDQAVRTPTAAPNOAPASAPAPPTAPAPQPRDQNLQIIEESRSE...
--RWYLPDPFVYHDD...QWVAILLQFLVAFELQLQVLMVLFVYLNHRYKSDVYHNERKAAHEVYHSGKPLPTSDQAVRTPTAAPNOAPASAPAPPTAPAPQPRDQNLQIIEESRSE...

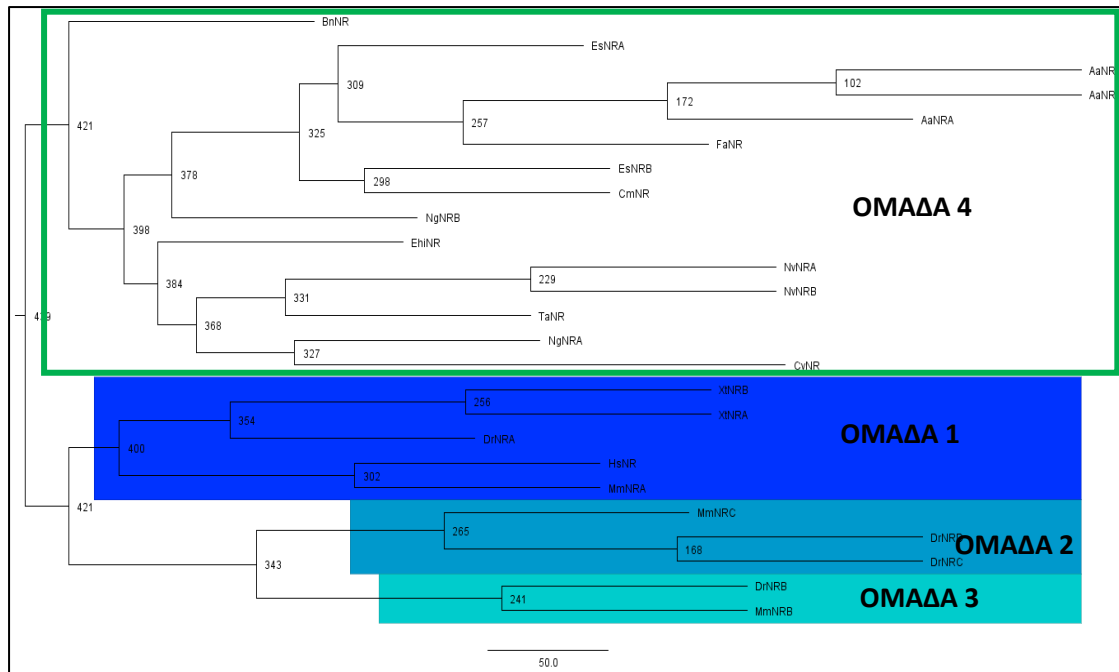
```

Εικόνα 43. Πολλαπλή στοίχιση πρωτεϊνικών αλληλουχιών που προέκυψαν από την αναζήτηση ορθολόγων με τη χρήση του BLAST και του Hmmer. Μέσα στο πράσινο πλαίσιο βρίσκεται το τμήμα της στοίχισης που αντιστοιχεί στις θέσεις 36-167 της ανθρώπινης πρωτεΐνης CASPR1 στο οποίο εντοπίζεται η επικράτεια FA58C. Στην επικράτεια αυτή κυρίως στοιχίζονται οι πρωτεΐνες από τον οργανισμό *Dictyostelium discoideum*. Για την οπτικοποίηση της στοίχισης χρησιμοποιήθηκε το πρόγραμμα Jalview (<http://www.jalview.org/>). Το χρώμα στα αμινοξικά κατάλοιπα υποδηλώνει το % ποσοστό ομοιότητας μεταξύ των αλληλουχιών, με την αυξανόμενη ένταση του χρώματος να αντιστοιχεί: σε άσπρο <= 40% ομοιότητα, ανοιχτό γκρι > 40% ομοιότητα, ανοιχτό μπλε >60% ομοιότητα, σκούρο μπλε >80% ομοιότητα. Στο όνομα κάθε πρωτεΐνης τα 2 ή 3 πρώτα γράμματα περιγράφουν τον οργανισμό από τον οποίο προέρχεται, σύμφωνα με τον πίνακα 15 του Παραρτήματος, ενώ οι επόμενοι 3 χαρακτήρες περιγράφουν την πρωτεΐνη CAS: Caspr1.

Από την παραπάνω εικόνα παρατηρούμε ότι οι 7 αλληλουχίες που προέρχονται από τον οργανισμό *Dictyostelium discoideum* στοιχίζονται κυρίως στο αμινοτελικό άκρο, όπου υπάρχει η επικράτεια FA58C της ανθρώπινης Caspr1 (πράσινο πλαίσιο). Το γεγονός αυτό σε συνδυασμό με τις παρατηρήσεις που πραγματοποιήθηκαν παραπάνω αποτελεί απόδειξη ότι οι 7 αυτές πρωτεϊνικές ακολουθίες δεν είναι πραγματικά ομόλογες με την ανθρώπινη Caspr1.

3.3.9 NogoR

Παρατηρούμε ότι από τις αναλύσεις που διεξήχθησαν με τα προγράμματα BLAST και HMMER προέκυψαν μια σειρά από υποψήφια ορθόλογα γονίδια και κατ' επέκταση πρωτεΐνες από διαφορετικές ομάδες των ευκαρυωτικών οργανισμών. Σημειώνεται ότι μόνο σε ορισμένα μετázωα (*M.musculus*, *X.tropicalis*, *D.rerio*) τα αποτελέσματα ήταν ξεκάθαρα και πληρούσαν εμφανώς όλες τις προϋποθέσεις που προαναφέρθηκαν για να γίνουν αποδεκτά. Για τους παραπάνω λόγους κατασκευάστηκε φυλογενετικό δέντρο, ώστε να εξεταστεί η πιθανή παρουσία ψευδώς θετικών αποτελεσμάτων. Το δέντρο αυτό παρατίθεται στη συνέχεια (εικόνα 44).



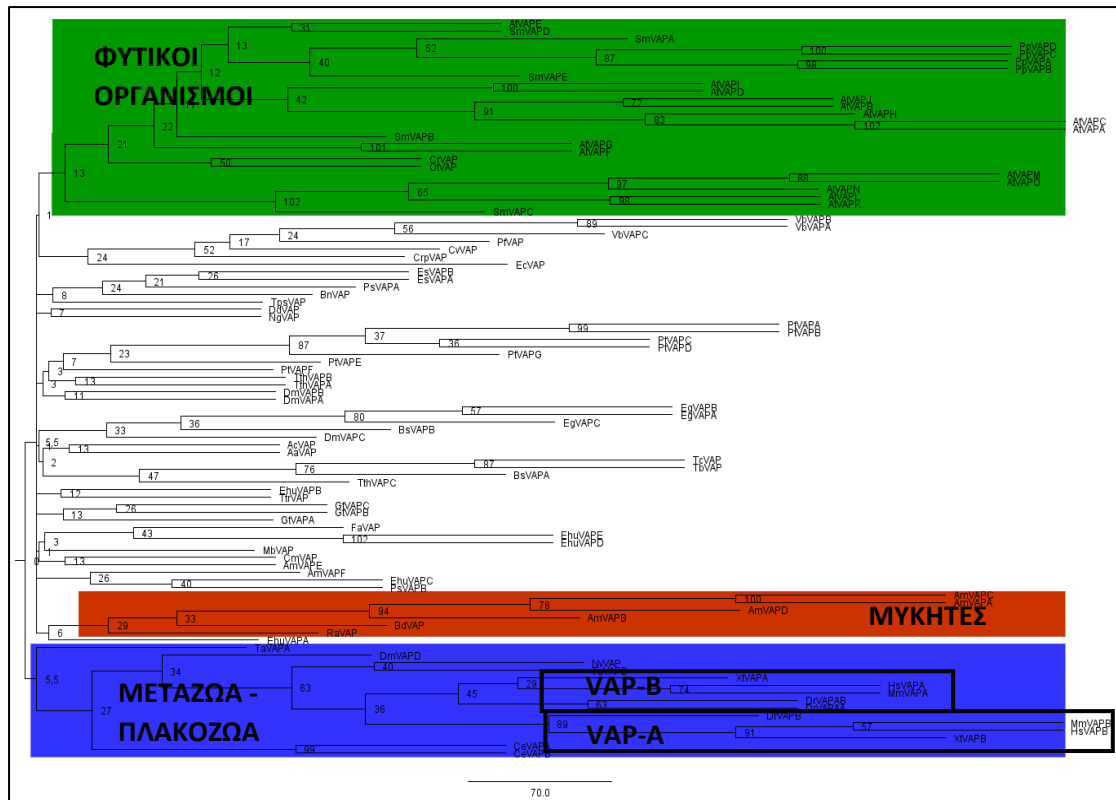
Εικόνα 44. Φυλογενετικό δέντρο για την πρωτεΐνη NogoR. Στην ομάδα 4 έχουν συγκεντρωθεί όλες οι πρωτεΐνες για τις οποίες τα αποτελέσματα ήταν αμφίβολα. Οι πρωτεΐνες αυτές είναι πλούσιες σε επαναλήψεις LRR και δεν αποτελούν πραγματικά ομόλογες της ανθρώπινης NogoR. Στην ομάδα 1 βρίσκονται οι πρωτεΐνες εκείνες οι οποίες είναι ομόλογες με την ανθρώπινη NogoR. Στις ομάδες 2 και 3 εντάσσονται πρωτεΐνες ομόλογες ως προς τις ανθρώπινες reticulon-4 receptor-like 1 precursor και reticulon-4 receptor-like 2 precursor οι οποίες αν και εμφανίζουν μεγάλο βαθμό ομοιότητας με τον υποδοχέα του RTN4A δεν είναι ορθόλογες με αυτόν. Στο όνομα κάθε πρωτεΐνης τα 2 ή 3 πρώτα γράμματα περιγράφουν τον οργανισμό από τον οποίο προέρχεται, σύμφωνα με τον πίνακα 15 του Παραρτήματος, ενώ οι επόμενοι 2 χαρακτήρες περιγράφουν την πρωτεΐνη NR: Nogo-receptor.

Από το δέντρο της εικόνας 44 φαίνεται ότι οι πρωτεΐνες χωρίζονται σε δύο μεγάλες ομάδες, των μεταζώων (ομάδες 1,2,3) και όλων των υπόλοιπων οργανισμών (ομάδα 4). Οι φυλογενετικές σχέσεις της ομάδας των μεταζώων υποστηρίζονται από υψηλό αριθμό bootstrap. Πιο συγκεκριμένα, παρατηρούμε ότι εντός της ομάδας των μεταζώων εμφανίζονται άλλες 3 υποομάδες. Στην ομάδα 1 εντάσσονται οι πρωτεΐνες οι οποίες είναι ομόλογες με την ανθρώπινη NogoR. Οι ομάδες 2 και 3 αποτελούνται από πρωτεΐνες οι οποίες είναι ομόλογες προς τις ανθρώπινες reticulon-4 receptor-like 1 precursor και reticulon-4 receptor-like 2 precursor αντίστοιχα. Πρόκειται για 2 πρωτεΐνες οι οποίες ταυτοποιήθηκαν λόγω του υψηλού βαθμού ομοιότητας τους με την ανθρώπινη NogoR, χωρίς όμως να ανήκουν στην ίδια πρωτεϊνική οικογένεια με αυτή. Συνεπώς, οι πρωτεΐνες των ομάδων 2 και 3 δεν είναι ορθόλογες με την ανθρώπινη NogoR. Στην ομάδα 4 που αποτελεί έναν ξεχωριστό κλάδο κατατάσσονται όλες οι πρωτεΐνες για τις οποίες δεν μπορούσαν να εξαχθούν ασφαλή συμπεράσματα όσον αφορά τη συγγένεια τους με την ανθρώπινη NogoR. Επιπλέον, παρατηρούμε ότι οι σχέσεις μεταξύ των πρωτεϊνών του κλάδου αυτού δεν είναι ισχυρές (μικρός αριθμός των bootstrap στους περισσότερους κλάδους) το οποίο πιθανά οφείλεται στο γεγονός ότι το μοναδικό κοινό στοιχείο που ενώνει τις πρωτεΐνες αυτές είναι η παρουσία επαναλήψεων λευκίνης στην αλληλουχία τους. Από τα δεδομένα

αυτά σε συνδυασμό με τα αποτελέσματα των αναλύσεων των BLAST και HMMER όπως αναφέρθηκαν παραπάνω οδηγούμαστε στο συμπέρασμα ότι όλες οι πρωτεΐνες της ομάδας 4 δεν εμφανίζουν πραγματική ομολογία με την ανθρώπινη NogoR.

3.3.10 VAPs

Οι αναλύσεις που προηγήθηκαν με τα προγράμματα BLAST και HMMER κατέδειξαν ότι η οικογένεια των πρωτεϊνών Var είναι αρκετά συντηρημένη, καθώς εμφανίζει αντιπροσώπους σε ολόκληρο το φάσμα των ευκαρυωτικών οργανισμών. Παρότι αρχικά η αναζήτηση των ορθόλογων γονιδίων έγινε με βάση τα γονίδια που κωδικοποιούν τις ανθρώπινες Var-A και Var-B, δεν ήταν τελικά δυνατή η ασφαλής και ξεκάθαρη κατηγοριοποίηση των πρωτεϊνών που προέκυψαν σε μία από τις δύο κατηγορίες. Ένα βασικό χαρακτηριστικό της οικογένειας το οποίο ήταν εμφανές ήδη από τα αποτελέσματα αυτά και επιβεβαιώθηκε και από το φυλογενετικό δέντρο το οποίο κατασκευάστηκε (εικόνα 45) είναι το γεγονός ότι σε κάθε οργανισμό υπάρχει διαφορετικός αριθμός γονιδίων που κωδικοποιούν πρωτεΐνες της οικογένειας των Var. Πιο συγκεκριμένα, σε έναν μεγάλο αριθμό οργανισμών από αυτούς που χρησιμοποιήθηκαν στην παρούσα διπλωματική εργασία φαίνεται να έχουν πραγματοποιηθεί ένας ή περισσότεροι εσωτερικοί διπλασιασμοί ενός αρχικού γονιδίου. Κάτι τέτοιο θα μπορούσαμε να πούμε ότι πιθανά είναι αναμενόμενο για τους φυτικούς οργανισμούς, στους οποίους συχνά παρατηρούνται εσωτερικοί διπλασιασμοί που έχουν ως σκοπό την αύξηση της αρμοστικότητας των οργανισμών αυτών (εικόνα 45 πράσινο πλαίσιο). Ένα ακόμη χαρακτηριστικό παράδειγμα πολλαπλών διπλασιασμών στην οικογένεια των Vars αποτελεί ο *C.elegans*, στον οποίο βρέθηκαν 30 Vars, οι οποίες όμως δεν χρησιμοποιήθηκαν όλες για την κατασκευή φυλογενετικού δέντρου για την αποφυγή σφαλμάτων. Από το φυλογενετικό δέντρο επίσης παρατηρούμε ότι πραγματικά ορθόλογες πρωτεΐνες των ανθρώπινων Var-A και Var-B υπάρχουν μόνο στους οργανισμούς *H.sapiens*, *M.musculus*, *X.tropicalis* και *D.rerio* (εικόνα 45 μαύρα πλαίσια). Συνεπώς, τα γονίδια που κωδικοποιούν τις πρωτεΐνες αυτές θα πρέπει να προέκυψαν από διπλασιασμό που συνέβη στον κοινό πρόγονο αυτών των οργανισμών.

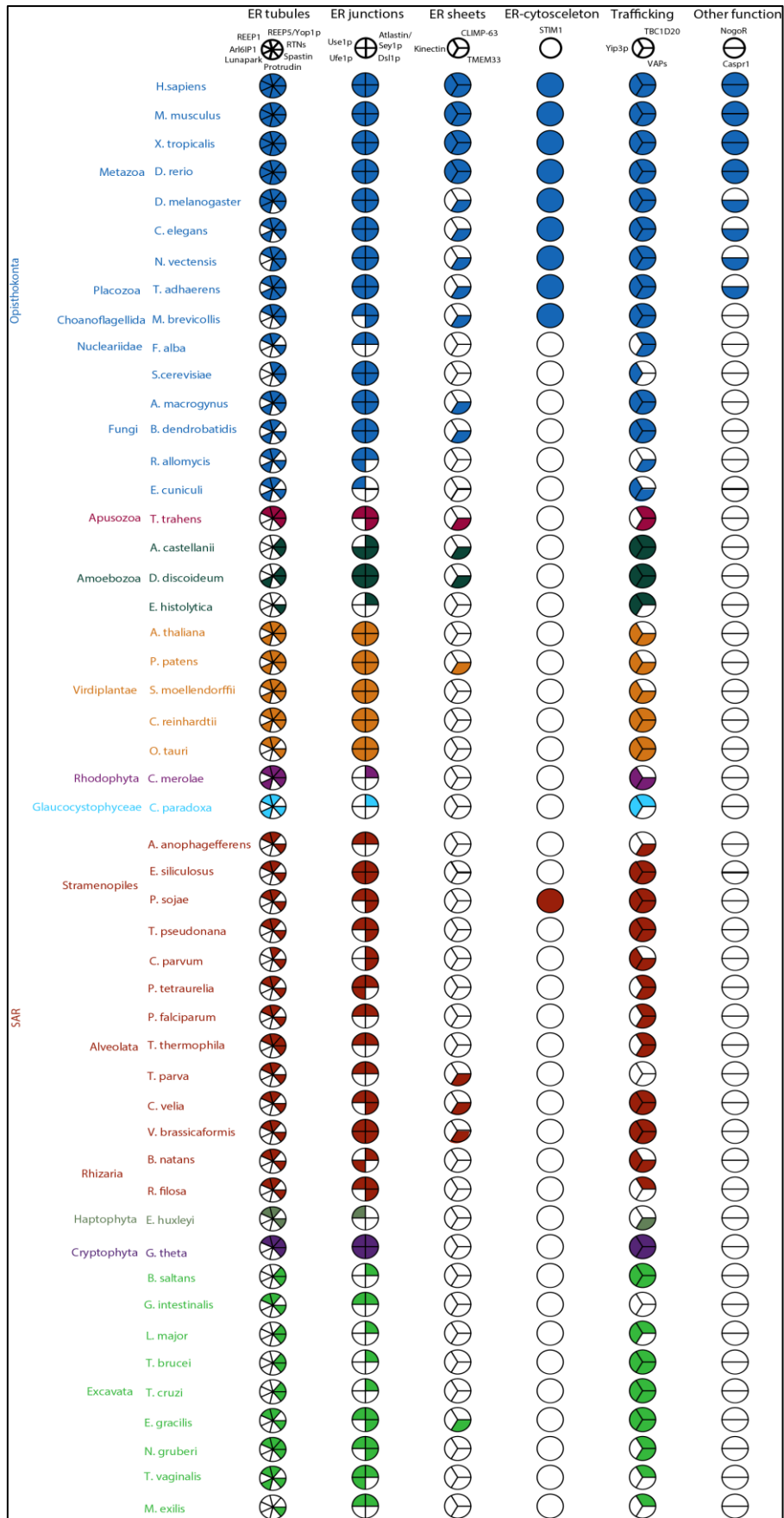


Εικόνα 45. Φυλογενετικό δέντρο των πρωτεϊνών μελών της οικογένειας των VAPs. Στο μπλε πλαίσιο βρίσκονται τα μετάρζωα και τα πλακόζωα, στο καφέ πλαίσιο οι μύκητες και στο πράσινο πλαίσιο τα φυτά. Στο όνομα κάθε πρωτεΐνης τα 2 ή 3 πρώτα γράμματα περιγράφουν τον οργανισμό από τον οποίο προέρχεται, σύμφωνα με τον πίνακα 15 του Παραρτήματος, ενώ οι επόμενοι 3 χαρακτήρες περιγράφουν την πρωτεΐνη VAP.

4 Συμπεράσματα

Το ενδομεμβρανικό σύστημα αποτελεί βασικό δομικό χαρακτηριστικό του ευκαρυωτικού κυττάρου που το διαφοροποιεί από το προκαρυωτικό κύτταρο και το κύτταρο των αρχαίων. Για το λόγο αυτό η μελέτη του συγκεκριμένου συστήματος εμφανίζει μεγάλο ενδιαφέρον από εξελικτική και όχι μόνο σκοπιά. Μέσα στο πλαίσιο αυτό εντάσσεται και η συγκεκριμένη διπλωματική εργασία, σκοπός της οποίας ήταν η πραγματοποίηση εκτενούς συγκριτικής γονιδιωματικής και φυλογενετικής ανάλυσης για 23 πρωτεΐνες του ενδομεμβρανικού συστήματος. Κοινό χαρακτηριστικό των πρωτεϊνών αυτών αποτελεί η παρουσία τους στις μεμβράνες του ενδομεμβρανικού συστήματος. Με εξαίρεση τις πρωτεΐνες NogoR και Dsl1p, όλες οι υπόλοιπες πρωτεΐνες είναι μεμβρανικές. Όπως περιγράφηκε αναλυτικά και στην εισαγωγή πολλές από τις πρωτεΐνες αυτές είναι δυνατό να επιτελούν περισσότερους από έναν βιολογικούς ρόλους στο κύτταρο, όπως συμβαίνει με τα *reticulons*, τις *Vaps* και τη *spastin*. Παρόλα αυτά όλες οι πρωτεΐνες είναι δυνατό να χωριστούν σε κατηγορίες ανάλογα με τον λειτουργικό τους ρόλο στο κύτταρο (πίνακας 1).

Για τις πρωτεΐνες αυτές πραγματοποιήθηκε εκτενής αναζήτηση ορθόλογων αντιπροσώπων σε 48 οργανισμούς αντιπροσωπευτικούς ολόκληρου του φάσματος των ευκαρυωτών. Με αφετηρία πρωτεϊνικές αλληλουχίες προερχόμενες τόσο από τον *Homo sapiens* όσο και από τον *Saccharomyces cerevisiae*, με τη χρήση των προγραμμάτων BLASTp και Hmmer, αλλά και μεθόδων φυλογενετικής ανάλυσης προέκυψαν τα αποτελέσματα του πίνακα 9 που συνοψίζονται και στην εικόνα 46.



Εικόνα 46. Συγκεντρικά αποτελέσματα όλων των αναλύσεων που προηγήθηκαν. Η ύπαρξη χρώματος στις «πίτες» (<https://sourceforge.net/projects/coulson/>) δηλώνει την παρουσία ορθόλογης πρωτεΐνης, ενώ η απουσία χρώματος δηλώνει την απουσία πρωτεΐνης.

Από τον πίνακα 9 παρατηρούμε ότι από όλες τις πρωτεΐνες οι οποίες συμμετέχουν στη διαμόρφωση της δομής των αγωγών του ενδοπλασματικού δικτύου μόνο οι 2 βασικές (reticulons, REEP5/Yop1p) παρουσιάζουν υψηλό βαθμό συντήρησης σε όλο το φάσμα των ευκαρυωτικών οργανισμών. Πιο συγκεκριμένα, οι πρωτεΐνες αυτές προσδίδουν την απαραίτητη καμπυλότητα στους αγωγούς του σωληνοειδούς ενδοπλασματικού δικτύου (Voeltz et al., 2006). Όλα τα μέλη της οικογένειας των reticulons προέκυψαν από ένα κοινό προγονικό γονίδιο, στον κοινό πρόγονο των ευκαρυωτικών οργανισμών. Στη διάρκεια της εξέλιξης φαίνεται να πραγματοποιήθηκαν αρκετοί εσωτερικοί διπλασιασμοί σε πολλούς οργανισμούς και κυρίως στα φυτά. Ο διπλασιασμός που οδήγησαν στην εμφάνιση των 4 πρωτεϊνών μελών της οικογένειας των reticulons στον άνθρωπο πρέπει να πραγματοποιήθηκαν στον κοινό πρόγονο των οργανισμών *H.sapiens*, *M.musculus*, *X.tropicalis* και *D.rerio* (εικόνα 34). Η πρωτεϊνική οικογένεια των REEP5/Yop1p ανήκει στην ευρύτερη οικογένεια των πρωτεϊνών REEPs, τα μέλη της οποίας εμφανίζουν υψηλό βαθμό ομοιότητας. Αυτός είναι και ένας βασικός λόγος για τον οποίο δεν κατέστη δυνατή η επίλυση των φυλογενετικών σχέσεων των πρωτεϊνών REEP5 και REEP1 που χρησιμοποιήθηκαν στα πλαίσια της συγκεκριμένης μελέτης. Παρόλα αυτά από τις αναλύσεις συγκριτικής γονιδιωματικής είναι εμφανής η ευρεία παρουσία πρωτεϊνικών μελών της οικογένειας REEP5/Yop1p σε όλους τους ευκαρυωτικούς οργανισμούς. Συγκεντρωτικά λοιπόν συμπεραίνουμε ότι ο σχηματισμός των αγωγών του ενδοπλασματικού δικτύου αποτελεί μια «αρχαία» λειτουργία η οποία στον κοινό πρόγονο των ευκαρυωτικών οργανισμών επιτελείτο από τους πρωτεϊνικούς προγόνους των reticulons και REEP5/Yop1p.

Στη λειτουργία αυτή όμως φαίνεται να συμμετέχουν και οι πρωτεΐνες lunapark, protrudin και Arl6IP1. Οι πρωτεΐνες αυτές είναι παρούσες μόνο σε ανώτερους ευκαρυωτικούς οργανισμούς (πίνακας 9), οι οποίοι εμφανίζουν υψηλότερο βαθμό πολυπλοκότητας. Το γεγονός αυτό αντικατοπτρίζεται και από το ρόλο των lunapark και protrudin, οι οποίες αποτελούν ρυθμιστές της πυκνότητας του σωληνοειδούς ενδοπλασματικού δικτύου καθώς και της αναλογίας των αγωγών και των ασκών του ενδοπλασματικού δικτύου. Αυτός μπορεί να είναι και ένας λόγος που εξηγεί τη συγκεκριμένη εξάπλωση σε ανώτερους μόνο ευκαρυώτες. Πιο συγκεκριμένα, η lunapark φαίνεται να εμφανίστηκε στα φυτά, η protrudin στα Choanoflagellida και η Arl6IP1 στα μετάζωα. Για την τελευταία δεν είναι γνωστές πολλές λεπτομέρειες για τον ακριβή της ρόλο. Σημειώνεται εδώ ότι δεν κατέστη δυνατό να διαπιστωθεί κατά πόσο οι πρωτεΐνες spastin και REEP1 έχουν συντηρηθεί έξω από τα μετάζωα και για το λόγο αυτό δεν μπορούν να συναχθούν περεταίρω συμπεράσματα για αυτές.

Τέσσερις πρωτεΐνες (Reticulons, CLIMP-63, Kinectin, TMEM33) οι οποίες συμμετέχουν στη διαμόρφωση της δομής των ασκών του ενδοπλασματικού δικτύου μελετήθηκαν στα πλαίσια της συγκεκριμένης διπλωματικής εργασίας. Από αυτές τα reticulons φαίνεται να έχουν την ευρύτερη εξάπλωση στους ευκαρυωτικούς οργανισμούς (πίνακας 9). Οι πρωτεΐνες CLIMP-63 και kinectin εμφανίζονται μόνο σε ορισμένα μετάζωα

(πίνακας 9). Οι πρωτεΐνες αυτές όμως δεν είναι απαραίτητες για το σχηματισμό *per se* των ασκών του ενδοπλασματικού δικτύου (Urade et al., 2014). Ορθόλογες πρωτεΐνες της TMEM33 είναι παρούσες εκτός από τα μετάζωα, στους μύκητες, τις αμοιβάδες και σε λίγους κατώτερους ευκαρυώτες (*alveolata* και *excavata*) (πίνακας 9). Ο ακριβής ρόλος της πρωτεΐνης αυτής όμως δεν είναι ακόμα γνωστός. Μια ακόμα πρωτεΐνη της οποίας η εξελικτική ιστορία θα έπρεπε να μελετηθεί για την εξαγωγή ασφαλέστερων συμπερασμάτων είναι και η p180. Από τα μέχρι στιγμής όμως αποτελέσματα συνάγεται ότι τα *reticulons* είναι ο κύριος καθοριστής του σχήματος των ασκών του ενδοπλασματικού δικτύου σε οργανισμούς εκτός των μεταζώων, πιθανά σε συνεργασία με την TMEM33. Παράλληλα, η παρουσία των πρωτεϊνών Climp-63 και kinectin στα μετάζωα πιθανά να λειτουργεί ως «δικλείδα ασφαλείας», λόγω του μεγαλύτερου βαθμού πολυπλοκότητας των οργανισμών αυτών.

Η ομοτυπική σύντηξη μεταξύ γειτονικών αγωγών του ενδοπλασματικού δικτύου με αποτέλεσμα την επέκταση και τη δημιουργία διακλαδώσεων στο σωληνοειδές δίκτυο στους ευκαρυωτικούς οργανισμούς, μεσολαμβάνεται από τις πρωτεΐνες *Atlastin* και *Sey1p* (Hu and Rapoport, 2016). Οι πρωτεΐνες *Atlastin* και *Sey1p* παρουσιάζουν ομολογία μόνο ως προς τη λειτουργία. Για το λόγο αυτό, όπως επιβεβαιώθηκε και από τη συγκεκριμένη μελέτη κατά κανόνα κάθε οργανισμός εμφανίζει μια από τις δύο αυτές πρωτεΐνες, με τις *Atlastins* να υπάρχουν κατά κανόνα στους ανώτερους ευκαρυώτες και τις *Sey1p* στους κατώτερους (πίνακας 9). Το γεγονός αυτό υποδηλώνει ότι προγονική πρωτεΐνη της *Sey1p* ήταν εκείνη που επιτελούσε τη λειτουργία της ομοτυπικής σύντηξης στον κοινό πρόγονο των ευκαρυωτικών οργανισμών και ότι η *Atlastin* προέκυψε αρκετά μεταγενέστερα. Παράλληλα, τελευταίες μελέτες έχουν καταδείξει ότι η ομοτυπική σύντηξη μεμβρανών στο σακχαρομύκητα επιτελείται και από τις πρωτεΐνες *Ufe1p*, *Use1p* και *Dsl1p*, πράγμα το οποίο δε συμβαίνει στα κύτταρα των θηλαστικών και της *Drosophila melanogaster* (Hu and Rapoport, 2016). Οι πρωτεΐνες αυτές εμφανίζουν αντιπροσώπους σε όλους τους ευκαρυώτες (πίνακας 9), χωρίς όμως να είναι γνωστό σε ποιους από αυτούς επιτελούν τη λειτουργία της μεμβρανικής ομοτυπικής σύντηξης.

Η καθολική παρουσία των πρωτεϊνών *Ufe1p*, *Use1p* και *Dsl1p* στους ευκαρυωτικούς οργανισμούς μπορεί να εξηγηθεί από τη συμμετοχή τους στις διαδικασίες της μεταφοράς από το σύμπλεγμα Golgi προς το ενδοπλασματικό δίκτυο, ως μέλη των συμπλόκων *Ufe1p/Use1p/Sec20* και *Dsl1p/Dsl3p/Tip20p* (Andag and Schmitt, 2003; Rogers et al., 2013). Άλλες πρωτεΐνες που εμπλέκονται στις διαδικασίες της κυστιδιακής μεταφοράς είναι οι *reticulons*, *spastin*, *TBC1D20*, *Yip3p* και οι VAPs (Andag and Schmitt, 2003; Lev et al., 2008; Sidjanin et al., 2016; Solowska and Baas, 2015). Από τον πίνακα 9 φαίνεται ότι και οι 3 τελευταίες πρωτεΐνες εμφανίζουν υψηλό βαθμό συντήρησης σε όλους τους ευκαρυωτικούς οργανισμούς. Επιπλέον, με βάση τον πίνακα 8 τα *reticulons* αλληλεπιδρούν με τις *spastin*, *TBC1D20*, και *Yip3p*. Κάθεμια από αυτές τις πρωτεΐνες με τη σειρά της αλληλεπιδρά με άλλες πρωτεΐνες οι οποίες συμμετέχουν στην ενδομεμβρανική κυκλοφορία, όπως οι

ESCRT-III, CHMP1B και IST1 για την spastin και οι Rab πρωτεΐνες για τις TBC1D20, και Yip3p (Andag and Schmitt, 2003; Sidjanin et al., 2016; Solowska and Baas, 2015). Συνεπώς, από όλα τα παραπάνω είναι εμφανές ότι η ενδομεμβρανική κυκλοφορία είναι μια λειτουργία αρκετά πολύπλοκη η οποία απαιτεί τη συμμετοχή πολλών διαφορετικών πρωτεϊνών και πρωτεϊνικών συμπλόκων. Επιπλέον, είναι μια λειτουργία απαραίτητη για τα ευκαρυωτικά κύτταρα λόγω του μεγάλου μεγέθους τους σε σύγκριση με τα προκαρυωτικά κύτταρα. Όλες οι πρωτεΐνες που σχετίζονται με τη λειτουργία αυτή και μελετήθηκαν στα πλαίσια της συγκεκριμένης εργασίας εμφανίζουν αντιπροσώπους σε ολόκληρο το φάσμα των ευκαρυωτικών οργανισμών. Τα αποτελέσματα αυτά σε συνδυασμό με τα αποτελέσματα της εικόνας 1 σχετικά με την προέλευση αρκετών πρωτεϊνών που λαμβάνουν μέρος στην ενδομεμβρανική κυκλοφορία (Koumandou and Field, 2011), αποδεικνύουν ότι πρόκειται για μια πολύ παλιά διαδικασία η οποία εξ' αρχής εμφάνιζε υψηλό βαθμό πολυπλοκότητας. Η εξέλιξη αυτής της διαδικασίας τοποθετείται παράλληλα με την εξέλιξη των ευκαρυωτικών κυττάρων.

Αρκετές πρωτεΐνες όπως η STIM1, η Spastin, η CLIMP-63 και η REEP1 διαθέτουν θέσεις αλληλεπίδρασης με τους μικροσωληνίσκους. Από αυτές όμως μόνο η πρώτη επιβεβαιωμένα συμμετέχει στην αλληλεπίδραση ανάμεσα στο ενδοπλασματικό δίκτυο και τον κυτταροσκελετό, προωθώντας τη μετακίνηση του σωληνοειδούς ενδοπλασματικού δικτύου μέσω του μηχανισμού TAC (Tip Attachment Complex) (English et al., 2009; Sandoz and van der Goot, 2015; Solowska and Baas, 2015). Η πρωτεΐνη αυτή είναι συντηρημένη στα μετάζωα, τα πλακόζωα και τα *Choanoflagellida*, σε περιορισμένο δηλαδή εύρος οργανισμών (πίνακας 9). Συνεπώς, θα πρέπει να υπάρχουν και άλλες πρωτεΐνες οι οποίες να επιτελούν τη διαδικασία αυτή στους υπόλοιπους ευκαρυώτες. Μια τέτοια όμως αναζήτηση θα ξεπερνούσε τα όρια της συγκεκριμένης μελέτης.

Η πρωτεΐνη NogoR είναι υποδοχέας της πρωτεΐνης NogoA (ή αλλιώς RTN4A), ο οποίος εντοπίζεται στην πλασματική μεμβράνη των νευρώνων. Η πρόσδεση της πρωτεΐνης RTN4A στον υποδοχέα NogoR μπορεί να οδηγήσει στην αναστολή της αύξησης των νευρώνων (Yang and Strittmatter, 2007). Από τα αποτελέσματα του πίνακα 9 παρατηρούμε ότι ο υποδοχέας αυτός φαίνεται να υπάρχει μόνο στα ανώτερα μετάζωα (*H.sapiens*, *M.musculus*, *X.tropicalis*, *D. rerio*). Οι οργανισμοί αυτοί σε αντίθεση με σχεδόν όλους τους υπόλοιπους οργανισμούς της μελέτης (με εξαίρεση τη *D. Melanogaster*), διαθέτουν πλήρως ανεπτυγμένο νευρικό σύστημα με νευρώνες. Συνεπώς, είναι απόλυτα λογικό ο υποδοχέας NogoR να εμφανίζεται μόνο σε αυτούς τους οργανισμούς.

Η πρωτεΐνη Caspr1 (contactin-associated protein) ανήκει σε μια οικογένεια διαμεμβρανικών πρωτεϊνών που παίζουν ρόλο στη διαμόρφωση και σταθερότητα των εμμύελων νευραξόνων (Zou et al., 2017). Επιπλέον, η πρωτεΐνη αυτή αλληλεπιδρά με την RTN4A μεσολαβώντας στον εντοπισμό των καναλιών καλίου στα paranodes των νευραξόνων (Yang and Strittmatter, 2007). Από τον πίνακα 9 παρατηρούμε ότι η παρουσία της Caspr1 είναι απολύτως επιβεβαιωμένη στους οργανισμούς *M.musculus*, *X.tropicalis* και

D.rerio οι οποίοι διαθέτουν πλήρως ανεπτυγμένο νευρικό σύστημα. Στους υπόλοιπους οργανισμούς (τα υπόλοιπα μετόζωα και τα Placozoa) φαίνεται να υπάρχει κάποιο μέλος της οικογένειας των πρωτεϊνών Caspr χωρίς να μπορεί να προσδιοριστεί ορθολογία με την Caspr1. Το γεγονός αυτό οφείλεται στον υψηλό βαθμό ομοιότητας των πρωτεϊνών αυτών. Τα μετόζωα διαθέτουν λιγότερο ή περισσότερο ανεπτυγμένο νευρικό σύστημα, γεγονός που θα μπορούσε να δικαιολογήσει την παρουσία των πρωτεϊνών Caspr σε αυτά. Κάτι τέτοιο όμως δεν συμβαίνει με τον οργανισμό *Trichoplax adhaerens* ο οποίος δε διαθέτει άξονες, συνάψεις ή μύες και γενικότερα νευρικό σύστημα (Jorgensen, 2014). Συνεπώς, μια πρωτεΐνη Caspr σε αυτόν τον οργανισμό θα έπρεπε να επιτελεί διαφορετική λειτουργία.

Από όλα τα παραπάνω συμπεράσματα είναι δυνατός ο προσδιορισμός της προέλευσης κάθε πρωτεΐνης που μελετήθηκε στα πλαίσια της συγκεκριμένης εργασίας, με εξαίρεση τις πρωτεΐνες REEP1 και Spastin (πίνακας 10).

Πίνακας 10. Εξελικτική προέλευση των πρωτεϊνών που μελετήθηκαν στην παρούσα διπλωματική. Οι πρωτεΐνες έχουν είτε i) αρχαία προέλευση που σημαίνει ότι σήμερα είναι παρούσες σε όλες τις σύγχρονες γενεαλογίες των ευκαρυωτικών οργανισμών, είτε ii) εμφανίζονται μόνο στα μετόζωα, είτε iii) εμφανίζονται στα holozoa (metazoa, placozoa, choanoflagellida), είτε δεν κατέστη δυνατό να προσδιοριστεί η προέλευση τους (αβέβαιη).

| Λειτουργίες | Πρωτεΐνες | Προέλευση της πρωτεΐνης | |
|-----------------------------------|-------------------|-------------------------|--------------|
| Δομή ενδοπλασματικού δικτύου (ΕΔ) | Αγωγοί | Reticulons | αρχαία |
| | | REEP5/Yop1p | αρχαία |
| | | REEP1 | αβέβαιη |
| | | Lunapark | αρχαία |
| | | Protrudin | holozoa |
| | | Arl6IP1 | μόνο μετόζωα |
| | | Spastin | αβέβαιη |
| | Κόμβοι αγωγών ΕΔ | Atlastin/Sey1p | αρχαία |
| | | Ufe1p/Use1p/Dsl1p | αρχαία |
| | Ασκοί | Reticulons | αρχαία |
| | | CLIMP-63 | μόνο μετόζωα |
| | | Kinectin | μόνο μετόζωα |
| | | TMEM33 | αρχαία |
| Αλληλεπίδραση ΕΔ-μικροσωληνίσκων | STIM1 | holozoa | |
| | Spastin | αβέβαιη | |
| | CLIMP-63 | μόνο μετόζωα | |
| | REEP1 | αβέβαιη | |
| Ενδομεμβρανική κυκλοφορία | Reticulons | αρχαία | |
| | Spastin | αβέβαιη | |
| | Ufe1p/Use1p/Dsl1p | αρχαία | |
| | TBC1D20 | αρχαία | |
| | Yip3p | αρχαία | |
| | VAPs | αρχαία | |
| Άλλες λειτουργίες | NogoR | μόνο μετόζωα | |
| | Caspr | holozoa | |

Συμπερασματικά από όλα τα παραπάνω είναι εμφανές ότι πολλά και σημαντικά γονίδια και κατ' επέκταση πρωτεΐνες που σχετίζονται με τη διαμόρφωση της δομής του ενδοπλασματικού δικτύου, την κυστιδιακή μεταφορά και άλλες λειτουργίες σχετιζόμενες με το ενδομεμβρανικό σύστημα είναι συντηρημένα σε ολόκληρο το φάσμα των ευκαρυωτικών οργανισμών. Πιο συγκεκριμένα, από τον πίνακα 10 παρατηρούμε ότι οι πρωτεΐνες Reticulons, REEP5/Yop1p, Lunapark, Atlantin/Sey1p, Ufe1p, Use1p και Dsl1p που συμμετέχουν στη διαμόρφωση της δομής των αγωγών του ενδοπλασματικού δικτύου εμφανίζουν «αρχαία» προέλευση. Επιπλέον, οι πρωτεΐνες Reticulons, Ufe1p, Use1p, Dsl1p, TBC1D20, Yip3p και VAPs που συμμετέχουν στα μονοπάτια της ενδομεμβρανικής κυκλοφορίας εμφανίζουν «αρχαία» προέλευση. Συνεπώς φαίνεται ότι οι τόσο η δομή των αγωγών του ενδοπλασματικού δικτύου όσο και οι μηχανισμοί της ενδομεμβρανικής κυκλοφορίας είναι «αρχαίοι», δηλαδή υπήρχαν ήδη στον τελευταίο κοινό πρόγονο των ευκαρυωτικών οργανισμών (LECA) και εμφάνιζαν υψηλό βαθμό πολυπλοκότητας. Κάτι τέτοιο όμως ήταν ήδη γνωστό για την ενδομεμβρανική κυκλοφορία.

Συνολικά όλα τα αποτελέσματα της συγκεκριμένης μελέτης έρχονται να στηρίξουν την υπόθεση της «αρχαίας» προέλευσης του ενδομεμβρανικού συστήματος. Πιο συγκεκριμένα, το ενδομεμβρανικό σύστημα αποτελεί βασικό παράγοντα διαφοροποίησης των ευκαρυωτών από τους προκαρυώτες και τα αρχαία. Συνεπώς, η εμφάνιση αυτού του συστήματος πρέπει να συνέβη στα αρχικά στάδια της εξέλιξης των ευκαρυωτικών οργανισμών. Η εξέλιξη του συστήματος αυτού τοποθετείται παράλληλα με την εξέλιξη του ευκαρυωτικού κυττάρου. Τέλος, η πληθώρα των πρωτεϊνών που εμφανίζουν αυτόν τον υψηλό βαθμό συντήρησης καταδεικνύουν την παρουσία υψηλού βαθμού πολυπλοκότητας στον τελευταίο κοινό πρόγονο των ευκαρυωτικών οργανισμών (LECA).

Βιβλιογραφία

- Ajawanawong, P., 2017. Molecular Phylogenetics: Concepts for a Newcomer. *Adv. Biochem. Eng. Biotechnol.* 160, 185–196. https://doi.org/10.1007/10_2016_49
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Andag, U., Schmitt, H.D., 2003. Dsl1p, an Essential Component of the Golgi-Endoplasmic Reticulum Retrieval System in Yeast, Uses the Same Sequence Motif to Interact with Different Subunits of the COPI Vesicle Coat. *J. Biol. Chem.* 278, 51722–51734. <https://doi.org/10.1074/jbc.M308740200>
- Anwar, K., Klemm, R.W., Condon, A., Severin, K.N., Zhang, M., Ghirlando, R., Hu, J., Rapoport, T.A., Prinz, W.A., 2012. The dynamin-like GTPase Sey1p mediates homotypic ER fusion in *S. cerevisiae*. *J. Cell Biol.* 197, 209–217. <https://doi.org/10.1083/jcb.201111115>
- Baum, B.R., 1989. PHYLIP: Phylogeny Inference Package. Version 3.2. Joel Felsenstein. *Q. Rev. Biol.* 64, 539–541. <https://doi.org/10.1086/416571>
- Chadrin, A., Hess, B., San Roman, M., Gatti, X., Lombard, B., Loew, D., Barral, Y., Palancade, B., Doye, V., 2010. Pom33, a novel transmembrane nucleoporin required for proper nuclear pore complex distribution. *J. Cell Biol.* 189, 795–811. <https://doi.org/10.1083/jcb.200910043>
- Chang, J., Lee, S., Blackstone, C., 2013. Protrudin binds atlastins and endoplasmic reticulum-shaping proteins and regulates network formation. *Proc. Natl. Acad. Sci.* 110, 14954–14959. <https://doi.org/10.1073/pnas.1307391110>
- Chen, S., Novick, P., Ferro-Novick, S., 2012. ER network formation requires a balance of the dynamin-like GTPase Sey1p and the Lunapark family member Lnp1p. *Nat. Cell Biol.* 14, 707–716. <https://doi.org/10.1038/ncb2523>
- Chiurchiù, V., Maccarrone, M., Orlicchio, A., 2014. The role of reticulons in neurodegenerative diseases. *NeuroMolecular Med.* 16, 3–15. <https://doi.org/10.1007/s12017-013-8271-9>
- Dilcher, M., Veith, B., Chidambaram, S., Hartmann, E., Schmitt, H.D., Von Mollard, G.F., 2003. Use1p is a yeast SNARE protein required for retrograde traffic to the ER. *EMBO J.* 22, 3664–3674. <https://doi.org/10.1093/emboj/cdg339>
- Eddy, S.R., 2018. HMMER User's Guide Biological. <https://doi.org/10.1002/9780471650126.dob0323.pub2>
- Eddy, S.R., 1998. Profile hidden Markov models. *Bioinformatics* 14, 755–63.
- Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. <https://doi.org/10.1093/nar/gkh340>
- Edgar, R.C., 2004. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5. <https://doi.org/10.1186/1471-2105-5-113>
- English, A.R., Zurek, N., Voeltz, G.K., 2009. Peripheral ER structure and function. *Curr. Opin. Cell Biol.* 21, 596–602. <https://doi.org/10.1016/j.ceb.2009.04.004>
- Farhan, H., Hauri, H.-P., 2009. Membrane Biogenesis: Networking at the ER with Atlastin. *Curr. Biol.* 19, R906–R908. <https://doi.org/10.1016/j.cub.2009.08.029>
- Field, M.C., Kelly, S., O'Reilly, A., Barlow, L.D., Nankissoor, N.N., Venkatesh, D., Boehm, C., Dacks, J.B., 2017. Evolution of the endomembrane systems of trypanosomatids – conservation and specialisation. *J. Cell Sci.* 130, 1421–1434. <https://doi.org/10.1242/jcs.197640>
- Friedman, J.R., Voeltz, G.K., 2011. The ER in 3D: a multifunctional dynamic membrane network. *Trends Cell Biol.* 21, 709–717. <https://doi.org/10.1016/j.tcb.2011.07.004>
- Geng, J., Shin, M.E., Gilbert, P.M., Collins, R.N., Burd, C.G., 2005. *Saccharomyces cerevisiae* Rab-GDI Displacement Factor Ortholog Yip3p Forms Distinct Complexes with the Ypt1

- Rab GTPase and the Reticulon Rtn1p. *Eukaryot. Cell* 4, 1166–1174.
<https://doi.org/10.1128/EC.4.7.1166-1174.2005>
- Giraldo, J.F.E., 2011. Implementação em hardware de um acelerador híbrido viterbi-plan7/algoritmo das divergências para comparação de proteínas.
- Goyal, U., Blackstone, C., 2013. Untangling the web: Mechanisms underlying ER network formation. *Biochim. Biophys. Acta - Mol. Cell Res.* 1833, 2492–2498.
<https://doi.org/10.1016/j.bbamcr.2013.04.009>
- Groden, J., Thliveris, A., Samowitz, W., Carlson, M., Gelbert, L., Albertsen, H., Joslyn, G., Stevens, J., Spirio, L., Robertson, M., Sargeant, L., Krapcho, K., Wolff, E., Burt, R., Hughes, J.P., Warrington, J., McPherson, J., Wasmuth, J., Le Paslier, D., Abderrahim, H., Cohen, D., Leppert, M., White, R., 1991. Identification and characterization of the familial adenomatous polyposis coli gene. *Cell* 66, 589–600.
[https://doi.org/10.1016/0092-8674\(81\)90021-0](https://doi.org/10.1016/0092-8674(81)90021-0)
- Gurel, P.S., Hatch, A.L., Higgs, H.N., 2014. Connecting the Cytoskeleton to the Endoplasmic Reticulum and Golgi. *Curr. Biol.* 24, R660–R672.
<https://doi.org/10.1016/j.cub.2014.05.033>
- Haas, A.K., Yoshimura, S. -i., Stephens, D.J., Preisinger, C., Fuchs, E., Barr, F.A., 2007. Analysis of GTPase-activating proteins: Rab1 and Rab43 are key Rabs required to maintain a functional Golgi complex in human cells. *J. Cell Sci.* 120, 2997–3010.
<https://doi.org/10.1242/jcs.014225>
- Hashimoto, Y., Shirane, M., Matsuzaki, F., Saita, S., Ohnishi, T., Nakayama, K.I., 2014. Protrudin Regulates Endoplasmic Reticulum Morphology and Function Associated with the Pathogenesis of Hereditary Spastic Paraplegia. *J. Biol. Chem.* 289, 12946–12961.
<https://doi.org/10.1074/jbc.M113.528687>
- Hu, J., Rapoport, T.A., 2016. Fusion of the endoplasmic reticulum by membrane-bound GTPases. *Semin. Cell Dev. Biol.* 60, 105–111.
<https://doi.org/10.1016/j.semcdb.2016.06.001>
- Hurt, C.M., Björk, S., Ho, V.K., Gilsbach, R., Hein, L., Angelotti, T., 2014. REEP1 and REEP2 proteins are preferentially expressed in neuronal and neuronal-like exocytotic tissues. *Brain Res* 30, 1–21. <https://doi.org/10.1016/j.brainres.2013.12.008>
- Ingle, E., Williams, J.H., Walker, C.E., Tsai, S., Colley, S., Sayer, M.S., Tilbrook, P.A., Sarna, M., Beaumont, J.G., Klinken, S.P., 1999. A novel ADP-ribosylation like factor (ARL-6), interacts with the protein-conducting channel SEC61 β subunit. *FEBS Lett.* 459, 69–74.
[https://doi.org/10.1016/S0014-5793\(99\)01188-6](https://doi.org/10.1016/S0014-5793(99)01188-6)
- Jorgensen, E.M., 2014. Animal Evolution: Looking for the First Nervous System. *Curr. Biol.* 24, R655–R658. <https://doi.org/10.1016/j.cub.2014.06.036>
- Klinger, C.M., Klute, M.J., Dacks, J.B., 2013. Comparative Genomic Analysis of Multi-Subunit Tethering Complexes Demonstrates an Ancient Pan-Eukaryotic Complement and Sculpting in Apicomplexa. *PLoS One* 8, 1–15.
<https://doi.org/10.1371/journal.pone.0076278>
- Koumandou, V.L., Dacks, J.B., Coulson, R., Field, M.C., 2007. Control systems for membrane fusion in the ancestral eukaryote; evolution of tethering complexes and SM proteins. *BMC Evol. Biol.* 17, 1–17. <https://doi.org/10.1186/1471-2148-7-29>
- Koumandou, L. V, Field, M.C., 2011. The Emergence of Cellular Complexity at the Dawn of the Eukaryotes: Reconstructing the Endomembrane System with In Silico and Functional Analyses. P. Pontarotti (ed.), *Evol. Biol. Concepts, Biodiversity, Macroevolution Genome Evol.* Springer-Verlag Berlin Heidelberg. 153–167.
<https://doi.org/10.1007/978-3-642-20763-1>
- Kurosaki, T., Baba, Y., 2010. Ca²⁺ signaling and STIM1. *Prog. Biophys. Mol. Biol.* 103, 51–58.
<https://doi.org/10.1016/j.pbiomolbio.2010.02.004>
- Lemus, L., Ribas, J.L., Sikorska, N., Goder, V., 2016. An ER-Localized SNARE Protein Is

- Exported in Specific COPII Vesicles for Autophagosome Biogenesis. *Cell Rep.* 14, 1710–1722. <https://doi.org/10.1016/j.celrep.2016.01.047>
- Lev, S., Halevy, D. Ben, Peretti, D., Dahan, N., 2008. The VAP protein family: from cellular functions to motor neuron disease. *Trends Cell Biol.* 18, 282–290. <https://doi.org/10.1016/j.tcb.2008.03.006>
- Lim, Y., Cho, I.-T., Schoel, L.J., Cho, G., Golden, J.A., 2015. Hereditary Spastic Paraplegia-Linked REEP1 Modulates ER- Mitochondria Contacts Youngshin. *Ann. Neurol.* 78, 679–696. <https://doi.org/10.1002/ana.24488>
- Maddison, W.P., Maddison, D.R., 2018. Mesquite: a modular system for evolutionary analysis. Version 3.6.
- Martincic, I., Peralta, M.E., Ngsee, J.K., 1997. Isolation and characterization of a dual prenylated Rab and VAMP2 receptor. *J. Biol. Chem.* 272, 26991–8.
- Meinl, E., Derfuss, T., Lington, C., 2010. Identifying targets for autoantibodies in CNS inflammation: Strategies and achievements. *Clin. Exp. Neuroimmunol.* 1, 47–60. <https://doi.org/10.1111/j.1759-1961.2009.00006.x>
- Murphy, S.E., Levine, T.P., 2016. VAP, a Versatile Access Point for the Endoplasmic Reticulum: Review and analysis of FFAT-like motifs in the VAPome. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* 1861, 952–961. <https://doi.org/10.1016/j.bbailip.2016.02.009>
- Ng, I.C., Pawijit, P., Teo, L.Y., Li, H., Lee, S.Y., Yu, H., 2016. Kinectin-dependent ER transport supports the focal complex maturation required for chemotaxis in shallow gradients. *J. Cell Sci.* 129, 2660–2672. <https://doi.org/10.1242/jcs.181768>
- Novarino, G., Ali G. Fenstermaker, Zaki, M.S., Hofree, M., Silhavy, J.L., Heiberg, A.D., Abdellateef, M., Rosti, B., Scott, E., Mansour, L., Masri, A., Kayserili, H., Al-Aama, J.Y., Ghada M. H. Abdel-Salam, A.K., Kara, M., Kara, B., Bozorgmehri, B., Ben-Omran, T., Mojahedi, F., Iman Gamal El Din Mahmoud, N.B., Bouhouche, A., Benomar, A., Hanein, S., Raymond, L., Forlani, S., Mascaro, M., Selim, L., Shehata, N., Al-Allawi, N., Bindu, P.S., Azam, M., Gunel, M., Caglayan, A., Bilguvar, K., Tolun, A., Issa, M.Y., Schroth, J., Spencer, E.G., Rosti, R.O., Akizu, N., Vaux, K.K., Johansen, A., Koh, A.A., Megahed, H., Durr, A., Brice, A., Stevanin, G., Gabriel, S.B., Ideker, T., Gleeson, J.G., 2014. Exome Sequencing Links Corticospinal Motor Neuron Disease to to Common Neurodegenerative Disorders. *Science (80-)*. 343, 506–511. <https://doi.org/10.1126/science.1247363>
- Oertle, T., Klinger, M., Stuermer, C.A.O., Schwab, M.E., 2003. A reticular rhapsody: phylogenic evolution and nomenclature of the RTN/Nogo gene family. *FASEB J.* 17, 1238–1247. <https://doi.org/10.1096/fj.02-1166hyp>
- Ohnishi, T., Shirane, M., Hashimoto, Y., Saita, S., Nakayama, K.I., 2014. Identification and characterization of a neuron-specific isoform of protrudin. *Genes to Cells* 19, 97–111. <https://doi.org/10.1111/gtc.12109>
- Otte, S., Belden, W.J., Heidtman, M., Liu, J., Jensen, O.N., Barlowe, C., 2001. Erv41p and Erv46p: new components of COPII vesicles involved in transport between the ER and Golgi complex. *J. Cell Biol.* 152, 503–18.
- Park, S.H., Zhu, P.-P., Parker, R.L., Craig Blackstone, 2010. Hereditary spastic paraplegia proteins REEP1, spastin, and atlastin-1 coordinate microtubule interactions with the tubular ER network 120, 1098–1110. <https://doi.org/10.1172/JCI40979DS1>
- Pascual-Caro, C., Espinosa-Bermejo, N., Pozo-Guisado, E., Martin-Romero, F.J., 2018. Role of STIM1 in neurodegeneration. *World J. Biol. Chem.* 9, 16–22. <https://doi.org/10.4331/wjbc.v9.i2.16>
- Patel, S.K., Indig, F.E., Olivieri, N., Levine, N.D., Latterich, M., 1998. Organelle Membrane Fusion: A Novel Function for the Syntaxin Homolog Ufe1p in ER Membrane Fusion. *Cell* 92, 611–620. [https://doi.org/10.1016/S0092-8674\(00\)81129-0](https://doi.org/10.1016/S0092-8674(00)81129-0)

- Rogers, J. V., Arlow, T., Inkellis, E.R., Koo, T.S., Rose, M.D., 2013. ER-associated SNAREs and Sey1p mediate nuclear fusion at two distinct steps during yeast mating. *Mol. Biol. Cell* 24, 3896–3908. <https://doi.org/10.1091/mbc.e13-08-0441>
- Rogers, J. V., McMahan, C., Baryshnikova, A., Hughson, F.M., Rose, M.D., 2014. ER-associated retrograde SNAREs and the Dsl1 complex mediate an alternative, Sey1p-independent homotypic ER fusion pathway. *Mol. Biol. Cell* 25, 3401–3412. <https://doi.org/10.1091/mbc.e14-07-1220>
- Sakabe, I., Hu, R., Jin, L., Clarke, R., Kasid, U.N., 2015. TMEM33: a new stress-inducible endoplasmic reticulum transmembrane protein and modulator of the unfolded protein response signaling. *Breast Cancer Res. Treat.* 153, 285–297. <https://doi.org/10.1007/s10549-015-3536-7>
- Sandoz, P.A., van der Goot, F.G., 2015. How many lives does CLIMP-63 have? *Biochem. Soc. Trans.* 43, 222–228. <https://doi.org/10.1042/BST20140272>
- Schmitt, H.D., 2010. Dsl1p/Zw10: common mechanisms behind tethering vesicles and microtubules. *Trends Cell Biol.* 20, 257–268. <https://doi.org/10.1016/j.tcb.2010.02.001>
- Schönthal, A.H., 2012. Endoplasmic Reticulum Stress: Its Role in Disease and Novel Prospects for Therapy. *Scientifica (Cairo)*. 2012, 1–26. <https://doi.org/10.6064/2012/857516>
- Seiler, S., Di Santo, S., Widmer, H.R., 2015. Non-canonical actions of Nogo-A and its receptors. *Biochem. Pharmacol.* 100, 28–39. <https://doi.org/10.1016/j.bcp.2015.08.113>
- Shibata, Y., Shemesh, T., Prinz, W.A., Palazzo, A.F., Kozlov, M.M., Rapoport, T.A., 2010. Mechanisms Determining the Morphology of the Peripheral ER. *Cell* 143, 774–788. <https://doi.org/10.1016/j.cell.2010.11.007>
- Sidjanin, D.J., Park, A.K., Ronchetti, A., Martins, J., Jackson, W.T., 2016. TBC1D20 mediates autophagy as a key regulator of autophagosome maturation. *Autophagy* 12, 1759–1775. <https://doi.org/10.1080/15548627.2016.1199300>
- Sklan, E.H., Serrano, R.L., Einav, S., Pfeiffer, S.R., Lambright, D.G., Glenn, J.S., 2007. TBC1D20 Is a Rab1 GTPase-activating Protein That Mediates Hepatitis C Virus Replication. *J. Biol. Chem.* 282, 36354–36361. <https://doi.org/10.1074/jbc.M705221200>
- Solowska, J.M., Baas, P.W., 2015. Hereditary spastic paraplegia SPG4: What is known and not known about the disease. *Brain* 138, 2471–2484. <https://doi.org/10.1093/brain/awv178>
- Spang, A., 2012. The DSL1 Complex: The Smallest but Not the Least CATCHR. *Traffic* 13, 908–913. <https://doi.org/10.1111/j.1600-0854.2012.01362.x>
- Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
- Stephany, C.É., Frantz, M.G., McGee, A.W., 2016. Multiple Roles for Nogo Receptor 1 in Visual System Plasticity. *Neuroscientist* 22, 653–666. <https://doi.org/10.1177/1073858415614564>
- Toyoshima, I., Yu, H., Steuer, E.R., Sheetz, M.P., 1992. Kinectin, a major kinesin-binding protein on ER. *J. Cell Biol.* 118, 1121–1131. <https://doi.org/10.1083/jcb.118.5.1121>
- Urade, T., Yamamoto, Y., Zhang, X., Ku, Y., Sakisaka, T., 2014. Identification and characterization of TMEM33 as a reticulon-binding protein. *Kobe J. Med. Sci.* 60, E57–65.
- Voeltz, G.K., Prinz, W.A., Shibata, Y., Rist, J.M., Rapoport, T.A., 2006. A Class of Membrane Proteins Shaping the Tubular Endoplasmic Reticulum. *Cell* 124, 573–586. <https://doi.org/10.1016/j.cell.2005.11.047>
- Wang, N., Rapoport, T.A., 2019. Reconstituting the reticular ER network – mechanistic implications and open questions. *J. Cell Sci.* 132, jcs227611. <https://doi.org/10.1242/jcs.227611>

- Xu, Y.Q., Sun, Z.Q., Wang, Y.T., Xiao, F., Chen, M.W., 2015. Function of Nogo-A/Nogo-A Receptor in Alzheimer's Disease. *CNS Neurosci. Ther.* 21, 479–485.
<https://doi.org/10.1111/cns.12387>
- Yamamoto, Y., Yoshida, A., Miyazaki, N., Iwasaki, K., Sakisaka, T., 2014. Arl6IP1 has the ability to shape the mammalian ER membrane in a reticulon-like fashion. *Biochem. J.* 458, 69–79. <https://doi.org/10.1042/BJ20131186>
- Yang, Y.S., Strittmatter, S.M., 2007. The reticulons: A family of proteins with diverse functions. *Genome Biol.* 8, 1–10. <https://doi.org/10.1186/gb-2007-8-12-234>
- Yoshida, H., 2007. ER stress and diseases. *FEBS J.* 274, 630–658.
<https://doi.org/10.1111/j.1742-4658.2007.05639.x>
- Zhang, D., Oliferenko, S., 2014. Tts1, the fission yeast homologue of the TMEM33 family, functions in NE remodeling during mitosis. *Mol. Biol. Cell* 25, 2970–2983.
<https://doi.org/10.1091/mbc.E13-12-0729>
- Zhao, Y., Zhang, T., Huo, H., Ye, Y., Liu, Y., 2016. Lunapark Is a Component of a Ubiquitin Ligase Complex Localized to the Endoplasmic Reticulum Three-way Junctions. *J. Biol. Chem.* 291, 18252–18262. <https://doi.org/10.1074/jbc.M116.737783>
- Zou, Y., Zhang, W., Liu, H., Li, X., Zhang, X., Ma, X., Sun, Y., Jiang, S., Ma, Q., Xu, D., 2017. Structure and function of the contactin-associated protein family in myelinated axons and their relationship with nerve diseases. *Neural Regen. Res.* 12, 1551.
<https://doi.org/10.4103/1673-5374.215268>
- Züchner, S., Wang, G., Tran-Viet, K.-N., Nance, M.A., Gaskell, P.C., Vance, J.M., Ashley-Koch, A.E., Pericak-Vance, M.A., 2006. Mutations in the Novel Mitochondrial Protein REEP1 Cause Hereditary Spastic Paraplegia Type 31. *Am. J. Hum. Genet.* 79, 365–369.
<https://doi.org/10.1086/505361>
- Νικολάου, Χ., 2015. Υπολογιστική Βιολογία.

Παράρτημα

Πίνακας 11. Πίνακας με τις βάσεις δεδομένων από τις οποίες βρέθηκαν τα πρωτεώματα των οργανισμών που χρησιμοποιήθηκαν στη συγκεκριμένη μελέτη.

| | Genome sources | BLAST | Hmmer |
|-----|------------------------------------|--|--|
| 1. | <i>T. adhaerens</i> | JGI | JGI |
| 2. | <i>F. alba</i> | Broad | Broad |
| 3. | <i>A. macrogynus</i> | Broad | Broad |
| 4. | <i>B. dendrobatidis</i> | JGI | JGI |
| 5. | <i>R. allomycis</i> | JGI | NCBI |
| 6. | <i>T. trahens</i> | Broad | Broad |
| 7. | <i>A. castellanii</i> | Broad | NCBI |
| 8. | <i>E. histolytica</i> | JCVI | JCVI originally, now on AmoebaDB |
| 9. | <i>A. thaliana</i> | NCBI | TAIR |
| 10. | <i>S. moellendorffii</i> | Phytozome | JGI |
| 11. | <i>O. tauri</i> | JGI | JGI |
| 12. | <i>C. merolae</i> | http://czon.jp/ | http://czon.jp/ |
| 13. | <i>C. paradoxa</i> | cyanophora.rutgers.edu | cyanophora.rutgers.edu |
| 14. | <i>A. anophagefferens</i> | JGI | JGI |
| 15. | <i>T. pseudonana</i> | JGI | JGI |
| 16. | <i>P. falciparum</i> | PlasmoDB | PlasmoDB |
| 17. | <i>C. velia</i> | CryptoDB | CryptoDB |
| 18. | <i>V. brassicaformis</i> | CryptoDB | CryptoDB |
| 19. | <i>B. natans</i> | JGI | JGI |
| 20. | <i>R. filosa</i> | NCBI | Ensembl |
| 21. | <i>E. huxleyi</i> | JGI | JGI |
| 22. | <i>G. theta</i> | JGI | JGI |
| 23. | <i>B. saltans</i> | Genedb (Sanger) | Sanger |
| 24. | <i>L. major</i> | TriTrypDB/Sanger | TriTrypDB |
| 25. | <i>T. brucei</i> | TriTrypDB | TriTrypDB |
| 26. | <i>T. cruzi</i> | TriTrypDB | TriTrypDB |
| 27. | <i>E. gracilis</i> (transcriptome) | Mark Field | Mark Field |
| 28. | <i>N. gruberi</i> | JGI | JGI |
| 29. | <i>M. musculus</i> | NCBI | NCBI |
| 30. | <i>X. tropicalis</i> | JGI | NCBI |
| 31. | <i>D. rerio</i> | NCBI | NCBI |
| 32. | <i>C. elegans</i> | NCBI | NCBI |
| 33. | <i>N. vectensis</i> | JGI | NCBI |
| 34. | <i>M. brevicollis</i> | JGI | NCBI |
| 35. | <i>E. cuniculi</i> | MicrosporidiaDB | NCBI |
| 36. | <i>D. discoideum</i> | Dictybase | Dictybase |
| 37. | <i>P. patens</i> | JGI | Phytozome |
| 38. | <i>C. reinhardtii</i> | Phytozome | Phytozome |
| 39. | <i>E. siliculosus</i> | OrcAE (https://bioinformatics.psb.ugent.be/orcae/overview/Ectsi) | JGI |
| 40. | <i>P. sojae</i> | JGI- v1.1 | JGI- v3.0 |

| | | | |
|-----|---|---|---|
| 41. | <i>C. parvum</i> | CryptoDB | CryptoDB |
| 42. | <i>T. parva</i> | NCBI | CryptoDB |
| 43. | <i>P. tetraurelia</i> | ParameciumDB | TGD |
| 44. | <i>T. thermophila</i> | TGD | TGD |
| 45. | <i>G. intestinalis</i> | GiardiaDB | GiardiaDB |
| 46. | <i>T. vaginalis</i> | TrichDB | TrichDB |
| 47. | <i>M. exilis</i> (<i>Monocercomonoides</i> sp.) | http://www.protistologie.cz/hampplab/data.html | http://www.protistologie.cz/hampplab/data.html |

Πίνακας 15. Συντομογραφίες των οργανισμών που χρησιμοποιήθηκαν για την κατασκευή των φυλογενετικών δέντρων.

| Οργανισμός | Συντομογραφία |
|---------------------------------------|---------------|
| <i>Homo sapiens</i> | Hs |
| <i>Mus musculus</i> | Ms |
| <i>Xenopus tropicalis</i> | Xt |
| <i>Danio rerio</i> | Dr |
| <i>Drosophila melanogaster</i> | Dm |
| <i>Caenorhabditis elegans</i> | Ce |
| <i>Nematostella vectensis</i> | Nv |
| <i>Trichoplax adhaerens</i> | Ta |
| <i>Monosiga brevicollis</i> | Mb |
| <i>Fonticula alba</i> | Fa |
| <i>Saccharomyces cerevisiae</i> | Sc |
| <i>Allomyces macrogynus</i> | Am |
| <i>Batrachochytrium dendrobatidis</i> | Bd |
| <i>Rozella allomycis</i> | Ra |
| <i>Encephalitozoon cuniculi</i> | Ec |
| <i>Thecamonas trahens</i> | Ttr |
| <i>Acanthamoeba castellanii</i> | Ac |
| <i>Dictyostelium discoideum</i> | Dd |
| <i>Entamoeba histolytica</i> | Ehi |
| <i>Arabidopsis thaliana</i> | At |
| <i>Physcomitrella patens</i> | Pp |
| <i>Selaginella moellendorffii</i> | Sm |
| <i>Chlamydomonas reinhardtii</i> | Cr |
| <i>Ostreococcus tauri</i> | Ot |
| <i>Cyanidioschyzon merolae</i> | Cm |
| <i>Cyanophora paradoxa</i> | Cyp |
| <i>Aureococcus anophagefferens</i> | Aa |
| <i>Ectocarpus siliculosus</i> | Es |
| <i>Phytophthora sojae</i> | Ps |
| <i>Thalassiosira pseudonana</i> | Tps |
| <i>Cryptosporidium parvum</i> | Crp |
| <i>Paramecium tetraurelia</i> | Pt |
| <i>Plasmodium falciparum</i> | Pf |
| <i>Tetrahymena thermophila</i> | Tth |
| <i>Theileria parva</i> | Tpa |
| <i>Chromera velia</i> | Cv |
| <i>Vitrella brassicaformis</i> | Vb |
| <i>Bigelowiella natans</i> | Bn |
| <i>Reticulomyxa filosa</i> | Rf |
| <i>Emiliana huxleyi</i> | Ehu |
| <i>Guillardia theta</i> | Gt |
| <i>Bodo saltans</i> | Bs |
| <i>Giardia intestinalis</i> | Gi |
| <i>Leishmania major</i> | Lm |
| <i>Trypanosoma brucei</i> | Tb |
| <i>Trypanosoma cruzi</i> | Tc |
| <i>Euglena gracilis</i> | Eg |
| <i>Naegleria gruberi</i> | Ng |

| | |
|---------------------------------|----|
| <i>Trichomonas vaginalis</i> | Tv |
| <i>Monocercomonoides exilis</i> | Me |