

**ΓΕΩΠΟΝΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ**

**ΣΧΟΛΗ ΤΡΟΦΙΜΩΝ, ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ ΚΑΙ ΑΝΑΠΤΥΞΗΣ**

**ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ ΤΜΗΜΑΤΟΣ  
ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ  
«ΒΙΟΛΟΓΙΑ ΣΥΣΤΗΜΑΤΩΝ»**

**Βιοϋπολογιστική ανάλυση της Λευκοθέας και αναπτυξιακό  
πρότυπο έκφρασης του γονιδίου *Lefkothea* στο *Arabidopsis***

**ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ**

**ΧΑΤΖΟΠΟΥΛΟΣ ΔΗΜΗΤΡΙΟΣ**

**ΕΠΙΒΛΕΠΟΥΣΑ ΚΑΘΗΓΗΤΡΙΑ: ΔΗΜΗΤΡΑ ΜΗΛΙΩΝΗ**

**ΕΠΙΚΟΥΡΗ ΚΑΘΗΓΗΤΡΙΑ Γ.Π.Α.**

**ΑΘΗΝΑ 2016**

**ΤΡΙΜΕΛΗΣ ΣΥΜΒΟΥΛΕΥΤΙΚΗ ΚΑΙ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ:**

ΕΠΙΚΟΥΡΗ ΚΑΘΗΓΗΤΡΙΑ **ΔΗΜΗΤΡΑ ΜΗΛΙΩΝΗ** (ΕΠΙΒΛΕΠΟΥΣΑ)

ΑΝΑΠΛΗΡΩΤΡΙΑ ΚΑΘΗΓΗΤΡΙΑ **ΚΑΛΛΙΟΠΗ ΠΑΠΑΔΟΠΟΥΛΟΥ**  
(ΠΑΝ/ΜΙΟ ΘΕΣΣΑΛΙΑΣ)

ΕΠΙΚΟΥΡΟΣ ΚΑΘΗΓΗΤΗΣ **ΣΤΑΜΑΤΗΣ ΡΗΓΑΣ** (Γ.Π.Α.)

**Βιοϋπολογιστική ανάλυση της Λευκοθέας και αναπτυξιακό πρότυπο  
έκφρασης του γονιδίου *Lefkothea* στο *Arabidopsis***

## ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΕΧΟΜΕΝΑ .....	2
ABSTRACT .....	4
ΠΕΡΙΛΗΨΗ .....	5
1. ΕΙΣΑΓΩΓΗ .....	6
Φαινοτυπικά χαρακτηριστικά του μεταλλάγματος <i>lefkothea 1-1</i> .....	12
Φαινοτυπικά χαρακτηριστικά του αλληλόμορφου <i>lefkothea 1-2</i> .....	17
Μορφολογία σπερμάτων και εμβρύων από το αγρίου τύπου φυτό <i>Col-0</i> και το αλληλόμορφο μετάλλαγμα <i>lefko 1-2</i> .....	20
Η πρωτεΐνη LEFKO1 περιέχει το τμήμα DUF860 ή PORR .....	20
2. ΥΛΙΚΑ & ΜΕΘΟΔΟΙ .....	21
Μηχανές ανάλυσης και πρόβλεψης τρισδιάστατης δομής πρωτεϊνών .....	21
i-tasser .....	21
Phyre2 .....	22
Μηχανές ανάλυσης και πρόβλεψης πυρηνικού εντοπισμού πρωτεϊνών .....	23
cNLS mapper .....	23
Suba3 .....	23
BLAST .....	24
Βάσεις Δεδομένων .....	24
GenBank .....	24
TAIR .....	25
Πρωτόκολλο απομόνωσης ολικού RNA από σπορόφυτα <i>Arabidopsis thaliana</i> .....	25
Πρωτόκολλο απομόνωσης ολικού RNA από έμβρυα <i>Arabidopsis thaliana</i> .....	26
Πρωτόκολλο απομόνωσης γονιδιωματικού DNA από <i>Arabidopsis</i> .....	28
Αλυσιδωτή αντίδραση πολυμεράσης (PCR) .....	29
Προγράμματα νευρωνικών δικτύων πρόβλεψης τοποθέτησης πρωτεϊνών .....	32
Προγράμματα Βιοπληροφορικής .....	32
3. ΑΠΟΤΕΛΕΣΜΑΤΑ .....	34
Το τμήμα PORR εντοπίζεται μόνο σε γονίδια φυτικών οργανισμών .....	34

Τα γονίδια του φυτού <i>Arabidopsis thaliana</i> που φέρουν το τμήμα PORR αποτελούν μια υπεροικογένεια .....	39
Δομή των 15 γονιδίων PORR με παρουσίαση ιντρονίων-εξονίων.....	43
Ομολογία των 15 PORR γονιδίων στο επίπεδο των αμινοξέων .....	47
Τοποθέτηση των 15 πρωτεϊνών PORR στα διάφορα υποκυτταρικά οργανίδια του <i>Arabidopsis</i> .....	49
Έκφραση του γονιδίου At5g62990.....	51
Ανάλυση με το efr browser για την έκφραση PORR γονιδίων σε διάφορους ιστούς.....	57
Φυλογένεια του γονιδίου lefkothea .....	<b>Error! Bookmark not defined.</b>
Έκφραση του γονιδίου At5g62990.....	<b>Error! Bookmark not defined.</b>
4. ΣΥΖΗΤΗΣΗ .....	58
Πρωτεΐνες πολλαπλής τοποθέτησης μέσα στο κύτταρο .....	58
Διαδική τοποθέτηση <i>de novo</i> σύνθεσης πρωτεϊνών .....	59
Οι πρωτεΐνες διπλής τοποθέτησης σε οργανίδια και πυρήνα συμμετέχουν στη σταθερότητα του DNA και την έκφραση γονιδίων .....	61
5. ΒΙΒΛΙΟΓΡΑΦΙΑ .....	65

## Abstract

Recent studies have shown that proteins containing the PORR (Plant Organelle RNA Recognition) domain are implicated in group II intron splicing mechanism of chloroplast encoded genes. *Lefkothea1* gene that contains the PORR domain is essential for embryo development. Plant organelles originate for endosymbiosis of bacteria in the primitive eukaryotic cell. During the course of evolution organellar targeting proteins encoded by the nuclear genome involved in fundamental biological processes including mRNA processing/splicing.

Based on genome scale searches for homologs of LEFKOTHEA1 this study reports the discovery and annotation of a complete repertoire of novel genes that belong to the LEFKOTHEA1 paralogous gene family of *Arabidopsis thaliana*. Through *in silico* analysis and bioinformatics this study investigates the quantity, structural features, expression characteristics, subcellular localizations and possible functional roles provides a scientific basis for further research on proteins with PORR domain in plant genomes.

## Περίληψη

Πρόσφατες δημοσιεύσεις έχουν δείξει πως πρωτεΐνες οι οποίες περιέχουν το τμήμα PORR (Plant Organelle RNA Recognition) εμπλέκονται στο μάτισμα των οργανιδιακών (π.χ. χλωροπλαστικών) γονιδίων τύπου II. Το γονίδιο *Lefkothea1* κωδικοποιεί μία πρωτεΐνη, η οποία περιέχει το τμήμα PORR και είναι ουσιαστική για την ομαλή ανάπτυξη του εμβρύου του φυτού *Arabidopsis thaliana*. Τα φυτικά οργανίδια προέρχονται από την ενδοσυμβίωση των βακτηρίων και των πρόδρομων ευκαρυωτικών κυττάρων. Κατά τη διάρκεια της εξέλιξης η τοποθέτηση των πρωτεϊνών στα οργανίδια ήταν σημαντική για βασικές βιολογικές διεργασίες όπως αυτή του ματίσματος και της ωρίμανσης του mRNA.

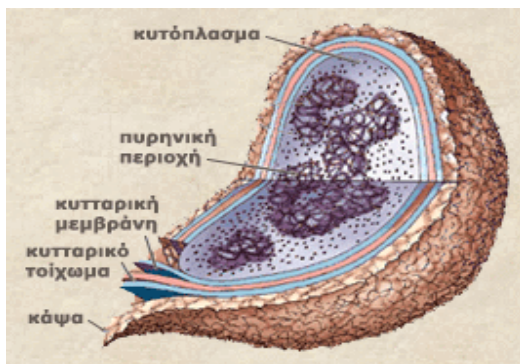
Με βάση γνωστά γονιδιώματα πολλών φυτικών ειδών, πραγματοποιήθηκε έρευνα για τα ομόλογα τη πρωτεΐνης LEFKOTHEA1 και βρέθηκε μια μεγάλη ομάδα γονιδίων που ανήκουν στα παράλογα γονίδια της οικογένειας των γονιδίων *Lefkothea1* στο *Arabidopsis thaliana*. Μέσω *in silico* analysis και βιοπληροφορικής η παρούσα μεταπτυχιακή διατριβή παρουσιάζει τις οικογένειες των γονιδίων τύπου *Lefkothea1* στα φυτικά είδη και αναλύει την ίδια την οικογένεια στο *Arabidopsis thaliana* σε επίπεδο υποκυτταρικής τοποθέτησης, έκφρασης, πρόβλεψης τριτοταγούς δομής και εξέλιξης.

Τα αποτελέσματα της εργασίας θα αποτελέσουν τη βάση για περαιτέρω έρευνα στην εξέλιξη και τη λειτουργία των γονιδίων-πρωτεϊνών με το τμήμα PORR στα φυτικά γονιδιώματα.

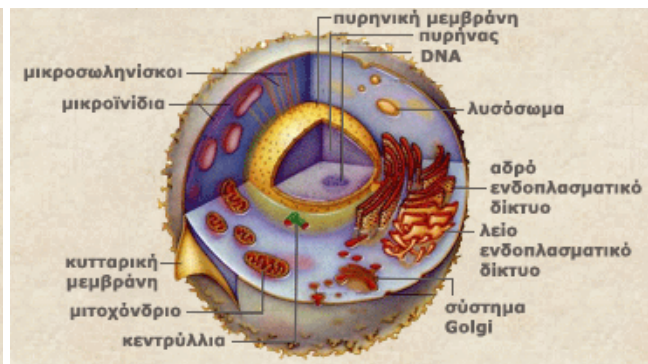
## 1. ΕΙΣΑΓΩΓΗ

Η ζωή, έχει επιχειρήσει πολλές φορές να απελευθερωθεί από τους ποικίλους περιορισμούς του περιβάλλοντός της. Με την καθιέρωση της φωτοσύνθεσης, λ.χ., απελευθερώθηκε από τον περιορισμό των πεπερασμένων προβιοτικών οργανικών ουσιών, και από μόνη της πλέον να αναλάβει τη σύνθεση των ενώσεων αυτών. Ένα άλλο, δεύτερο στη σειρά, μεγάλο άλμα και θεμελιώδες γεγονός της εξέλιξης μετά την προέλευση της κυτταρικής ζωής ήταν και η προέλευση-δημιουργία του «ευκαρυωτικού» κυττάρου, του κυττάρου δηλαδή που έχει διακριτό πυρήνα με μεμβράνη, σαν «κάρυο» (καρύδι). Και τούτο γιατί χωρίς αυτή τη μορφή ζωής δεν θα είχε επιτευχθεί η εξελικτική πρόοδος που χαρακτηρίζει την πλούσια βιόσφαιρά μας. Δεν θα υπήρχαν κοράλλια και πεταλούδες, όμορφα δάση και παιχνιδιάρικα δελφίνια, δεν θα υπήρχε ο άνθρωπος ή τα φυτά. Πιθανόν να υπήρχαν μόνο προκαρυωτικά κύτταρα χωρίς ξεχωριστό πυρήνα - βακτήρια και κυανοφύκη -, ένας κόσμος πρωτόγονος αλλά αρκετά έως πολύ προσαρμοσμένος στα διάφορα περιβάλλοντα. .

Το πώς κατάφερε το προκαρυωτικό κύτταρο να δώσει το πρώτο ευκαρυωτικό κύτταρο με πυρήνα, ειδικά οργανίδια, όπως μιτοχόνδρια και χλωροπλάστες (αν είναι φυτικό) ή ακόμη μαστίγια για την κίνησή του και ινίδια για την ακριβοδίκαιη κατανομή του γενετικού του υλικού στα θυγατρικά κύτταρα δεν είναι πλήρως γνωστό. Η εξελικτική αυτή πορεία χάνεται μέσα από την μη ύπαρξη τέτοιων χαμένων κρίκων (missing links) της ίδιας της εξέλιξης. Είναι όμως γενικά αποδεκτό ότι η βασική οργάνωση των ευκαρυωτικών κυττάρων, των κυττάρων δηλαδή όλων των ειδών εκτός των βακτηρίων και κυανοφυκών, βασίζεται στην προέλευση ενός πρώτου ευκαρυωτικού κυττάρου που έδωσε όλα τα άλλα.



Προκαρυωτικό κύτταρο



Ευκαρυωτικό (ζωικό) κύτταρο



Και τούτο επειδή όλα τα ευκαρυωτικά κύτταρα διαθέτουν ορισμένα όμοια πολύπλοκα οργανίδια όπως μιτοχόνδρια και χλωροπλάστες. Επιπλέον, η ομοιότητα της μιτωτικής συσκευής, του μηχανισμού δηλαδή διαίρεσης του ευκαρυωτικού πυρήνα, ορισμένων βιοχημικών κύκλων (κύκλος Embden-Meyerhof, κύκλος του Krebs), του μεταβολισμού της γλυκόζης και του μηχανισμού της πρωτεϊνοσύνθεσης, όλα όμοια στα ευκαρυωτικά κύτταρα (άλλα και αρκετά παρόμοια με τα προκαρυωτικά, όπως η πρωτεϊνοσύνθεση), μηδενίζουν την πιθανότητα ανεξάρτητης προέλευσης των διάφορων αρχικών ευκαρυωτικών οργανισμών.

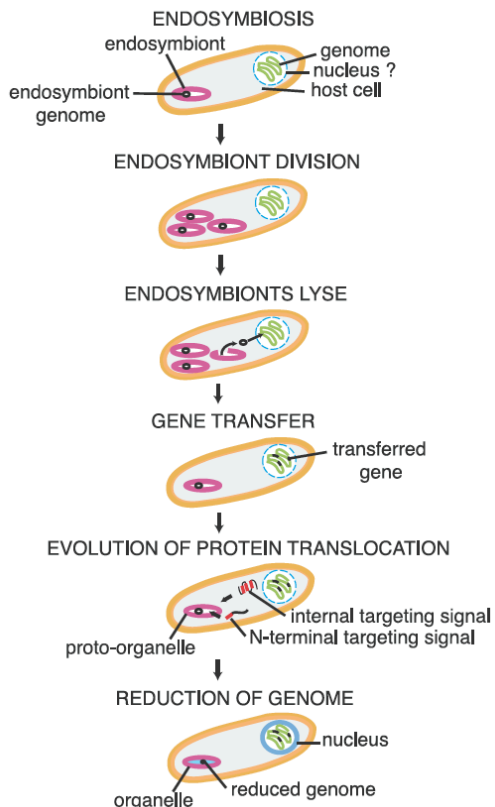
Το πώς όμως μορφοποιήθηκαν τα πολύπλοκα οργανίδια και οι βασικοί βιοχημικοί μηχανισμοί του ευκαρυωτικού κυττάρου είναι ένα άλλο ερώτημα, που φαίνεται να έχει συναντήσει και αυτό την απάντησή του συμβάλλοντας στην πληρέστερη κατανόηση της εξελικτικής ιστορίας της ζωής. Η αναζητούμενη απάντηση στο εν λόγω ερώτημα δομήθηκε, με δυαδικό τρόπο, τις τελευταίες δεκαετίες. Έτσι άρχισε η ραγδαία πρόοδος της κατανόησης της βιολογίας του κυττάρου και της εξελικτικής πορείας. Μια άποψη λοιπόν που φιλοδοξούσε να ερμηνεύσει την προέλευση των χλωροπλάστων και των μιτοχονδρίων ανήκει στον Bogorad και βασίζεται στο γεγονός ότι ορισμένες από τις πρωτεΐνες των οργανιδίων κωδικοποιούνται από γονίδια που εδράζουν στον πυρήνα. Γενετικό υλικό βέβαια έχουν και οι χλωροπλάστες και τα μιτοχόνδρια που είναι έξω από τον πυρήνα. Από την άλλη μεριά, ο προκαρυωτικός πυρήνας, το πυρηνοειδές, δεν διαθέτει πυρηνική μεμβράνη, με αποτέλεσμα τη διαφοροποίηση και του μηχανισμού πρωτεϊνοσύνθεσης στα ευκαρυωτικά συστήματα όχι τόσο στην διαδικασία αλλά στη θέση πραγματοποίησης της δηλαδή σε ένα ξεχωριστό διαμέρισμα, το κυτταρόπλασμα.

Σύμφωνα με τη θεωρία λοιπόν του Bogorad, όλο το DNA του ευκαρυωτικού κυττάρου ήταν κάποτε συγκεντρωμένο μόνο στον πυρήνα και ίσως να περιβαλλόταν από μια «πρωτόγονη» πυρηνική μεμβράνη σχηματίζοντας έναν πρωτόγονο πυρήνα, τον πρωτοπυρήνα. Αργότερα όμως ένα σύνολο γονιδίων περιβλήθηκε με μεμβράνη, ξέφυγε από τον πρωτοπυρήνα τα οποία με τη σειρά τους «μετασχηματίστηκαν» σε οργανίδια, σε μιτοχόνδρια και σε χλωροπλάστες. Η θεωρία αυτή είχε πολλά κενά, όπως η τυχαιοποίηση του συνόλου των γονιδίων που μεταφέρθηκαν όπως και άλλα.

Μια άλλη θεωρία που επεξεργάστηκε η Margulis βασισμένη σε μορφολογικές, βιοχημικές και παλαιοντολογικές έρευνες, ρίχνει περισσότερο φως στην προέλευση των εν λόγω οργανιδίων μέσα από μια συμβιωτική διαδικασία προκαρυωτικών οργανισμών που υπήρξε αρχικά σε ένα κύτταρο-δέκτη ή ξενιστή (host) το οποίο μπορούμε να θεωρήσουμε σαν αρχέγονο ευκαρυωτικό κύτταρο με μια υποτυπώδη πυρηνική μεμβράνη

(Εικόνα 1). Η συμβίωση αυτή προκάλεσε στη συνέχεια διαφοροποίηση του κυττάρου-δότη με διάφορους τρόπους και συνακόλουθο μετασχηματισμό του σε οργανίδια. Τα ισχυρά επιχειρήματα υπέρ της θεωρίας της συμβίωσης είναι πολλά και την καθιστούν επικρατέστερη της θεωρίας του Bogorad.

Μια τέτοια ισχυρή ένδειξη λοιπόν υπέρ της θεωρίας της συμβίωσης είναι η μεγάλη ομοιότητα που παρατηρείται στις χλωροφύλλες των ευκαρυωτικών κυττάρων και των προκαρυωτικών κυανοφυκών, ενώ έχουν αποκαλυφθεί και πολλές περιπτώσεις συμβίωσης ετερότροφων ευκαρυωτικών οργανισμών και κυανοφυκών. Το DNA εξάλλου των χλωροπλαστών της ευκαρυωτικής *Euglena*, που βρίσκεται πολύ χαμηλά στην εξελικτική ιεραρχία, είναι πολύ όμοιο με αυτό των κυανοφυκών. Μια άλλη πολύ ισχυρή μοριακή ένδειξη ή ακόμη και «απόδειξη» υπέρ της θεωρίας της συμβίωσης αφορά το γεγονός ότι το rRNA των χλωροπλαστών του καλαμποκιού έχει νουκλεοτιδικές αλληλουχίες πολύ όμοιες με το αντίστοιχο rRNA του βακτηρίου *E. coli*, μια κατάσταση που δεν ισχύει αν η σύγκριση αφορά το αντίστοιχο πυρηνικό RNA του καλαμποκιού.



### Εικόνα 1. Μοντέλο ενδοσυμβίωσης.

Γονιδιακή μεταφορά προς τον πυρήνα. Βακτηριακά endosymbionts πιθανώς να υπήρξε ως αντίγραφο του πληθυσμού (ροζ) μέσα στον ξενιστή (πορτοκαλί), όπως παρατηρείται στις σύγχρονες συμβιώσεις. Η λύση αυτών των βακτηρίων, ή/και κατά τη διάρκεια της διαφυγής του DNA κατά τη διαίρεση, θα μπορούσε να παρέχει την πηγή για γενετική μεταφορά προς τον πυρήνα. Το αποτέλεσμα είναι η γενετική ποικιλομορφία μετά την εξέλιξη μιας μηχανικής πρωτεϊνικής μετατόπισης προς το οργανίδιο και η απώλεια του γονιδίου που οδήγησε σε μείωση οργανιδιακών γονιδιωμάτων.

Για την ανακάλυψη και εν συνεχεία υιοθέτησης αυτής της θεωρίας υπάρχουν κάποιες ισχυρές αποδείξεις. Αυτές είναι οι εξής:

Τα μιτοχόνδρια και οι χλωροπλάστες περιέχουν DNA που είναι πιο κοντά στη μορφή του προκαρυωτικού γενώματος παρά του κυττάρου στο οποίο βρίσκεται. Αυτό διότι το DNA των οργανιδίων αυτών είναι κυκλικό και έχει μικρό μέγεθος.

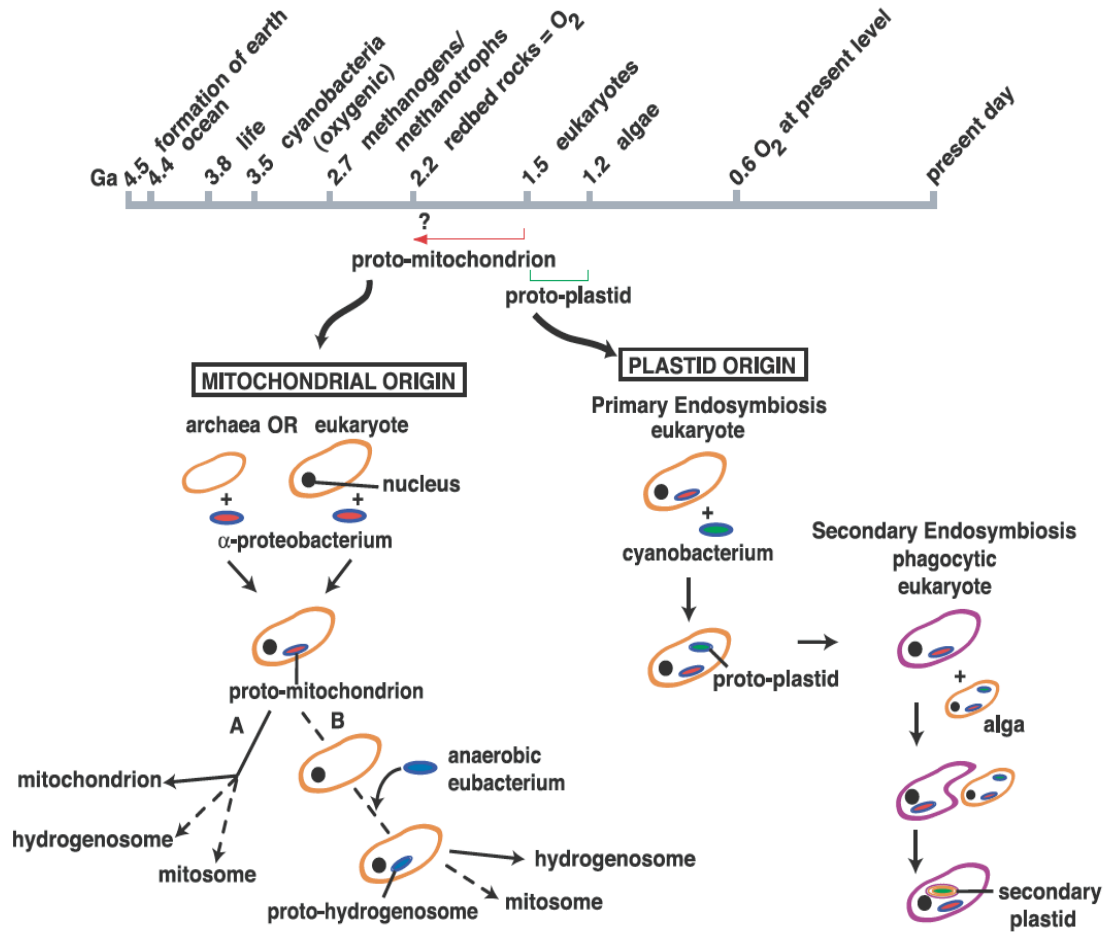
Τα οργανίδια αυτά αποτελούνται από δύο μεμβράνες, όπου η εσωτερική έχει δομή παρόμοια με αυτήν των προκαρυωτικών κυττάρων και η εξωτερική (μιτοχονδριακή) πιθανώς προέρχεται από τη μεμβράνη του φαγοσώματος του κυττάρου ξενιστή. Τα ριβοσώματα των μιτοχονδρίων και των χλωροπλαστών εμφανίζουν δομή παρόμοια με εκείνη των προκαρυωτικών κυττάρων (70s) και ευαισθησία σε αντιβιοτικά όπως ακριβώς και τα βακτήρια.

Το πρώτο αμινοξύ που τοποθετούν τα ριβοσώματα των οργανιδίων αυτών στην νεοσυντιθέμενη αλυσίδα είναι η N-φορμυλομεθειονίνη σε αντίθεση με τα ριβοσώματα του ενδοπλασματικού δικτύου των ευκαρυωτικών κυττάρων που ξεκινούν την πρωτεϊνοσύνθεση με το αμινοξύ μεθειονίνη.

Η ανάλυση των συχνοτήτων του DNA και φυλογενετικοί υπολογισμοί προτείνουν ότι το πυρηνικό DNA περιέχει γονίδια που πιθανότατα προέρχονται από πλαστίδια. Σε αντίθεση με το πυρηνικό DNA, το χλωροπλαστικό DNA μπορεί να μεταγραφεί με ακρίβεια και πιστότητα από την RNA πολυμεράση του E.coli, ενός βακτηρίου, και το mRNA που προκύπτει μπορεί να μεταφρασθεί τόσο από τα χλωροπλαστικά όσο και από τα βακτηριακά ριβοσώματα.

Η συμβίωση αυτή, που περιγράφηκε παραπάνω είναι το κλειδί της πολυκυτταρικότητας, καθώς μέσω αυτής τα ευκαρυωτικά κύτταρα κατάφεραν να ξεφύγουν από την αναερόβια κατάσταση στην οποία βρισκόντουσαν και να εκμεταλλευτούν το οξυγόνο που έμελλε να γίνει το πιο απαραίτητο συστατικό της ζωής. Η απελευθέρωση αυτή από τις αναερόβιες συνθήκες, βοήθησε την περαιτέρω εξάπλωση των ευκαρυωτών και την ανάπτυξη σχέσεων μεταξύ τους που οδήγησαν στην εμφάνιση της πολυκυτταρικότητας από μονοκύτταρους οργανισμούς.

Η προέλευση λοιπόν του ευκαρυωτικού κυττάρου ήταν αναμφίβολα μια μεγάλη επιτυχία της φύσης που συνέβη μόνο μία φορά (όπως έγινε και με την προέλευση του πρωτοκυττάρου) πριν από 0,9-1,5 δισ. χρόνια. Η διαδικασία που η φύση «επέλεξε» για να προοδεύσει εξελικτικά ήταν η συμβίωση και η συνεργασία που ενίσχυσαν τις δυνάμεις της.



**Εικόνα 2. Η εξέλιξη της ζωής στον πλανήτη γη.** Χρονοδιάγραμμα για την προέλευση της ζωής και τις μεγάλες εισβολές που δημιουργούν τα μιτοχόνδρια και τα πλαστίδια. Τα μιτοχόνδρια / - τα α-πρωτεοβακτήρια εμφανίζονται με κόκκινο χρώμα, hydrogenosomes / αναερόβια ευβακτήρια με σκούρο μπλε, πλαστίδια / κυανοβακτήρια με πράσινο, και ενδομεμβράνες τριγύρω πλαστιδίων δευτεροβάθμιας τάξης με κίτρινο. Οι διακεκομμένες γραμμές δείχνουν άλυτες σχέσεις. Α και Β, τα σενάρια Α ή Β. Ga, δισεκατομμύρια χρόνια πριν.

Η εξελικτική (χρονική) πορεία με τα δεδομένα που υπάρχουν μέχρι σήμερα δείχνει ότι πρώτα μεταφέρθηκε ένα α-πρωτεοβακτήριο το οποίο «μεταμορφώθηκε» σε μιτοχόνδριο μέσα στο ευκαρυωτικό κύτταρο και αργότερα (χρονικά) μια δεύτερη εισβολή στο πρώτο ευκαρυωτικό κύτταρο (με μιτοχόνδριο) από ένα κυανοβακτήριο δημιούργησε το πρώτο «πράσινο» ευκαρυώτη ικανό να παράγει οξυγόνο (Εικόνα 2).

Η εξέλιξη των μιτοχονδρίων και χλωροπλαστών κατά τη βακτηριακή ενδοσυμβίωση συνοδεύτηκε από μεγάλης κλίμακας μεταφοράς γονιδίων προς τον πυρήνα. Κατά συνέπεια, πολλές οργανισιακές πρωτεΐνες κωδικοποιούνται από πυρηνικά

γονίδια βακτηριακής καταγωγής που διατηρούν τη προγονική λειτουργία τους. Ωστόσο, κατά τη διάρκεια της μακράς συνεξέλιξης τους, τα μιτοχόνδρια και οι χλωροπλάστες με το κύτταρο ξενιστή τους, απέκτησαν χαρακτηριστικά τα οποία δεν είναι αντιπροσωπευτικά των βακτηριακών τους προγόνων. Επίσης πολλά από τα γονίδια βακτηριακής προέλευσης που εδράζουν στον πυρήνα, άρχισαν να δίνουν νέες λειτουργίες στον ευκαρυώτη-ξενιστή πέρα από τις ήδη υπάρχουσες λειτουργίες που προσδίδουν στα οργανίδια (μιτοχόνδρια-χλωροπλάστες). Η προέλευση των γονιδίων που προσδίδουν τέτοια χαρακτηριστικά μόνο προσφατά έχει ξεκινήσει να διευκρινίζεται.

Όπως για παράδειγμα, ένας αριθμός γονιδίων-πρωτεϊνών που έχουν διπλή τοποθέτηση σε μιτοχόνδριο και χλωροπλάστη χωρίς την ανάγκη του διπλασιασμού του συγκεκριμένου γονιδίου στο πυρηνικό γονιδίωμα.

Το σύνολο των μεταβολικών χαρακτηριστικών για το RNA, των φυτικών μιτοχονδρίων και χλωροπλάστων παρέχουν εντυπωσιακά παραδείγματα μη προκαρυωτικών χαρακτηριστικών τα οποία όμως αποκτήθηκαν κατά τη διάρκεια της εξέλιξης. Για παράδειγμα, και τα δύο οργανιδιακά γονιδιώματα είναι πλούσια σε ιντρόνια. Τα RNA και στα δύο οργανίδια είναι τροποποιημένα μετά από την επεξεργασία του RNA και μετα-μεταγραφικές τροποποιήσεις έχουν το κυρίαρχο ρόλο στον προσδιορισμό του γονιδίου σε σχέση με την αφθονία των προϊόντων που κωδικοποιούν. Πολλές πρωτεΐνες που συμμετέχουν σε τέτοιες διαδικασίες δεν προέρχονται από την ενδοσυμβίωση, αλλά προέκυψε στο πλαίσιο της πυρηνικής οργανιδιακής συνεξέλιξης. Για παράδειγμα, η 35πεπτιδιο (pentatricopeptide) επανάληψης (PPR) πρωτεϊνική οικογένεια βρίσκεται μόνο σε ευκαρυώτες, όπου εμπλέκεται σε διαδικασίες που σχετίζονται με το RNA στα μιτοχόνδρια και χλωροπλάστες. Τα τωρινά δεδομένα υποστηρίζουν ότι οι PPR πρωτεΐνες γενικά λειτουργούν ως πλατφόρμες RNA-αλληλεπίδρασης, αλλά φαίνεται να προέρχονται από το μοτίβο της tetratricopeptide επανάληψης (PPR), ένα πιο αρχαίο μοτίβο που συνδέει τη πρωτεΐνη πρόσδεσης.

Γενικότερα όμως, όσο αφορά τις μετα-μεταγραφικές τροποποιήσεις του RNA, όπως για παράδειγμα το μάτισμα, η εξελικτική πορεία δείχνει μια πολύ ενδιαφέρουσα κατάσταση. Έτσι παρουσιάζονται επίσης οριμένα δεδομένα τα οποία συγκλίνουν στην άποψη που προαναφέρθηκε. Ειδικότερα, είναι γνωστό ότι στους κατώτερους οργανισμούς που περιέχουν ιντρόνια, τα RNA αυτά λειτουργούν σαν ριβοένζυμα και μπορούν από μόνα τους να κάνουν αυτοσυρραφή (self-splicing) χωρίς την απαίτηση πρωτεϊνών του σπλισοσώματος (spliceosome) που υπάρχουν στους ευκαρυώτες. Στα οργανίδια που περιέχουν γονίδια με ιντρόνια τύπου II έχει δημιουργηθεί μια νέα κατάσταση.

Η οικογένεια πρωτεϊνών που καθορίζεται από την "περιοχή ή επικράτεια της άγνωστης λειτουργίας 860"(Domain of Unknown Function 860) ταιριάζει με αυτό το γενικό πρότυπο. Η DUF860 protein WTF1 είναι απαραίτητη για τη αποκοπή ιντρονίων της ομάδας II στους χλωροπλάστες. Η πρωτεΐνη αυτή συνδέεται in vivo με γενετικά προκαθορισμένες -πρωτεΐνες συνδέτες του RNA. Η WTF1 και η περιοχή DUF860 παρουσιάζουν δραστηριότητες δέσμωσης με RNA in vitro. Τα περισσότερα μέλη της οικογένειας DUF860 εντοπίζονται σε χλωροπλάστες ή μιτοχόνδρια (ενώ τα γονίδια εδράζουν στον πυρήνα), υποδηλώνοντας ότι οι πρωτεΐνες έχουν πολλαπλούς ρόλους στη γονιδιακή έκφραση και ροή της πληροφορία και ωρίμανσης του RNA στα δύο οργανίδια. Αν και η πρωτεΐνες DUF860 βρίσκονται μόνο σε φυτά, υπάρχει μια συσχέτιση με μια ομάδα υδρολασών ουβικιτίνης (UBH) που βρίσκονται σε ευκαρυωτικούς οργανισμούς. Ωστόσο, τα διαρθρωτικά (δομικά) μοντέλα δείχνουν ότι η περιοχή DUF860 υιοθετεί μια δομή που διαφέρει από αυτή των ενζύμων UBH, και ότι έχει μια επιφάνεια που θυμίζει μοτίβα ελικοειδής επανάληψης συνδεσης με RNA όπως τα PPR και PUM-HD μοτίβα (REF, REF).

### 1.1 Φαινοτυπικά χαρακτηριστικά του μεταλλάγματος *lefkothea1-1*

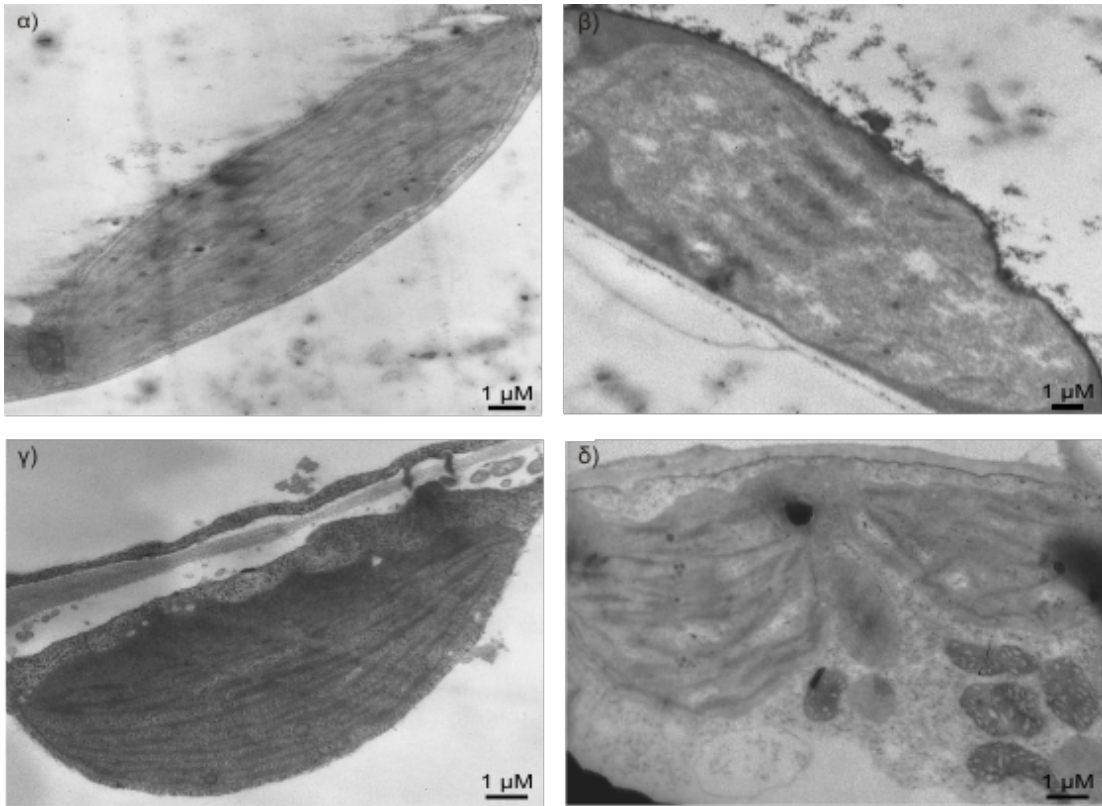
Στα λειτουργικά πλαίσια της ωρίμανσης του RNA κινείται και η πρωτεΐνη ΛΕΥΚΟΘΕΑ (LEFKOTHEA), όπως θα φανεί παρακάτω. Μετά από σάρωση μεταλλαγμάτων που πραγματοποιήθηκε στο Εργαστήριο Μοριακής Βιολογίας, η μετάλλαξη *lefkothea1* επιλέχθηκε μεταξύ άλλων μεταλλαγμάτων εξαιτίας του ιδιαίτερου φαινοτύπου που σχετίζεται με αποκλειστική λεύκανση των κοτυληδόνων (Εικόνα 3). Ο χαρακτηριστικός αυτός φαινότυπος αποτέλεσε το κίνητρο επιλογής του ονόματος "Λευκοθέα" που σύμφωνα με τη μυθολογία αποτελεί μία από τις λευκές νηρηίδες της θάλασσας.

Το μετάλλαγμα *lefkothea1-1* (*lefko1-1*) προέρχεται από χημική μεταλλαξογένεση με EMS (Ethane Methane Sulfonate) που δημιουργεί μεταλλάξεις σημείου. Στο *lefko1-1* η μετάλλαξη χαρτογραφήθηκε στο γονίδιο At5g62990, το οποίο έχει μία αμινοξική υποκατάσταση της γλυκίνης (373) σε ασπαρτικό οξύ (G<sub>373</sub>→D), εξαιτίας της αλλαγής της βάσης G στη θέση 1188bp σε A, κατωφορικά του κωδικονίου έναρξης της μετάφρασης (Αθανασιάδη, 2008).



**Εικόνα 3. Χαρακτηριστικές φαινοτυπικές ανωμαλίες του μεταλλάγματος *lefko1-1*.** Οι άσπρες κοτύλες και η μικρότερη ρίζα είναι από τα χαρακτηριστικά γνωρίσματα της μετάλλαξης. Α) Φαινότυπος μεταλλαγμένων σποροφύτων (3 δεξιά σπορόφυτα) και Β) Χαρακτηριστική χλώρωση *lefko1-1* κοτυληδόνων (κάτω σειρά) φυτών ηλικίας 5 ημερών.

Σε αντίθεση με τις κοτυληδόνες τα φύλλα του μεταλλάγματος *lefko1-1* είναι πράσινα και δεν έχουν κάποιο χαρακτηριστικό λεύκανσης. Βιομετρικές αναλύσεις αποκάλυψαν ότι η ανάπτυξη του μεταλλάγματος υστερεί εμφανώς στα πρώτα στάδια, μετά τη βλαστικότητα των σπερμάτων, συγκριτικά με τα φυτά αγρίου τύπου *Col-0* ιδιαίτερα στο μήκος της πρωτογενούς ρίζας. Ανάλογα με τις συνθήκες φωτισμού είναι δυνατό να εμφανιστεί χλώρωση στα πρώτα ζεύγη φύλλων αλλά μετά από κάποιο χρονικό διάστημα τα φύλλα πρασινίζουν πλήρως. Επειδή η χαρακτηριστική φαινοτυπική ανωμαλία της μετάλλαξης *lefko1-1* είναι οι λευκές κοτυληδόνες, είναι φανερό ότι η ανωμαλία της μετάλλαξης να προέρχεται από το έμβρυο μια και οι κοτύλες αποτελούν εμβρυακό όργανο. Είναι επίσης φανερό ότι λεύκανση πρέπει να προέρχεται από αλλαγές στους χλωροπλάστες. Οι τελευταίοι συνεργάζονται με τα μιτοχόνδρια για τη καλή ομοιοστάση του φυτικού κυττάρου. Από το γεγονός αυτό πηγάζει και η μικρότερη σε μέγεθος της πρωτογενούς ρίζας. Με τεχνικές ηλεκτρονικής μικροσκοπίας διερευνήθηκαν πιθανές μορφολογικές αλλοιώσεις των χλωροπλαστών (Εικόνες 4 και 5). Η παρατήρηση έγινε σε κύτταρα από κοτυληδόνες και πρώτα φύλλα φυτών *Col-0* και *lefko1-1* που αναπτύχθηκαν σε θρεπτικό μέσο με 1 % σακχαρόζη στους 22 °C με φωτοπερίοδο 16h φως και 8h σκοτάδι.



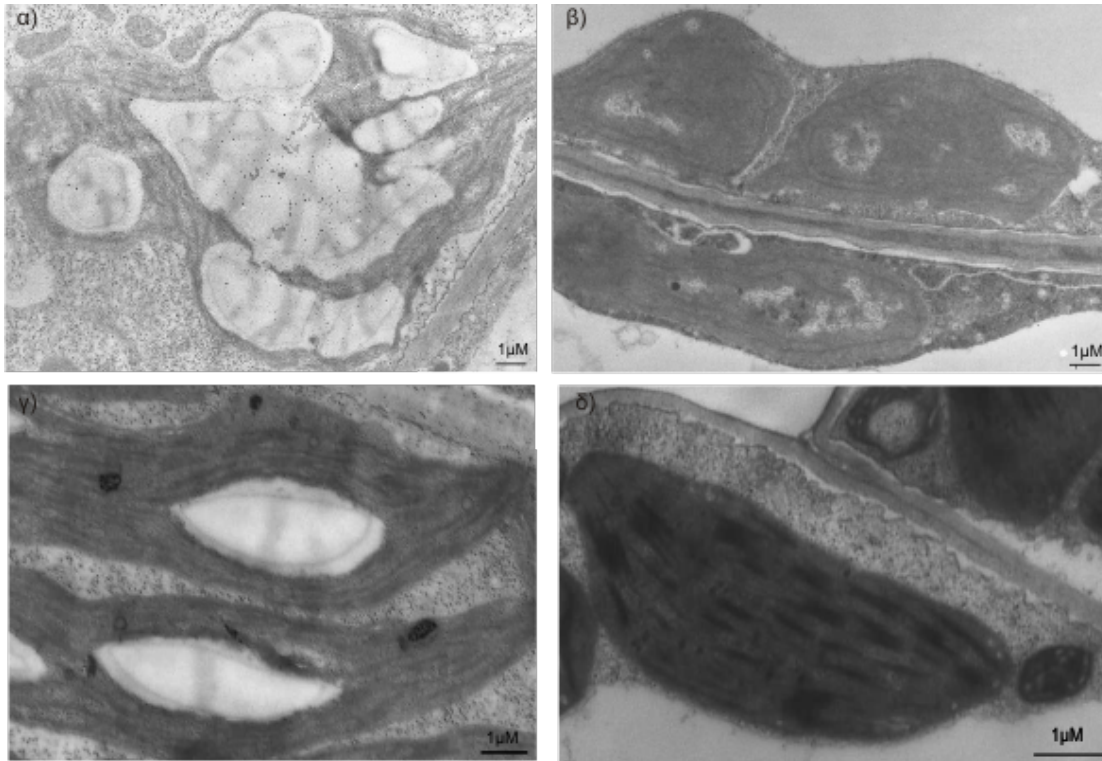
**Εικόνα 4. Παρατήρηση της μορφολογίας των χλωροπλαστών του μεταλλάγματος *lefko1-1* και των φυτών *Col-0*.** α) Χλωροπλάστης από κύτταρα κοτυληδόνας φυτού *Col-0* ηλικίας 5 ημερών, β) χλωροπλάστης από κύτταρα κοτυληδόνας φυτού *lefko1-1* ηλικίας 7 ημερών, γ) χλωροπλάστης από κύτταρα πρώτων φύλλων φυτών *Col-0* ηλικίας 14 ημερών, δ) χλωροπλάστης από κύτταρα πρώτων φύλλων φυτών *lefko1-1* ηλικίας 14 ημερών.

Οι Εικόνες 4 και 5 έδειξαν ότι ο φακοειδής σχηματισμός και η οργανωμένη δομή των θυλακοειδών και των grana στους χλωροπλάστες των φυτών αγρίου τύπου δεν παρατηρείται στα φυτά της μετάλλαξης *lefko1-1*. Συγκεκριμένα οι χλωροπλάστες είναι μικρότεροι και δεν είναι άρτια σχηματισμένοι με προβληματικοί οργάνωση των θυλακοειδών. Εντούτοις, η μορφολογία του χλωροπλάστη στα πρώτα φύλλα του μεταλλάγματος *lefko1-1* είναι βελτιωμένη συγκριτικά με τις κοτυληδόνες του χωρίς όμως και πάλι να φτάνει την αρτιότητα των χλωροπλαστών που εμφανίζουν τα φυτά αγρίου τύπου (Εικόνα 4). Προκειμένου να εξεταστεί αν οι φαινοτυπικές αυτές ανωμαλίες θα μπορούσαν να εξαλειφθούν εάν συμπληρωνόταν εξωγενώς πηγή άνθρακα, τα φυτά αναπτύχθηκαν σε θρεπτικό μέσο που περιείχε 3% σακχαρόζη. Για την παρατήρηση της μορφολογίας των χλωροπλαστών σε κύτταρα των κοτυληδόνων τα φυτά αναπτύχθηκαν



στο σκοτάδι εξαιτίας πιθανού εκφυλισμού των φωτοσυνθετικών χρωστικών στο φως. Για την αντίστοιχη παρατήρηση στα πρώτα φύλλα, τα φυτά αναπτύχθηκαν σε φυσιολογικό κύκλο φωτοπεριόδου, 16h φως και 8h σκοτάδι.

Τα αποτελέσματα έδειξαν ότι η μορφολογία των χλωροπλαστών σε κύτταρα *lefko1-1* κοτυληδόνων που αναπτύχθηκαν στο σκοτάδι είναι προβληματική (Εικόνα 5). Οι χλωροπλάστες δεν έχουν την τυπική οργάνωση των grana και δεν συσσωρεύουν άμυλο όπως οι χλωροπλάστες των φυτών αγρίου τύπου (Εικόνα 5α). Η αύξηση της σακχαρόζης στο θρεπτικό μέσο ανάπτυξης, βελτιώνει την εικόνα των χλωροπλαστών στα πρώτα φύλλα του μεταλλάγματος *lefko1-1* που εμφανίζουν βελτιωμένο πρότυπο οργάνωσης των θυλακοειδών (Εικόνα 5δ). Στις ίδιες συνθήκες στους χλωροπλάστες των πρώτων φύλλων των φυτών αγρίου τύπου παρατηρούνται αμυλόκοκκοι. Οι αμυλόκοκκοι αποτελούν αποθησαυριστικές δομές και αντιπροσωπεύουν το καθαρό κέρδος σε ενέργεια και σκελετούς άνθρακα από τη λειτουργία της φωτοσύνθεσης. Επομένως, η χαρακτηριστική αλλοίωση της μορφολογίας των χλωροπλαστών του φυτού *lefko1-1*, δε φαίνεται να είναι συνέπεια φωτοαναστολής -περίσσειας φωτός, αφού η εικόνα των κοτυληδόνων του δε βελτιώνεται στο σκοτάδι. Από την άλλη, φαίνεται να υπάρχει πρόβλημα θρεπτικής ανεπάρκειας εξαιτίας πιθανής κακής λειτουργίας του φωτοσυνθετικού μηχανισμού, αφού οι χλωροπλάστες εκμεταλλεύονται το συμπλήρωμα σακχαρόζης για να οργανώσουν καλύτερα τη λεπτή τους δομή. Στις αντίστοιχες συνθήκες οι χλωροπλάστες των φυτών αγρίου τύπου αποταμιεύουν την περίσσεια της σακχαρόζης με τη μορφή αμυλοκόκκων (Εικόνα 5γ).



**Εικόνα 5. Μορφολογία χλωροπλαστών μεταλλάγματος *Col-0* και *lefko1-1* φυτών**

α) χλωροπλάστης από κύτταρα κοτυληδόνων φυτών *Col-0* ηλικίας 5 ημερών που αναπτύχθηκαν στο σκοτάδι, β) χλωροπλάστης από κύτταρα κοτυληδόνων φυτών *lefko1-1* ηλικίας 7 ημερών που αναπτύχθηκαν στο σκοτάδι, γ) χλωροπλάστες από κύτταρα πρώτων φύλλων φυτών *Col-0* ηλικίας 14 ημερών σε θρεπτικό μέσο με 3 % σακχαρόζη και δ) χλωροπλάστες από κύτταρα πρώτων φύλλων φυτών *lefko1-1* 0 ηλικίας 14 ημερών σε θρεπτικό μέσο με 3 % σακχαρόζη.

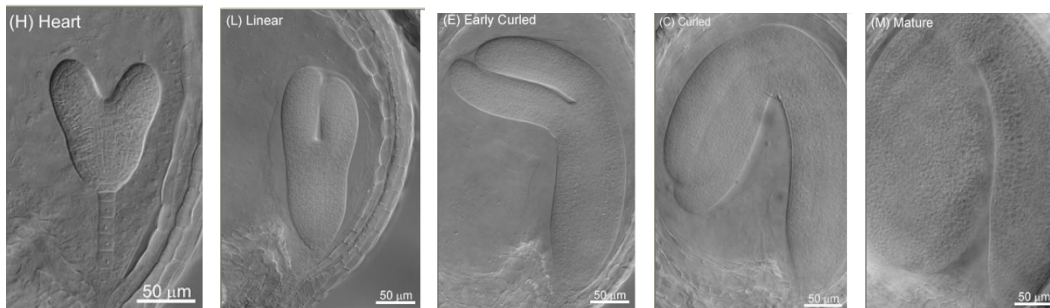
Η ποσοτικοποίηση των φωτοσυνθετικών χρωστικών ταυτίστηκε με το φαινότυπο της λεύκανσης των κοτυληδόνων του φυτού *lefko1-1* και τη σταδιακή αναστροφή του φαινοτύπου που παρατηρείται στα πρώτα φύλλα. Μετρήσεις των φωτοσυνθετικών χρωστικών, χλωροφύλλης α και β (Chl α και β) και των καροτενοειδών που αποτελούν χρωστικές προστασίας των κεραίων των φωτοσυστημάτων πραγματοποιήθηκαν στις κοτυληδόνες και στα πρώτα φύλλα φυτών *Col-0* και *lefko1-1*. Οι μετρήσεις αποκάλυψαν δραματική μείωση των χρωστικών της φωτοσύνθεσης, όσο και των καροτενοειδών στις κοτυληδόνες του μεταλλάγματος *lefko1-1*, σε σχέση με τα φυτά αγρίου τύπου. Όσον αφορά στα φύλλα του *lefko1-1* η περιεκτικότητα των χρωστικών είναι ελαφρά μειωμένη σε σύγκριση με τα φυτά αγρίου τύπου (Προσωπική επαφή Γ. Δάρας).

Η μορφολογία των χλωροπλαστών βελτιώνεται στα πρώτα φύλλα σε σχέση με τις κοτυληδόνες, φαινόμενο που επιβεβαιώνεται και με τη ποσοτική ανάλυση των χρωστικών

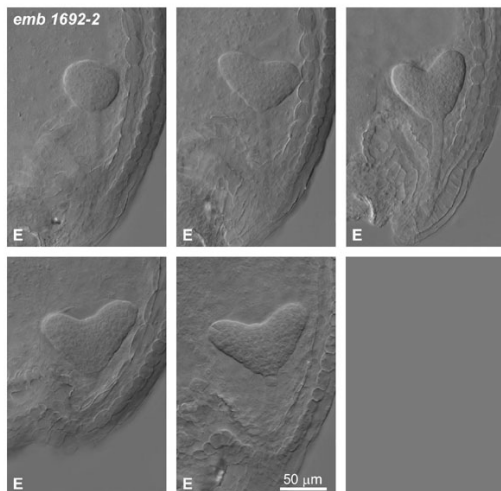
της φωτοσύνθεσης. Το παραπάνω γεγονός υποδεικνύει ότι η μετάλλαξη *lefko1-1* επηρεάζει τη μορφογένεση ιστών με εμβρυακές καταβολές, όπως είναι οι κοτυληδόνες. Οπότε η πρωτεΐνη Lefko1 φαίνεται να έχει κρίσιμο ρόλο στην ανάπτυξη του εμβρύου, κάτι το οποίο επιβεβαιώνεται και από τη μελέτη του μεταλλάγματος *lefkothea1-2*.

## 1.2 Φαινοτυπικά χαρακτηριστικά του αλληλόμορφου *lefkothea1-2*

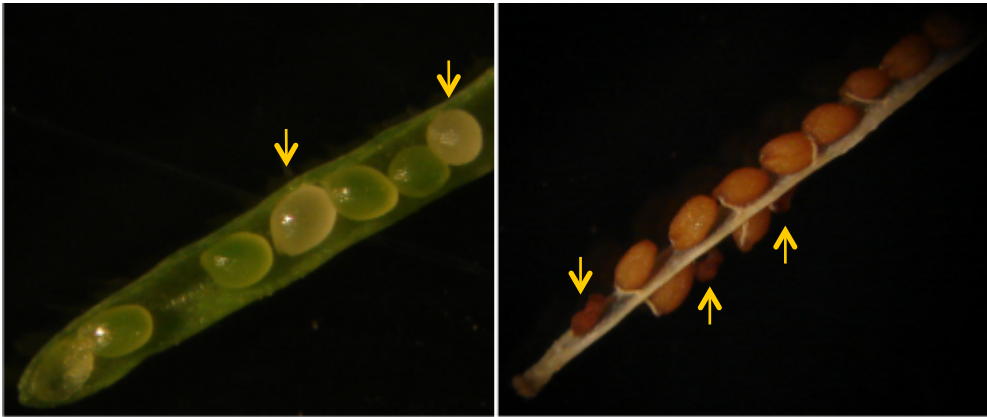
Βιοπληροφορική αναζήτηση και άλλων αλληλομόρφων μεταλλαγμάτων οδήγησε στον εντοπισμό του αλληλόμορφου *lefkothea1-2*, που οφείλεται σε T-DNA ένθεση στο γονίδιο At5g62990 (Εικόνα 6), ανωφορικά 30bp από το κωδικόνιο έναρξης. Η ίδια σειρά T-DNA είχε προηγουμένα μελετηθεί (Tzafrir, 2003; Tzafrir, 2004) και είχε χαρακτηριστεί ως *emb1692*. Η σειρά αυτή διατίθεται από τον ερευνητή Meinke, ο οποίος την εντόπισε στα πλαίσια μιας γενικότερης έρευνας για μεταλλάγματα που εμπλέκονται στη εμβρυογένεση. Έχει μελετηθεί η μορφολογία των εμβρύων της σειράς με μικροσκοπία Nomarski και παρουσιάζεται στη βάση δεδομένων [www.seedgenes.org](http://www.seedgenes.org) (Εικόνα 6). Η μετάλλαξη αυτή είναι θνησιγόνος. Περαιτέρω παρατήρηση των καρποταξιών του T-DNA αλληλομόρφου *lefko1-2* επιβεβαίωσαν τη Μενδελιανή διάσχιση του θνησιγόνου φαινότυπου σε κατάσταση ομοζυγωτίας (Εικόνα 7).



Nomarski Images



**Εικόνα 6. Φωτογραφίες μικροσκοπίας Nomarski από τη βάση δεδομένων [www.seedgenes.org](http://www.seedgenes.org).** Α. Μικροφωτογραφίες εμβρύων στα στάδια καρδιάς (heart), τορπίλης (linear or torpedo), μέχρι ώριμο (mature) αγρίου τύπου φυτού Col-0. Β. Μικροφωτογραφίες εμβρύων από στάδια σφαιρικό, καρδιάς του μεταλλάγματος *emb1692/lefko1-2*. Η ανάπτυξη του μεταλλάγματος σταματάει στο στάδιο της καρδιάς

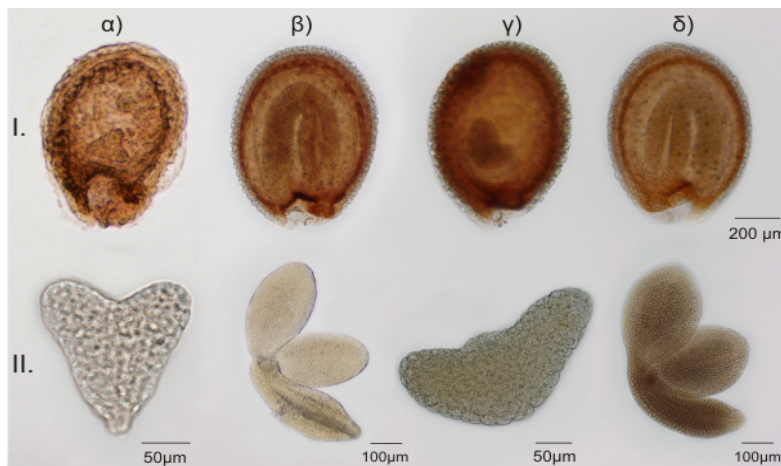


**Εικόνα 7. Καρποταξίες ετερόζυγων ως προς την ένθεση φυτών *lefko1-2*.** Παρατηρείται ανομοιομορφία των αναπτυσσόμενων σπερμάτων. Με βέλη σημειώνονται τα μη φυσιολογικά σπέρματα. Φαίνονται τα σπέρματα από την ημίζυγη κατάσταση για την ένθεση. Τα άσπρα και τα συρρικνωμένα σπέρματα υποδηλώνουν την ομόζυγη κατάσταση για την T-DNA ένθεση (3:1) που οδηγεί σε θάνατο

Τα ώριμα σπέρματα εξετάστηκαν περισσότερο για να ελεγχθεί αν παρουσιάζουν κάποια ανωμαλία στη δομή ή στην ανάπτυξη των εμβρύων που εμπεριέχουν, σαν κι αυτές που παρουσιάζουν άλλα μεταλλάγματα όπως τα *sco 1-2*, *sco 1-3* (Ruppel and Hangarter, 2007) και το *raspberry3* (Arya, 2002). Η επεξεργασία ώριμων σπερμάτων ετερόζυγων ως προς την ένθεση φυτών *lefko1-2* με NaOH έδειξε πως το  $\frac{1}{4}$  των εμβρύων, που αντιστοιχούν στα ομόζυγα ως προς την ένθεση σπέρματα, έφτανε ως το καρδιόσχημο στάδιο ανάπτυξης (Εικόνα 8γ), ενώ τα υπόλοιπα  $\frac{3}{4}$  (Εικόνα 8δ) ολοκλήρωναν την ανάπτυξή τους φυσιολογικά, σχηματίζοντας κοτυληδόνας και ριζίδιο. Μολονότι η ανάπτυξη των ομόζυγων ως προς την ένθεση εμβρύων σταματά στο καρδιόσχημο στάδιο, το έμβryo ακόμη και στη φάση αυτή, δεν έχει την οργανωμένη και συνεκτική δομή που παρουσιάζει το αντίστοιχης ηλικίας Col-0 έμβryo (Εικόνα 8α). Στο τελικό, όμως, στάδιο ανάπτυξης των εμβρύων, η εικόνα που παρουσιάζουν τα Col-0 έμβρυα (Εικόνα 8β) είναι

όμοια με τα έμβρυα που φέρουν την ένθεση σε ετεροζυγωτή κατάσταση (Εικόνα 8). Έτσι είναι φανερό ότι το γονίδιο είναι ουσιαστικό για την ανάπτυξη του οργανισμού και ότι η λειτουργία της πρωτεΐνης που κωδικοποιείται από το γονίδιο At5g62990 δεν καλύπτεται από κάποιο άλλο γονίδιο. Γενικά στα φυτά, ορισμένες λειτουργίες (γονίδια) δεν καλύπτονται από κάποιες άλλες που κωδικοποιούνται συνήθως από μέλη πολυγονιδιακής οικογένειας.

Σύμφωνα με τα παραπάνω και επειδή η μοριακή και λειτουργική ανάλυση του μεταλλάγματος *lefko1-2* δεν καθίσταται εύκολη λόγω της θνησιμότητας του εμβρύου, ο αλληλόμορφος *lefko1-1* αποτελεί το πλέον κατάλληλο γενετικό υπόβαθρο για τη διεξοδική μελέτη της λειτουργίας αφού ναι μεν η αμινοξική υποκατάσταση ( $G_{373} \rightarrow D$ ) ευθύνεται για σημαντικές φαινοτυπικές ανωμαλίες, τα φυτά όμως μπορούν να αναπτυχθούν, γονιμοποιούνται και δίνουν απογόνους.



**Εικόνα 8. Μορφολογία σπερμάτων και εμβρύων από το αγρίου τύπου φυτό *Col-0* και το αλληλόμορφο μετάλλαγμα *lefko1-2*** I. Σπέρματα από α) *Col-0* φυτό στη φάση ανάπτυξης καρδιόσχημου εμβρύου, β) *Col-0* φυτό στη φάση του ώριμου εμβρύου γ), δ) από καρποταξία ετεροζυγωτού ως προς την ένθεση φυτού. Στο γ) φαίνεται σπέρμα από ομόζυγο ως προς την T-DNA ένθεση έμβρυο (1/4), ενώ στο δ) σπέρμα που περιέχει έμβρυο με την ένθεση T-DNA σε ετεροζυγική κατάσταση ή χωρίς ένθεση (3/4). II. Παρουσιάζονται τα αντίστοιχα έμβρυα από τα σπέρματα που βρίσκονται άνωθέν τους.

### 1.3 Η πρωτεΐνη LEFKO1 περιέχει το τμήμα DUF860 ή PORR

Σε πρωτεϊνικό επίπεδο, το γονίδιο *LEFKO1* (At5g62990) αποτελείται από τρεις κύριες δομές. Το N-τελικό άκρο που περιλαμβάνει το πεπτιδίο συνθήματος για τοποθέτηση στον χλωροπλάστη, το οποίο εκτείνεται από το πρώτο έως το 60ο αμινοξύ), το τμήμα DUF 860 (*Domain of Unknown Function 860*) μεγέθους 319aa (61 – 380) και το C-τελικό άκρο μεγέθους 114aa (381–494aa) (Εικόνα 9). Το τμήμα DUF860 παρουσιάζεται αποκλειστικά σε γονίδια φυτικών οργανισμών.



**Εικόνα 9. Σχηματική απεικόνιση των δομών του γονιδίου *LEFKO1* και η θέση των μεταλλάξεων.** Το γονίδιο *LEFKO1* αποτελείται από το N-άκρο, το τμήμα DUF860/PORR και το C-άκρο. Η μετάλλαξη *lefko1-1* βρίσκεται στο τμήμα DUF860/PORR. Η μετάλλαξη *lefko1-2* είναι η ένθεση T-DNA 30bp ανωφορικά του κωδικονίου έναρξης του γονιδίου και απεικονίζεται με το ανεστραμμένο τρίγωνο.

Πρόσφατα, για πρώτη φορά στο φυτό *Zea mays* αναγνωρίσθηκε και μελετήθηκε η πρωτεΐνη που περιέχει το τμήμα DUF860 (*domain of unknown function*) (Kroeger *et al.*, 2009). Η πρωτεΐνη WTF1 (**W**hat's **t**his **f**actor 1) εντοπίστηκε να ανοσοκαθίζει με τα ιντρόνια της ομάδας II χλωροπλαστικών RNA (splicing intron group II). Σχηματίζει ομοδιμερές σύμπλοκο με την πρωτεΐνη RNC1 (RNase III-domain protein) και φαίνεται ότι είναι απαραίτητη για την ομαλή λειτουργία του ματίσματος των χλωροπλαστικών ιντρονίων (Kroeger *et al.*, 2009). Πειραματικά δεδομένα έδειξαν ότι το τμήμα DUF860 έχει ένα μέχρι τώρα άγνωστο μοτίβο *in vitro* σύνδεσης με το RNA (Kroeger *et al.*, 2009).

Συνεπώς το DUF860 μετονομάστηκε σε PORR (**P**lant **O**rganelle **R**NA **R**ecognition motif, μοτίβο αναγνώρισης RNA φυτικών οργανιδίων). Το *Arabidopsis thaliana* έχει 15 ομόλογα γονίδια που φέρουν το τμήμα DUF860, μεταξύ των οποίων είναι το γονίδιο *Lefko1* (At5g62990) και το ορθόλογο At4g01037 του *Zea mays* WTF1 στο *Arabidopsis*. Τα περισσότερα μέλη της οικογένειας DUF860 φαίνεται να στοχεύουν στους χλωροπλάστες ή στα μιτοχόνδρια, γεγονός που υποδηλώνει ότι οι πρωτεΐνες με αυτό το τμήμα πιθανόν να έχουν πολλαπλούς ρόλους στο μεταβολισμό του RNA και στα δύο οργανίδια. Παρόλα αυτά, από βιοπληροφορικές αναλύσεις που πραγματοποιήθηκαν, υπάρχουν ενδείξεις ότι παράλληλα τοποθετούνται και στον πυρήνα των κυττάρων και έτσι προκύπτουν σημαντικές ερωτήσεις σε σχέση με τη λειτουργία τους, την εξέλιξη των μηχανισμών επικοινωνίας μεταξύ οργανιδίων και πυρήνα σε συνάρτηση με την ανάπτυξη του φυτού.

Έτσι γίνεται φανερό ότι μια εμπειρισταωμένη βιοπληροφορική ανάλυση είναι ουσιαστική για την συγκριτική γονιδιωματική των γονιδίων στο *Arabidopsis*, της *in silico* τοποθέτησης στα διάφορα οργανίδια του κυττάρου καθώς και της δομικής ομοιότητας ή διαφοράς σε επίπεδο πρωτεϊνών ή και γονιδίων. Η παιρετέρω ανάλυση σε επίπεδο διαφορετικών γονιδιωμάτων των φυτών από τα αρχαία μέχρι τα πιο εξελιγμένα θα φωτίσει τη λειτουργική εξελικτική πορεία της πρωτεΐνης και ενδεχομένως να δώσει αρκετά στοιχεία για τη συμβιωτική θεωρία ή ακόμη για την εξελικτική πορεία της δομικής συμπεριφοράς της εξαγωγής των ιντρονίων (αυτοσυραφή σε σχέση με τη χρήση πρωτεϊνών όπως στα ιντρονια των ευκαρυωτών τύπου I).

## 2. ΥΛΙΚΑ & ΜΕΘΟΔΟΙ

### 2.1 Μηχανές ανάλυσης και πρόβλεψης τρισδιάστατης δομής πρωτεϊνών

#### 2.1.1 *i-tasser*

Το I-TASSER (Iterative Threading **ASSE**mblly **R**efinement, βελτίωση με Επαναληπτική Συνέλευση Threading) είναι μια ιεραρχική μέθοδο για τη δομή των

πρωτεϊνών και πρόβλεψη της λειτουργίας αυτών. Τα διαρθρωτικά πρότυπα προσδιορίζονται από τη PDB (Protein Data Bank) με πολλαπλή LOMETS (Local Meta-Threading-Server) προσέγγιση. Πλήρους μήκους ατομικά μοντέλα κατασκευάζονται και στη συνέχεια με πρότυπο το επαναληπτικό θραύσμα προσομοιώσεων. Τελικά η λειτουργία του στόχου προέρχεται από το σπείρωμα των 3D μοντέλων μέσω της βάσης δεδομένων της λειτουργίας των πρωτεϊνών BioLiP (Ligand-protein binding database). I-TASSER (ως «Zhang-Server»), κατετάγη ως ο No1 server για πρόβλεψη δομής των πρωτεϊνών κατά τα τελευταία CASP7 (Critical Assessment of Techniques for Protein Structure Prediction), CASP8, CASP9, πειράματα CASP10, και CASP11. Επίσης, κατατάσσεται ως η καλύτερη πρόβλεψη για λειτουργία με CASP9.

### 2.1.2 Phyre2

Το Phyre2 (**P**rotein **H**omology/**A**nalog**Y** **R**ecognition **E**ngine ) είναι μια σουίτα εργαλείων που διατίθενται στο διαδίκτυο για τη πρόβλεψη και ανάλυση πρωτεϊνικών δομών, τη λειτουργία τους και μεταλλάξεις που μπορεί να οδηγήσουν για ταυτοποίηση της λειτουργικής δομής. Το επίκεντρο του Phyre2 είναι να παρέχει σε βιολόγους ένα εργαλείο βιοπληροφορικής. Το Phyre2 αντικαθιστά το Phyre, η αρχική έκδοση του οποίου έχει ήδη δημοσιευτεί. Σε αυτήν την αναβαθμισμένη έκδοση, χρησιμοποιείται προηγμένη μέθοδος ανίχνευσης ομολογίας για την κατασκευή 3D μοντέλων, προβλέπουν θέσεις σύνδεσης και ανάλυσης αποτελεσμάτων των αμινοξικών παραλλαγών (π.χ., nonsynonymous SNPs [nsSNPs] single nucleotide polymorphism- πολυμορφισμοί ενός νουκλεοτιδίου) για την αλληλουχία της πρωτεΐνης του χρήστη. Οι χρήστες οδηγούνται μέσω αποτελεσμάτων με ένα απλό interface σε ένα επίπεδο λεπτομέρειας που καθορίζουν οι ίδιοι. Το πρωτόκολλο αυτό θα καθοδηγήσει τους χρήστες από την υποβολή της πρωτεϊνικής αλληλουχίας για την ερμηνεία της δευτεροταγής και τριτοταγής δομής των μοντέλων τους, τη σύνθεσή τους και το ποιοτικό μοντέλο. Μια σειρά από πρόσθετα διαθέσιμα εργαλεία περιγράφεται ώστε να βρεθεί μια δομή πρωτεΐνης σε ένα γονιδίωμα. Η υποβολή μεγάλου αριθμού αλληλουχιών ταυτόχρονα είναι δυνατή αυτόματα όμως οι πρωτεΐνες αυτές είναι δύσκολο να μοντελοποιηθούν.



## 2.2 Μηχανές ανάλυσης και πρόβλεψης πυρηνικού εντοπισμού πρωτεϊνών

### 2.2.1 cNLS mapper

Το cNLS Mapper προβλέπει με ακρίβεια τα σήματα πυρηνικού εντοπισμού (NLSs, Nuclear Localization Signals) ειδικά για το μονοπάτι της ιμφορτίνης (importin) α/β υπολογίζοντας τις NLS βαθμολογίες (επίπεδα δραστηριοτήτων NLS), αλλά όχι με τη συμβατική αναζήτηση ομοιότητας αλληλουχίας ή με τη στρατηγική μηχανικής μάθησης. Οι βαθμολογίες NLS υπολογίζονται σε τέσσερα NLS προφίλ, καθένα από τα οποία αντιπροσωπεύει μια συνεισφορά για κάθε αμινοξύ σε κάθε θέση μέσα σε μια τάξη-δομή NLS στο σύνολο των δραστηριοτήτων NLS. Τα προφίλ που παράγονται από μια εκτεταμένη ανάλυση αντικατάστασης αμινοξέος για κάθε κατηγορία NLS στην ζύμη. Έχει βρεθεί ότι κάθε υπόλειμμα μέσα σε NLS, στις περισσότερες των περιπτώσεων, συμβάλλει προσθετικά και ανεξάρτητα σε ολόκληρη τη δραστηριότητα. Έτσι, τα NLS μπορεί να υπολογιστούν με διαδοχική θετική ή αρνητική συμβολή βαθμολογιών που περιλαμβάνονται σε προφίλ NLS. Η πρόσθετη ιδιότητα των NLS μοτίβων επέτρεψε το σχεδιασμό ισχυρών αναστολών πεπτιδίων ειδικά για το μονοπάτι εισαγωγής στον πυρήνα με βάση την ιμφορτίνη α/β επιλέγοντας κατά σειρά αμινοξέα με υψηλή βαθμολογία συνεισφοράς σε ένα προφίλ NLS.

### 2.2.2 Suba3

Το Suba παρέχει ένα ισχυρό εργαλείο για τη διερεύνηση της ενδοκυττάριας κατανομής στο φυτό *Arabidopsis* μέσα από την ενοποίηση διαφορετικών συνολικών δεδομένων και μέσω παροχής προσβάσιμης διαδικτυακής συνεργασίας για την κατασκευή ισχυρών ερωτήσεων με βάση το χρήστη, που προκύπτουν σε μια one-stop-shop για την πρωτεϊνική εντοπισμού σε αυτό το εργασιακό μοντέλο.

Το SUBA3 (**SUB**cellular database of **A**rabidopsis ) κατέχει μεγάλης κλίμακας πρωτεομικά σύνολα εντοπισμού με ή χωρίς GFP στα διάφορα κυτταρικά διαμερίσματα του *Arabidopsis*. Περιέχει επίσης συλλογικές βιοπληροφορικές προβλέψεις εντοπισμού για τις υποκυττάριας πρωτεΐνες. Στοιχεία αλληλεπίδρασης πρωτεΐνης-πρωτεΐνης επιτρέπουν την αναζήτηση για την αλληλεπίδραση ζεύγων πρωτεϊνών.

Πληροφορίες υποκυτταρικού εντοπισμού μπορούν να συμβάλουν στην κατανόηση της λειτουργίας των πρωτεϊνών, πρωτεΐνες απελευθέρωσης και της βιολογικής μεταξύ τους σχέσης. Ενώ μια ποικιλία τεχνολογιών απασχολούνται αυτή τη στιγμή ώστε να

προσδιοριστεί η υποκυτταρική τοποθέτηση των πρωτεϊνών, πολλές από τις πληροφορίες αυτές δεν είναι διαθέσιμες σε έναν ολοκληρωμένο ιστοτόπο. Σε μια προσπάθεια για μια σαφέστερη εικόνα των πειραματικών δεδομένων και για να γίνει γενικότερα μια καλύτερη κατανόηση της υποκυττάριας διαμερισματοποίησης έχουν συγκεντρωθεί διάφορες πηγές δεδομένων για την κατασκευή του SUBA3. Η βάση δεδομένων έχει μια προσβάσιμη διαδικτυακή απεικόνιση που επιτρέπει σε προχωρημένα συνδυαστικά ερωτήματα που πρέπει να αναληφθούν και να συνδυαστούν σχετικά με τα δεδομένα που παρέχονται.

## 2.3 BLAST

Στη βιοπληροφορική, BLAST **B**asic **L**ocal **A**lignment **S**earch **T**ool (βασικό εργαλείο ανάλυσης τοπικής στοίχισης) είναι ένας αλγόριθμος για τη σύγκριση πρωτογενών βιολογικών πληροφοριών αλληλουχίας, όπως τις αλληλουχίες αμινοξέων των διαφορετικών πρωτεϊνών ή αλληλουχιών νουκλεοτιδίων DNA. Μια αναζήτηση BLAST επιτρέπει στον ερευνητή να συγκρίνει μια ακολουθία με μια βιβλιοθήκη ή βάση δεδομένων αλληλουχιών, και να προσδιορίσει τη βιβλιοθήκη αλληλουχιών που μοιάζουν με την αλληλουχία ερωτήματος μέσα σε ένα ορισμένο όριο.

Διαφορετικοί τύποι BLAST είναι διαθέσιμοι ανάλογα με τις αλληλουχίες. Για παράδειγμα, μετά την ανακάλυψη ενός προηγούμενως άγνωστο γονίδιο στον ποντικό, ένας επιστήμονας μπορεί να εκτελέσει τυπικά μια αναζήτηση BLAST του ανθρώπινου γονιδιώματος για να βρεθεί αν οι άνθρωποι φέρουν ένα παρόμοιο γονίδιο. Το BLAST θα προσδιορίσει αλληλουχίες στο ανθρώπινο γονιδίωμα που μοιάζουν με το γονίδιο ποντικού που βασίζεται στην ομοιότητα της αλληλουχίας. Ο αλγόριθμος BLAST και το πρόγραμμα έχουν σχεδιαστεί από τον Stephen Altschul, Warren Gis, Webb Miller, Eugene Myers, και τον David J. Lipman στο Εθνικό Ινστιτούτο Υγείας και δημοσιεύθηκε στο περιοδικό *Journal of Molecular Biology* το 1990 και παρατίθεται πάνω από 50.000 φορές (REF). Η συγκριτική ανάλυση δεν περιορίζεται στα ζώα αλλά επεκτείνεται σε όλους τους οργανισμούς. Είναι από τα βασικότερα βιοπληροφορικά εργαλεία που αξιοποιείται πλέον από διαφορετικούς κλάδους της επιστήμης της Βιολογίας

## 2.4 Βάσεις Δεδομένων

### 2.4.1 GenBank

Η βάση δεδομένων αλληλουχιών GenBank είναι προσβάσιμη από όλους. Παρέχει μια σχολαστική συλλογή όλων των διαθέσιμων νουκλεοτιδικών αλληλουχιών και

μπορεί να δημιουργήσει πρωτεϊνικές μεταφράσεις των αλληλουχιών αυτών. Αυτή η βάση δεδομένων παράγεται και διατηρείται από το National Center for Biotechnology Information (NCBI). Το National Center for Biotechnology Information αποτελεί μέρος του Εθνικού Ινστιτούτου Υγείας των Ηνωμένων Πολιτειών. Η GenBank και οι συνεργάτες του λαμβάνουν ακολουθίες που παράγονται σε εργαστήρια σε όλο τον κόσμο από περισσότερους από 100.000 διακριτούς οργανισμούς. Στα περισσότερα από 30 χρόνια από την ίδρυσή της, η GenBank έχει γίνει η πιο σημαντική και η πιο ισχυρή βάση δεδομένων για την έρευνα σε όλα σχεδόν τα πεδία της βιολογίας, των οποίων τα δεδομένα είναι προσβάσιμα ελεύθερα και προβάλλονται από εκατομμύρια ερευνητών σε όλο τον κόσμο. Η GenBank συνεχίζει να αυξάνεται με εκθετικό ρυθμό, διπλασιάζοντας τη βάση δεδομένων κάθε 18 μήνες. Το Φεβρουάριο του 2013, περιείχε πάνω από 150 δισεκατομμύρια νουκλεοτιδικές βάσεις σε περισσότερες από 162 εκατομμύρια σειρές. Είναι δομημένο με τέτοιον τρόπο ώστε να γίνεται άμεση υποβολή πληροφοριών από επιμέρους εργαστήρια, καθώς και από μεγάλες υποβολές μεγάλης κλίμακας από κέντρα αλληλούχισης.

#### 2.4.2 TAIR

Το TAIR (The **A**rabidopsis **I**nformation **R**esource) διατηρεί μια βάση δεδομένων με γενετικά και μοριακά δεδομένα βιολογίας για το μοντέλο φυτό *Arabidopsis thaliana*. Διαθέσιμες πληροφορίες από το TAIR περιλαμβάνει την πλήρη αλληλουχία του γονιδιώματος μαζί με τη γονιδιακή δομή, πληροφορίες για το προϊόν του γονιδίου, την έκφραση του γονιδίου, τα αποθέματα του DNA (κλώνων) και των σπόρων, χάρτες γονιδιώματος, γενετικοί και φυσικοί δείκτες, δημοσιεύσεις, καθώς και πληροφορίες σχετικά με την έρευνα για το *Arabidopsis* στην ευρύτερη κοινότητα. Γονιδιακά δεδομένα λειτουργίας ενημερώνονται κάθε εβδομάδα από τις τελευταίες δημοσιευμένες ερευνητικές βιβλιογραφίες και από δεδομένα δημοσίων υποβολών.

### 2.5 Απομόνωση ολικού RNA από σπορόφυτα *Arabidopsis thaliana*

Φρέσκο φυτικό υλικό (ολόκληρα φυτά *Arabidopsis* ή διάφορα όργανά του) ομογενοποιείται παρουσία υγρού αζώτου με κατάλληλο γουδί λειοτρίβησης.

- Το δείγμα μεταφέρεται σε eppendorf tube των 2ml. (Για κάθε 2gr ιστού προστίθεται 1ml διαλύματος απομόνωσης και 1ml φαινόλης). Έντονη ανάμιξη του δείγματος.
- Φυγοκέντρηση (max speed) για 10min.
- Το υπερκείμενο μεταφέρεται σε νέο tube και ακολουθεί καθαρισμός του δείγματος με ίσο όγκο φαινόλης.
- Επαναλαμβάνεται φυγοκέντρηση κάτω από τις ίδιες συνθήκες.
- Το υπερκείμενο μεταφέρεται σε νέο tube και ακολουθεί προσθήκη ίσου όγκου φαινόλης-χλωροφόρμιου-ισοαμυλικής αλκοόλης (25:24:1). Έντονη μίξη.
- Επαναλαμβάνεται η ίδια φυγοκέντρηση.
- Μεταφορά του υπερκειμένου σε νέο tube χωρίς να διαταραχθεί η ενδιάμεση φάση και προσθήκη ίσου όγκου φαινόλης-χλωροφόρμιου-ισοαμυλικής αλκοόλης (25:24:1). Έντονη μίξη.
- Επαναλαμβάνεται η ίδια φυγοκέντρηση.
- Ξανά μεταφορά του υπερκειμένου σε νέο tube χωρίς να διαταραχθεί η ενδιάμεση φάση και προσθήκη ίσου όγκου χλωροφόρμιου-ισοαμυλικής αλκοόλης (24:1). Έντονη μίξη.
- Επαναλαμβάνεται η ίδια φυγοκέντρηση.
- Για την κατακρήμνιση των νουκλεϊνικών οξέων, στο υπερκείμενο προστίθεται 1/10 του όγκου 3M CH<sub>3</sub>COONa pH 4.8 και 2.5 όγκοι παγωμένης αιθανόλης 100% και μεταφέρεται το δείγμα στους -20 °C overnight.
- Ακολουθεί φυγοκέντρηση στους 4 °C σε max speed για 10-30 min..
- Το ίζημα διαλύεται σε αποστειρωμένο ddH<sub>2</sub>O (με θέρμανση στους 65°C όταν χρειάζεται)

## 2.6 Απομόνωση ολικού RNA από έμβρυα *Arabidopsis thaliana*

Το πρωτόκολλο που ακολουθεί αποτελεί τροποποίηση του πρωτοκόλλου από την αναφορά των Onate-Sanchez *et al.*, 2008.

Έμβρυα από το κατάλληλο στάδιο του φυτού *Arabidopsis thaliana*, απομονώνονται με τη βοήθεια λαβίδας και στερεοσκοπίου

Στα έμβρυα προστίθεται το διάλυμα απομόνωσης (0.4 M LiCl, 0.2 M Tris pH:8, 25 mM EDTA, 1% SDS ) όσο το δυνατό λιγότερο όγκος, μέχρι 150μl. Μέσα στο διάλυμα τα έμβρυα μπορούν να αποθηκευτούν και στους -80°C

- Με μικρογουδί ομογενοποιούνται στα έμβρυα
- Σε 150μl διάλυμα ομογενοποίησης (έμβρυα και διάλυμα απομόνωσης) προστίθενται 150μl χλωροφόρμιο και ακολουθεί ισχυρή ανακίνηση για 30".
- Φυγοκέντρηση για 5' σε θερμοκρασία δωματίου.
- Μεταφέρεται το υπερκείμενο σε νέο erpendorf tube και προστίθενται 300μl φαινόλης, ακολουθεί ισχυρή ανακίνηση για 1'. Στη συνέχεια προστίθενται 150μl χλωροφόρμιο κα ισχυρή ανακίνηση για 3'.
- Φυγοκέντρηση για 5' σε θερμοκρασία δωματίου.
- Μεταφέρεται το υπερκείμενο σε νέο erpendorf tube (~250μl) και προστίθεται 1/3 του όγκου 8M LiCl (83μl). Ακολουθεί ελαφριά ανακίνηση. Το δείγμα αφήνεται στους -20°C για 1h.
- Φυγοκέντρηση για 30' στους 4°C
- Απομακρύνεται το υπερκείμενο και το ίζημα επαναιωρείται σε 30μl ddH<sub>2</sub>O.
- Ακολουθεί αντίδραση DNase I (στα 60μl) όπως αναφέρεται παρακάτω.
- Μετά την αντίδραση DNase I προστίθενται στο δείγμα 440μl ddH<sub>2</sub>O, 7μl 3M NaAc pH:5.2 (Sodium Acetate) και 250μl 100% αιθανόλης.
- Φυγοκέντρηση για 10' στους 4°C. (Με αυτόν τον τρόπο απομακρύνονται τα συστατικά της αντίδρασης DNase I και οι πολυσακχαρίτες).
- Το υπερκείμενο μεταφέρεται σε νέο erpendorf tube και εκεί προστίθενται 43μl 3M NaAc pH:5.2 και 750μl 100% αιθανόλης. Ανακινείται ελαφρά και αφήνεται overnight στους -20 °C.
- Φυγοκέντρηση για 30' στους 4°C.
- Απομακρύνεται προσεχτικά το υπερκείμενο, ακολουθεί μια γρήγορη φυγοκέντρηση για να απομακρυνθεί τυχόν υπερκείμενο που έχει μείνει και το ίζημα επαναιωρείται στα 20-25μl με ddH<sub>2</sub>O.

## 2.7 Αντίδραση DNase

Απαραίτητο βήμα για να καταστραφεί το DNA που απομονώνεται κατά την απομόνωση RNA,

- προσθέτουμε τα ακόλουθα αντιδραστήρια σε ένα αποστειρωμένο erpendorf:

--	--

Nucleic Acids	75 $\mu$ l
RQ1 buffer	15 $\mu$ l
RQ1 DNase (1u/ $\mu$ l)	5 $\mu$ l
RNase out	1 $\mu$ l
H2O	54 $\mu$ l
TOTAL	150 $\mu$ l

- 1 ώρα στους 37 °C
- Ανεβάζουμε τον όγκο στα 400  $\mu$ l
- Ακολουθεί καθαρισμός με με φαινόλη/Sevag και καθίζηση με αιθανόλη

## 2.8 Απομόνωση γονιδιωματικού DNA από Arabidopsis

Φρέσκο φυτικό υλικό ομογενοποιείται παρουσία υγρού αζώτου σε κατάλληλο γουδί λειοτρίβισης ή μέσα σε erpendorf tube με μικρογουδί.

- Το ρυθμιστικού διάλυμα CTAB (2% CTAB [hexadecyltrimethylammonium bromide] 100 mM TrisHCl [pH=8], 20 mM EDTA, 1.4 M NaCl) τοποθετείται στο υδατόλουτρο στους 65°C.
- Προσθέτουμε έναν όγκο (200 $\mu$ l) ζεστού CTAB και τοποθετούμε τα erpendorf σε υδατόλουτρο στους 65°C για 10-30 λεπτά.
- Προστίθεται ίσος όγκος SEVAG [χλωροφόρμιου:ισοαμυλικής αλκοόλης (24:1)] και το μίγμα αναδεύεται ελαφρά.
- Φυγοκέντρηση για 3 λεπτά στις μέγιστες στροφές (13.000 στρ/λεπτό).
- Μεταφορά του υπερκειμένου σε καθαρό erpendorf tube.
- Προσθέτουμε 0,7 του όγκου (140 $\mu$ l) ισοπροπανόλης, αναδεύουμε με το χέρι και το αφήνουμε σε ηρεμία για 10 λεπτά σε θερμοκρασία δωματίου.
- Φυγοκέντρηση του δείγματος για 15 λεπτά σε θερμοκρασία δωματίου στις μέγιστες στροφές.
- Το ίζημα που προκύπτει ξεπλένεται προσθέτωντας 0,5 ml 100% αιθανόλης και στη συνέχεια ακολουθεί φυγοκέντρηση για 2 λεπτά στις 13.000 στροφές.
- Απομακρύνση του υπερκειμένου.
- Επαναιώρηση του ιζήματος σε 100 $\mu$ l H<sub>2</sub>O ή TE (10 mM Tris-HCl, pH 7.4, 1 mM EDTA).

## 2.9 Ποσοτικοποίηση νουκλεϊνικών με φασματοφωτόμετρο

Προσδιορισμός της συγκέντρωσης των νουκλεϊνικών οξέων.  
Ο υπολογισμός της συγκέντρωσης των νουκλεϊνικών οξέων έγινε με μέτρηση της οπτικής πυκνότητας (Optical Density, OD) χρησιμοποιώντας το φασματοφωτόμετρο U-1100 της Hitachi. Τα νουκλεϊνικά οξέα απορροφούν σε μήκος κύματος 260nm και η συγκέντρωσή τους υπολογίζεται από τον τύπο των Beer-Lambert. Σε μήκος κύματος 280nm απορροφούν οι πρωτεΐνες εξαιτίας κυρίως της περιέχουσας τρυπτοφάνης, ενώ στα 240nm απορροφούν τα φαινολικά παράγωγα.

Επομένως ο λόγος  $OD_{260}/OD_{280}$  αντιπροσωπεύει την καθαρότητα του διαλύματος των νουκλεϊκών οξέων ως προς τις πρωτεΐνες. Ο ιδανικός λόγος ισούται με 1.8 για DNA και 2 για RNA.

Η συγκέντρωση των νουκλεϊκών οξέων υπολογίζεται από τον τύπο των Beer-Lambert:

$$c(\mu\text{g/ml}) = OD(260) * a * DF(\text{dilution factor})$$

Όπου α ορίζεται ως η συγκέντρωση των νουκλεϊκών οξέων που παρατηρείται όταν  $O.D.260=1$  και συγκεκριμένα για δείγμα DNA,  $a=50 \mu\text{g/ml}$  ενώ για RNA,  $a=40\mu\text{g/ml}$ , D.F. είναι ο συντελεστής αραίωσης και ισούται με την αραιωση που έχει πραγματοποιηθεί για να γίνει η φωτομέτρηση.

## 2.10 Αλυσιδωτή αντίδραση πολυμεράσης (PCR)

Ένα γενικό πρωτόκολλο για την ενίσχυση τμημάτων DNA είναι το ακόλουθο:

Συστατικά Αντίδρασης	Πυκνό Διάλυμα	Όγκος	Τελική Συγκέντρωση
Ρυθμιστικό Διάλυμα PCR <sup>(1)</sup>	10 x	5μl	1 x
dNTPs	2mM	5μl	200μM
Ευθύς εκκινητής	3μM	5μl	300nM
Ανάστροφος εκκινητής	3μM	5μl	300nM
DNA	-	2 μl	10ng gDNA

DyNAzyme™ EXT (FINNZYMES) (Θερμοανθεκτική πολυμεράση)	EXT ( <sup>3</sup> )	1unit/μl	1μl	1 unit
ddH <sub>2</sub> O	-	-	έως τα 50μl	-
Τελικός όγκος αντίδρασης 50μl				

(<sup>1</sup>) Ρυθμιστικό διάλυμα PCR με συγκέντρωση 15mM MgCl<sub>2</sub> [10x Optimized DyNAzyme Ext Buffer (F-514)]

(<sup>3</sup>) Η DyNAzyme™ EXT Πολυμεράση της FINNZYMES (F-512S) έχει το χαρακτηριστικό να προσθέτει αδενίνη στο 3' άκρο του DNA χωρίς να έχουμε προσθέσει αδενίνη ως μήτρα. Αυτό το γνώρισμα διευκολύνει την κλωνοποίηση των προϊόντων της PCR σε φορείς όπως ο pGEM κ.α.

Κάθε ένα από τα συστατικά της αντίδρασης τοποθετούνται σε σωλήνα χωρητικότητας 0,2ml κατάλληλο για αντιδράσεις PCR.

➤ Οι θερμοκύκλοι επιλέγονται με βάση το πρότυπο που παρουσιάζεται ακολούθως:

1 x	Αποδιάταξη του DNA μήτρα για 2 λεπτά στους 94 <sup>0</sup> C
20-30 x	Αποδιάταξη του DNA στους 94 <sup>0</sup> C για 30''  Υβριδισμός εκκινητών με τις συμπληρωματικές αλληλουχίες του DNA, στο αντίστοιχο Tm ( <sup>1</sup> ) για 30''  Επιμήκυνση (πολυμερισμός του DNA) στους 72 <sup>0</sup> C για 2' έως 8' ( <sup>2</sup> )
1 x	Τελική επιμήκυνση προϊόντος στους 72 <sup>0</sup> C για 10'
	Διατήρηση στους 10 °C για 5''

(<sup>1</sup>) Η θερμοκρασία υβριδισμού εξαρτάται από τη θερμοκρασία τήξης Tm του εκκινητή που υπολογίζεται από τον τύπο:  $69,3 + 0,41 \cdot GC\% - 650/\text{αριθμό βάσεων εκκινητή}$ .



<sup>(2)</sup> Ο χρόνος επιμήκυνσης εξαρτάται από το μέγεθος του τμήματος DNA που ενισχύεται με βάση τις προδιαγραφές της θερμοανθεκτικής πολυμεράσης και οι οποίες είναι 1,3kb -1.5kb/1min.

➤ Όταν τελειώσουν οι αντιδράσεις PCR, αναλύονται σε πηκτή αγαρόζης 0,8%-3%, ανάλογα με το μέγεθος των τμημάτων DNA και το διαχωρισμό που θέλουμε να επιτύχουμε.

➤ Οι εκκινητές που χρησιμοποιήθηκαν στις αντιδράσεις PCR των οποίων τα προϊόντα προορίζονται για κλωνοποιήσεις σχεδιάστηκαν με θέση αναγνώρισης ενζύμου περιορισμού στο 5' άκρο, η οποία θα μας διευκολύνει κατά την υποκλωνοποίηση του τμήματος σε άλλους φορείς. Η θέση αναγνώρισης επιλέχτηκε ώστε να είναι μοναδική τόσο μέσα στο τμήμα DNA που επιθυμούμε να κλωνοποιήσουμε όσο και σε άλλους φορείς.

## 2.11 Ηλεκτροφόρηση σε πηκτή αγαρόζης

Ανάλυση νουκλεϊκών οξέων σε πηκτή αγαρόζης

Ηλεκτροφόρηση είναι η τεχνική εκείνη, κατά την οποία μόρια εξαναγκάζονται να κινηθούν διαμέσου ενός πορώδους υλικού, με εφαρμογή ηλεκτρικού πεδίου. Επιθυμητή ποσότητα αγαρόζης τοποθετείται σε κατάλληλο όγκο νερού (w/v) και θερμαίνεται σε φούρνο μικροκυμάτων για μερικά λεπτά (έως την πλήρη διάλυση της αγαρόζης)

- Ακολούθως προστίθεται ποσότητα πυκνού TAE, ώστε η τελική του συγκέντρωση να είναι 1X και αναδεύεται. (50 x TAE ρυθμιστικού διαλύματος : 24.2gr (w/v) Tris-base, 100ml 0.5M EDTA pH8 (Ethylene- Diamine- Tetracetic Acid disodium salt), 57ml CH<sub>3</sub>COOH for 1lt of buffer)

> Τελευταία προστίθεται βρωμιούχο αιθίδιο σε τελική συγκέντρωση 0.5μg/ml

> Το διάλυμα αφήνεται να κρυώσει λίγο και έπειτα τοποθετείται σε καλούπι της συσκευής ηλεκτροφόρησης, στο οποίο έχουν προσαρτηθεί τα αντίστοιχα χτενάκια για τη δημιουργία θέσεων προσθήκης διαλύματος νουκλεϊκών οξέων (πηγαδάκια).

> Μετά τη στερεοποίηση του πηκτώματος αγαρόζης, αφαιρούνται τα χτενάκια και τοποθετείται το δοχείο-καλούπι στη συσκευή ηλεκτροφόρησης, η οποία πληρείται με ρυθμιστικό διάλυμα.

> Στα προς ηλεκτροφόρηση δείγματα προστίθεται κατάλληλη μπλε χρωστική και έπειτα ποσότητα αυτών τοποθετείται στις θέσεις-πηγαδάκια του πηκτώματος αγαρόζης.

- > Εφαρμόζεται συνεχής τάση 50-120V.
- > Το πήκτωμα αγαρόζης εξετάζεται σε υπεριώδη ακτινοβολία.

## 2.12 Προγράμματα νευρωνικών δικτύων πρόβλεψης τοποθέτησης πρωτεϊνών

PSORT: <http://wolfsort.org/>

Predotar: <http://urgi.versailles.inra.fr/predotar/predotar.html>

TargetP: <http://www.cbs.dtu.dk/services/TargetP/>

ChloroP: <http://www.cbs.dtu.dk/services/ChloroP/>

MitoProt: <http://ihg.gsf.de/ihg/mitoprot.html>

BaCello: <http://gpcr.biocomp.unibo.it/bacello/>

SubLoc: [www.bioinfo.tsinghua.edu.cn/SubLoc/](http://www.bioinfo.tsinghua.edu.cn/SubLoc/)

Cello: <http://cello.life.nctu.edu.tw/>

Nucleo: <http://pprowler.itee.uq.edu.au/Nucleo-Release-1.0/>

NUC-Ploc: <http://www.csbio.sjtu.edu.cn/bioinf/Nuc-PLoc/>

Y-Loc: <http://abi.inf.uni-tuebingen.de/Services/YLoc/webloc.cgi>

NetNES: <http://www.cbs.dtu.dk/services/NetNES/>

## 2.13 Προγράμματα Βιοπληροφορικής

<http://www.oxfordjournals.org/nar/database/c>

<http://www.seedgenes.org>

<http://www.arabidopsis.org>

<http://www.softberry.ru>

<http://www.ncbi.nlm.nih.gov/>

<http://au.expasy.org/>

<http://affymetrix.arabidopsis.info>

<https://www.genevestigator.com/gv/index.jsp>

<http://www.bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi>

<http://jsp.weigelworld.org/expviz/expviz.jsp>

<http://atted.jp/>

<http://www.ebi.ac.uk/microarray-as/ae/>

<http://aranet.mpimp-golm.mpg.de/>

<http://www.jgi.doe.gov/>

<http://www.1001genomes.org>

<http://bar.utoronto.ca/eplant/>

<http://mpss.udel.edu>

<http://www.plantgdb.org/AtGDB>

<http://mips.helmholtz-muenchen.de/plant/genomes.jsp>

<http://mfold.bioinfo.rpi.edu/>

<http://mfold.burnet.edu.au/>

## 2.14 Εκκινητές

Όνομασία	Αλληλουχία	Tm
----------	------------	----

<b>GUSATGR</b>	5' AA-CTCGAG-CCGCAAATCTCTCAGAAATTCAT 3'	57.6 °C
<b>GUSTGAR</b>	5' AA-GTCGAC-CCATGATTCCACAGCACTATTTGC 3'	61.0 °C
<b>GUSTGA25R</b>	5' AT-CTCGAG-AGGGAGGAAGAAGCTGGAGAAACA 3'	62.7 °C
<b>90.2</b>	5' TACGACACCTACGCCACATT 3'	57.3 °C
<b>90.4</b>	5' AGAATAGCTCTGGATGTCTC 3'	55.3 °C
<b>04120F</b>	5' GCCAAGAAGTTGTTATCTCTGCC 3'	57 °C
<b>04120R</b>	5' GCTCGACCTGTTGTCGCCAACG 3'	57 °C

### 3. ΑΠΟΤΕΛΕΣΜΑΤΑ

#### 3.1 Το τμήμα PORR εντοπίζεται μόνο σε γονίδια φυτικών οργανισμών

Ερευνώντας για άλλες πρωτεΐνες τύπου LEFKO1 εκτός του *Arabidopsis thaliana* μέσω του προγράμματος BlastP (Basic Local Alignment Search Tool, NCBI), βρέθηκε ότι υπάρχουν ομόλογες πρωτεΐνες μόνο στους φυτικούς οργανισμούς που περιέχουν το τμήμα PORR (Εικόνες 10). Στις βάσεις δεδομένων και ιδιαίτερα στο TAIR ([www.arabidopsis.org](http://www.arabidopsis.org)) και NCBI, το PORR είχε προηγουμένως χαρακτηριστεί ως

DUF860 (Domain of Unknown Function) ή λανθασμένα ως υδρολάση του καρβόξυ-άκρου της ουμπικουϊνίνης. Πρόσφατη αναφορά ανέτρεψε αυτό τον χαρακτηρισμό, αφού παρουσιάστηκε ότι το τμήμα PORR συνδέεται με οργανιδιακό RNA και συμμετέχει στο μάτισμα των οργανιδιακών ιντρονίων κατηγορίας II, ενώ παράλληλα η ομολογία σε αμινοξικό επίπεδο με τις υδρολάσεις του καρβόξυ-άκρου της ουμπικουϊνίνης είναι ελάχιστη (Kroeger *et al.*, 2009). Τα περισσότερα μέλη της οικογένειας PORR/DUF860 προβλέπεται ότι τοποθετούνται στα μιτοχόνδρια ή στους χλωροπλάστες, είτε στον πυρήνα, οπότε η λειτουργία τους μάλλον θα πρέπει να σχετίζεται με βασικές διεργασίες μεταβολισμού του RNA στα υποκυτταρικά αυτά διαμερίσματα που περιέχουν νουκλεϊνικά οξέα. Η αμινοξική στοίχιση έδειξε ότι η ομολογία των πρωτεϊνών LEFKOTHEA στους διάφορους οργανισμούς μπορεί να διαχωριστεί σε δύο επίπεδα. Στο πρώτο επίπεδο παρουσιάζεται μια ελάχιστη έως μηδενική ομολογία στο άμινο N—άκρο καθώς και στο καρβόξυ C-άκρο. Είναι γνωστό ότι στο N-ακρο των πρωτεϊνών που κατευθύνονται στα μιτοχόνδρια και χλωροπλάστες (πεπτιδίο οδηγός) δεν υπάρχει σημαντικός βαθμός ομολογίας. Το δεύτερο επίπεδο, στην περιοχή PORR, είναι διακριτή η ομολογία των πρωτεϊνών. Στην περιοχή αυτή φαίνεται καθαρά μια διαφορετική ομολογία με υψηλή συντήρηση στο N-άκρο καθώς και στο C-άκρο της περιοχής PORR. Στο κεντρικό τμήμα της περιοχής PORR ο βαθμός συντήρησης δεν ακολουθείται (Εικόνα 10). Γίνεται φανερό ότι η λειτουργικότητα της πρωτεΐνης θα εντοπίζεται κυρίως στα δύο άκρα της περιοχής PORR. Υπάρχουν υψηλά συντηρημένες μικροακολουθίες αμινοξέων (σε μαύρο πλαίσιο) καθώς και αλλαγές αμινοξέων με παρόμοιες βιοχημικές-βιοφυσικές ιδιότητες (γκρι πλαίσιο) (Εικόνα 10).

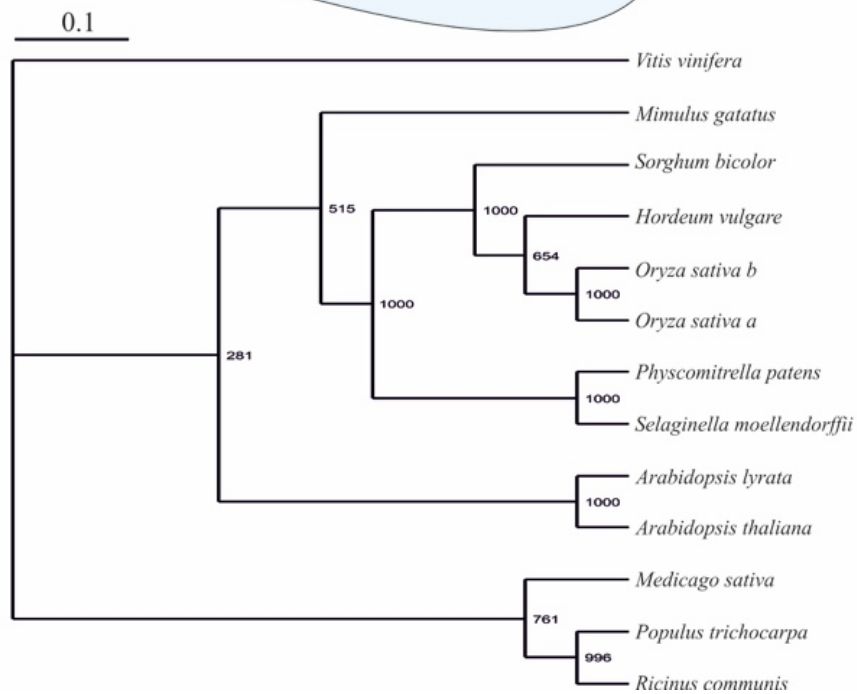
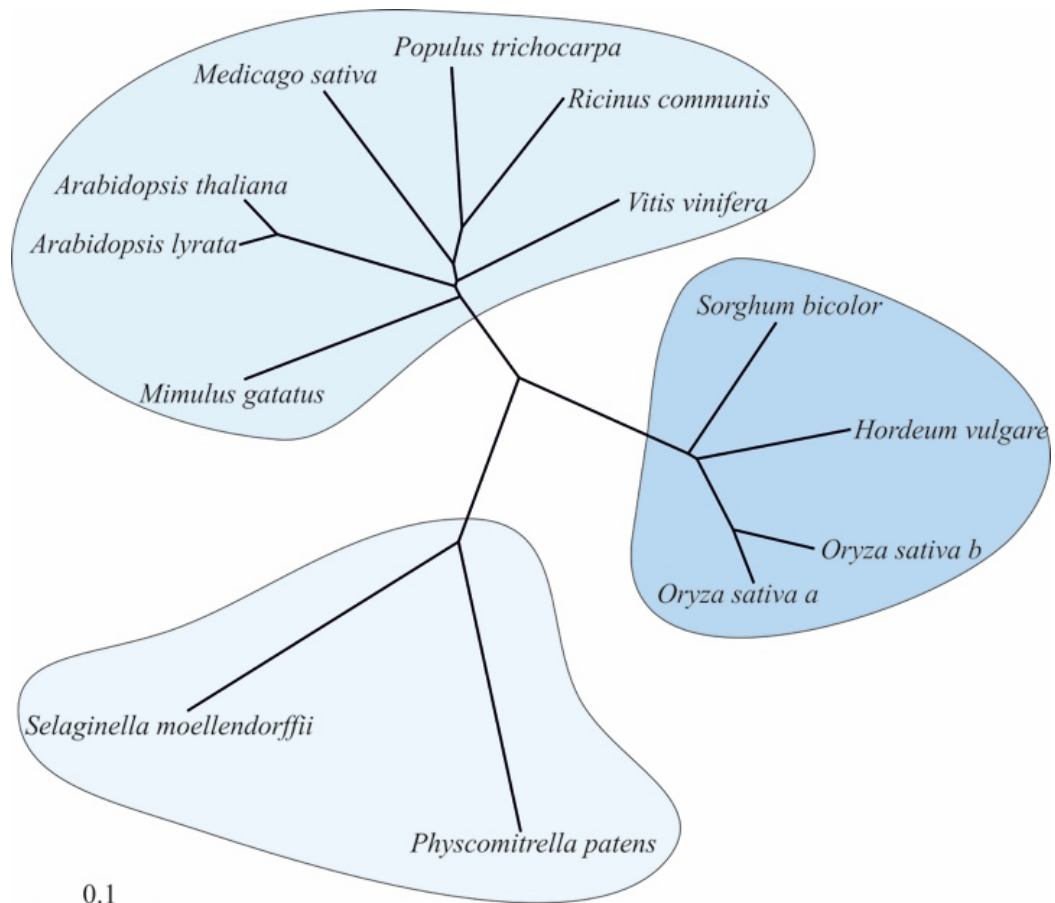
Οι πρωτεΐνες ομόλογες προς τη LEFKOTHEA στα μονοκότυλα (π.χ. ρύζι-*Oryza*), δικότυλα (π.χ. λεύκα-*Populus*), στα βρύα (*Physcomitrella*) και στις πτέρρες (*Selaginella*). Δηλαδή υπάρχουν σε όλα τα φυτά (ανώτερα-κατώτερα) όμως απουσιάζουν από τους μονοκύτταρους ευκαρυωτικούς φωτοσυνθετικούς οργανισμούς όπως *Chlamydomonas* (Εικόνα 10).

Oryza sativa a : -----MEMAIAAVVLGGG--GEADVRRGGVGAAGAAATMPDGRGRGWPREKEAVVLPPLTTSFELSVSHLHGRR : 70  
Oryza sativa b : -----ATTASS--SLSFLS--SQPH--RRVRLP-----PPPTAAPTPTAVRCTFA : 47  
Hordeum vulgare : -----VILSGKNLFIHSAMAP--SLSVLP--APFLTFLHP--RRLRHQT-----AAPARTNPQTLTAVRCAA : 60  
Sorghum bicolor : -----MUVVAP--ACSFLL--APSLR--ARFRPH-----FMNSTAKPLSHTLIRCAA : 46  
Arabidopsis lyrata : -----ALITQLVPS--SPDK-LF-----FFSFFLPSIPLIERTSNL : 34  
Arabidopsis thaliana : -----MMALITQLVPS--SPEK-LF-----CFSFFLPSIPLIERSNV : 36  
Vitis vinifera : -----VALSLFVS--CKDKHTI-----FLNSNPLPQPWPIYNTIN : 35  
Populus trichocarpa : -----VALYLSSE--CTGCFL-----KYTFPSSTPSELHAN-- : 31  
Ricinus communis : -----MQLLSLSLP--GNCCFLS-----KSSFLPSAPRIYL-- : 30  
Medicago sativa : -----ALCFHELY--TNNGVFT-----ELNSTFRGTPESSFVW-- : 34  
Mimulus guttatus : -----MSIARFLISTH--HIDAPFT-----LLERFPLPKHIFSGS : 38  
Phycomitrella patens : MAMALRASMAVGGFRSSDRSLSNFVSAP-----GAAESVALGRQLFLGCTQ : 48  
Selaginella moellendorffii : -----MAVALFLESTLFLNS-----KKQLRASSSSTFFSGAGG : 35  
  
Oryza sativa a : KGRHRERRRASAGDGGG--GLEGETRCWRQTEPRFRAPNRLRFARKLVLLLSKDRFELLRVLRCHRHTGCPPHRR : 148  
Oryza sativa b : FAAAAP--SAASAASARS--IPPEKLVRCALDRQAAR--ANRLRFARKLVLLLSKDRFELLRVLRCHRHTGCPPHRR : 121  
Hordeum vulgare : K-----AASPAPAP--LPPKLVRCALDRQAAR--ASRLRFARKLVLLLSKDRFELLRVLRCHRHTGCPPHRR : 126  
Sorghum bicolor : TTAAGREGPPFPKLPFPFPKLVRCALDRQAAR--ASRLRFARKLVLLLSKDRFELLRVLRCHRHTGCPPHRR : 122  
Arabidopsis lyrata : SRTSSKKLGLSLSN--ISCSARIVRSPSLDRHVVM--QNRVRFCKMNTLLSHPRFELIKLRCSSHTLTP--N : 107  
Arabidopsis thaliana : TKTSKKLGLSLSN--ISCSARIVRSPSLDRHVVM--QNRVRFCKMNTLLSHPRFELIKLRCSSHTLTP--N : 109  
Vitis vinifera : RNSDKNRKPKNLSIS--VCSSTLKIYRSPSLDRHVVM--QNKIRFVKIKLTLSSHPRFELIKLRCSSHTLTP--N : 108  
Populus trichocarpa : --PSLTCRKRPRYKRLS--ISCSARIVRSPSLDRHVVM--QNRVRFCKMNTLLSHPRFELIKLRCSSHTLTP--N : 103  
Ricinus communis : --LNSKSRKRVHSLS--ISCVSTKIVRDAKLDKRVHM--QNKIRFVKIKLTLSSHPRFELIKLRCSSHTLTP--D : 100  
Medicago sativa : --RKLTKCNLKHVSVS--ISCVSTKIVRDRDLDRHVVM--QNRVRFCKMNTLLSHPRFELIKLRCSSHTLTP--N : 106  
Mimulus guttatus : --TNTRKLGAFAMITCSSSSPKIVRDRDLDRHVVM--QNKIRFVKIKLTLSSHPRFELIKLRCSSHTLTP--N : 110  
Phycomitrella patens : SLRCFVFPNCRVSPNGKRLLTAVRQKDYLDKDAVIGGKRLRVCKIRDLVVKQFGVQVSRDLRIGRFRCLGID--G : 125  
Selaginella moellendorffii : NRRNFIAIPSTRHRVQRLPVVTATVRETQDYLDKVVVERKRVVNVKRNLLVVKQFGVQVSRDLRIGRFRCLGID--A : 112  
  
Oryza sativa a : RFLVPEVRYRPALEFRLEFQAPTS-HLISPN--LSTLPAALTPAEEAALADLAALRGSSAPRLLAKNHRLLLLAP-GRS : 224  
Oryza sativa b : RFLVPEVRYRPALEFRLEFQAPTS-HLISPN--LSTLPAALTPAEEAALADLAALRGSSAPRLLAKNHRLLLLAP-GRS : 197  
Hordeum vulgare : RFLVPEVRYRPALEFRLEFQAPTS-HLISPN--LSTLPAALTPAEEAALADLAALRGSSAPRLLAKNHRLLLLAP-GRS : 201  
Sorghum bicolor : RFLVPEVRYRPALEFRLEFQAPTS-HLISPN--LSTLPAALTPAEEAALADLAALRGSSAPRLLAKNHRLLLLAP-GRS : 201  
Arabidopsis lyrata : GFLVPEVRYRPALEFRLEFQAPTS-HLISPN--LSTLPAALTPAEEAALADLAALRGSSAPRLLAKNHRLLLLAP-GRS : 184  
Arabidopsis thaliana : FLAFLVPEVRYRPALEFRLEFQAPTS-HLISPN--LSTLPAALTPAEEAALADLAALRGSSAPRLLAKNHRLLLLAP-GRS : 186  
Vitis vinifera : RFLVPEVRYRPALEFRLEFQAPTS-HLISPN--LSTLPAALTPAEEAALADLAALRGSSAPRLLAKNHRLLLLAP-GRS : 185  
Populus trichocarpa : PCSFLVPEVRYRPALEFRLEFQAPTS-HLISPN--LSTLPAALTPAEEAALADLAALRGSSAPRLLAKNHRLLLLAP-GRS : 180  
Ricinus communis : RFLVPEVRYRPALEFRLEFQAPTS-HLISPN--LSTLPAALTPAEEAALADLAALRGSSAPRLLAKNHRLLLLAP-GRS : 177  
Medicago sativa : RFLVPEVRYRPALEFRLEFQAPTS-HLISPN--LSTLPAALTPAEEAALADLAALRGSSAPRLLAKNHRLLLLAP-GRS : 183  
Mimulus guttatus : RFLVPEVRYRPALEFRLEFQAPTS-HLISPN--LSTLPAALTPAEEAALADLAALRGSSAPRLLAKNHRLLLLAP-GRS : 188  
Phycomitrella patens : KRVVADLRKFPVFBVVEEGTAKYFRF-----FVAKKQYREARLRK--EVEDVAVTRIKLNNNSV-DKS : 191  
Selaginella moellendorffii : RRRVADLRKFPVFBVVEEGTAKYFRF-----FVAKKQYREARLRK--EVEDVAVTRIKLNNNSV-DKS : 136  
  
Oryza sativa a : LLVSKLHAPDLGLANDEF-ATGCPHEDDLTFYNTSHCHLGLVDPFPPPPPLFP-----F : 282  
Oryza sativa b : LLVSKLHAPDLGLANDEF-ATGCPHEDDLTFYNTSHCHLGLVDPFPPPPPLFP-----F : 255  
Hordeum vulgare : LLVSKLHAPDLGLANDEF-ATGCPHEDDLTFYNTSHCHLGLVDPFPPPPPLFP-----L : 260  
Sorghum bicolor : LLVSKLHAPDLGLANDEF-ATGCPHEDDLTFYNTSHCHLGLVDPFPPPPPLFP-----L : 256  
Arabidopsis lyrata : LLLSKLHAPDLGLANDEF-SRLNDYEDKRTVDTSYRAELVSWDPELANQMS-----P : 242  
Arabidopsis thaliana : LLLSKLHAPDLGLANDEF-SRLNDYEDKRTVDTSYRAELVSWDPELANQMS-----P : 244  
Vitis vinifera : LLLSKLHAPDLGLANDEF-SRLNDYEDKRTVDTSYRAELVSWDPELANQMS-----P : 243  
Populus trichocarpa : LLVSKLHAPDLGLANDEF-SRLNDYEDKRTVDTSYRAELVSWDPELANQMS-----A : 237  
Ricinus communis : LLLSKLHAPDLGLANDEF-SRLNDYEDKRTVDTSYRAELVSWDPELANQMS-----R : 234  
Medicago sativa : LLVSKLHAPDLGLANDEF-SRLNDYEDKRTVDTSYRAELVSWDPELANQMS-----R : 241  
Mimulus guttatus : LLLSKLHAPDLGLANDEF-SRLNDYEDKRTVDTSYRAELVSWDPELANQMS-----R : 245  
Phycomitrella patens : LLVSKLHAPDLGLANDEF-SRLNDYEDKRTVDTSYRAELVSWDPELANQMS-----R : 271  
Selaginella moellendorffii : LLVSKLHAPDLGLANDEF-SRLNDYEDKRTVDTSYRAELVSWDPELANQMS-----L : 159  
  
Oryza sativa a : RPAAPSDRILIRRRRPH--LIRKGLNLRRAIDVLLRHSRHSPEEPLDEGASL--MERRACAVVREVLAMTV : 357  
Oryza sativa b : RPAAPSDRILIRRRRPH--LIRKGLNLRRAIDVLLRHSRHSPEEPLDEGASL--MERRACAVVREVLAMTV : 330  
Hordeum vulgare : RPAAPSDRILIRRRRPH--LIRKGLNLRRAIDVLLRHSRHSPEEPLDEGASL--MERRACAVVREVLAMTV : 335  
Sorghum bicolor : RPAAPSDRILIRRRRPH--LIRKGLNLRRAIDVLLRHSRHSPEEPLDEGASL--MERRACAVVREVLAMTV : 331  
Arabidopsis lyrata : EV-VDRGLIVRIRKPKR--LIRKGLNLRRAIDVLLRHSRHSPEEPLDEGASL--MERRACAVVREVLAMTV : 318  
Arabidopsis thaliana : E-VDRGLIVRIRKPKR--LIRKGLNLRRAIDVLLRHSRHSPEEPLDEGASL--MERRACAVVREVLAMTV : 319  
Vitis vinifera : EV-VDRGLIVRIRKPKR--LIRKGLNLRRAIDVLLRHSRHSPEEPLDEGASL--MERRACAVVREVLAMTV : 319  
Populus trichocarpa : EK-VSRELIVRIRKPKR--LIRKGLNLRRAIDVLLRHSRHSPEEPLDEGASL--MERRACAVVREVLAMTV : 313  
Ricinus communis : AD-ASLDLIVRIRKPKR--LIRKGLNLRRAIDVLLRHSRHSPEEPLDEGASL--MERRACAVVREVLAMTV : 310  
Medicago sativa : E-HLEDLIVRIRKPKR--LIRKGLNLRRAIDVLLRHSRHSPEEPLDEGASL--MERRACAVVREVLAMTV : 317  
Mimulus guttatus : DVDKSHGLIVRIRKPKR--LIRKGLNLRRAIDVLLRHSRHSPEEPLDEGASL--MERRACAVVREVLAMTV : 323  
Phycomitrella patens : EEIEEESLARTRERQSKKTLNPKCCVFRADLRKQELRSPSPADTKD-LRASVPAEKAALVIVPELHLSLS : 350  
Selaginella moellendorffii : QEDYEEITVSGRVKPKR--LIRKGLNLRRAIDVLLRHSRHSPEEPLDEGASL--MERRACAVVREVLAMTV : 235  
  
Oryza sativa a : EKKTLVDHLTHFRDFCFNRLRMLVRRHPEFYYSVIGLRSVLELAEFLDGRVLEDMVAARDSEELVRECKRMR : 437  
Oryza sativa b : EKKTLVDHLTHFRDFCFNRLRMLVRRHPEFYYSVIGLRSVLELAEFLDGRVLEDMVAARDSEELVRECKRMR : 410  
Hordeum vulgare : EKKTLVDHLTHFRDFCFNRLRMLVRRHPEFYYSVIGLRSVLELAEFLDGRVLEDMVAARDSEELVRECKRMR : 415  
Sorghum bicolor : EKKTLVDHLTHFRDFCFNRLRMLVRRHPEFYYSVIGLRSVLELAEFLDGRVLEDMVAARDSEELVRECKRMR : 411  
Arabidopsis lyrata : EKKTLVDHLTHFRDFCFNRLRMLVRRHPEFYYSVIGLRSVLELAEFLDGRVLEDMVAARDSEELVRECKRMR : 398  
Arabidopsis thaliana : EKKTLVDHLTHFRDFCFNRLRMLVRRHPEFYYSVIGLRSVLELAEFLDGRVLEDMVAARDSEELVRECKRMR : 399  
Vitis vinifera : EKKTLVDHLTHFRDFCFNRLRMLVRRHPEFYYSVIGLRSVLELAEFLDGRVLEDMVAARDSEELVRECKRMR : 399  
Populus trichocarpa : EKKTLVDHLTHFRDFCFNRLRMLVRRHPEFYYSVIGLRSVLELAEFLDGRVLEDMVAARDSEELVRECKRMR : 393  
Ricinus communis : EKKTLVDHLTHFRDFCFNRLRMLVRRHPEFYYSVIGLRSVLELAEFLDGRVLEDMVAARDSEELVRECKRMR : 390  
Medicago sativa : EKKTLVDHLTHFRDFCFNRLRMLVRRHPEFYYSVIGLRSVLELAEFLDGRVLEDMVAARDSEELVRECKRMR : 397  
Mimulus guttatus : EKKTLVDHLTHFRDFCFNRLRMLVRRHPEFYYSVIGLRSVLELAEFLDGRVLEDMVAARDSEELVRECKRMR : 403  
Phycomitrella patens : EKKTLVDHLTHFRDFCFNRLRMLVRRHPEFYYSVIGLRSVLELAEFLDGRVLEDMVAARDSEELVRECKRMR : 429  
Selaginella moellendorffii : EKKTLVDHLTHFRDFCFNRLRMLVRRHPEFYYSVIGLRSVLELAEFLDGRVLEDMVAARDSEELVRECKRMR : 315  
  
Oryza sativa a : RARRGG--LLALAD-DSDEDEDEEEDGEEQDSVQVDSGSW--LLEDGGIG-EDWEEVG----DLG--EGSDDDDADAELDAL : 506  
Oryza sativa b : RARRGG--LLALAD-DSDEDEDEEEDGEEQDSVQVDSGSW--LLEDGGIG-EDWEEVG----DLG--EGSDDDDADAELDAL : 479  
Hordeum vulgare : RARRGG--VFPVDDGSDSDEDEEDNVAEGRSSEVDGEGFLFEECSVAGEDWNEVG----DGGEGEDGDDDEPESDAM : 488  
Sorghum bicolor : RARRGG--LLTFDG-DSDEDEDEEAEAEEGSLEVDDDFGLFEDCGIIG-EDWEEVG----DGGR-SEAYEYDAELGET : 481  
Arabidopsis lyrata : RRRRRK----GDKRVEEYSNGDNRDETIDDESDDL EYDGFENLFDSEDSG--VEYHFEEDD-BDEARVNGGS : 467  
Arabidopsis thaliana : RRRRRK----GDKRVEEYSNGDNRDETIDDESDDL EYDGFENLFDSEDSG--VEYHFEEDD-BDEARVNGGS : 469  
Vitis vinifera : RRRRRK----SINNEVGG--YDDTIDDD--IDGTRKGFDSLLDFEDIDGDYEDFGSDNEEE--DDEYDLSRVG : 463  
Populus trichocarpa : RRRRRK----GKYG-NDTGDSNDVDEND--ELLDFA--DYDGFENVFESEDSG--FDYFFIDKDDHASECFADYGEN : 460  
Ricinus communis : RRRRRK----IVMNTNDRSDD--DASD--EFDLID--DYDGFENVFESEDSG--FVYDNDYDLDHASCETASHGHE : 456  
Medicago sativa : RRRRRK----HARIDKEFGRLSVGNQNFDDSDIDYDNDIENDFK--GYDGFEDIEFEDL--FEAEDYDLDHASECFADYGEN : 468  
Mimulus guttatus : RRRRRK----RIKRYVNGVQNFHQLVDEYCG--EGGLDLDLQVDDAS--SKDQVVDDESEDE : 460  
Phycomitrella patens : RRRRRK----DDDEVDGDDNDDDDDDDDDD--L : 457  
Selaginella moellendorffii : RRRRRK----EGSDEDDGDSWSDDEMSPETREFLEERRRRRVMIALRPKVPCCGPTFPAMKAAALLTAAATVTL : 391  
  
Oryza sativa a : EEPVVKKAAEGLVDTGSEL--DA----- : 529  
Oryza sativa b : EEPVVKKAAEGLVDTGSEL--DA----- : 502  
Hordeum vulgare : EEPVVKKAAEGLVDTGSEL--DA----- : 511  
Sorghum bicolor : EEPVVKKAAEGLVDTGSEL--DA----- : 504  
Arabidopsis lyrata : IEPVSRKLFSSGSSSDEAKGAVESW----- : 492  
Arabidopsis thaliana : IEPVSRKLFSSGSSSDEAKGAVESW----- : 494  
Vitis vinifera : VNEWT--ADAHVVNDEEKGPSEFW----- : 485  
Populus trichocarpa : GELTADTSFDGTSDDGGSS--EFW----- : 483  
Ricinus communis : GEPVTVD--GSNEEGRSS--EFW----- : 475  
Medicago sativa : GEPVTVD--TAGPFFQNGSDTEEK--EFW----- : 493  
Mimulus guttatus : EEPVVKKAAEGLVDTGSEL--DA----- : 479  
Phycomitrella patens : VGVAGMASVQPKKQVPTAPK--EFW----- : 481  
Selaginella moellendorffii : IGLGIARDGGERKCCGSSG--ECPEDHPGQEEHANELVAVSPLARRPTEDGRTRSVHAYFRITIGSLFLKRRNVAK : 470

**Εικόνα 10. Πολλαπλή ευθυγράμμιση αμινοξικών ακολουθιών των ορθόλογων πρωτεϊνών LEFKO1 από φυτικούς οργανισμούς.** Ο κόκκινος αστερίσκος δείχνει το συντηρημένο αμινοξύ Γλυκίνη-373, που στη μετάλλαξη *lefko1-1* έχει αντικατασταθεί από το Ασπαρτικό οξύ ( $G_{373} \rightarrow D$ ). Σε μαύρο πλαίσιο τα αμινοξέα που παρμένουν, σε γκρι πλαίσιο αλλαγές αμινοξέων με παραπλήσιες βιοχημικές-βιοφυσικές ιδιότητες. Οι οργανισμοί που χρησιμοποιήθηκαν είναι οι εξής (κατά ίδια σειρά προβολής στην Εικόνα 10): *Oryza sativa a*, *oryza sativa b*, *Hordeum vulgare*, *Sorghum bicolor*, *Arabidopsis lyrata*, *Arabidopsis thaliana*, *Vitis vinifera*, *Populus trichocarpa*, *Ricinus communis*, *medicago sativa*, *Mimulus gatatus*, *Physcomitrella patens*, *Selaginella moellendorffii*. Οι αριθμοί στα δεξιά αναφέρονται τις θέσεις των αμινοξέων στις διάφες πρωτεΐνες. Παύλες έχουν ενσωματωθεί από το πρόγραμμα για καλύτερη στοίχιση των αλληλουχιών. Χρησιμοποιείται μονογράμματα για τα αμινοξέα.

Αξιοσημείωτο είναι ότι οι ορθόλογες πρωτεΐνες LEFKOTHEA από μονοκοτυλήδονα και δικοτυλήδονα φυτά σε αντίθεση με τα αρχέγονα βρυόφυτα και πτέρες *Selaginella moellendorffii* και *Physcomitrella patens* διατηρούν συντηρημένη τη γλυκίνη-373, γεγονός που υποδηλώνει το σημαντικό λειτουργικό ρόλο του συγκεκριμένου αμινοξέως (Εικόνα 10)

Στο φυλογενετικό δένδρο που δημιουργήθηκε από την πολλαπλή ευθυγράμμιση των ορθόλογων αμινοξικών ακολουθιών διακρίνονται τρεις κύριοι κλάδοι (Εικόνα 11). Ο πρώτος (I) κλάδος περιέχει πρωτεΐνες ορθόλογες της LEFKOTHEA από δικοτυλήδονα φυτά, όπως το *Arabidopsis* (*Arabidopsis thaliana* και *Arabidopsis lyrata*), το *Mimulus gattatus*, τη Μηδική (*Medicago sativa*), τη λεύκα (*Populus trichocarpa*), τη ρετινολαδιά (*Ricinus communis*) και το αμπέλι (*Vitis vinifera*). Ο δεύτερος (II) κλάδος αποτελείται από ορθόλογες πρωτεΐνες που προέρχονται από μονοκοτυλήδονα φυτά, όπως το κριθάρι (*Hordeum vulgare*), το γλυκό σόργο (*Sorghum bicolor*) και το ρύζι (*Oryza sativa*). Ο τρίτος (III) κλάδος περιλαμβάνει ορθόλογες πρωτεΐνες από τα αρχέγονα βρυόφυτα και πτέρες *Selaginella moellendorffii* και *Physcomitrella patens*. Συνεπώς στο βασίλειο των φυτών οι ορθόλογες πρωτεΐνες ομαδοποιούνται ανάλογα με την εξελικτική τους διαδικασία. Από το μήκος της διχάλας φαίνεται και ο βαθμός ομολογίας ή εξέλιξης, Για παράδειγμα το μήκος που χωρίζεται στη διχάλα *Arabidopsis thaliana*, *Arabidopsis lyrata* είναι πολύ μικρό ενώ αντίθετα το μήκος της διχάλας του *Medicago* και *Populus* είναι σχετικά μεγαλύτερο (Εικόνα 11). Η παρουσία γονιδίων LEFKOTHEA στα φυτά υποδεικνύει ότι η εμφάνιση του γονιδίου έχει γίνει με την εξέλιξη των φωτοσυνθετικών οργανισμών με αγωγό ιστό ή πολυκύταρων φωτοσυνθετικών οργανισμών.



Εικόνα 11. Φυλογενετικό δένδρο Neighbor-Joining χωρίς ρίζα (Unrooted) υπολογισμένο με βάση τη πολλαπλή ευθυγράμμιση μεταξύ των ορθόλογων πρωτεϊνών LEFKOTHEA από



**διάφορους φυτικούς οργανισμούς.** Unrooted Neighbor-Joining εξελικτικό δένδρο όπως προέρχεται από τη πολλαπλή ευθυγράμμιση των αμινοξικών ακολουθιών πρωτεϊνών LEFKO1 από διάφορους φυτικούς οργανισμούς. Κάθε διαφοροποίηση αντιπροσωπεύει διαφορετικούς εξελικτικούς κλάδους. Το μήκος κάθε γραμμής ή το μήκος της διχάλας του κλάδου αντιπροσωπεύει την απόσταση (εξελικτική) της πρωτεΐνης. Το 0.1 αντιπροσωπεύει την φυλογενετική απόσταση Το κάτω δενδρόγραμμα αναφέρεται σε φυλογενετικές απόστάσεις. Οι αριθμοί στις διχάλες αναφέρονται στις τιμές της επαναληψιμότητας σε bootstrap από 1000 επαναλήψεις.

Το φυλογενετικό δένδρο (Εικόνα 11 κάτω) δείχνει επίσης μια σχετική ανακατονομή των γονιδίων όσο αφορά την φυλογενετικής τους σχέση. Φαίνεται ότι το αμπέλι έχει μικρή σχέση με όλα τα υπόλοιπα. Επίσης φαίνεται να υπάρχει μια σχετικά καλή φυλογένεια των μονοκότυλων φυτών όσο αφορά την πρωτεΐνη LEFKOTHEA με τα κατώτερα φυτά βρύα και πτέρρες.

### **3.2 Τα γονίδια του φυτού *Arabidopsis thaliana* που φέρουν το τμήμα PORR αποτελούν μια υπεροικογένεια**

Προκειμένου να βρεθούν κάποια στοιχεία για τη λειτουργία της πρωτεΐνης LEFKOTHEA, έγινε σύγκριση της αμινοξικής αλληλουχίας της (blastp) με όλους τους οργανισμούς που εμπεριέχονται στη βάση δεδομένων του NCBI. Η σύγκριση αυτή έδειξε πως η πρωτεΐνη παρουσιάζει ομολογία με ποικίλες άλλες πρωτεΐνες φυτικών οργανισμών (*Physcomitrella patens*, *Oryza sativa*, *Sorghum bicolor*, *Populus sp.*, *Arabidopsis* κ.ά). Εκτός αυτού, το ενδιαμέσο κύριο τμήμα της πρωτεΐνης του γονιδίου που φέρει τη μετάλλαξη *lefkothea*, παρουσιάζει ομολογία με τον τομέα PORR που απαντάται μόνο σε φυτικούς οργανισμούς.

Με έναν δεύτερο έλεγχο blastp, ολόκληρης της αμινοξικής αλληλουχίας της πρωτεΐνης, αλλά και του τομέα DUF860/PORR στον οργανισμό *Arabidopsis thaliana* μόνο και με τη χρήση των προγραμμάτων Clustalx, GeneDoc, Treeview βγήκε το παρακάτω δενδρόγραμμα. Η αναζήτηση ομόλογων γονιδίων του *Lefko1* στο *Arabidopsis thaliana* είχε ως αποτέλεσμα να εντοπιστούν επιπλέον 14 γονίδια που έχουν το τμήμα PORR (Plant Organelle RNA Recognition motif). Τα 15 μέλη της υπεροικογένειας παρόλο που έχουν σημαντική ομολογία μεταξύ τους είναι φανερό ότι ταυτόχρονα είναι αρκετά διαφορετικά όπως φαίνεται άλλωστε από την πολλαπλή ευθυγράμμιση και το φυλογενετικό δένδρο. Η σκιαγραφημένη περιοχή αντιστοιχεί στο υψηλά συντηρημένο τμήμα PORR, τα υπόλοιπα

όμως τμήματα της πρωτεΐνης στο άμινο και καρβόξυ-τελικό άκρο είναι ελάχιστα συντηρημένα. Η ευθυγράμμιση επιπλέον αποκαλύπτει ότι (i) η Γλυκίνη στη θέση 373 που στο μετάλλαγμα *lefko1-1* έχει τροποποιηθεί σε Ασπαρτικό οξύ ( $G_{373} \rightarrow D$ ), διατηρείται σε αρκετά μέλη της υπεροικογένειας και (ii) ενώ το σινιάλο NES στη θέση 373-392 της πρωτεΐνης *Lefko1* είναι συντηρημένο μεταξύ των ομόλογων γονιδίων, αντίθετα το μονομερές NLS της πρωτεΐνης *Lefko1* στη θέση 396 έως 405 δεν φαίνεται να είναι συντηρημένο μεταξύ των ομόλογων γονιδίων.

Στο φυλογενετικό δένδρο παρατηρείται ότι το γονίδιο *Lefko1* (*At5g62990*) εντάσσεται σε έναν κλάδο 4 γονιδίων μαζί με το γονίδιο *At4g01037* ορθόλογο του γονιδίου *WTF1* (*What's This Factor*) από το καλαμπόκι (*Zea mays*) που έχει το χαρακτηρισμένο τμήμα PORR. Το μετάλλαγμα *wtf1* στο καλαμπόκι έχει φαινοτυπικά χαρακτηριστικά παρόμοια της μετάλλαξης *lefko1-1* και συγκεκριμένα τα φύλλα είναι υποκίτρινα έως πλήρως χλωρωτικά. Το τμήμα PORR της πρωτεΐνης *WTF1* έχει αναφερθεί ότι αποτελεί συστατικό των ριβονουκλεοπρωτεϊνικών συμπλόκων που συμμετέχουν στο μάτισμα των ιντρονίων κατηγορίας II (group II intron splicing) που έχουν ορισμένα γονίδια του χλωροπλάστη (Kroeger *et al.*, 2009). Το γονίδιο *RPD1* (*ROOT PRIMORDIUM DEFECTIVE 1*; *At4g33495*) αποτελεί επίσης μέλος της υπεροικογένειας με το τμήμα PORR. Οι αλληλόμορφοι *rpd1-1* και *rpd1-2* που έχουν προέλθει από αμινοξικές υποκαταστάσεις εξαιτίας μεταλλαξογένεσης EMS παρουσιάζουν ευαισθησία στη θερμοκρασία και παρουσιάζουν ανωμαλίες στον σχηματισμό επιγενετικών ριζών (adventitious roots), ενώ συνολικά φαίνεται να επηρεάζεται ο πολλαπλασιασμός των κυττάρων (Konishi και Sugiyama, 2006). Στον αλληλόμορφο *rpd1-3* που έχει προκύψει από ένθεση T-DNA έχει διακοπεί η έκφραση του γονιδίου επομένως αναστέλλεται η ανάπτυξη στο σφαιρικό στάδιο του εμβρύου (globular stage) (Konishi και Sugiyama, 2006).

### 3.3. Δομή και χρωμοσωμική τοποθέτηση των γονιδίων PORR στο *Arabidopsis thaliana*

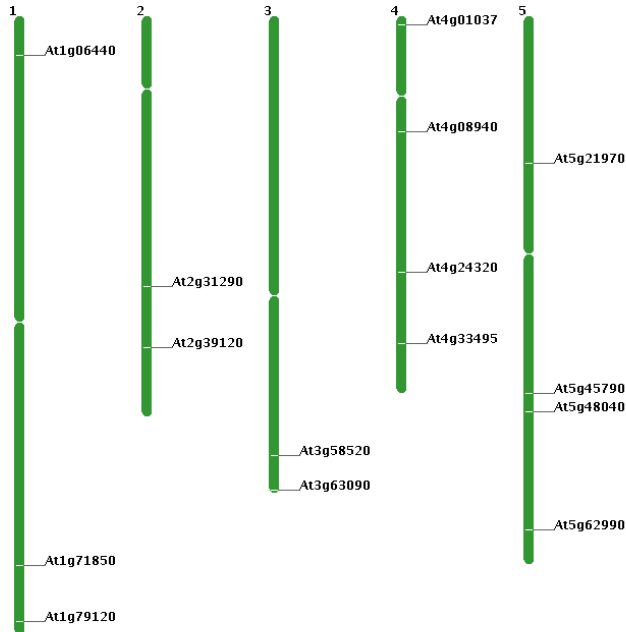
<b>Protein</b>	Amino acid residues	MW (Da)	pI	Chromosome
At5g62990	494	56.79	7.79	5
At4g01037	528	61.23	6.63	4
At1g06440	390	44.7	9.52	1
At5g45790	439	51.9	9.75	5
At4g33495	409	48.28	10.03	4
At5g48040	422	49.72	8.12	5
At3g58520	418	48.82	9.17	3
At2g39120	387	45.2	9.32	2
At4g24320	395	45.69	10.4	4
At1g71850	470	53.88	9.66	1
At1g79120	413	47.32	10.26	1
At2g31290	415	48.67	9.57	2
At3g63090	404	47.81	9.81	3
At4g08940	395	46.58	10.03	4
At5g21970	449	52.98	7.14	5

**Εικόνα 12. Οικογένεια γονιδίων PORR στο *Arabidopsis*.** Στοιχεία αυτών κατά σειρά: αριθμός αμινοξέων, μοριακό βάρος, pI, αριθμός χρωμοσώματος που εδράζουν τα γονίδια.

Στην Εικόνα 12 βλέπουμε την οικογένεια των 15 γονιδίων PORR στο *Arabidopsis* σε σχέση με το μέγεθος τους. Μπορούμε να βγάλουμε το συμπέρασμα ότι τα μεγέθη αυτών είναι περίπου παρόμοια. Η μικρότερη πρωτεΐνη PORR έχει 387 αμινοξέα δηλαδή 45.2 kDa και η μεγαλύτερη 528 αμινοξέα δηλαδή 61.23 kDa. Η LEFKOTHEA παρουσιάζει ένα μοριακό βάρος που κατατάσσεται στα υψηλά. Φαίνεται ότι ο διπλασιασμός που προέκυψε κατά τη διάρκεια της εξέλιξης στην ουσία πρόσθεσε ή αφαίρεσε διάφορα αμινοξέα τροποποιώντας, ενδεχομένως, και τη λειτουργία του κάθε μέλους της οικογένειας. Δηλαδή να πρόσθεσε επιπλέον λειτουργίες σε ορισμένα μέλη της οικογένειας ή να αφαίρεσε.

Είναι αναμενόμενο για πρωτεΐνες οι οποίες έχουν δομική/λειτουργική συνάφεια με τα νουκλεϊνικά να είναι θετικά φορτισμένες. Όντως η βιοπληροφορική ανάλυση έδειξε ότι τα περισσότερα μέλη της οικογένειας είναι θετικά φορτισμένες πρωτεΐνες. Μόνο η πρωτεΐνη At4g01037 οποία έχει και το μεγαλύτερο μέγεθος είναι σχετικά όξινη με pI 6.63, ενώ η πρωτεΐνη At1g79120 είναι η πιο βασική με pI 10.26 (Εικόνα 12). Η LEFKOTHEA παρουσιάζει ένα σχετικά βασικό φορτίο με pI 7.79. Έτσι η *in silico* ανάλυση του pI προσθέτει ακόμη περισσότερο στη λειτουργική δράση των μελών της οικογένειας PORR. Έντεκα από τα δεκαπέντε μέλη της οικογένειας έχουν pI πάνω από 9.0 δηλαδή έχουν ένα

ισχυρά βασικό φορτίο το οποίο απαιτείται για ένωση τους με τα αρνητικά φορτισμένα μόρια του RNA για να επιτελέσουν τις διάφορες λειτουργίες τους.



**Εικόνα 13. Θέση των 15 γονιδίων PORR στα χρωμοσώματα στο Arabidopsis.** Φαίνεται μια περίπου ισοκατανομή των 15 γονιδίων στα πέντε χρωμοσώματα του Arabidopsis.

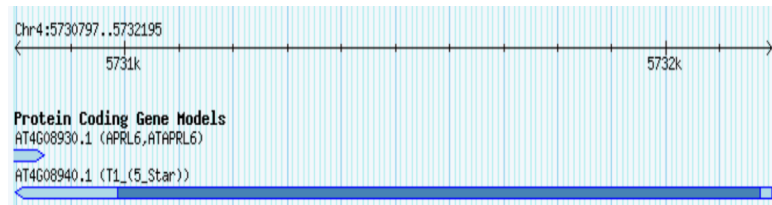
Ένα βασικό ερώτημα που τίθεται είναι για την κατανομή των γονιδίων της οικογένειας PORR στα διάφορα χρωμοσώματα του Arabidopsis. Φαίνεται ότι κάθε χρωμόσωμα έχει από 2 έως 4 γονίδια τα οποία όμως είναι διάσπαρτα σε ολόκληρο το μέγεθος του χρωμοσώματος (Εικόνα 13). Αυτό φαίνεται ιδιαίτερ για το χρωμόσωμα 4. Η πιο κοντική απόσταση δύο μελών είναι εντοπισμένη στο χρωμόσωμα 5, για τα γονίδια At5g45790 και At5g48040 (Εικόνα 13) τα οποία έχουν ένα παραπλήσιο μοριακό βάρος (περίπου 50 kDa) όπως φαίνεται στην Εικόνα 12. Η κατανομή των γονιδίων στα πέντε χρωμοσώματα του Arabidopsis μαρτυρά ότι ο διπλασιασμός για τα δύο αυτά γονίδια έχει γίνει πρωταρχικά στην εξέλιξη και αργότερα επήλθε μια σειρά από διπλασιασμούς και μεταφορά στα υπόλοιπα χρωμοσώματα. Δεν μπορούμε όμως να παραβλεψουμε ότι το Arabidopsis περιέχει και μεγάλες διπλασιασμένες περιοχές των χρωμοσωμάτων

### 3.4 Δομή των 15 γονιδίων PORR του Arabidopsis με παρουσίαση ιντρονίων εξονίων.

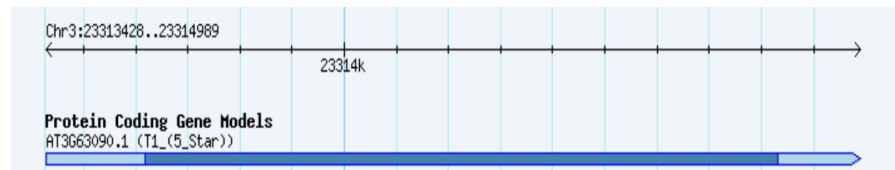
At5g21970



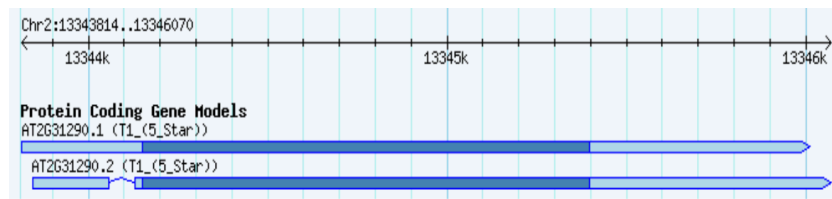
At4g08930



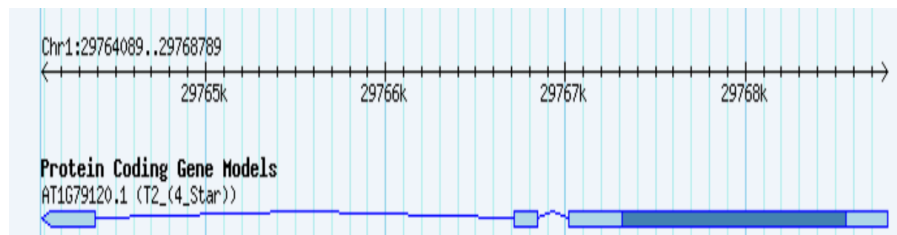
At3g63090



At2g31290



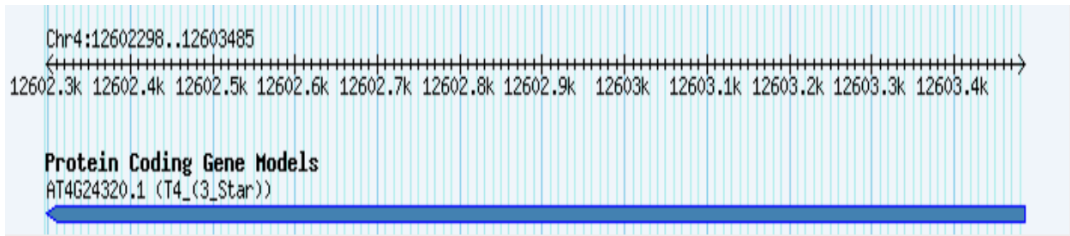
At1g79120



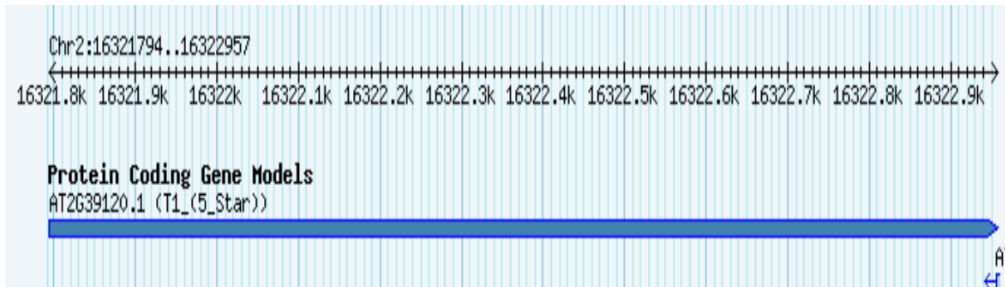
At1g71850



At4g24320



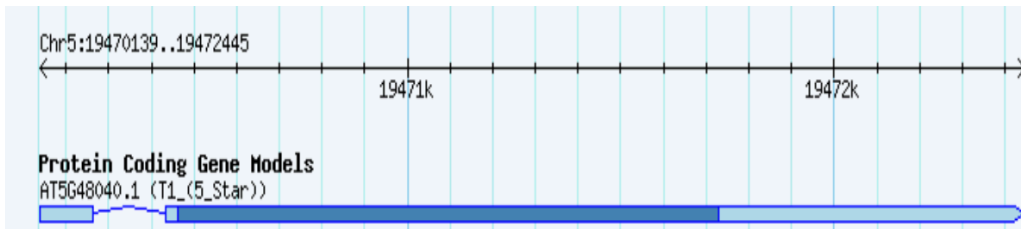
At2g39120



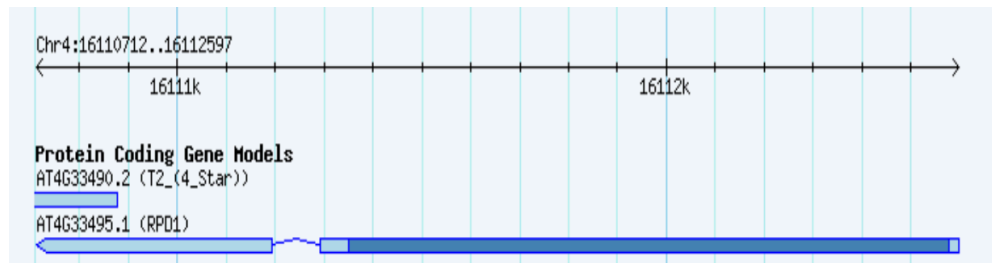
At3g58510



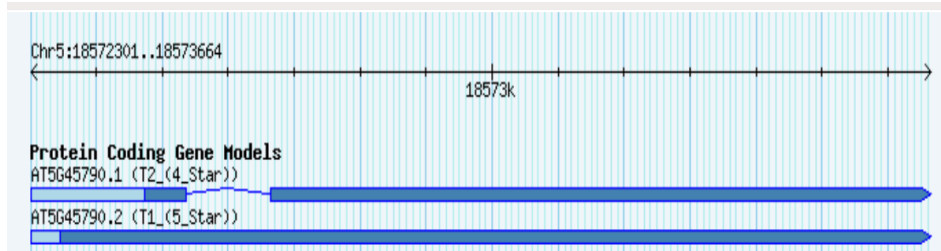
At5g48040



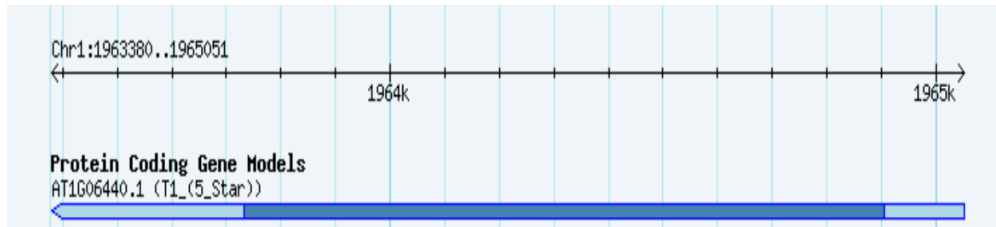
At4g33495



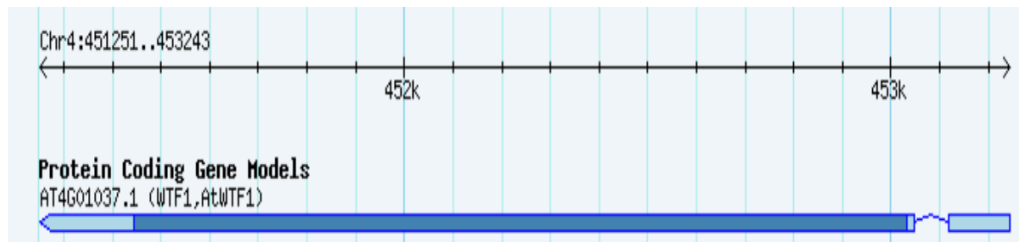
At5g45790



At1g06440



At4g01037



At5g62990



**Εικόνα 14. Δομική ανάλυση των 15 γονιδίων PORR με εξώνια-εσώνια.** Μπλέ σκούρο κωδικές αλληλουχίες εξωνίων, μπλε ανοιχτό 5' ή 3' μη-μεταφραζόμενες περιοχές (UTR, untranslated regions) εξωνίων. Γραμμή με γωνία κλίσης τα ιντρόνια. Το βέλος δείχνει τη φορά της μεταγραφής. Το μέγεθος του γονιδίου φαίνεται. Οι αριθμοί αναφέρονται στα γονίδια και στη θέση τους στο αντίστοιχο χρωμόσωμα.

Μια άλλη βασική παρατήρηση για τα γονίδια της οικογένειας PORR του *Arabidopsis* είναι η δομή τους. Η εξαγωγή των πληροφοριών από την ανάλυση όπως φαίνεται στις Εικόνες 14 και 15 δείχνει ότι 1. Τα περισσότερα γονίδια δεν έχουν ιντρόνια (At2g31290, At3g63090, At4g089402, At1g71850, At4g24320, At2g39120, At1g06440). Τα γονίδια At5g21970, At5g48040, At4g01037, At5g62990 (LEFKOTHEA) και το At2g31290 έχουν ένα ιντρόνιο στη 5' μη-μεταφραζόμενη περιοχή. Το τελευταίο γονίδιο

μπορεί να δίνει ένα δεύτερο μήνυμα χωρίς την αποκοπή ιντρονίου χωρίς φυσικά να αλλάζει το πρωτεϊνικό προϊόν. 3. Το γονίδιο At4g33495 έχει ένα ιντρόνιο στη 3' μη-μεταφραζόμενη περισοχή, 4. Το γονίδιο At1g79120 έχει δύο ιντρόνια στη 3' μη-μεταφραζόμενη περιοχή. 5. Τα γονίδια που περιέχουν ιντρόνια στη κωδική τους περιοχή παρουσιάζουν ιδιαίτερο ενδιαφέρον. Το γονίδιο At5g45790 μπορεί να δώσει δύο μηνύματα. Ένα από το οποίο να έχει αφαιρεθεί το ιντρόνιο και το δεύτερο να περιέχει το ιντρόνιο δηλαδή να έχει γίνει παράληψη ιντρονίου (skip intron). Το πρωτεϊνικό προϊόν που παράγεται από τα δύο αυτά μηνύματα είναι διαφορετικό.

Ενδιαφέρον επίσης παρουσιάζει το γονίδιο At3g58520 το οποίο έχει τα εξής χαρακτηριστικά. Παρουσιάζεται σε δύο μορφές μνημάτων όπου στην πρώτη περίπτωση έχει ένα ιντρόνιο και στη δεύτερη έχει δύο ιντρόνια. Το μέρος του πρώτου εξωνίου φαίνεται να αποτελεί μέρος ιντρονίου στη δεύτερη περίπτωση που υπάρχουν δύο ιντρόνια. Επίσης, μέρος του ιντρονίου της πρώτης περίπτωσης αποτελεί εξώνιο για τη δεύτερη περίπτωση δηλαδή γίνεται εναλλακτικό μάτισμα αλλά και αλλαγή ιντρονικής/εξωνικής πειοχής. Οι ισομορφές των πρωτεϊνών που παράγονται από το γονίδιο αυτό είναι διαφορετικές.

introns	0	0/1	1	1/2	2
At5g21970					
At4g08940					
At3g63090					
At2g31290					
At1g79120					
At1g71850					
At4g24320					
At2g39120					
At3g58520					
At5g48040					
At4g33495					
At5g45790					
At1g06440					
At4g01037					
At5g62990					

**Εικόνα 15. Κατηγοριοποίηση των 15 PORR γονιδίων με βάση το πιθανό αριθμό ιντρονίων.** Το At3g58510 φαίνεται να δημιουργεί δύο mRNA τα οποία μπορεί να προκύπτουν από εναλλακτικό

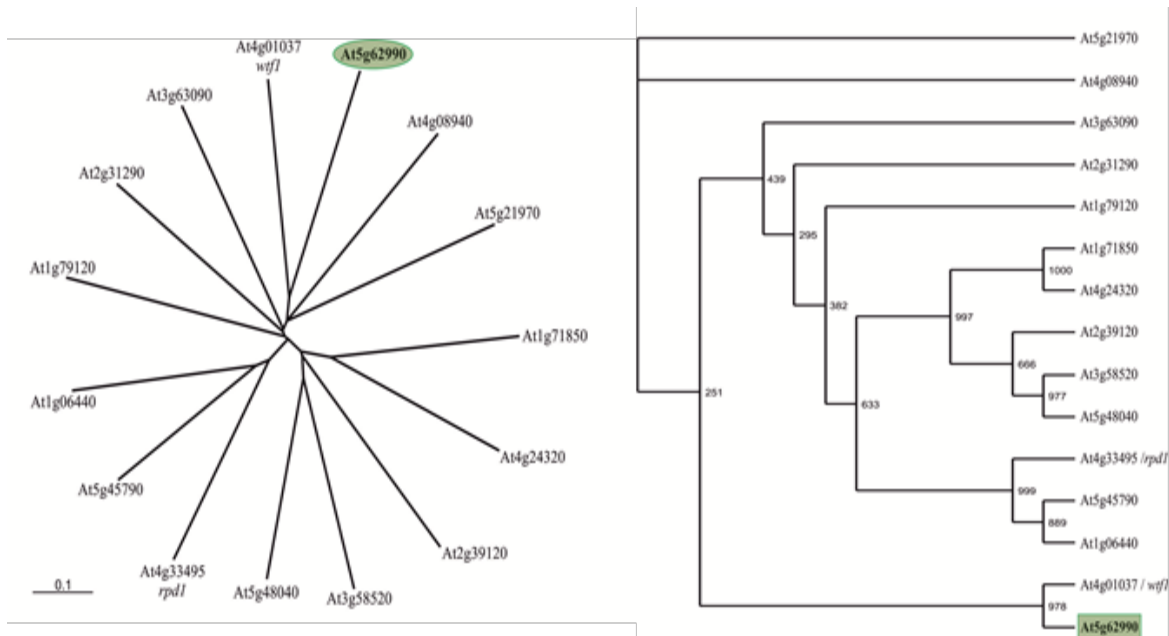




**Εικόνα 16. Πολλαπλή ευθυγράμμιση αμινοξικών ακολουθιών των πρωτεϊνών της οικογένειας PORR του *Arabidopsis*.** Ο κόκκινος αστερίσκος δείχνει το συντηρημένο αμινοξύ Γλυκίνη-373, που στη μετάλλαξη *lefko1-1* έχει αντικατασταθεί από το Ασπαρτικό οξύ (G<sub>373</sub>→D). Σε μαύρο πλαίσιο τα αμινοξέα που παρμένουν, σε γκρι πλαίσιο αλλαγές αμινοξέων με παραπλήσιες βιοχημικές-βιοφυσικές ιδιότητες. Οι αριθμοί στα δεξιά αναφέρονται τις θέσεις των αμινοξέων στις διάφορες πρωτεΐνες. Παύλες έχουν ενσωματωθεί από το πρόγραμμα για καλύτερη στοίχιση των αλληλουχιών. Χρησιμοποιείται μονογράμματα για τα αμινοξέα. Με πράσινη και κίτρινη σκιαφράφιση το NES και το NLS της πρωτεΐνης LEFKOTHEA.

Η ομολογία των πρωτεϊνών της οικογένειας PORR του *Arabidopsis* με πολλαπλή ευθυγράμμιση των αμινοξέων δείχνει ότι η οικογένεια έχει μια μερική ομολογία (Εικόνα 16A). Εντοπίζεται ότι στο άμινο άκρο δεν υπάρχει καθόλου ομολογία μια και η περιοχή αυτή οδηγεί την ώριμη πλέον πρωτεΐνη στα διάφορα υοκυτταρικά οργανίδια δεν αποτελεί μέρος της ώριμης πρωτεΐνης (συνήθως αποκόπτεται κατά την είσοδο στα οργανίδια και δεν αποτελεί μέρος της λειτουργικής πρωτεΐνης). Το ίδιο περίπου συμβαίνει και στο καρβόξυ άκρο. Μάλλον η περιοχή αυτή δεν αποτελεί μέρος της λειτουργικής δράσης της πρωτεΐνης ή οι αλλαγές που ενυπάρχουν στο μόριο προσδιορίζουν και τη λειτουργία της πρωτεΐνης.

Η περιοχή PORR δείχνει μια σχετική ομολογία. Υπάρχουν μόνο μερικές μικρές αλληλουχίες αμινοξέων που έχουν μια υψηλή ομοιότητα ενώ το μεγαλύτερο μέρος της περιοχής αυτής έχει μια σχετικά χαμηλή ομοιότητα (Εικόνα 16A). Υπάρχουν και αρκετά μέρη της περιοχής αυτής που δεν έχουν καμία ομοιότητα ή πολύ χαμηλή. Οι ομόλογες περιοχές είναι διάσπαρτες μέσα στην περιοχή PORR. Οι περιοχές αυτές είναι οι αλληλουχίες γύρω από το 115, από το 150 και από το 280 αμινοξύ της LEFKOTHEA. Μια μεγάλη σχετικά περιοχή με συντήρηση είναι από το 300-390 αμινοξύ της LEFKOTHEA. Στις αλληλουχίες αυτές φαίνεται να εντοπίζεται η λειτουργική δράση των πρωτεϊνών PORR και της συγγένειας τους με το RNA. Το NES βρέθηκε μόνο στη LEFKOTHEA (Εικόνα 16A).



**Εικόνα 16B. Φυλογενετικό δένδρο Neighbor-Joining χωρίς ρίζα (Unrooted) και με ρίζα (Rooted) των ομόλογων πρωτεϊνών LEFKO1 του φυτού *Arabidopsis thaliana***

Η σκιαγράφηση με πράσινο χρώμα αντιστοιχεί στην πρωτεΐνη LEFKO1 με κωδικό At5g62990 και η πρωτεΐνη At4g01037 είναι ορθόλογη της *wfl* από το ρύζι. Στο δένδρο με ρίζα στα τμήματα των διακλαδώσεων παρουσιάζονται οι τιμές της μεθόδου bootstrap. Οι αριθμοί στις διχάλες αναφέρονται στις τιμές της επαναληψιμότητας σε bootstrap από 1000 επαναλήψεις.

### 3.6 Τοποθέτηση των 15 πρωτεϊνών PORR στα διάφορα υποκυτταρικά οργανίδια του *Arabidopsis*.

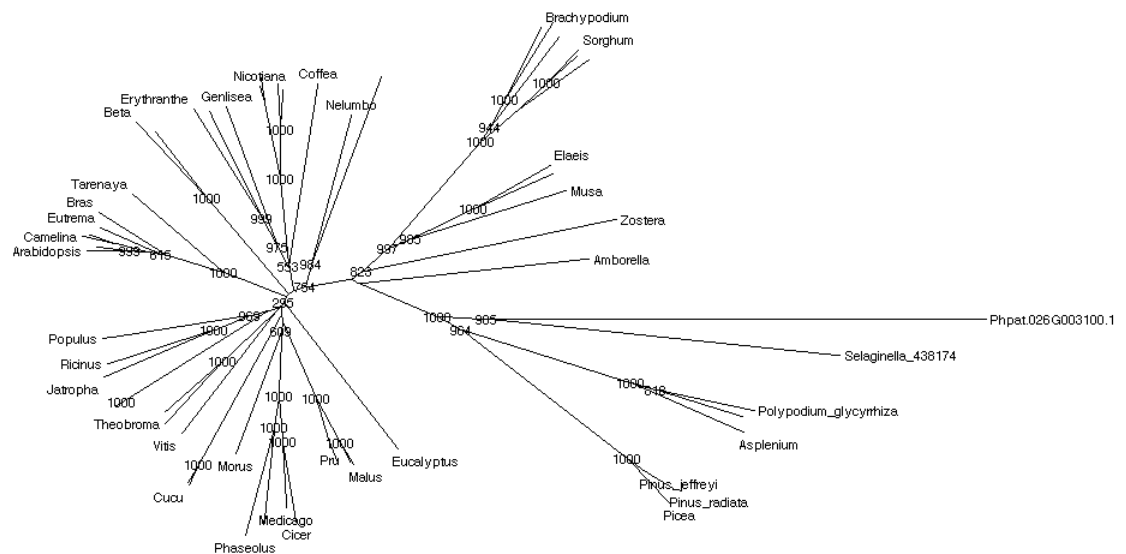
Για να βρεθεί η τοποθέτηση των μελών της οικογένειας PORR στο *Arabidopsis* χρησιμοποιήθηκαν διάφορα προγράμματα και από την ανάλυση πειραματικών δεδομένων. Με ανάλυση MS/MS (mass spectroscopy, φασματογραφία μάζας) βρέθηκαν ότι οι πρωτεΐνες που πηγάζουν από τα γονίδια At5g62990 και At4g01037 εντοπίζονται στα πλαστίδια (Takabayashi et al., 2013) όπως ακριβώς βρέθηκε και με την *in silico* ανάλυση, ενώ οι πρωτεΐνες που πηγάζουν από τα γονίδια At5g48040 και At2g31290 εντοπίζονται στη πλασματική μεμβράνη (Mitra et al., 2009; Li et al., 2012) παρ' όλο που η βιοπληροφορική ανάλυση έδειξε ότι εντοπίζονται στα μιτοχόνδρια (Εικόνα 17). Ενδεχομένως οι πλασματικές μεμβράνες να απομονώθηκαν μαζί με μιτοχόνδρια.



**Εικόνα 17. Τοποθέτηση των 15 PORR πρωτεϊνών σε υποκυτταρικά οργανίδια του *Arabidopsis*.** Τα MS/MS και GFP πάρθηκαν από πειραματικά δεδομένα, τα υπόλοιπα από βάσεις δεδομένων.

### 3.7 Φυλογένεια του γονιδίου *lefkothea*

Το γονίδιο *lefkothea* βρίσκεται σε ανώτερους φυτικούς οργανισμούς και όχι σε κατώτερα φυτά, φύκη, πρωτεοβακτήρια, κυανοβακτήρια, μύκητες και ζώα. Η φυλογενετική ανάλυση με το δένδρόγραμμα που παρουσιάζεται στην Εικόνα 18 δείχνει ότι μπορεί να γίνει ένας σχετικά καλός διαχωρισμός. Τα δικότυλα κάνουν ένα ιδιαίτερο κλάδο ο οποίος είναι πολύ διαφορετικός από αυτό των μονοκότυλων φυτών. Ο κλάδος των τελευταίων δείχνει να είναι πολύ κοντά στο κλάδο των κατώτερων φυτών όπως *Physcomitrellea* και *Selaginela*. Φαίνεται ότι στα δύο αυτά κατώτερα φυτά οι πρωτεΐνες *LEFKOTHEA* είναι πολύ απομακρυσμένες και έχουν μια μικρή σχετική ομοιότητα ή ομολογία. Άλλωστε αυτό φαίνεται και από την Εικόνα 10 (πολλαπλής ευθυγράμμισης αμινοξέων) όπου ένα μεγάλος μέρος της περιοχής PORR λείπει από την αντίστοιχη πρωτεΐνη του *Selaginela*.

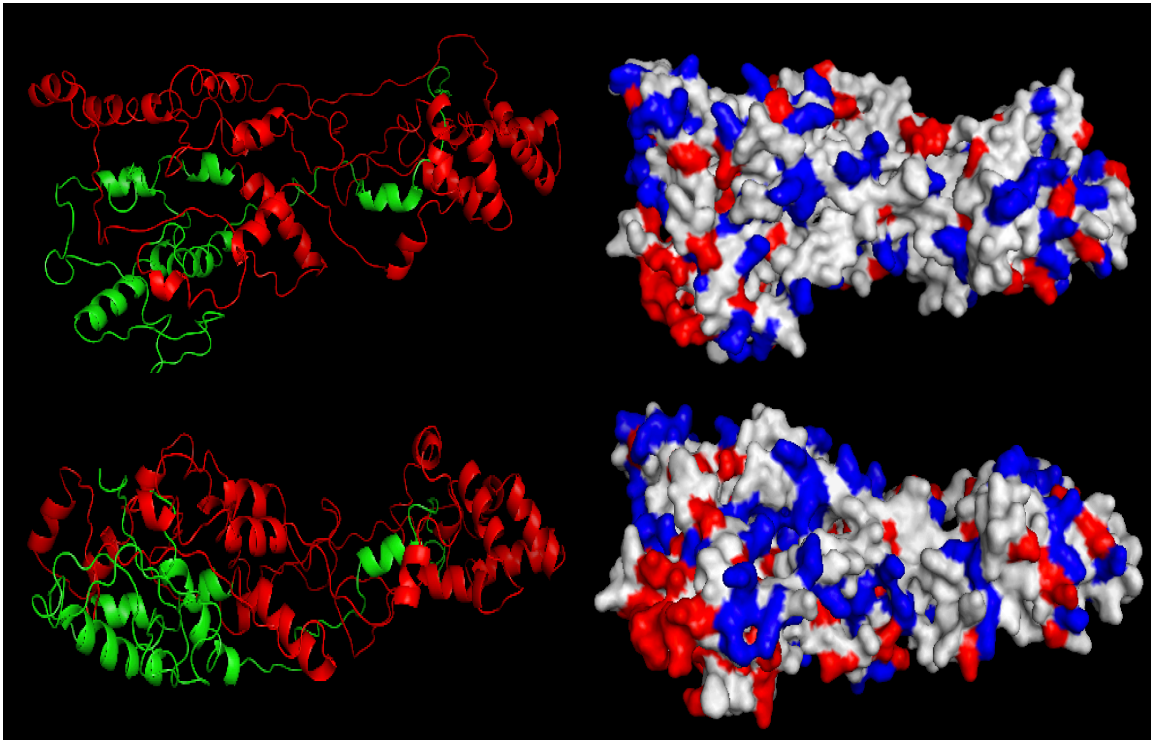


**Εικόνα 18. Φυλογένεια του γονιδίου *Lefkothea* στα διάφορα φυτά.** Οι αριθμοί στις διχάλες αναφέρονται στις τιμές της επαναληψιμότητας σε bootstrap από 1000 επαναλήψεις. Τα κλαδογράμματα γίνονται φανερά.



### 3.8 Τρισδιάστατη Μοντελοποίηση (3D) δομής της LEFKOTHEA PORR με τη χρήση βιοπληροφορικής.

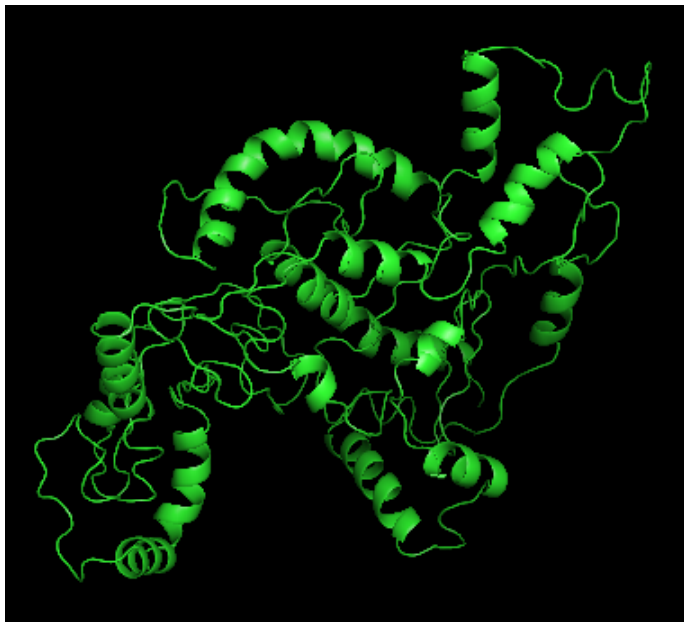
Για την ανάλυση της δομής της πρωτεΐνης LEFKOTHEA χρησιμοποιήθηκε το πρόγραμμα PyMol. Η ανάλυση έδειξε ότι το μεγαλύτερο μέρος της πρωτεΐνης βρίσκεται στη διαμόρφωση της α-έλικας και υπάρχει μια αύλακα που θα μπορούσε να ενωθεί με το RNA. Επίσης φαίνεται ότι υπάρχει μια ισοκατανομή όσο αφορά τα θετικά φορτισμένα αμινοξέα. Το ίδιο συμβαίνει και με τα αρνητικά φορτισμένα αμινοξέα εκτός μιας περιοχής που υπάρχουν συγκεντρωμένα τα αρνητικά φορτισμένα. Φαίνεται επίσης ότι τα θετικά φορτισμένα αμινοξέα υπερτερούν από αυτά των αρνητικών. Φαίνεται επίσης ότι στην αύλακα υπάρχει περισσότερο θετικό φορτίο απ' ότι αρνητικό.



Εικόνα 19. Τρισδιάστατη δομή (3D μοντέλο) της πρωτεΐνης LEFKOTHEA με τη χρήση του προγράμματος PyMol. Αριστερά, η παρουσίαση με α-δομές και στροφές κάτω από δύο διαφορετικές γωνίες και δεξιά, μοντέλο με πληρότητα (space filled) επίσης κάτω από διαφορετικές

γωνίες. Με μπλε τα φορτισμένα θετικά αμινοξέα, με κόκκινο τα αρνητικά φορτισμένα αμινοξέα και με άσπρο τα ουδέτερα.

Επειδή η ανάλυση της πολλαπλής ευθυγράμμισης των αμινοξέων της Εικόνας 10 έδειξε ότι η ομόλογη προς της LEFKOTHEA πρωτεΐνη από το *Selaginella* έχει σχετικά μικρή περιοχή PORR αποτελούμενη από 260 περίπου αμινοξέα σχέση με τη LEFKOTHEA του *Arabidopsis* που αποτελείται από 340 αμινοξέα, έγινε ανάλυση της τρισδιάστατης μοντελοποίησης της. Η ανάλυση της τρισδιάστατης δομής από τη *Selaginella* έδειξε ότι δεν υπάρχει μεγάλη ομοιότητα με αυτή της πρωτεΐνης στο *Arabidopsis*. Φαίνεται ότι η περιοχή αυτή είναι η μικρότερη δομικά ικανή να επιτελέσει τη λειτουργία της σύνδεσης με RNA και της ικανότητας συρραφής των εξωνίων και απάλειψης από τα ιντρόνια. Ενδεχομένως στους κατώτερους φυτικούς οργανισμούς (*Selaginella*), η δομή αυτή να είναι αρχέγονη για τη λειτουργικότητα της πρωτεΐνης, ή να έχει εξελιχθεί στο *Arabidopsis* έτσι ώστε να μπορεί να επιτελεί περισσότερες λειτουργίες ή να έχει πολλαπλούς στόχους RNA όσο αφορά το μάτισμα διαφορετικών ιντρονίων (τύπου II και I). Δεν μπορεί να αποκλειστεί και η πιθανότητα η πρωτεΐνη LEFKOTHEA να είναι cargo protein (μεταφορική πρωτεΐνη) και να μεταφέρει τα ώριμα mRNA έξω από το πυρήνα για να μεταφραστούν σε πρωτεΐνες που ενδεχομένως αρκετές από αυτές να στοχεύουν προς το χλωροπλάστη για την κανονικότητα και τη λειτουργική επάρκεια του χλωροπλάστη. Άλλωστε η έλειψη της πληροφορίας της πρωτεΐνης οδηγεί σε θνησιμότητα, ανεπάρκεια του χλωροπλάστη ή μη επάρκεια στην περίπτωση του μεταλλάγματος *lefko1-1* με φαινότυπο τις λευκές κοτύλες.





**Εικόνα 20. Τρισδιάστατο μοντέλο της ομόλογης πρωτεΐνης προς τη LEFKOTHEA του *Selaginella*.** Αριστερά, η παρουσίαση με  $\alpha$ -δομές και στροφές κάτω από δύο διαφορετικές γωνίες και δεξιά, μοντέλο με πληρότητα (filled) επίσης κάτω από διαφορετικές γωνίες. Με μπλε τα φορτισμένα θετικά αμινοξέα, με κόκκινο τα αρνητικά φορτισμένα αμινοξέα και με άσπρο τα ουδέτερα.

### 3.9 Έκφραση του γονιδίου *At5g62990*

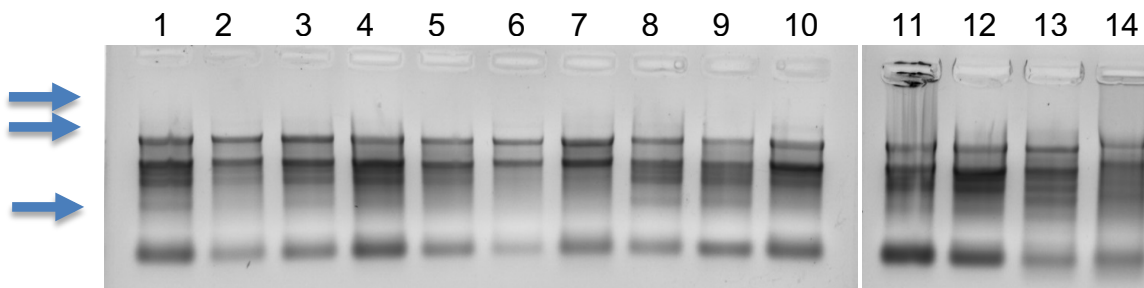
Προκειμένου να μελετηθεί το πρότυπο της έκφρασης του γονιδίου *At5g62990* καθ'ολη τη διάρκεια της ανάπτυξης του φυτού *Arabidopsis thaliana* απομονώθηκε ολικό RNA από διάφορους ιστούς κι έγιναν αντιδράσεις αντίστροφης μεταγραφής (RT – PCR). Η ποσότητα του RNA εκτιμήθηκε με φασματοφωτόμετρο και η ακεραιότητα του σε ηλεκτροφόρηση αгарόζης 1% (Εικόνα 21). Τα rRNA υποδεικνύουν ότι η μεθοδολογία της απομόνωσης του RNA που ακολουθήθηκε είχε σαν αποτέλεσμα ικανής ποσότητας και καλής ποιότητας RNA. Χρησιμοποιήθηκε ο εκκινητής GUSTGAR για την αντίστροφη μεταγραφή και οι εκκινητές 90.2 και 90.4 για το PCR που ακολούθησε στο cDNA. Κατά τη μελέτη της έκφρασης με τη μέθοδο του ημιποσοτικού RT-PCR, απομονώθηκε ολικό RNA από νεαρά σπορόφυτα ανεπτυγμένα σε φως και σκοτάδι, από ώριμα φύλλα, βλαστούς, άνθη και καρποταξίες τριών αναπτυξιακών σταδίων. Ως γονίδιο αναφοράς-ελέγχου χρησιμοποιήθηκε το γονίδιο GAPDH (*At3g04120* - με εκκινητές τους 04120F και 04120R) που κωδικοποιεί την αφυδρογονάση της γλυκεριναλδεϋδης-3P που συμμετέχει στο μονοπάτι της γλυκόλυσης και της νεογλυκογένεσης. Οι κύκλοι του PCR που χρησιμοποιήθηκαν ήταν οι ενδεδιγμένοι κατά τη διάρκεια της λογαριθμικής αύξησης του προϊόντος. Τα προϊόντα της ανάυσης του PCR ηλεκτροφορήθηκαν σε 1.6% αгарόζη (Εικόνα 22).

Το γονίδιο *At5g62990* εμφανίζει υψηλό επίπεδο έκφρασης στα νεαρά σπορόφυτα και στα όργανα εμβρυακής προέλευσης (κοτυληδόνες, πρωτογενής ρίζα). Αντίθετα, στα ώριμα φύλλα και βλαστούς παρατηρούνται σχετικά χαμηλά επίπεδα έκφρασης. Από το στάδιο του άνθους ως εκείνο της εξέλιξής του σε ώριμη καρποταξία, παρατηρείται μια ελαφρά κλιμακούμενη μείωση της έκφρασης.

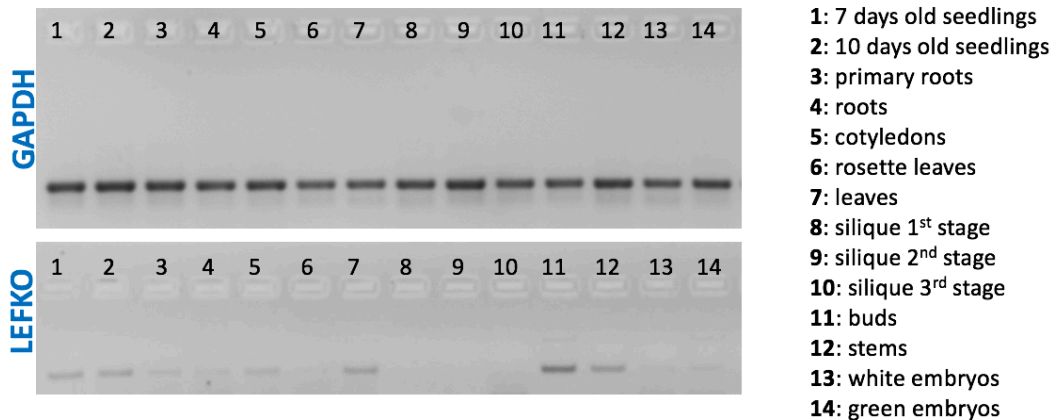
Μια ακόμη πηγή μελέτης της έκφρασης του γονιδίου *At5g62990*, αποτέλεσε η ιστοσελίδα <http://www.genevestigator.ethz.ch>, για την έκφραση τόσο στα διάφορα όργανα του φυτού όσο και στα στάδια ανάπτυξής του. Τα αποτελέσματα στηρίζονται σε αναλύσεις που έχουν γίνει με τη βοήθεια μικροσυστοιχιών (microarrays) και παρουσιάζονται στην Εικόνα 23. Η ανάλυση των μικροσυστοιχιών δίνει παρόμοια αποτελέσματα με αυτά που

προέκυψαν από το RT-PCR, δηλαδή έκφραση σε όλα τα στάδια ανάπτυξης του φυτού, με έμφαση στη βλάστηση, στο σχηματισμό του άνθους και στην εμβρυογένεση.

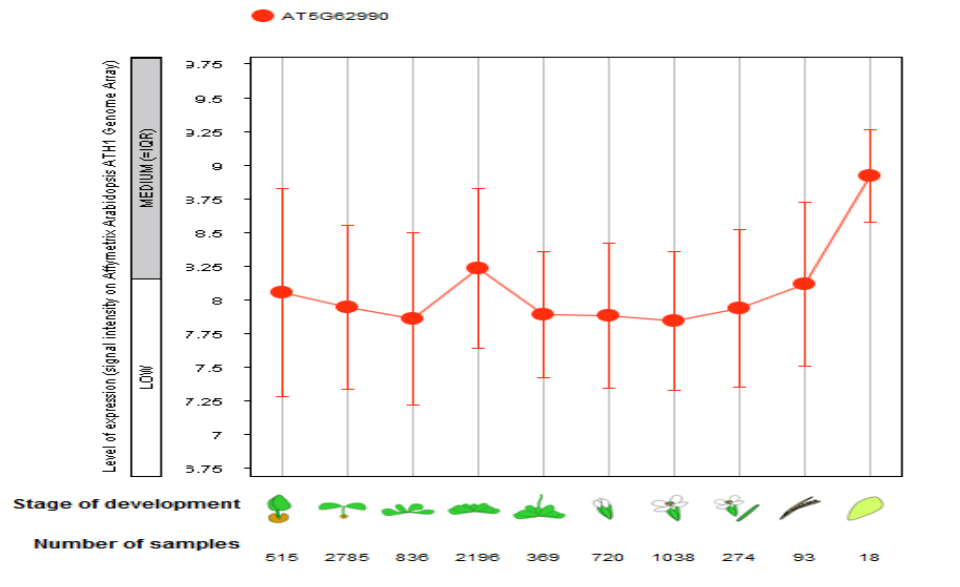
Περαιτέρω ανάλυση της έκφρασης μετά τη χρήση του αλγορίθμου κανονικότητας GCOS (GeneChip Operating System) επαλήθευσε ότι η έκφραση του γονιδίου είναι σχετικά χαμηλή σε όλα τα στάδια ανάπτυξης του οργανισμού και σε όλα τα όργανα. Υπάρχει μια μικρή αύξηση της ποσότητας του mRNA σε αποξηραμένα σπέρματα δηλαδή στα έμβρυα, που όμως μετά τα πρώτα στάδια της έκπτυξης υπάρχει μια σχετική ελάττωση (Εικόνα 24). Φαίνεται ότι η δραστηριότητα του γονιδίου στα έμβρυα είναι ζωτικής σημασίας όπως ακριβώς έχει δείχθει και από τη θνησιγόνο μετάλλαξη όπου το έμβryo παγώνει στο στάδιο της καρδιάς.



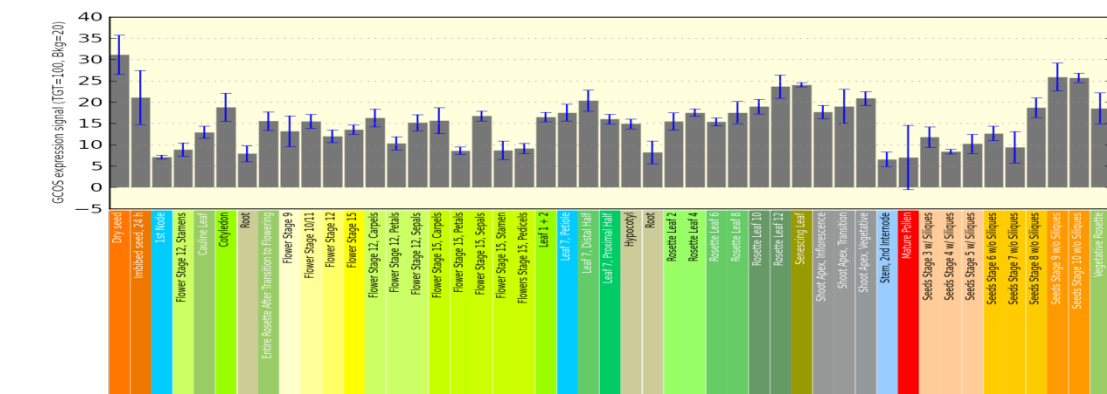
**Εικόνα 21. Ηλεκτροφόρηση ολικού RNA σε 1% αгарόζη.** 1. 7 days old seedlings, 2. 10 days old seedlings, 3. Primary roots, 4. roots, 5. Cotyledons, 6. Rosette leaves, 7. Leaves, 8. Siliques 1st stage, 9. Siliques 2nd stage, 10. Siliques 3rd stage, 11. Buds, 12. Stems, 13. White embryos, 14. Green embryos. Τα βέλη υποδεικνύουν τα 26S, 18S και 5S rRNA και δείχνει την ακεραιότητα του RNA (χωρίς να έχει υποστεί πέψη από RNases του κυττάρου).



**Εικόνα 22. Πρότυπο έκφρασης του *lefkothea* κατά τη διάρκεια της ανάπτυξης.** Το δείγμα 1 είναι από RNA απομονωμένο από 7 ημερών αναπτυσσόμενους βλαστούς, το 2 από 10 χρονών βλαστούς, το 3 από νεαρή καρποταξία, το 4 από ρίζα, το 5 από κοτυληδόνες, το 6 από φύλλα ροζέτας, το 7 από φύλλα, το 8 από περικάρπιο 1<sup>ου</sup> σταδίου, το 9 από περικάρπιο 2<sup>ου</sup> σταδίου, το 10 από περικάρπιο 3<sup>ου</sup> σταδίου, το 11 από μπουμπούκια, το 12 από μίσχους, το 13 από άσπρα έμβρυα και το 14 από πράσινα έμβρυα. Ως γονίδιο αναφοράς χρησιμοποιήθηκε το γονίδιο GAPDH.



**Εικόνα 23. Ανάλυση με το eFP browser για την έκφραση του γονιδίου *LEFKOTHEA* σε διάφορους ιστούς.** Η ανάλυση δείχνει ότι η έκφραση του γονιδίου είναι χαμηλή και αύξηση της έκφρασης παρουσιάζεται στα έμβρυα-σπέρματα.



**Εικόνα 24. Ανάλυση της έκφρασης του γονιδίου *LEFKOTHEA* στα διάφορα όργανα και αναπτυξιακά στάδια του *Arabidopsis*.** Γενικά η έκφραση του γονιδίου είναι χαμηλή με μια μικρή αύξηση όταν το σπέρμα είναι αποξηραμένο.

## 4. ΣΥΖΗΤΗΣΗ

### 4.1 Πρωτεΐνες πολλαπλής τοποθέτησης μέσα στο κύτταρο

Στα ευκαρυωτικά κύτταρα, πολλές κυτταρικές λειτουργίες διπλασιάστηκαν εξαιτίας της εξέλιξης και αυτό είχε ως αποτέλεσμα να πραγματοποιούνται παρόμοιες λειτουργίες σε διαφορετικά υποκυτταρικά διαμερίσματα και από διαφορετικές ισομορφές ενζύμων (Krause and Krupinska, 2009). Για μεγάλο χρονικό διάστημα, υποστηριζόταν ο κανόνας ότι ένα γονίδιο, αντιστοιχούσε σε μια πρωτεΐνη και μια πρωτεΐνη σε ένα υποκυτταρικό διαμέρισμα μέσα στο κύτταρο. Όμως η ανακάλυψη πρωτεϊνών με διπλή τοποθέτηση σε δύο διαμερίσματα του κυττάρου ανέτρεψε τον παραπάνω ισχυρισμό (Natsoulis *et al.*, 1986). Στα φυτικά κύτταρα, η πρώτη πρωτεΐνη διπλής τοποθέτησης που ανακαλύφθηκε, ήταν η ρεδοκτάση της γλουταθειόνης (Glutathione Reductase) στο *Pisum sativum* και βρέθηκε ότι τοποθετείται στα μιτοχόνδρια αλλά και στους χλωροπλάστες (Creissen *et al.*, 1995). Από τότε, ο αριθμός των πρωτεϊνών με αποδεδειγμένη διπλή ή πολλαπλή τοποθέτηση αυξάνεται σταθερά (Karniely *et al.*, 2005) παρέχοντας συμπληρωματικές γνώσεις στη ρύθμιση αυτής της λειτουργίας.

Έχουν χαρακτηριστεί δύο βασικοί μηχανισμοί διπλής τοποθέτησης των πρωτεϊνών στα οργανίδια (Peeters and Small, 2001; Mackenzie, 2005). Ο πρώτος μηχανισμός (twin presequences) στηρίζεται στην ύπαρξη δυο διαδοχικών σε σειρά πεπτιδίων συνθήματος που αναγνωρίζονται εκλεκτικά από το σύστημα υποδοχέων του κάθε οργανιδίου στο οποίο πρόκειται να τοποθετηθεί το πολυπεπτίδιο που συντέθηκε στο κυτταρόπλασμα. Απαραίτητη προϋπόθεση στο μηχανισμό αυτό είναι ότι πρέπει να κωδικοποιηθούν δυο ανεξάρτητες ισομορφές της πρωτεΐνης ώστε η κάθε μια να φέρει στο N-τελικό άκρο ένα από τα πεπτίδια συνθήματος. Ο δεύτερος μηχανισμός (ambiguous presequence) εξασφαλίζει την είσοδο του πολυπεπτιδίου σε δυο οργανίδια, εξαιτίας της ιδιότητας του μοναδικού πεπτιδίου συνθήματος που εντοπίζεται στο N-τελικό άκρο, να αναγνωρίζεται από τους υποδοχείς του συστήματος εισόδου και των δυο οργανιδίων. Οι δυο αυτοί μηχανισμοί φαίνεται ότι καθορίζουν την ταυτόχρονη είσοδο μιας πρωτεΐνης στα

Plant protein	Localization methods*	Function in organelles	Function in nucleus	
DHFR	P-gold N-gold M-act	Dihydrofolate reductase-thymidylate synthase	Dihydrofolate reductase-thymidylate synthase	
ΜΙΤΟ	MFP1	Localization at stroma side of thylakoids and co-purifying with nucleoids	MAR-binding filament-like protein 1, component of nuclear matrix	. Στα
οργ	SEBF	RNA binding?	Transcriptional repressor of the <i>PR10a</i> gene	ουσία
ΠΕΠ	STEP1, cp31	RNA-binding protein	Telomere-binding protein	
	Whirly1	RNA-binding factor involved in splicing factor	Transcription factor; telomere binding	
	AtCDT1	Division, interaction with ARC6	Replication	ύ του
ΜΟΥ	LEM1	Ribosomal protein (PRPS9)	Required for embryogenesis	ρίζεται
ΑΠΨ	AtLig1	DNA ligase	DNA ligase	χρονη
ΤΟΠ	NtWin4	Unknown regulatory function	Transcriptional repressor	ης για
ΤΟ C	ATXR5	Unknown	Cell cycle and DNA replication interaction with PCNA	οποίο
ΚΑΙ ;	At2g44940	Unknown	AP2/EREBP transcription factor	).
	AtIPT3	Farnesylated: N-gfp Nonfarnesylated: P-gfp	Cytokinin biosynthesis	άζουν
	APL (At1g79430)	Unknown	Transcription factor	

δυναμική τοποθέτηση συμμετέχοντας στο μεταβολισμό των νουκλεϊνικών οξέων μπορεί να είναι ένα μέσο για τη διασφάλιση της επικοινωνίας μεταξύ των διαμερισμάτων του κυττάρου που έχουν γενετικό υλικό.

## 4.2 Δυναμική τοποθέτηση *de novo* σύνθεσης πρωτεϊνών

Ένας κοινός μηχανισμός για την παραγωγή πολλαπλών πρωτεϊνών από ένα γονίδιο είναι η χρησιμοποίηση διαφορετικής αρχής της μεταγραφής ή της μετάφρασης ή η χρησιμοποίηση διαφορετικών εξωνίων από εναλλακτικό μάτισμα (Πίνακας 2). Ένα παράδειγμα διαφορετικής έναρξης της μεταγραφής και της μετάφρασης, με αποτέλεσμα τη δυναμική τοποθέτηση στον πυρήνα και τα μιτοχόνδρια στο *Arabidopsis thaliana* αποτελεί η πρωτεΐνη DNA ligase 1 (Sunderland *et al.*, 2006). Αυτή η πρωτεΐνη εμφανίζει δυναμική τοποθέτηση και είναι χαρακτηριστικό παράδειγμα της κυριαρχίας του μιτοχονδριακού πεπτιδίου συνθήματος έναντι του πυρηνικού σινιάλου τοποθέτησης (NLS, Nuclear Localization Signal) (Sunderland *et al.*, 2006). Χαρακτηριστικό επίσης παράδειγμα τοποθέτησης στον πυρήνα και στους χλωροπλάστες αποτελεί η πρωτεΐνη cp31 η οποία εξαιτίας διαφορετικών κωδικονίων έναρξης που προκύπτουν από εναλλακτικό μάτισμα, σχηματίζει εναλλακτικές πρωτεϊνικές ισομορφές που συνδέονται με RNA.

**Εικόνα 25. Πρωτεΐνες με δυαδική τοποθέτηση στον πυρήνα και στα φυτικά οργανίδια (Krause and Krupinska, 2009).** Παρουσιάζονται πρωτεΐνες με δυαδική τοποθέτηση στον πυρήνα και στα φυτικά οργανίδια, η μέθοδος μελέτης της τοποθέτησής τους και η λειτουργία τους στο χλωροπλάστη και τον πυρήνα.

Η μετά-μεταφραστική τροποποίηση των πρωτεϊνών, όπως για παράδειγμα η φωσφορυλίωση, αποτελεί εναλλακτικό ρυθμιστικό μηχανισμό ελέγχου υποκυτταρικής τοποθέτησης πρωτεϊνών με δυαδικά μοτίβα στόχευσης (Karniely *et al.*, 2005). Πρόσφατα, περιγράφηκε ένας νέος τύπος τροποποίησης ο οποίος ελέγχει τη κατανομή της πρωτεΐνης phosphateisopentenyl-transferase του φυτού *Arabidopsis thaliana* (AtIPT3) στα πλαστίδια και τον πυρήνα. Αυτή η πρωτεΐνη περιέχει ένα μοτίβο μετα-μεταφραστικής τροποποίησης τύπου “farsenylation” δηλαδή σύνδεσης ενός ισοπρενοειδούς με μια κυστεΐνη στο καρβόξυ τελικό άκρο της πρωτεΐνης. Πειράματα παροδικού μετασχηματισμού επιδερμικών κυττάρων κρεμμυδιού (*Allium cepa*) αποκάλυψαν ότι η μεταφραστική σύντηξη της πρωτεΐνης με το γονίδιο *GFP* οδηγεί το πολυπεπτίδιο που τοποθετείται στον πυρήνα. Αντίθετα όταν η πρωτεΐνη δεν υποστεί την τροποποίηση τύπου “farsenylation” τοποθετείται στα πλαστίδια (Galichet *et al.*, 2008). Εκτός από τις παραπάνω μετά-μεταφραστικές τροποποιήσεις, η τοποθέτηση των πρωτεϊνών επηρεάζεται και από την αλληλεπίδραση με άλλες πρωτεΐνες (Πίνακας 2). Τέτοιο παράδειγμα αποτελεί το γονίδιο του καπνού *Tobacco Stress-Induced Gene-1* (*Tsi1*) το οποίο δεσμεύεται με την πρωτεΐνη protein Tsi1 (Ham *et al.*, 2006).

### 4.3 Οι πρωτεΐνες διπλής τοποθέτησης σε οργανίδια και πυρήνα συμμετέχουν στη σταθερότητα του DNA και την έκφραση γονιδίων

Παρόλο που το σενάριο διπλής υποκυτταρικής τοποθέτησης πρωτεΐνης αρχικά είχε παραβλεφθεί, ο Ian Small και οι συνεργάτες του από το 1998 είχε θεωρήσει πιθανό το γεγονός η ίδια πρωτεΐνη να βρίσκονται ταυτόχρονα σε δύο υποκυτταρικά διαμερίσματα που περιέχουν DNA προκειμένου να συμμετέχει στο μεταβολισμό του DNA ή και του RNA (Small *et al.*, 1998). Αν και η πληθώρα κοινών δραστηριοτήτων που σχετίζεται με τη σταθερότητα του γονιδιώματος και την έκφραση των γονιδίων, επιβάλλει την ύπαρξη πρωτεϊνών ταυτόχρονα στον πυρήνα ή σε υποκυτταρικά οργανίδια όπως τα πλαστίδια και τα μιτοχόνδρια, οι πρώτες αναφορές περιέγραφαν αποκλειστικά τη δυαδική τοποθέτηση πρωτεϊνών στα μιτοχόνδρια και τους χλωροπλάστες (Silva-Filho *et al.*, 2003). Αντίθετα, οι πρωτεΐνες που τοποθετούνται στο πυρήνα και ταυτόχρονα σε ένα ή και στα δύο φυτικά οργανίδια δεν έχουν μέχρι στιγμής μελετηθεί διεξοδικά. Στις περισσότερες των περιπτώσεων, οι πρωτεΐνες διπλής τοποθέτησης, είχαν αρχικά περιγραφεί ότι τοποθετούνται μόνο σε ένα υποκυτταρικό διαμέρισμα και αργότερα χαρακτηρίστηκε η δυαδική τοποθέτησή τους σε οργανίδιο που περιέχει DNA επιβεβαιώνοντας την αρχική υπόθεση των Small *et al.* (Πίνακας 2).

Για παράδειγμα η πρωτεΐνη MFP1 (MAR-binding filament-like protein 1) είχε αρχικά θεωρηθεί λανθασμένα ότι τοποθετείται αποκλειστικά στον πυρήνα. Η πρωτεΐνη MFP1 εντοπίζεται μόνο στα φυτά και εμφανίζει μη ειδική δέσμευση με το DNA. Με τη κατασκευή MFP1-GFP εντοπίστηκαν στίγματα στη περιφέρεια του πυρήνα των κυτάρων φυτών καπνού, *Nicotiana benthamiana*, και εξαιτίας αυτού του προτύπου υποκυτταρικής τοποθέτησης η πρωτεΐνη χαρακτηρίστηκε ως συστατικό μέρος του πυρήνα (Gindullis *et al.*, 1999). Παρόλα αυτά, πιο λεπτομερής ανάλυση αποκάλυψε ότι η πρωτεΐνη MFP1 τοποθετείται και στους χλωροπλάστες, όπου συνδέεται με τα θυλακοειδή του στρώματος του χλωροπλάστη και απομονώνεται μαζί με τα νουκλεοειδή του χλωροπλάστη (Jeong *et al.*, 2003). Οι ανοσολογικές αναλύσεις με ειδικό αντίσωμα επιβεβαίωσαν ότι η πρωτεΐνη βρίσκεται σε αφθονία στους χλωροπλάστες, αλλά και στο πυρήνα των κυτάρων (Samaniego *et al.*, 2008). Με βάση τα χαρακτηριστικά υποκυτταρικής τοποθέτησης, αυτή η πρωτεΐνη φαίνεται μάλλον ότι σχετίζεται με τη δομική οργάνωση του DNA στον πυρήνα και στα πλαστίδια.

Ένα άλλο παράδειγμα πρωτεΐνης με όμοιες λειτουργίες στον πυρήνα και σε οργανίδια που περιέχουν DNA, είναι αυτό της DNA-λιγάσης-1 του φυτού *Arabidopsis thaliana* (Sunderland *et al.*, 2006) (Πίνακας 3). Η έναρξη της μετάφρασης από το πρώτο

κωδικόνιο έναρξης παράγει μια μιτοχονδριακή ισομορφή, ενώ η έναρξη της μετάφρασης από το δεύτερο κωδικόνιο έναρξης οδηγεί στη σύνθεση ισομορφής που τοποθετείται στον πυρήνα (Sunderland *et al.*, 2006). Επειδή το φυτό *Arabidopsis* έχει μόνο μία λιγάση, φαίνεται ότι η πρωτεΐνη αυτή έχει παρόμοιους ρόλους στον πυρήνα και στα οργανίδια. Έτσι φαίνεται καθαρά ότι ο οργανισμός ελατώνει το περιεχόμενο του DNA του πυρήνα με τη χρήση μόνο ενός γονιδίου με δυαδική τοποθέτηση αλλά αυστηρά την ίδια λειτουργία και στα δύο υποκυτταρικά οργανίδια.

Μία ειδική κατηγορία πρωτεϊνών με δυαδική τοποθέτηση στα οργανίδια και τον πυρήνα είναι οι πρωτεΐνες που δεσμεύονται στα τελομερή. Αυτές οι πρωτεΐνες παίζουν σημαντικό ρόλο στη δομή και τη σταθερότητα του γονιδιώματος μέσω της σύνδεσης με μονόκλωνο ή δίκλωνο DNA στα τελομερή. Ορισμένες από αυτές τις πρωτεΐνες έχουν χαρακτηριστεί ως πλαστιδιακές και δεσμεύουν RNA (Kwon *et al.*, 2004) όπως οι cp31 και cp33 (Ohta *et al.*, 1995). Η μεταφραστική σύντηξη με GFP συντημένης μορφής της πρωτεΐνης cp31, που ονομάζεται STEP1 (single-stranded telomere-binding protein 1), τοποθετείται στον πυρήνα κυττάρων του φυτού *Solanum tuberosum*. Επίσης, η πρωτεΐνη SEBF (silencing element binding factor) που απομονώθηκε ως πρωτεΐνη δέσμησης μονόκλωνου DNA και είναι ομόλογη της cp31 ή PR10a (pathogenesis-related 10a) ανιχνεύθηκε ανοσολογικά ταυτόχρονα στο πυρήνα και στον χλωροπλάστη (Boyle *et al.*, 2001). Προτείνεται λοιπόν ότι η SEBF είναι ένας πυρηνικός ρυθμιστής που αποθηκεύεται στο χλωροπλάστη μέχρι να επέλθουν συγκεκριμένες συνθήκες οπότε θα απελευθερωθεί εκτός του χλωροπλάστη. Άλλη πρωτεΐνη με τα παρόμοια χαρακτηριστικά δέσμησης στα τελομερή η οποία όμως διαθέτει και πεπτιδίο συνθήματος για στόχευση στον χλωροπλάστη είναι η Whirly1 (Krause *et al.*, 2005; Yoo *et al.*, 2007).

Οι μεταγραφικοί παράγοντες αποτελούν σημαντική ομάδα πρωτεϊνών με πιθανή δυαδική τοποθέτηση στον πυρήνα και στα οργανίδια. *In silico* ανάλυση γονιδίων του *Arabidopsis* που κωδικοποιούν για μεταγραφικούς παράγοντες, αποκάλυψε ότι 48 πρωτεΐνες πιθανώς να τοποθετούνται στα πλαστίδια (Wagner *et al.*, 2006). Μία διαφορετική προσέγγιση *in silico* ανάλυσης προέβλεψε ότι 78 μεταγραφικοί παράγοντες πιθανώς τοποθετούνται στα πλαστίδια ενώ 12 στα μιτοχόνδρια (Schwacke *et al.*, 2007). Από την ομάδα των μιτοχονδριακών μεταγραφικών παραγόντων, η πρωτεΐνη APL (altered phloem development-like) μέσω σύντηξης με GFP πρόσφατα αποδείχθηκε πειραματικά ότι τοποθετείται εκτός του πυρήνα και στα μιτοχόνδρια (Carrie *et al.*, 2009). Στην συγκεκριμένη περίπτωση, η λειτουργική δράση (στόχευση σε γονίδια για ρύθμιση της έκφρασης) είναι λογικό να είναι διαφορετική όσο αφορά τα γονίδια-στόχους.



Το πιο χαρακτηριστικό και μελετημένο παράδειγμα διπλής τοποθέτησης μεταγραφικού παράγοντα είναι η πρωτεΐνη Whirly1. Η Whirly1 ανήκει σε μια μικρή οικογένεια μεταγραφικών παραγόντων που συνήθως αποτελείται από μόνο 2 μέλη στα περισσότερα αγγειόσπερμα (Desveaux *et al.*, 2005). Αρχικά μελετήθηκε η λειτουργία της στον πυρήνα όπου δεσμεύει μονόκλωνο DNA συμμετέχοντας στη μεταγραφική ενεργοποίηση του γονιδίου *PR10a* (Desveaux *et al.*, 2000) και στην ομοίωση του μήκους των τελομερών (Yoo *et al.*, 2007; Kwon *et al.*, 2004). Παρόλα αυτά, η σύντηξη με GFP σε μετασχηματισμένους πρωτοπλάστες μεσόφυλλου του φυτού της πατάτας, αποκάλυψε υποκυτταρική τοποθέτηση στους χλωροπλάστες, εγείροντας ζήτημα ταυτόχρονης δυαδικής τοποθέτησης της πρωτεΐνης *in vivo* στον πυρήνα αλλά και στους χλωροπλάστες. Παράλληλα, πειράματα ανοσολογικής ανίχνευσης έδειξαν καθαρά ότι η φυσιολογική μορφή της πρωτεΐνης Whirly1 τοποθετείται στον πυρήνα και τους χλωροπλάστες του ίδιου κυττάρου σε φύλλα του φυτού *Hordeum vulgare* (Grabowski *et al.*, 2008). Συνεπώς η Whirly1 αποτελεί μέχρι στιγμής τη μοναδική φυτικής προέλευσης πρωτεΐνη που τοποθετείται στον πυρήνα και τους χλωροπλάστες του ίδιου κυττάρου. Η πρωτεΐνη Whirly1 του *Zea mays* έχει παρατηρηθεί ότι συνδέεται *in vitro* με πλαστιδιακό DNA (Prikyl *et al.*, 2008). Παρόλα αυτά, τα αποτελέσματα της ανάλυσης του μεταλλάγματος *why1* θέτουν υπό αμφισβήτηση τη συμμετοχή της πρωτεΐνης στην αντιγραφή του DNA ή τη επιδιόρθωσή του, αφού ουσιαστικά δεν μεταβάλλονται τα επίπεδα του πλαστιδιακού DNA (Prikyl *et al.*, 2008).

Όμως στα μεταλλάγματα *why1* παρατηρήθηκε ασυνήθιστα υψηλή συγκέντρωση ριβοσωμάτων, η οποία οδήγησε σε φαινότυπο λεύκανσης (Prikyl *et al.*, 2008). Το *Arabidopsis thaliana* έχει 2 πρωτεΐνες Whirly (Whirly1 και Whirly3), που διαφέρουν ως προς την τοποθέτησή τους στο χλωροπλάστη. Η διαγονιδιακή κατασκευή *Whirly1-GFP* αποκάλυψε τη τοποθέτηση της πρωτεΐνης στον χλωροπλάστη σε δομές που εμφανίζουν στίγματα, ενώ η κατασκευή *Whirly3-GFP* εμφανίζει στον χλωροπλάστη διάχυτο πρότυπο (Krause *et al.*, 2005). Με βάση αυτά τα δεδομένα φαίνεται ότι τα ομόλογα γονίδια *Whirly* που τοποθετούνται στα πλαστίδια έχουν διαφορετικές μεταξύ τους λειτουργίες. Στο καλαμπόκι η μεγαλύτερη ποσότητα της πλαστιδιακής Whirly ανιχνεύθηκε ως διαλυτή πρωτεΐνη στο στρώμα των χλωροπλάστων ενώ ένα μικρό μέρος εντοπίστηκε σε μεμβρανικό κλάσμα των θυλακοειδών (Prikyl *et al.*, 2008). Η υπόθεση που διαμορφώνεται είναι πως ότι η πρωτεΐνη τοποθετείται στα πλαστίδια και υπό συγκεκριμένες συνθήκες απελευθερώνεται για να τοποθετηθεί στον πυρήνα όπου δρα ως

μεταγραφικός παράγοντας. Βέβαια δεν μπορεί να αποκληθεί επιπρόσθετη λειτουργία της πρωτεΐνης στα πλαστίδια.

Μία άλλη πρωτεΐνη που θα μπορούσε να βρίσκεται στα πλαστίδια πριν πραγματοποιήσει το ρόλο της στην ενεργοποίηση της έκφρασης πυρηνικών γονιδίων, είναι η ATXR5 (*Arabidopsis trithorax related protein 5*) στο *Arabidopsis* (Raynaud *et al.*, 2006). Η ATXR5 ανήκει στην οικογένεια πρωτεϊνών που περιέχει το τμήμα SET και σχετίζεται με τον επιγενετικό έλεγχο της γονιδιακής έκφρασης (Kouzarides *et al.*, 2002; Lachner *et al.*, 2002). Η αποθήκευση της στα πλαστίδια και η μετέπειτα απελευθέρωσή της, χρήζει πειραματικής επιβεβαίωσης, για παράδειγμα μέσω μελέτης της λειτουργίας της όταν η πρωτεΐνη βρίσκεται είτε μόνο στον πυρήνα ή είτε στα πλαστίδια. Αυτό θα μπορούσε να επιτευχθεί μέσω λειτουργικής συμπληρωματικότητας σε γενετικό υπόβαθρο μεταλλαγμάτων στα οποία δεν υπάρχει η πρωτεΐνη, όπου μια ανασυνδυασμένη μορφή της πρωτεΐνης τοποθετείται αποκλειστικά στον πυρήνα ή στα πλαστίδια, αλλά όχι ταυτόχρονα και στα δύο υποκυτταρικά διαμερίσματα.

Strategy	Target organelle	Protein and organism
<b>I. Formation of multiple proteins</b>		
A. Multiple transcription starts	M, N	DNA ligase 1, <i>Arabidopsis</i>
B. Multiple translation starts	M, N	DNA ligase 1, <i>Arabidopsis</i>
C. Alternative splicing	P, N	cp31, <i>Arabidopsis</i>
<b>II. Post-translational modification of single protein</b>		
A. Farnesylation	P, N	AtIPT3, <i>Arabidopsis</i>
B. Protein-protein interaction	M, N	Apn1, <i>Saccharomyces</i>
	M, N	HMG-CoA synthase interacting with PPAR $\alpha$ , humans
<b>III. Organelle-nucleus shuttle</b>		
A. Release from organelle surface	P, N	pBrrp, <i>Arabidopsis</i>
	P, N	Tsip1, Tobacco
B. Membrane transporters	M, N	MDDX28?, human
C. Vesicle-mediated release	M, Cyt	Hsp60, mHsp70, Humans
D. Protein shuttle via membrane contact sites	M, N	?

**Εικόνα 26. Μηχανισμοί υποκυτταρικής τοποθέτησης μίας πρωτεΐνης στα οργανίδια και τον πυρήνα (Krause and Krupinska, 2009).** Με βάση μελετημένες πρωτεΐνες, παρουσιάζονται προτεινόμενοι μηχανισμοί τοποθέτησης στον πυρήνα (N) και στους χλωροπλάστες/ πλαστίδια (P), κυτταρόπλασμα (cyt) ή στα μιτοχόνδρια (M).

## 5. ΒΙΒΛΙΟΓΡΑΦΙΑ

**Boyle B and Brisson N.** 2001. Repression of the defense gene PR-10a by the single-stranded DNA binding protein SEBF. *Plant Cell* **13**, 2525–2537

**Carrie C, Kühn K, Murcha MW, Duncan O, Small ID, O'Toole N, Whelan J.** 2009. Approaches to defining dual-targeted proteins in Arabidopsis. *Plant J.* **57**, 1128-1139.

**Creissen G, Reynolds H, Xue Y, Mullineaux P.** 1995. Simultaneous targeting of pea glutathione reductase and of a bacterial fusion protein to chloroplasts and mitochondria in transgenic tobacco. *Plant Journal* **8**, 167–175

**Desveaux D, Després C, Joyeux A, Subramaniam R and Brisson N.** 2000. PBF-2 is a novel single-stranded DNA binding factor implicated in PR-10a gene activation in potato. *The Plant Cell* **12**, 1477–1489

**Desveaux D, Maréchal A, Brisson N.** 2005. Whirly transcription factors: defense gene regulation and beyond. *Trends in Plant Science* **10**, 95–102

**Galichet A , Hoyerová K, Kamínek M and Gruissem W.** 2008. Farnesylation directs AtIPT3 subcellular localization and modulates cytokinin biosynthesis in Arabidopsis. *Plant Physiology* **146**, 1155–1164

**Gindullis F and Meier I.** 1999. Matrix attachment region binding protein MFP1 is localized in discrete domains at the nuclear envelope. *Plant Cell* **11**, 1117–1128

**Grabowski E, Miao Y, Mulisch M and Krupinska K.** 2008. Single-stranded DNA-binding protein Whirly1 in barley leaves is located in plastids and the nucleus of the same cell. *Plant Physiology.* **147**, 1800–1804

**Ham BK, Park JM, Lee SB, Kim MJ, Lee IJ, Kim KJ, Kwon CS, and Paek KH.** 2006. Tobacco Tsp1, a DnaJ-type Zn finger protein, is recruited to and potentiates T si1-Mediated Transcriptional Activation. *The Plant Cell* **18(8)**: 2005–2020

**Jeong SY, Rose A and Meier I.** 2003. MFP1 is a thylakoid-associated, nucleoid binding protein with a coiled-coil structure. *Nucleic Acids Research* **31**, 5175–5185

**Karniely S and Pines O.** 2005. Single translation – dual destination: mechanisms of dual protein targeting in eukaryotes. *EMBO Reports* **6**, 420–425

**Kouzarides, T.** 2002. Histone methylation in transcriptional control. *Current Opinion in*

*Genetics & Development* **12**, 198–209

**Krause K, Kilbienski I, Mulisch M, Rödiger A, Schäfer A and Krupinska K.** 2005. DNA-binding proteins of the Whirly family in *Arabidopsis thaliana* are targeted to the organelles. *FEBS Letters* **579**, 3707–3712

**Kroeger T, Watkins K, Friso G, J. van Wijk K and Barkan A.** 2009. A plant-specific RNA-binding domain revealed through analysis of chloroplast group II intron splicing. *Proceedings of National Academy of Science USA* **106**:11, 4537-4542

**Kwon C, and Chung IK.** 2004. Interaction of an *Arabidopsis* RNA binding protein with plant single-stranded telomeric DNA modulates telomerase activity. *Journal of Biological Chemistry* **279**: 12812–12818

**Lachner M. and Jenuwein T.** 2002. The many faces of histone lysine methylation. *Current Opinion in Cell Biology* **14**, 286–298

**Li, B., Takahashi, D., Kawamura, Y., and Uemura, M.** (2012). Comparison of plasma membrane proteomic changes of *Arabidopsis* suspension-cultured cells (T87 Line) after cold and ABA treatment in association with freezing tolerance development. *Plant and Cell Physiology* **53**, 543-554.

**Mitra, S. K., Walters, B. T., Clouse, S. D., & Goshe, M. B.** (2009). An efficient organic solvent based extraction method for the proteomic analysis of *Arabidopsis* plasma membranes. *Journal of Proteome Research*, *8*(6), 2752-2767.

**Narsai, R., Law, S. R., Carrie, C., Xu, L., & Whelan, J.** (2011). In-depth temporal transcriptome profiling reveals a crucial developmental switch with roles for RNA processing and organelle metabolism that are essential for germination in *Arabidopsis*. *Plant physiology*, *157*(3), 1342-1362.

**Natsoulis G, Hilger F, Fink GR.** 1986. The HTS1 gene encodes both the cytoplasmic and mitochondrial histidine tRNA synthetases of *Saccharomyces cerevisiae*. *Cell* **46**, 235–243

**Ohta M, Sugita M, Sugiura M.** 1995. Three types of nuclear genes encoding chloroplast RNA-binding proteins (cp29, cp31 and cp33) are present in *Arabidopsis thaliana*: presence of cp31 in chloroplasts and its homologue in nuclei/cytoplasm. *Plant Mol. Biol.* **27**:529-539.

**Onate-Sanchez L.** 2008. DNA-free RNA isolation protocols for *Arabidopsis thaliana*, including seeds and siliques. *BMC Research Notes* **1**, 93

**Prikryl J, Watkins KP, Friso G, van Wijk KJ, Barkan A.** 2008. A member of the Whirly family is a multifunctional RNA- and DNA-binding protein that is essential for chloroplast biogenesis. *Nucleic Acids Res.* **36**:5152-5165.

**Raynaud C, Sozzani R, Glab N, Domenichini S, Perennes C, Cella R, Kondorosi E, Bergounioux C.** 2006. Two cell-cycle regulated SET-domain proteins interact with proliferating cell nuclear antigen (PCNA) in *Arabidopsis*. *Plant Journal* **47**, 395–407

**Silva-Filho MC.** 2003. One ticket for multiple destinations: dual targeting of proteins to distinct subcellular locations. *Current Opinion in Plant Biology* **6**, 589–595

**Small I, Wintz H, Akashi K, Mireau H.** 1998. Two birds with one stone: genes that encode products targeted to two or more compartments. *Plant Molecular Biology* **38**, 265–277

**Samaniego R, de la Torre C and Díaz de la Espina SM.** 2008. Characterization, expression and subcellular distribution of a novel MFP1 (matrix attachment regionbinding filament-like protein 1) in onion. *Protoplasma* **233**, 31–38

**Sato N, Ohshima K, Watanabe A, Ohta N, Nishiyama Y, Joyard J and Douce R.** 1998. Molecular Characterization of the PEND Protein, a Novel bZIP Protein Present in the Envelope Membrane That Is the Site of Nucleoid Replication in Developing Plastids. *The Plant Cell* **10**: 859-872

**Schwacke R, Fischer K, Ketelsen B, Krupinska K and Krause K.** 2007. Comparative survey of plastid and mitochondrial targeting properties of transcription factors in *Arabidopsis* and rice. *Molecular Genetics & Genomics* **277**, 631–646

**Sunderland PA, West CE, Waterworth WM, Bray CM.** 2006. An evolutionarily conserved translation initiation mechanism regulates nuclear or mitochondrial targeting of DNA ligase 1 in *Arabidopsis thaliana*. *Plant Journal* **47**, 356–367

**Takabayashi, A., Kadoya, R., Kuwano, M., Kurihara, K., Ito, H., Tanaka, R., & Tanaka, A.** (2013). Protein co-migration database (PCoM-DB) for *Arabidopsis* thylakoids and *Synechocystis* cells. *SpringerPlus* **2**, 148.

**Wagner R and Pfannschmidt T.** 2006. Eukaryotic transcription factors in plastids – bioinformatic assessment and implications for the evolution of gene expression machineries in plants. *Gene* **381**, 62–70

**Yoo HH, Kwon C, Lee MM, Chung IK.** 2007. Single-stranded DNA binding factor AtWHY1 modulates telomere length homeostasis in Arabidopsis. *Plant Journal* **49**, 442–451