

ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΤΜΗΜΑ ΒΙΟΧΗΜΕΙΑΣ & ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ



ΒΙΟΠΛΗΡΟΦΟΡΙΚΗ ΚΑΙ
ΕΞΕΛΙΚΤΙΚΗ ΑΝΑΛΥΣΗ ΤΩΝ
ΜΕΤΑΦΟΡΕΩΝ ΣΑΚΧΑΡΩΝ
ΤΗΣ MAJOR FACILITATOR
SUPERFAMILY

27/06/2018

ΔΙΑΚΟΓΕΩΡΓΙΟΥ ΑΛΕΞΑΝΔΡΑ

UNIVERSITY OF THESSALY
DEPARTMENT OF BIOCHEMISTRY & BIOTECHNOLOGY

BIOINFORMATIC &
EVOLUTIONARY ANALYSIS
OF THE MAJOR FACILITATOR
SUPERFAMILY SUGAR
TRANSPORTERS

27/06/2018

DIAKOGEORGIU ALEXANDRA

Η παρούσα διπλωματική εργασία εκπονήθηκε στο εργαστήριο Βιοπληροφορικής, του τμήματος Βιοχημείας και Βιοτεχνολογίας (TBB), της Σχολής Επιστημών Υγείας του Πανεπιστημίου Θεσσαλίας(ΠΘ).

Υπεύθυνος καθηγητής:

Αμούτζιας Γρηγόριος, Επίκουρος καθηγητής Βιοπληροφορικής στη Γενωμική, TBB, Π.Θ.

Τριμελής επιτροπή:

- Αμούτζιας Γρηγόριος, Επίκουρος Καθηγητής Βιοπληροφορικής στη Γενωμική, TBB, Π.Θ.
- Φλεμετάκης Εμμανουήλ, Αναπληρωτής Καθηγητής Γεωπονικής Βιοτεχνολογίας, Γεωπονικό Πανεπιστημίο Αθηνών
- Φριλίγγος Ευστάθιος, Καθηγητής Βιολογικής Χημείας, Τμήμα Ιατρικής, Πανεπιστήμιο Ιωαννίνων

ΕΥΧΑΡΙΣΤΙΕΣ

Αρχικά, θα ήθελα να ευχαριστήσω θερμά τον υπεύθυνο καθηγητή, κ. Γρηγόριο Αμούτζια για όσα με δίδαξε και με συμβούλεψε καθ' όλη τη διάρκεια της εκπόνησης της διπλωματικής μου εργασίας, καθώς και τους συναδέλφους και φίλους μου Ντουντούμη Χρύσα και Φλιάτουρα Βασιλική για τις συμβουλές και τη βοήθεια τους. Ιδιαίτέρως ένα μεγάλο ευχαριστώ στον Νικολαΐδη Μάριο, που στάθηκε δίπλα μου από την αρχή μέχρι και την ολοκλήρωση της πτυχιακής αυτής εργασίας τόσο σαν φίλος όσο και σαν δάσκαλος. Τέλος, θα ήθελα να ευχαριστήσω την αδερφή μου Διακογεωργίου Μαρία και τον σύντροφο της Τζαφάλια Νέστορα για την ψυχολογική τους στήριξη.

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΕΧΟΜΕΝΑ	5
ΠΕΡΙΕΧΟΜΕΝΑ ΕΙΚΟΝΩΝ	6
ΠΕΡΙΕΧΟΜΕΝΑ ΠΙΝΑΚΩΝ	7
ΠΕΡΙΛΗΨΗ	8
ABSTRACT	9
1. ΕΙΣΑΓΩΓΗ	10
1.1 Μεταφορά μικρών μορίων διαμέσου της κυτταρικής μεμβράνης	10
1.1.1 Παθητική Διάχυση	10
1.1.2 Διευκολυνόμενη διάχυση	11
1.1.2.1 Διευκολυνόμενη διάχυση με πρωτεΐνες-φορείς	11
1.1.2.2 Διευκολυνόμενη διάχυση με πρωτεΐνες-διαύλους	12
1.1.3 Ενεργός μεταφορά	12
1.1.3.1 Ενεργός μεταφορά υποβοηθούμενη από την υδρόλυση ATP	13
1.1.3.2 Ενεργός μεταφορά ωθούμενη από τη διαβάθμιση ιόντων	14
1.2. Μεταφορείς σακχάρων	15
1.2.1 Sodium-Glucose Symporters	16
1.2.2 SWEETs	17
1.2.3 Glucose transporters	17
1.2.4 Δομή μεταφορέων σακχάρων	18
1.2.5 The Major Facilitator Superfamily	19
1.3 Οι φυτικοί μεταφορείς σακχάρων	22
2. ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ	24
2.1. Λήψη πρωτεωμάτων	24
2.2. Προγράμματα	25
2.2.1 Perl	25
2.2.2 Python	25
2.2.3 MEME Suite 4.12.0	25
2.2.4 HMMER 3.1	25
2.2.5 SeaView 4.4.2	26
2.2.6 TreeDyn & ETE Toolkit	26

2.2.8 USEARCH	26
2.2.9 TIMETREE (THE TIMESCALE OF LIFE)	26
3. ΑΠΟΤΕΛΕΣΜΑΤΑ-ΣΥΖΗΤΗΣΗ	27
3.1 Αναγνώριση μοτίβων της οικογένειας 2.A.1.1	27
3.2 Έλεγχος μονοφυλίας της οικογένειας 2.A.1.1	28
3.3 Δημιουργία HMMs για τον εντοπισμό των 2.A.1.1	30
3.4 Φυλογενετική ανάλυση των 2.A.1.1 από αντιπροσωπευτικά πρωτεώματα	30
4. ΣΥΜΠΕΡΑΣΜΑΤΑ	40
5. ΒΙΒΛΙΟΓΡΑΦΙΑ	41

ΠΕΡΙΕΧΟΜΕΝΑ ΕΙΚΟΝΩΝ

<i>Εικόνα 1. Διαπερατότητα των μεμβρανών.</i>	10
<i>Εικόνα 2. Μοντέλο του μηχανισμού της διευκολυνόμενης διάχυσης της γλυκόζης.</i>	11
<i>Εικόνα 3. Μοντέλο ενός διαύλου ιόντων.</i>	12
<i>Εικόνα 4. Μοντέλο λειτουργίας αντλίας Na^+-K^+.</i>	13
<i>Εικόνα 5. Μηχανισμός ενεργούς μεταφοράς γλυκόζης.</i>	14
<i>Εικόνα 6. Μηχανισμοί μεταφοράς στους ευκαρυώτες.</i>	15
<i>Εικόνα 7. Διανομή των μεταφορέων γλυκόζης στα ανθρώπινα επιθηλιακά κύτταρα του εντέρου.</i>	16
<i>Εικόνα 8. Δομή των μεταφορέων γλυκόζης.</i>	18
<i>Εικόνα 9. Δομές των τριών ειδών ευκαρυωτικών μεταφορέων σακχάρων.</i>	19
<i>Εικόνα 10. Πρότυπη MFS αναδίπλωση.</i>	21
<i>Εικόνα 11. Φυλογενετικό δένδρο της οικογένειας MST.</i>	23
<i>Εικόνα 12. Μοτίβα των 2.A.1.1 αντιπροσώπων</i>	27
<i>Εικόνα 13. Ενδεικτική παρουσίαση μερικών αποτελεσμάτων του MAST της 2.A.1.1 έναντι της βιβλιοθήκης της TCDB</i>	28
<i>Εικόνα 14. Φυλογενετικό δένδρο των οικογενειών 2.A.1.1, 2.A.1.9 και 2.A.1.19</i>	29
<i>Εικόνα 15. Οι 7 υποοικογένειες της οικογένειας 2.A.1.1 στα φυτά.</i>	31
<i>Εικόνα 16. Φυλογενετικό δέντρο της Subfamily 1 (TMT).</i>	32
<i>Εικόνα 17. Φυλογενετικό δέντρο της Subfamily2 (INT).</i>	33
<i>Εικόνα 18. Φυλογενετικό δέντρο της Subfamily3 (ESL).</i>	34

<i>Εικόνα 19. Φυλογενετικό δέντρο της Subfamily4 (VGT).</i>	35
<i>Εικόνα 20. Φυλογενετικό δέντρο της Subfamily5 (pGLcT/SGB1).</i>	36
<i>Εικόνα 21. Φυλογενετικό δέντρο της Subfamily6 (PMT).</i>	37
<i>Εικόνα 22. Φυλογενετικό δέντρο της Subfamily7 (STP).</i>	38
<i>Εικόνα 23. Ο αριθμός ομολόγων των 7 υποοικογενειών σε 12 αντιπροσωπευτικά πρωτεώματα φυτικών οργανισμών.</i>	39

ΠΕΡΙΕΧΟΜΕΝΑ ΠΙΝΑΚΩΝ

<i>Πίνακας 1. Οι δέκα οικογένειες που περιλαμβάνει η MFS</i>	20
<i>Πίνακας 2. Τα πρωτεώματα που χρησιμοποιήθηκαν στις αναλύσεις μας.</i>	24
<i>Πίνακας 3. Οι 7 φυτικές υποοικογένειες της 2.A.1.1 και ο αριθμός των ομολόγων που εντοπίστηκαν</i>	32

ΠΕΡΙΛΗΨΗ

Τα σάκχαρα αποτελούν τις κύριες πηγές ενέργειας όλων των οργανισμών και για αυτό οι μεταφορείς τους είναι από τις σημαντικότερες πρωτεΐνες-φορείς στα κύτταρα. Στους ευκαρυώτες συναντάμε τρεις μεγάλες κατηγορίες μεταφορέων σακχάρων, τους μεταφορείς γλυκόζης της υπεροικογένειας MFS, τους συµμεταφορείς νατρίου-γλυκόζης και τους SWEETs. Σκοπός αυτής της διπλωµατικής εργασίας ήταν η εξελικτική και φυλογενετική ανάλυση της οικογένειας φυτικών μεταφορέων γλυκόζης της MFS. Αρχικά αναπτύξαµε µια γενική βιοπληροφορική µέθοδο εντοπισµού τους σε όλους τους ευκαρυώτες, βασισµένου σε Hidden Markov Models και ικανού να τις διαχωρίσει από άλλες συγγενικές οικογένειες μεταφορέων. Στη συνέχεια, αναλύσαµε 12 αντιπροσωπευτικά πρωτεώµατα φυτικών οργανισµών από χλωρόφυτα, βρυόφυτα, γυµνόσπερµα και αγγειόσπερµα για τον εντοπισµό και την φυλογενετική ανάλυσή αυτών των μεταφορέων. Η εξελικτική ανάλυση έδειξε ότι χωρίζονται σε 7 υποοικογένειες, όπου οι δύο εμφανίστηκαν τουλάχιστον πριν από 950 εκατοµύρια χρόνια, ενώ οι υπόλοιπες 5 τουλάχιστον πριν από 500 εκατοµύρια χρόνια. Παραδόξως, οι πιο αρχέγονες υποοικογένειες (VGT & pGLcT/SGB1) δεν είχαν τον µεγαλύτερο αριθµό οµολόγων αλλά αντιθέτως αρκετά µικρό, γεγονός που υποδηλώνει κάποιου είδους εξελικτική πίεση που ευνοεί την παρουσία ενός ή λίγων µόνο οµολόγων για αυτές.

ABSTRACT

Sugars are the main sources of energy for all organisms, therefore their transporters are very important protein-carriers in the cell. In eukaryotes, there are three major classes of sugar transporters, the glucose transporters of the MFS superfamily, the sodium-glucose symporters and the SWEETs. The purpose of this work was the evolutionary and phylogenetic analysis of plant sugar transporters, which belong to the MFS superfamily. At first, we developed a generic bioinformatics method, based on Hidden Markov Models, that may detect these transporters in all eukaryotes and distinguish them from other closely related transporters. Next, we used this methodology to detect and analyze the MFS sugar transporters of 12 representative plant proteomes, from chlorophyta, bryophyta, gymnosperms and angiosperms. The evolutionary analysis divided the detected transporters into 7 subfamilies. Two of these appeared at least 950 million years ago, whereas the other 5 appeared at least 500 million years ago. Paradoxically, the oldest two subfamilies (VGT & pGLcT/SGB1) did not contain the largest number of homologs. This finding suggests a certain type of evolutionary pressure that favours the existence of one or only a few homologs per proteome, for these two ancient subfamilies.

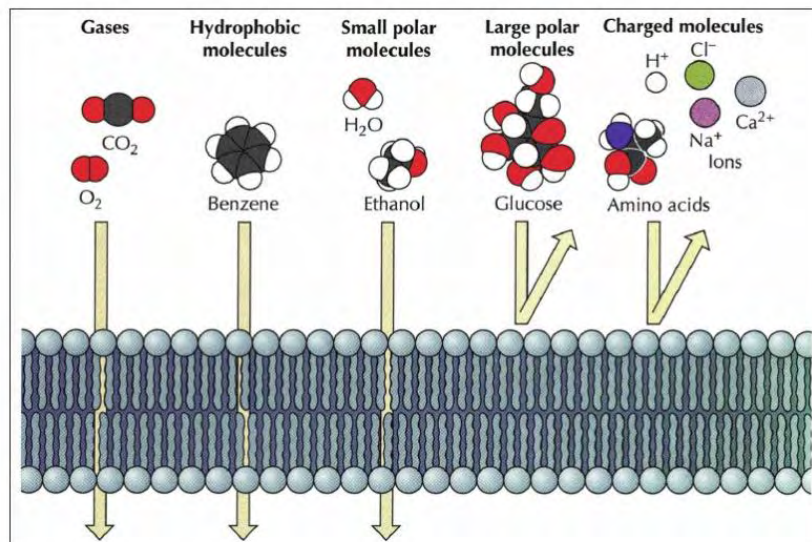
1. ΕΙΣΑΓΩΓΗ

1.1 Μεταφορά μικρών μορίων διαμέσου της κυτταρικής μεμβράνης

Αν και μερικά μόρια είναι ικανά να διαπεράσουν την κυτταρική μεμβράνη με παθητική διάχυση (passive diffusion), στην πλειοψηφία τους δεν μπορούν και η μεταφορά τους γίνεται με δευκολυνόμενη διάχυση (facilitated diffusion) (Cooper and Hausman 2011).

1.1.1 Παθητική Διάχυση

Η παθητική διάχυση είναι μη επιλεκτική διαδικασία και αποτελεί τον πιο απλό τρόπο μετάβασης των μορίων από την μία μεριά της κυτταροπλασματικής μεμβράνης στην άλλη (Cooper and Hausman 2011). Το μόριο εισέρχεται μέσα στη διπλοστιβάδα φωσφολιπιδίων και φθάνει στη απέναντι πλευρά της μέσω διάχυσης, όπου και απελευθερώνεται στο υδατικό διάλυμα που υπάρχει εκεί. Η κατεύθυνση της μεταφοράς του μορίου εξαρτάται από τη σχετική συγκέντρωση του στις δυο πλευρές της μεμβράνης και η μεταφορά του πραγματοποιείται πάντοτε σύμφωνα με τη διαβάθμιση αυτής. Μόνο μικρά και σχετικά υδρόφοβα μόρια έχουν τη δυνατότητα παθητικής διάχυσης με σημαντική ταχύτητα, όπως αέρια, υδρόφοβα μόρια και μικρά πολικά μόρια χωρίς ηλεκτρικό φορτίο, ενώ μεγαλύτερα πολικά μη φορτισμένα μόρια, όπως η γλυκόζη, ή φορτισμένα μόρια οποιουδήποτε μεγέθους είναι αδύνατον να πραγματοποιήσουν παθητική διάχυση, οπότε μεταφέρονται μέσω διευκολυνόμενης διάχυσης (Εικόνα 1) (Chien 2013).



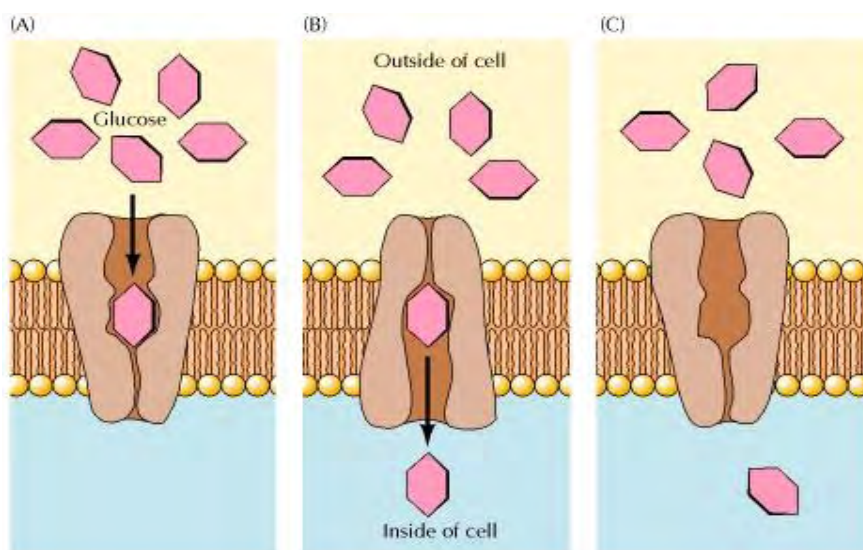
Εικόνα 1. Διαπερατότητα των μεμβρανών (Chien 2013).

1.1.2 Διευκολυνόμενη διάχυση

Η διευκολυνόμενη διάχυση πραγματοποιείται, επίσης, σύμφωνα με την συγκέντρωση των μορίων στις δύο πλευρές της μεμβράνης και χωρίς παροχή εξωτερικής πηγής ενέργειας (Cooper and Hausman 2011). Τα μόρια μεταφέρονται μέσω πρωτεϊνών, ώστε να μην έρχονται σε επαφή με τον υδρόφοβο πυρήνα της φωσφολιπιδικής διπλοστιβάδας. Οι πρωτεΐνες της διευκολυνόμενης διάχυσης χωρίζονται σε πρωτεΐνες-φορείς και πρωτεΐνες-διαύλους. Οι πρωτεΐνες-φορείς (carrier proteins) δεσμεύουν το μόριο προς μεταφορά από τη μια μεριά της πλασματικής μεμβράνης και αλλάζοντας τη διαμόρφωσή τους, μεταφέρουν το μόριο στην άλλη μεριά. Αντίθετα, οι πρωτεΐνες-διαύλοι (channel proteins) αποτελούν ανοικτούς πόρους και επιτρέπουν την ελεύθερη διέοδο μορίων ανάλογα μόνο με το μέγεθος και το φορτίο τους.

1.1.2.1 Διευκολυνόμενη διάχυση με πρωτεΐνες-φορείς

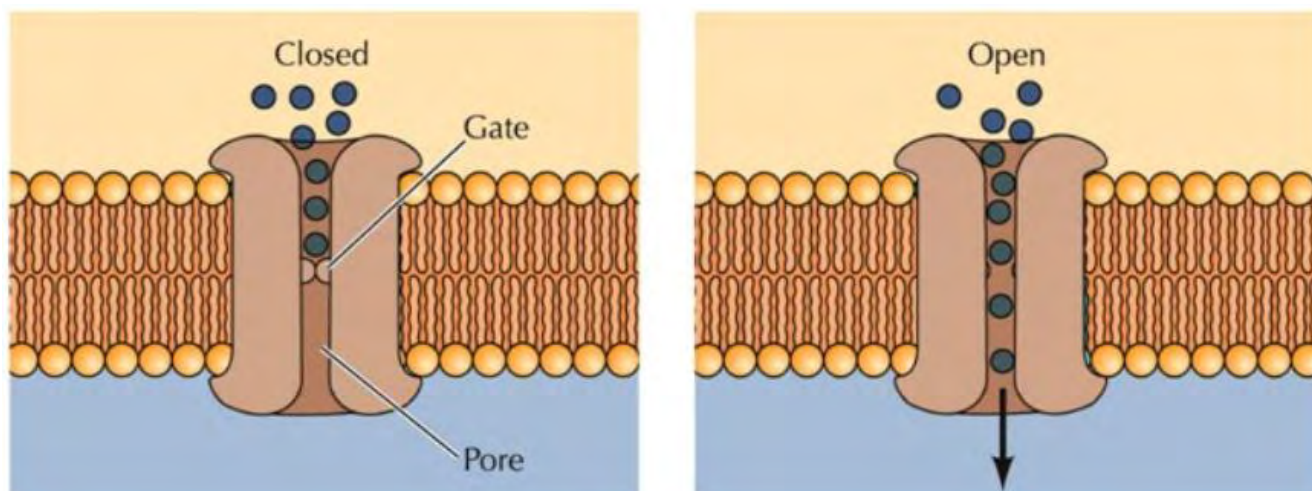
Χαρακτηριστικό παράδειγμα δράσης των πρωτεϊνών-φορέων αποτελεί το μοντέλο του μηχανισμού διευκολυνόμενης διάχυσης της γλυκόζης, (Cooper and Hausman 2011) που φαίνεται και στην Εικόνα 2 (Cooper 2000). Σύμφωνα με αυτό, ο μεταφορέας της γλυκόζης απαντάται σε δύο διαμορφώσεις και, ανάλογα με την διαμόρφωση που θα λάβει, το κέντρο δέσμευσης της θα εκτίθεται είτε στο εσωτερικό είτε στο εξωτερικό του κυττάρου. Στην πρώτη διαμόρφωση (A), η γλυκόζη εισέρχεται στο κέντρο δέσμευσης, στη συνέχεια ο μεταφορέας αλλάζει διαμόρφωση (B) και απελευθερώνει την γλυκόζη στο εσωτερικό του κυττάρου. Τέλος, ο μεταφορέας θα επιστρέψει στην αρχική του διαμόρφωση (C).



Εικόνα 2. Μοντέλο του μηχανισμού της διευκολυνόμενης διάχυσης της γλυκόζης (Cooper 2000).

1.1.2.2 Διευκολυνόμενη διάχυση με πρωτεΐνες-διαύλους

Οι καλύτερα μελετημένοι είναι οι διαύλοι ιόντων, των οποίων ο ρόλος είναι η μεταφορά ιόντων διαμέσου των μεμβρανών (Cooper and Hausman 2011). Οι διαύλοι αυτοί απαντώνται σε όλα τα κύτταρα, η μεταφορά μέσω αυτών είναι ιδιαίτερος επιλεκτική, πραγματοποιείται με πολύ μεγάλη ταχύτητα και δε βρίσκονται μόνιμα σε ανοιχτή διαμόρφωση. Για να μεταβούν σε ανοιχτή διαμόρφωση, απαιτούν κάποιο σήμα, όπως την πρόσδεση σηματοδοτικού μορίου (διαύλοι ελεγχόμενοι από προσδέτη ή προσδετοελεγχόμενοι) είτε την αλλαγή του ηλεκτρικού δυναμικού της κυτταροπλάσματικής μεμβράνης (διαύλοι ελεγχόμενοι από διαφορά δυναμικού ή τασεοελεγχόμενοι). Στην Εικόνα 3 (<http://slideplayer.com/slide/8411421/>) παρουσιάζεται το μοντέλο ενός διαύλου ιόντων, κατά το οποίο όταν ο διάυλος βρίσκεται στην κλειστή διαμόρφωση, η διέλευση ιόντων παρεμποδίζεται από μία πύλη, ενώ στην ανοιχτή διαμόρφωση, κατά την οποία η πύλη αυτή είναι ανοικτή, πραγματοποιείται ταχύτατη μεταφορά ιόντων μέσω του πόρου. Ο πόρος όμως αυτός που σχηματίζει ο διάυλος επιτρέπει τη διέλευση μόνο συγκεκριμένων ιόντων με κατάλληλο μέγεθος και φορτίο.



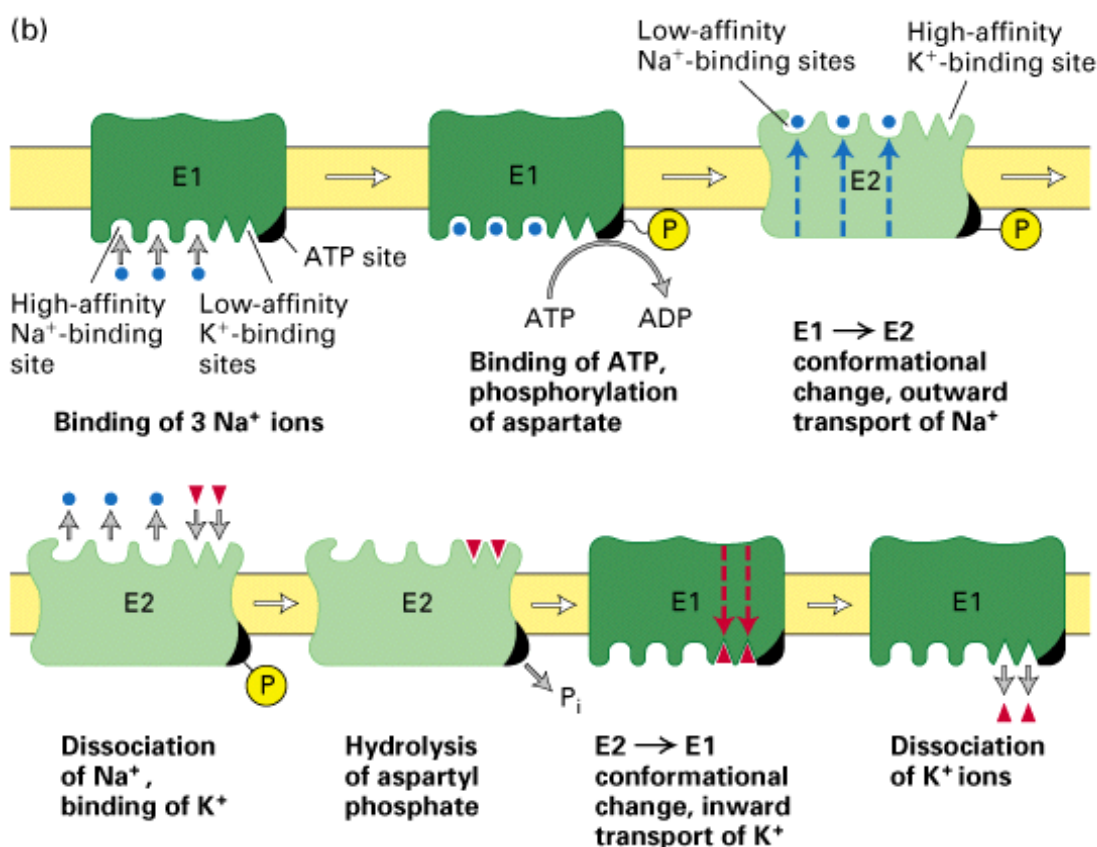
Εικόνα 3. Μοντέλο ενός διαύλου ιόντων (<http://slideplayer.com/slide/8411421/>).

1.1.3 Ενεργός μεταφορά

Στους δύο προηγούμενους τύπους μεταφοράς, παθητική και διευκολυνόμενη διάχυση, η μεταφορά των μορίων γίνεται πάντοτε με κατεύθυνση που καθορίζεται από τις ηλεκτροχημικές διαβαθμίσεις μεταξύ των δύο πλευρών της μεμβράνης (Cooper and Hausman 2011). Ωστόσο, είναι απαραίτητο κάποιες φορές το κύτταρο να μεταφέρει μόρια αντίθετα από τη διαβάθμιση συγκέντρωσης τους. Για να το πετύχει αυτό εκτελεί ενεργό μεταφορά (active transport). Η ενεργός μεταφορά απαιτεί την ύπαρξη μίας συζευγμένης αντίδρασης παραγωγής ενέργειας, με την οποία θα πραγματοποιηθεί η μη ευνοούμενη ενεργειακά μεταφορά.

1.1.3.1 Ενεργός μεταφορά υποβοηθούμενη από την υδρόλυση ATP

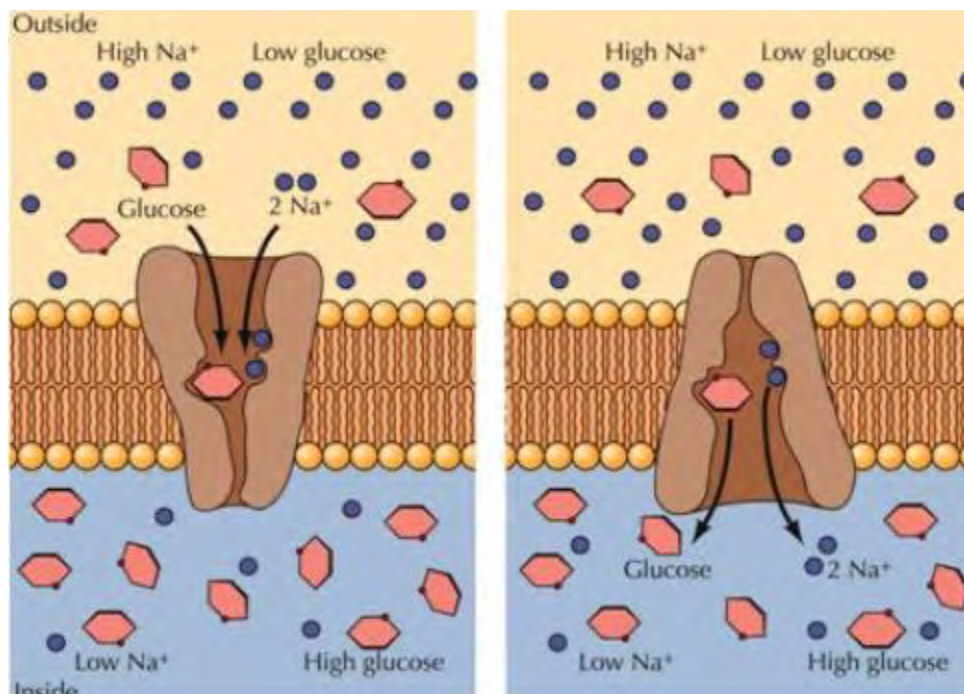
Η ενεργός μεταφορά που ωθείται άμεσα από την υδρόλυση του ATP, πραγματοποιείται χαρακτηριστικά από τις αντλίες ιόντων (ion pumps) (Cooper and Hausman 2011). Παράδειγμα τέτοιας αντλίας είναι η ATPάση $\text{Na}^+\text{-K}^+$, η οποία χρησιμοποιεί την ενέργεια που παράγεται από την υδρόλυση ATP για να μεταφέρει ιόντα Na^+ και K^+ αντίθετα από τις ηλεκτροχημικές τους διαβαθμίσεις. Το μοντέλο λειτουργίας της παρουσιάζεται στην Εικόνα 4 (<http://www.zoology.ubc.ca/~gardner/ATPase%20pumps.htm>). Άλλα παραδείγματα είναι η αντλία Ca^+ , που χρησιμοποιεί ATP για τη μεταφορά ιόντων ασβεστίου από το κυτταρόπλασμα στο εξωτερικό του κυττάρου ή στον αυλό του ενδοπλασματικού δικτύου, και οι μεταφορείς τύπου ABC, που χρησιμοποιούν την ενέργεια υδρόλυσης ATP για μεταφορά προς μία κατεύθυνση είτε για εισροή είτε για εκροή μορίων.



Εικόνα 4. Μοντέλο λειτουργίας αντλίας $\text{Na}^+\text{-K}^+$ (<http://www.zoology.ubc.ca/~gardner/ATPase%20pumps.htm>). Αρχικά, τρία μόρια Na^+ δεσμεύονται σε θέσεις που είναι εκτεθειμένες στο εσωτερικό του κυττάρου και επάγεται φωσφορυλίωση της αντλίας εξαρτούμενη από ATP. Η φωσφορυλίωση εκθέτει τις θέσεις δέσμευσης Na^+ στην επιφάνεια του κυττάρου και μειώνει την συγγένεια δέσμευσης, με αποτέλεσμα τα ιόντα Na^+ να απελευθερωθούν από το κύτταρο. Συγχρόνως δύο ιόντα K^+ δεσμεύονται σε θέσεις υψηλής συγγένειας, οι οποίες είναι εκτεθειμένες στην επιφάνεια του κυττάρου. Η δέσμευσή τους προκαλεί αποφωσφορυλίωση της αντλίας, η οποία επιστρέφει στην αρχική της διαμόρφωση και απελευθερώνει τα ιόντα K^+ στο εσωτερικό του κυττάρου.

1.1.3.2 Ενεργός μεταφορά ωθούμενη από τη διαβάθμιση ιόντων

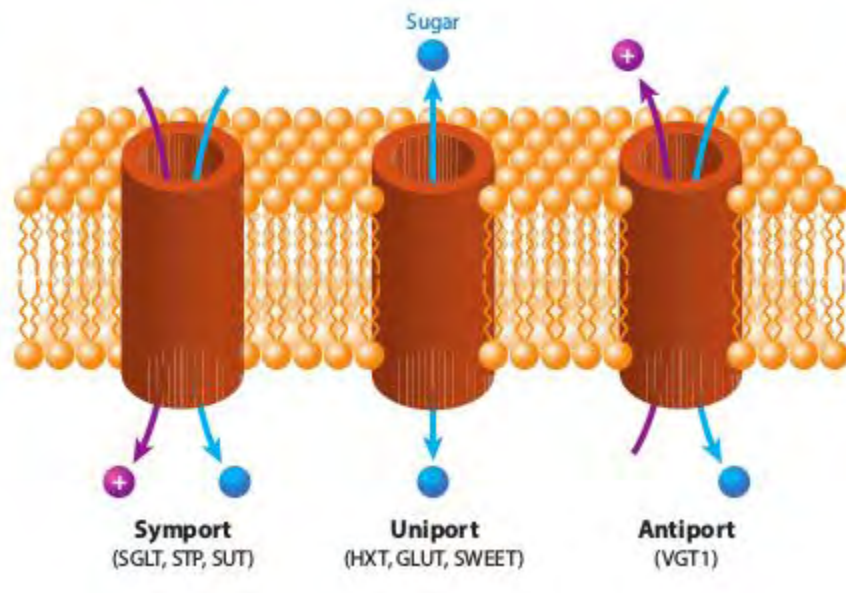
Σε αντίθεση με τις αντλίες ιόντων και τους μεταφορείς τύπου ABC, άλλοι μεταφορείς για να μεταφέρουν μόρια προς τη μη ευνοϊκή ενεργειακά κατεύθυνση χρησιμοποιούν συζευγμένη μεταφορά αυτών με ένα δεύτερο μόριο, το οποίο μεταφέρεται ευνοϊκά προς εκείνη την κατεύθυνση (Cooper and Hausman 2011). Για παράδειγμα, η πρόσληψη γλυκόζης από τον εντερικό αυλό στα επιθηλιακά κύτταρο του εντέρου πραγματοποιείται από έναν συµμεταφορέα που μεταφέρει συζευγμένα ένα μόριο γλυκόζης με δύο ιόντα Na^+ . Ο μηχανισμός της ενεργούς αυτής μεταφοράς παρουσιάζεται στην Εικόνα 5 (<http://slideplayer.com/slide/8411421/>).



Εικόνα 5. Μηχανισμός ενεργούς μεταφοράς γλυκόζης (<http://slideplayer.com/slide/8411421/>). Η μεταφορά της γλυκόζης από τον εντερικό αυλό στο εσωτερικό των επιθηλιακών κυττάρων ωθείται από την διαβάθμιση των ιόντων Na^+ . Συγκεκριμένα, η ενεργειακά ευνοούμενη μεταφορά Na^+ παρέχει την ενέργεια για την πρόσληψη της γλυκόζης αντίθετα από τη διαβάθμιση της συγκέντρωσης της. Ο μεταφορέας δεσμεύει και μεταφέρει ταυτόχρονα ένα μόριο γλυκόζης με δύο μόρια Na^+ στο εσωτερικό του κυττάρου.

Υπάρχουν τρεις κατηγορίες μεταφορέων της μεμβράνης: α) οι συµμεταφορείς (symporters), που μεταφέρουν δύο μόρια προς την ίδια κατεύθυνση, β) οι μονομεταφορείς (uniporters), που μεταφέρουν μόνο ένα μόριο είτε εσωτερικά είτε εξωτερικά της μεμβράνης και γ) οι αντιμεταφορείς (antiporters), οι οποίοι μεταφέρουν δύο μόρια προς αντίθετες κατευθύνσεις (Cooper and Hausman 2011). Στην Εικόνα 6 (Chen *et al.* 2015) παρουσιάζονται οι τρεις παραπάνω κατηγορίες μεταφορέων,

ενώ στις παρενθέσεις παρατίθενται μερικά παραδείγματα πρωτεϊνών-φορέων που ανήκουν σε κάθε κατηγορία.

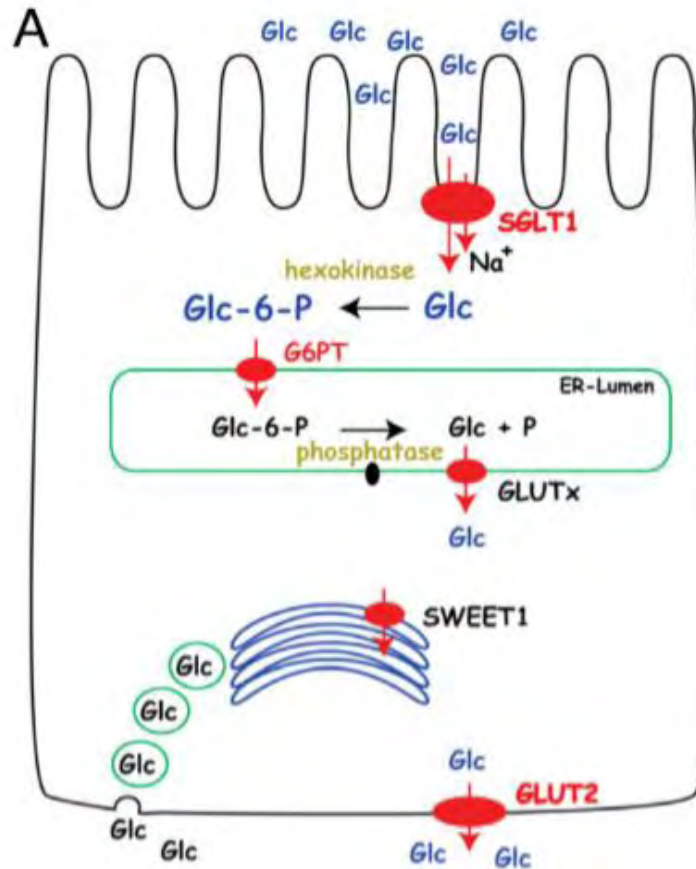


Εικόνα 6. Μηχανισμοί μεταφοράς στους ευκαρυώτες (Chen *et al.* 2015).

1.2. Μεταφορείς σακχάρων

Τα σάκχαρα αποτελούν τις σημαντικότερες οργανικές ενώσεις που λαμβάνονται από τους ζωντανούς οργανισμούς, και προέρχονται από φωτοσυνθετικούς οργανισμούς και κυρίως φυτά (Chen *et al.* 2015). Σε αυτά, μεταφέρονται από τα φύλλα στους υπόλοιπους φυτικούς ιστούς, και αφού αποπολυμεριστούν, εισέρχονται κυρίως στα κύτταρα των ριζών, των σπόρων και των φρούτων. Στους πολυκύτταρους οργανισμούς τα διαλυτά σάκχαρα χρησιμοποιούνται ως πηγές ανθρακικών σκελετών για τη βιοσύνθεση πολλών άλλων κυτταρικών ενώσεων, οσμολύτες, μόρια-σινιάλα, παροδικές αποθήκες ενέργειας και ως μόρια-μεταφορείς. Η μεταφορά τους διακυτταρικά πραγματοποιείται από πρωτεΐνες μεταφορείς, οι οποίες είναι υπεύθυνες για την πρόσληψη ή την απέκκριση τους από τα κύτταρα ή τα υποκυτταρικά διαμερίσματα. Μέχρι σήμερα έχουν χαρακτηριστεί τρεις κατηγορίες ευκαρυωτικών μεταφορέων σακχάρων, οι μεταφορείς γλυκόζης (GLUTs), οι συμμεταφορείς νατρίου-γλυκόζης (SGLTs) και οι SWEETs. Στα θηλαστικά, οι GLUTs διανέμονται σε όλα τα κύτταρα και καταλύουν την διευκολυνόμενη διάχυση της γλυκόζης σύμφωνα με την βαθμίδα συγκέντρωσης της (Deng and Yan 2016). Από την άλλη, οι SGLTs κυρίως εκφράζονται στο φλοιό του εντέρου και του νεφρού και καταλύουν τη μεταφορά της γλυκόζης αντίθετα από τη συγκέντρωσή της, εκμεταλλευόμενοι τα μόρια νατρίου (Yan 2015; Wright, Loo, and Hirayama 2011). Τέλος, οι SWEETs

είναι υπεύθυνοι για την εκροή και την ενδοκυτταρική διακίνηση των σακχάρων στα φυτά, ενώ μένει να αποδειχθεί το ίδιο και στα θηλαστικά (Chen *et al.* 2010). Έτσι, οι τρεις αυτές οικογένειες μεταφορέων



Εικόνα 7 Διανομή των μεταφορέων γλυκόζης στα ανθρώπινα επιθηλιακά κύτταρα του εντέρου (Chen *et al.* 2010).

γλυκόζης έχουν διακριτές φυσιολογικές λειτουργίες και μηχανισμούς, οι οποίοι παρουσιάζονται στην Εικόνα 7 (Chen *et al.* 2010) σε ένα απλοποιημένο μοντέλο.

1.2.1 Sodium-Glucose Symporters

Οι SGLTs είναι οι καλύτερα μελετημένοι μεταφορείς (César-Razquin *et al.* 2015; Thorens and Mueckler 2010) και ανήκουν στην οικογένεια SSF (sodium-solute family), η οποία είναι υπεύθυνη για συζευγμένες διεργασίες μεταφοράς, από τη διοχέτευση νερού και ουρίας έως την ανίχνευση της γλυκόζης στους χολινεργικούς νευρώνες (Wright, Loo, and Hirayama 2011). Συγκεκριμένα, κάθε μέλος αυτής της οικογένειας είναι ειδικό για συγκεκριμένο υπόστρωμα, το οποίο μπορεί να είναι αμινοξύ, νευροδιαβιβαστής ή ιώδιο. Φυλογενετικές αναλύσεις υποδεικνύουν ότι οι SGLTs ανήκουν στην υπεροικογένεια APC (amino acid/polyamine/organocation superfamily) (Höglund *et al.* 2011). Ο

SGLT1 είναι συµµεταφορέας νατρίου-γλυκόζης, µεταφέροντας ένα µόριο γλυκόζης µε δύο µόρια νατρίου, ενώ ο SGLT2 συµµεταφέρει µαζί µε το µόριο γλυκόζης, ένα µόνο µόριο νατρίου (Wright, Loo, and Hirayama 2011).

1.2.2 SWEETs

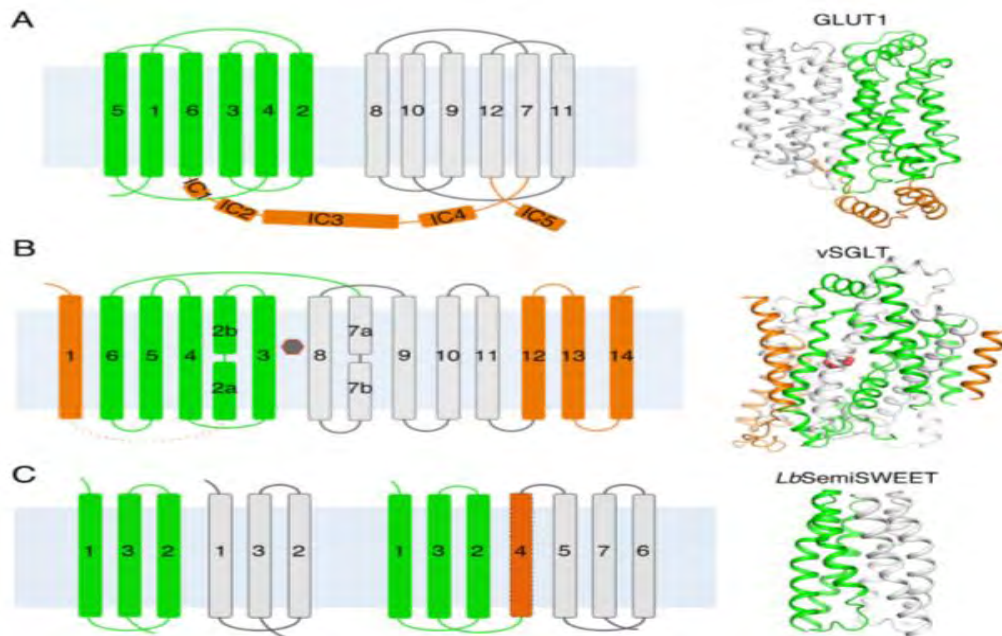
Οι SWEETs (César-Razquin *et al.* 2015; Thorens and Mueckler 2010) αποτελούν την πιο πρόσφατα ανακαλυφθείσα οικογένεια µεταφορέων σακχάρων, η οποία προέρχεται από τα αρχαιοβακτήρια και πέρασε στα φυτά και στον άνθρωπο (Frommer, Sosso, and Chen 2013). Δεν έχει αποδειχθεί ακόµα ότι ανήκουν στους µονοµεταφορείς (Chen *et al.* 2015). Χαρακτηρίζονται για την ικανότητα τους να µεταφέρουν ποικίλους µονοσακχαρίτες και δισακχαρίτες, για τη µεσολάβησή τους τόσο στην εισροή όσο και στην εκροή και για τις µικρές συγγένειες τους για τα σάκχαρα. Η χαµηλή αυτή συγγένεια δείχνει ότι οι µεταφορείς αυτοί ειδικεύονται σε υψηλές ταχύτητες ανακύκλωσης, παρά σε αποτελεσµατική µετακίνηση σε χαµηλές συγκεντρώσεις υποστρώµατος.

1.2.3 Glucose transporters

Οι GLUTs ανήκουν στην υπεροικογένεια MFS (Major facilitator superfamily) (César-Razquin *et al.* 2015; Thorens and Mueckler 2010). Ταξινοµούνται σε τρεις κλάσεις (Mike Mueckler and Thorens 2013), την Κλάση 1 στην οποία ανήκουν οι GLUT1-4 και ο GLUT14, στην Κλάση 2, στην οποία ανήκουν οι GLUT5,7,9 και 11 και στην Κλάση 3, στην ανήκουν οι GLUT6, 8, 10, 12 και 13, µεταφορείς που οι περισσότεροι ακόµα δεν έχουν χαρακτηριστεί. Ο GLUT1, ο οποίος ανήκει στους µονοµεταφορείς, ήταν ένας από τους πρώτους µεµβρανικούς µεταφορείς που κλωνοποιήθηκαν και χαρακτηρίστηκαν. Αυτός ο µεταφορέας γλυκόζης βρίσκεται σε όλους τους ιστούς, αλλά λειτουργεί κυρίως στα ερυθροκύτταρα και στον αιµατεγκεφαλικό φραγµό (M. Mueckler *et al.* 1985; Mike Mueckler and Thorens 2013). Ο GLUT2 είναι ο κύριος µεταφορέας γλυκόζης στα ηπατοκύτταρα και στο έντερο και είναι ο µόνος GLUT που καταλύει τη ροή γλυκόζης µε δύο κατευθύνσεις ανάλογα µε την κατάσταση τροφοδοσίας ή νηστείας. Ο GLUT3 εκφράζεται κυρίως στους νευρώνες και ονοµάζεται «νευρωνικός µεταφορέας γλυκόζης» (Nagamatsu *et al.* 1992). Ο GLUT4 εκφράζεται σε µεγάλο βαθµό στους σκελετικούς µύες και στον λιπώδη ιστό. Μετά από τη σηµατοδότηση ινσουλίνης, µετεγκαθίσταται από τα ενδοκυτταρικά κυστιδία στη πλασµατική µεµβράνη, για αυτό και ονοµάζεται «µεταφορέας γλυκόζης που ανταποκρίνεται στην ινσουλίνη» (Govers 2014; Huang and Czech 2007). Ο GLUT5 είναι κυρίως υπεύθυνος για την απορρόφηση της φρουκτόζης στα επιθηλιακά κύτταρα του εντέρου, αλλά εκφράζεται και σε άλλα όργανα, όπως το ήπαρ και ο εγκέφαλος (Mike Mueckler and Thorens 2013). Οι GLUT7 και GLUT11 µεταφέρουν και γλυκόζη και φρουκτόζη. Ο GLUT9 είναι µεταφορέας ουρικού οξέος.

1.2.4 Δομή μεταφορέων σακχάρων

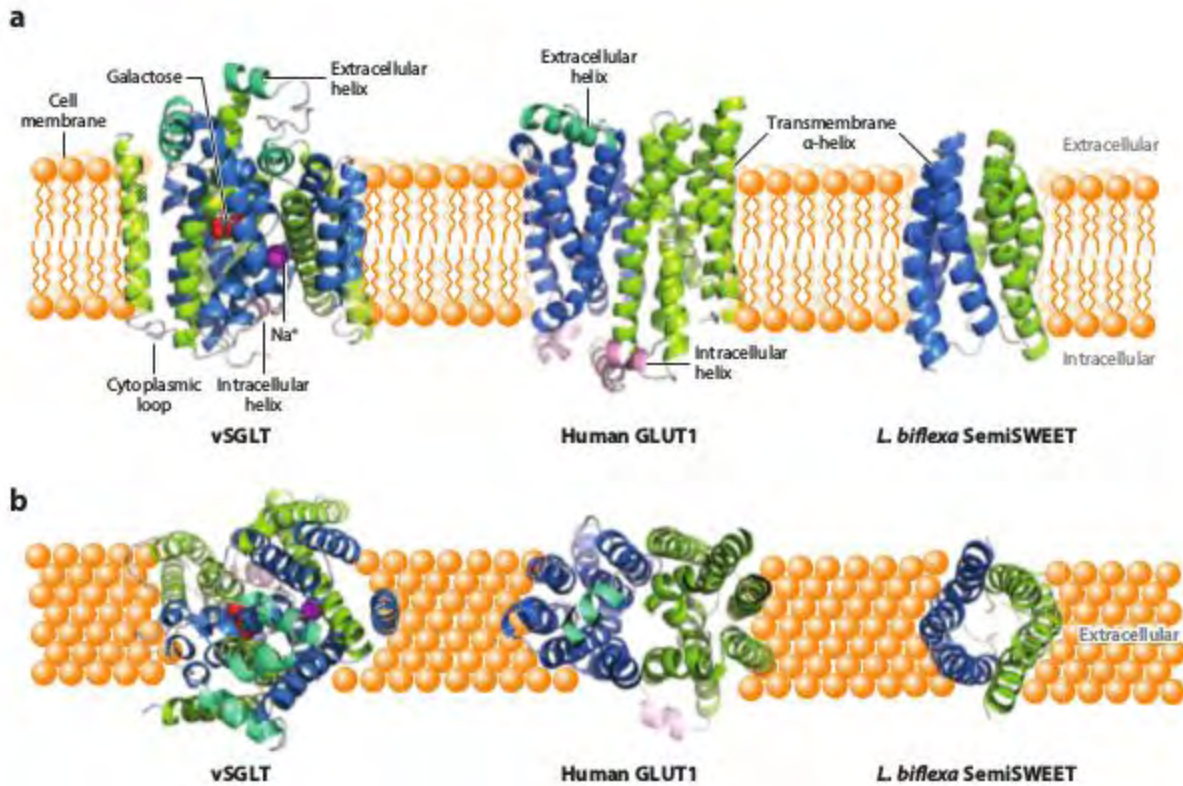
Ανάλογα με την υπερικογένεια που ανήκουν οι μεταφορείς, θα έχουν και την αντίστοιχη τυπική δομή της υπερικογένειας. Άρα, οι GLUTs έχουν την τυπική δομή της MFS (Shi 2013) και οι SGLTs της APC (Shi 2013). Οι SWEETs στους ευκαρυωτικούς οργανισμούς διαθέτουν 7 διαμεμβρανικές περιοχές ή TMs (transmembrane segments), ενώ τα βακτηριακά ομόλογα τους διαθέτουν μόλις 3 TMs και ονομάζονται SemiSWEETs (Deng and Yan 2016). Είναι οι μικρότεροι μεταφορείς που γνωρίζουμε μέχρι σήμερα και λειτουργούν ως παράλληλο διμερές. Οι SWEETs διαθέτουν επιπρόσθετα την TM4, η οποία συνδέει τις δυο παράλληλες bundles τριπλής έλικας (Shi 2013). Όσον αφορά τους SGLTs, οι APC πρωτεΐνες χαρακτηρίζονται από την «LeuT» αναδίπλωση, η οποία περιέχει 10 TMs οργανωμένες σε ένα ζευγάρι «5+5» ανάστροφες επαναλήψεις, ενώ και τα δύο άκρα της αλυσίδας βρίσκονται στην κυτοπλασματική μεριά. Σε κάθε επανάληψη, η πρώτη διαμεμβρανική περιοχή είναι τυπι-



Εικόνα 8. Δομή των μεταφορέων γλυκόζης (Shi 2013). (Α) Δομή των GLUTs. Διαθέτουν 12 διαμεμβρανικές έλικες, η οποίες μοιράζονται σε δύο domains, το N (πράσινο) και το C (λευκό) και μία ενδοκυτταρική έλικα που ονομάζεται ICH (πορτοκαλί). Δεξιά φαίνεται η δομή του ανθρώπινου GLUT1 (PDB: 4PYP). (Β) Δομή των SGLTs. Οι TM2-11 συνιστούν «515» ανάστροφες επαναλήψεις (πράσινο και λευκό) της LeuT- αναδίπλωσης του vSGLT. Το μαύρο εξάγωνο υποδηλώνει το υπόστρωμα του. Δεξιά φαίνεται η κλειστή διαμόρφωση του vSGLT (PDB: 3DH4). (Γ) Δομή των SWEETs. Στα αριστερά φαίνεται η τοπολογία των SemiSWEETs και η προβλεπόμενη τοπολογία του ανθρώπινου SWEET1. Και οι δύο είναι διατεταγμένες σε μοτίβο 1-3-2 και χρωματισμένες με πράσινο και λευκό. Στα δεξιά φαίνεται η δομή του LbSemiSWEET (PDB: 4QNC). Σε όλες τις δομές δίνεται το λειτουργικό διμερές των SemiSWEETs. Κάθε μονομερές περιέχει τρεις έλικες.

κά μία διακεκομμένη έλικα, που αποτελεί σύνηθες χαρακτηριστικό των δευτερογενών μεταφορέων και επιτελεί σημαντικό ρόλο στην αναγνώριση και τη μεταφορά του υποστρώματος. Οι GLUTs χαρακτηριστικά έχουν 12TMs και μία ενδοκυτταρική έλικα ICH, η οποία παίζει σημαντικό ρόλο στην λειτουργία

για του μεταφορέα. Στην Εικόνα 9 (Chen *et al.* 2015) παρατίθενται τρία παραδείγματα δομών μεταφορέων, ένα για κάθε υπεροικογένεια, όπως αυτές φαίνονται στο πλάι (α) και σε κάτοψη (β).



Εικόνα 9. Δομές των τριών ειδών ευκαρυωτικών μεταφορέων σακχάρων (Chen *et al.* 2015).

Συγκρίνοντας τις δομές των μεταφορέων σακχάρων MFS, SGLTs και SWEETs διαπιστώνουμε ότι οι θέσεις δέσμευσης σακχάρων έχουν διαφορετικές διαμορφώσεις αλλά και κοινά χαρακτηριστικά (Chen *et al.* 2015). Ένα κοινό χαρακτηριστικό είναι, για παράδειγμα, ότι οι θέσεις πρόσδεσης βρίσκονται κοντά στο κέντρο της μεμβρανικής διπλοστοιβάδας και είναι τοποθετημένες στην διεπαφή δύο επαναλήψεων. Επίσης, τα αρωματικά κατάλοιπα παίζουν κρίσιμο ρόλο είτε παρέχοντας αλληλεπίδραση στοίβαξης με το υπόστρωμα είτε παρεμποδίζοντας το υπόστρωμα να αποδεσμευτεί.

1.2.5 The Major Facilitator Superfamily

Η MFS είναι η μεγαλύτερη υπεροικογένεια δευτερογενών μεταφορέων που εντοπίζεται στην βίωση και διανέμεται σχεδόν σε όλα τα φύλα (Reddy *et al.* 2012). Σύμφωνα με την TCDB (Transporter Classification Database; www.tcdb.org), περιλαμβάνει 10 οικογένειες πρωτεϊνών που

καταλύουν τόσο δευτερογενή μεταφορά όσο και άλλες διαδικασίες. Στον Πίνακα 1 παρουσιάζονται οι οικογένειες αυτές ονομαστικά και το σύνολο των πρωτεϊνών που περιλαμβάνουν.

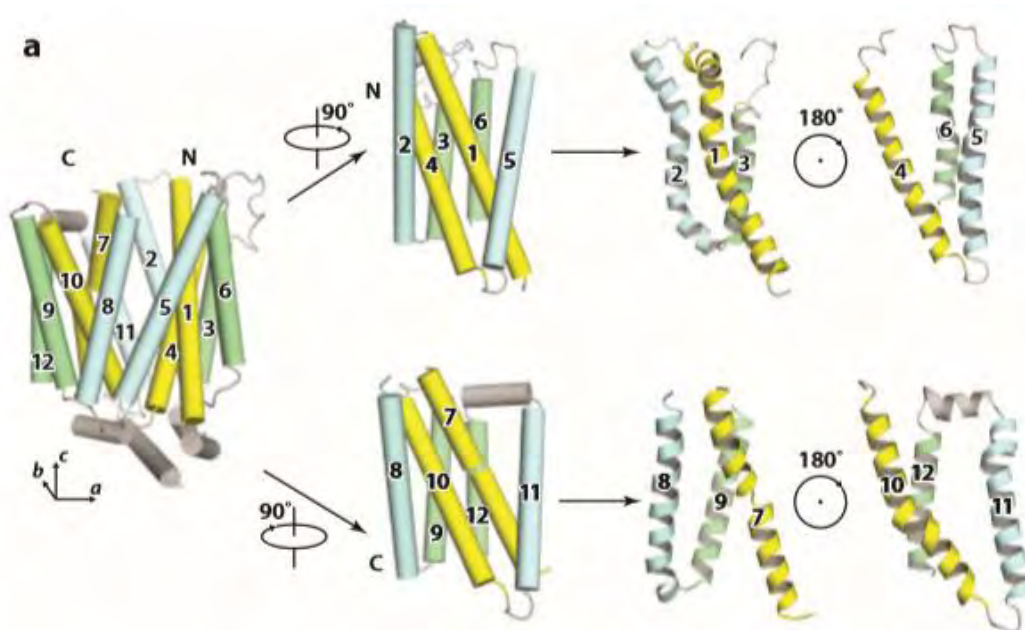
TCID	Όνομα οικογένειας	Σύνολο πρωτεϊνών
2.A.1	The Major Facilitator Superfamily (MFS)	880
2.A.2	The Glycoside-Penchenide-Hexuronide (GPH):Cation Symporter Family	38
2.A.12	The ATP:ADP Antiporter (AAA) Family	21
2.A.17	The Proton-dependent Oligopeptide Transporter (POT/PTR) Family	43
2.A.48	The Reduced Folate Carrier (RFC) Family	6
2.A.57	The Equilibrative Nucleoside Transporter (ENT) Family	33
2.A.60	The Organo Anion Transporter (OAT) Family	29
2.A.71	The Folate-Biopterin Transporter (FBT) Family	11
2.A.100	The Ferroportin (Fpn) Family	6
9.B.111	The 6 TMS Lysyl tRNA Synthetase (LysS) Family	3

Πίνακας 1. Οι δέκα οικογένειες που περιλαμβάνει η MFS

Στην παρούσα διπλωματική εργασία στόχος μας ήταν η φυλογενετική ανάλυση της Sugar Porter Family (TCID: 2.A.1.1). Η MFS (TCID: 2.A.1) μέχρι σήμερα περιέχει 87 οικογένειες, πρώτη εκ των οποίων είναι η SP Family, και περιλαμβάνει μονομεταφορείς, συμμεταφορείς και αντιμεταφορείς. Πολλές από τις οικογένειες έχουν προσδιοριστεί φυλογενετικά και δεν περιέχουν κανένα λειτουργικά χαρακτηρισμένο μέλος και ονομάζονται UMFs (Unknown Major Facilitators) (Reddy *et al.* 2012). Κάθε μία από τις καλά χαρακτηρισμένες οικογένειες μεταφέρει διαφορετικό σύνολο ενώσεων, όπως απλούς μονοσακχαρίτες, ολιγοσακχαρίτες, αμινοξέα, μικρά πεπτίδια, βιταμίνες, συμπράγοντες ενζύμων, φάρμακα, χρωμοφόρα, νουκλεοβάσεις, νουκλεοζίτες, νουκλεοτίδια, χηλικά σύμπλοκα σιδήρου, οργανικά και ανόργανα ανιόντα και κατιόντα. Εξελικτικά οι μεταφορείς αυτής της υπεροικογένειας φαίνεται ότι προήλθαν από μία μονή 2 TMS hairpin δομή, η οποία τριπλασιάστηκε σε 6 TMS μονάδα, που στη συνέχεια διπλασιάστηκε σε 12 TMS πρωτεϊνική τοπολογία. Ωστόσο, υπάρχουν και οικογένειες που εμφανίζουν τοπολογία 14 TMS, από τις οποίες οι δύο επιπλέον έλικες συνήθως εντοπίζονται κεντρικά, ανάμεσα στις δυο επαναλαμβανόμενες μονάδες των 6TMS και ίσως έχουν προέλθει από έναν ενδογονιδιακό (intragenic) διπλασιασμό μιάς αρχέγονης φουρκέτας (hairpin). Επιπρόσθετα, οι οικογένειες με 12 TMS δε γνωρίζουμε πως προήλθαν, αλλά

ίσως προσδίδουν βέλτιστη λειτουργικότητα. Έχει αποδειχθεί ότι υπάρχουν ομόλογα με 13, 14, 15 ή 16 TMS, αλλά και με 24TMS. Τα τελευταία ομόλογα προήλθαν από πλήρη διπλασιασμό και παρουσιάζουν διπλές λειτουργίες.

Όλοι οι MFS μεταφορείς μοιράζονται μία κοινή και χαρακτηριστική αναδίπλωση στον πυρήνα, που ονομάζεται MFS fold και παρουσιάζεται στην Εικόνα 10 (Yan 2015). Μια τυπική MFS fold περιλαμβάνει δύο domains, που αποτελούνται συνήθως από 6 διαδοχικές TMs και ονομάζονται N και C domains. Οι δυο αυτές περιοχές παρουσιάζουν διπλή ψευδοσυμμετρία που σχετίζεται με τον άξονα ψ . Σε κάθε domain, οι 6TMs οργανώνονται σε ζευγάρι «3+3» ανεστραμμένων επαναλήψεων. Οι TMs της κάθε επανάληψης παίζουν σημαντικό ρόλο στη λειτουργία και τη δομή της πρωτεΐνης. Συγκεκριμένα, κάθε πρώτη έλικα κάθε δεματίου τριών ελίκων (TMs 1,4,7 και 10) τοποθετείται στο κέντρο του μεταφορέα και όλες μαζί δομούν μονοπάτι μεταφοράς. Η πλειοψηφία των καταλοίπων εκεί έχουν ως ρόλο τη σύνδεση του υποστρώματος και τη συμμεταφορά. Εξωτερικά του πυρήνα βρίσκονται οι TMs 2, 5, 8 και 11, οι οποίες μεσολαβούν την επαφή των N και C domains, ενώ κάποια κατάλοιπα τους ίσως συμμετέχουν στη δέσμευση του υποστρώματος και στη συμμεταφορά του. Εξωτερικά των TMs 1, 4, 7 και 10 τοποθετούνται οι TMs 3, 6, 9 και 12 συμβάλλοντας στην δομική σταθερότητα του μεταφορέα. Τέλος, παρατηρώντας τη δομή σαν σύνολο, είναι ενδιαφέρον ότι οι αντίστοιχες TMs κάθε ενός από τις δύο επαναλήψεις τριών TMs, πάντα τοποθετούνται η μία δίπλα στην άλλη και έχουν αντίθετες κατευθύνσεις.



Εικόνα 10. Πρότυπη MFS αναδίπλωση (Yan 2015). Φαίνονται οι δύο διακριτές περιοχές N και C, οι οποίες σχηματίζουν την χαρακτηριστική δομή των 12 διαμεμβρανικών περιοχών και σχετίζονται με περιστροφή περίπου 180 μοιρών γύρω από τον άξονα ψ . Κάθε μία τέτοια περιοχή αποτελείται από δύο ανάστροφες επαναλήψεις τριών διαμεμβρανικών ελίκων. Οι αντίστοιχες έλικες σε κάθε μονάδα έχουν το ίδιο χρώμα.

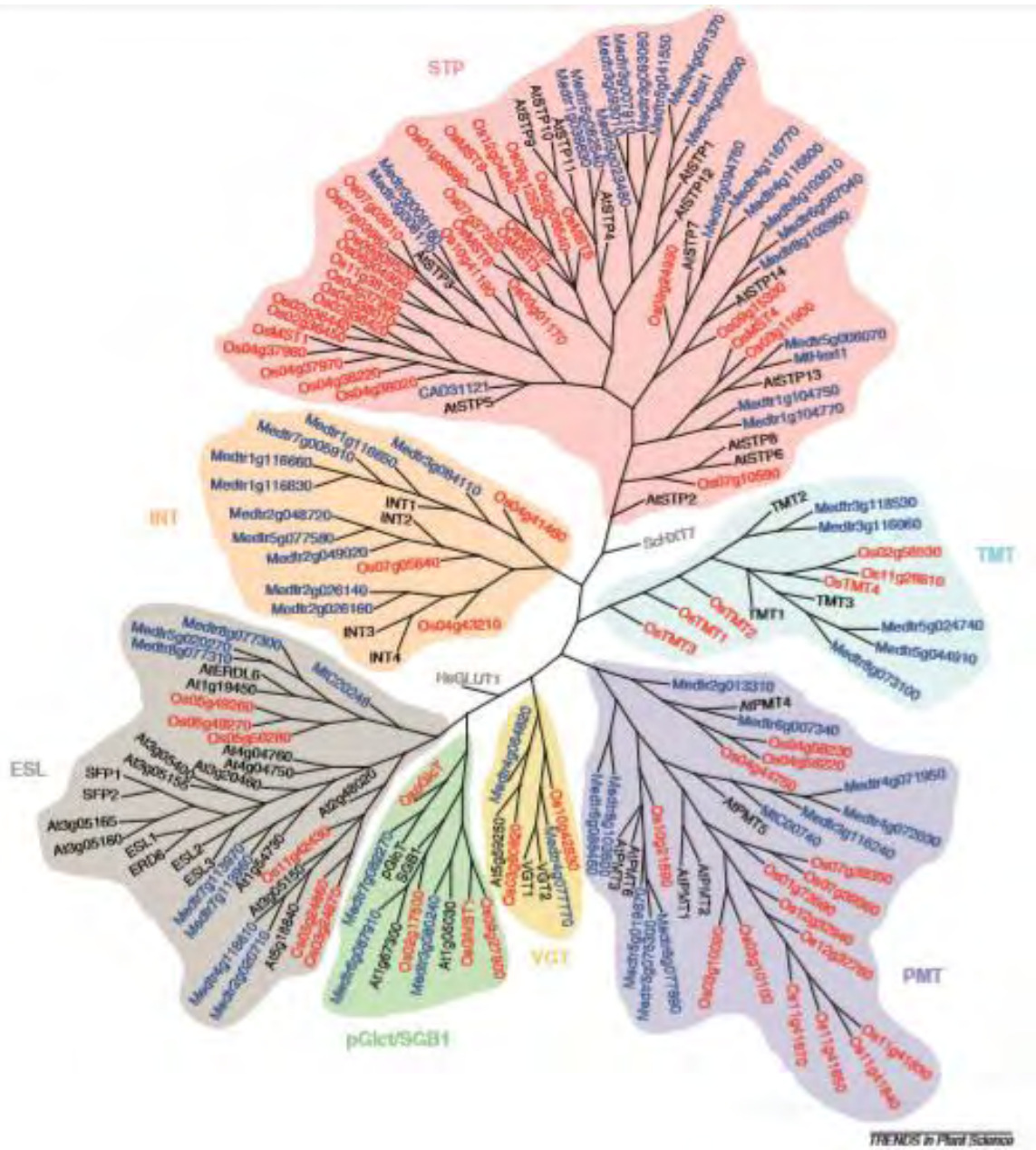
1.3 Οι φυτικοί μεταφορείς σακχάρων

Τα σάκχαρα έχουν ποικίλους ρόλους στα φυτά (Wang *et al.* 2016). Λειτουργούν ως μόρια αποθήκευσης ενέργειας και μεταφοράς άνθρακα. Άλλος ρόλος τους είναι η μεταφορά μηνυμάτων, δηλαδή δρουν σαν ορμόνες.

Η μεταφορά των σακχάρων πραγματοποιείται από πρωτεΐνες μεταφορείς, οι οποίοι μεταφέρουν τα σάκχαρα είτε με παθητική, αλλά ελεγχόμενη διάχυση είτε με ενεργό μεταφορά εξαρτώμενη από ενέργεια (Büttner 2007). Οι μεταφορείς εξοζών και σουκρόζης ανήκουν σε τρεις μεγάλες οικογένειες, i) την οικογένεια μεταφορέων σουκρόζης (SUT ή SUC), ii) την οικογένεια monosaccharide transporter (-like) (MST-like), των οποίων οι μεταφορείς είναι μέλη της MFS και λειτουργούν ως συμμεταφορείς πρωτονίων-σακχάρων (Doidy *et al.* 2012), iii) την οικογένεια μονοσακχαριτών SWEETs, οι οποίοι είναι μονομεταφορείς γλυκόζης και φέρουν 7 διαμεμβρανικές περιοχές. Με βάση μια προηγούμενη φυλογενετική ανάλυση (Εικόνα 11) που έγινε στα *Arabidopsis*, *Medicago truncatula* και στο ρύζι, οι MSTs χωρίζονται σε επτά υποοικογένειες, τις STP (sugar transport protein), PMT (polyol/monosaccharide transporter), INT (inositol transporter), VGT (vacuolar glucose transporter), TMT (tonoplast membrane transporter), pGLcT (plastidic glucose transporter) και ESL (early-responsive to dehydration six-like) (Doidy *et al.* 2012).

Οι STPs λειτουργούν ως συμμεταφορείς πρωτονίων/εξοζών και έχουν υψηλή συγγένεια για τη γλυκόζη. Και μέλη της PMT φαίνεται να εμπλέκονται στην πρόσληψη εξοζών διαμέσου της πλασματικής μεμβράνης, ενώ υποστρώματα τους αποτελούν εξόζες, πεντόζες και πολυόλες, όπως η μαννιτόλη και η σορβιτόλη. Επίσης, έχουν ρόλο στην αλληλεπίδραση των φυτών με τους μύκητες. Οι VGTs και οι TMTs είναι μεταφορείς των χυμοτοπίων και ρόλος τους είναι η είσοδος των σακχάρων σε αυτά για παροδική ή μόνιμη αποθήκευση. Οι ESLs που επίσης ανιχνεύονται στα κενοτόπια, εμπλέκονται στην ανασύσταση των υδατανθράκων. Οι pGLcTs τοποθετούνται στους χλωροπλάστες και ρόλος τους είναι η εκροή της γλυκόζης από αυτούς. Τέλος, οι INT έχουν ρόλο στην μεταγωγή σήματος (Büttner 2007).

Σκοπός αυτής της εργασίας ήταν να επαναλάβουμε την φυλογενετική ανάλυση των MSTs με περισσότερα (12) πρωτεύματα φυτικών οργανισμών από διαφορετικές εξελικτικές γραμμές, όπως τα χλωρόφυτα, βρυόφυτα, γυμνόσπερμα, και βασικά αγγειόσπερμα, ώστε να κατανοήσουμε πόσο αρχέγονες είναι οι επιμέρους υποοικογένειες.



Εικόνα 11. Φυλογενετικό δένδρο της οικογένειας MST (Doidy *et al.* 2012).

2. ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ

2.1. Λήψη πρωτεωμάτων

Οι οργανισμοί, των οποίων τα πρωτεώματα χρησιμοποιήθηκαν για την παρούσα ερευνητική εργασία, παρουσιάζονται στον Πίνακα 2. Δίνεται το όνομα του αρχείου σε fasta format το οποίο περιέχει το κάθε πρωτέωμα, καθώς και η πηγή από την οποία λήφθηκε αυτό και τον αριθμό των πρωτεϊνών που περιέχει. Να σημειωθεί ότι τα πρωτεώματα των *Arabidopsis thaliana* και *Homo sapiens* περιέχουν πολλαπλές ισομορφές για την ίδια πρωτεΐνη και για αυτό φιλτραρίστηκαν ώστε να περιέχουν μόνο τις μεγαλύτερες πρωτεΐνες-ισομορφές.

Οργανισμός	Όνομα αρχείου fasta	Σύνολο πρωτεϊνών	Πηγή πρωτεώματος
<i>Amborella trichopoda</i>	proteome.Amborella_trichopoda.fasta	27313	https://plants.ensembl.org/index.html
<i>Arabidopsis thaliana</i>	proteome.Arabidopsis_thaliana_longest_pep.fasta	27416	https://www.arabidopsis.org/
<i>Chlamydomonas reinhardtii</i>	proteome.Chlamydomonas_reinhardtii.fasta	14489	https://plants.ensembl.org/index.html
<i>Escherichia coli</i>	proteome.Escherichia_coli_K-12.fasta	4140	http://www.uniprot.org/
<i>Homo sapiens</i>	proteome.Homo_sapiens_longest_pep.fasta	23070	http://www.uniprot.org/
<i>Lotus japonicus</i>	proteome.Lotus_japnicus.fasta	48105	http://www.kazusa.or.jp/lotus/
<i>Medicago truncatula</i>	proteome.Medicago_truncatula.fasta	57585	https://plants.ensembl.org/index.html
<i>Oryza sativa japonica</i>	proteome.Oryza_sativa_japonica.fasta	42377	https://plants.ensembl.org/index.html
<i>Ostreococcus lucimarinus</i>	proteome.Ostreococcus_lucimarinus.fasta	7603	https://plants.ensembl.org/index.html
<i>Ostreococcus tauri</i>	proteome.Ostreococcus_tauri.fasta	7951	https://bioinformatics.psb.ugent.be/plaza/versions/pico-plaza/
<i>Physcomitrella patens</i>	proteome.Physcomitrella_patens.fasta	38354	https://plants.ensembl.org/index.html
<i>Pinus sylvestris</i>	proteome.Pinus_sylvestris.fasta	36106	https://bioinformatics.psb.ugent.be/plaza/versions/gymno-plaza/download/index
<i>Saccharomyces cerevisiae</i>	proteome.Saccharomyces_cerevisiae.fasta	5917	http://www.ensembl.org/index.html
<i>Selaginella moellendorffii</i>	proteome.Selaginella_moellendorffii.fasta	34825	https://plants.ensembl.org/index.html
<i>Volvox carteri</i>	proteome.Volvox_carteri.fasta	15544	https://bioinformatics.psb.ugent.be/plaza/versions/pico-plaza/

Πίνακας 2. Τα πρωτεώματα που χρησιμοποιήθηκαν στις αναλύσεις μας.

2.2. Προγράμματα

Για την πραγμάτωση της εργασίας, χρησιμοποιήθηκαν τα παρακάτω υπολογιστικά εργαλεία και γλώσσες προγραμματισμού.

2.2.1 Perl

Η Perl είναι μία σχετικά απλή γλώσσα προγραμματισμού, κατάλληλη για μορφοποίηση και ανάλυση δεδομένων.

2.2.2 Python

Η Python είναι μία ισχυρή γλώσσα προγραμματισμού, παρόμοια με τις Perl, Ruby, Java ή Scheme. (“BeginnersGuide/Overview - Python Wiki” n.d.) Διαθέτει μεγάλη βιβλιοθήκη που υποστηρίζει πολλές δυνατότητες, όπως η σύνδεση σε web servers, η αναζήτηση σε αρχεία με χρήση μεταχαρακτήρων και η ανάγνωση και επεξεργασία αρχείων. Τέλος, μπορεί να τρέξει παντού, συμπεριλαμβανομένου σε MacOSX, Windows, Linux, Unix, και ανεπίσημες κατασκευές ακόμα και σε Adroid ή iOS.

2.2.3 MEME Suite 4.12.0

Το MEME Suite διαθέτει εργαλεία για ανάλυση αλληλουχιών με βάση συγκεκριμένα μοτίβα (“Introduction - MEME Suite” n.d.). Τα εργαλεία που χρησιμοποιήθηκαν στην παρούσα εργασία ήταν το MEME και το MAST. Το MEME (Multiple Em for Motif Elicitation) ανακαλύπτει νέα, χωρίς κενά μοτίβα (επαναλαμβανόμενα, σταθερού μήκους μοτίβα) των αλληλουχιών που του εισάγουμε, και διαχωρίζει τα μοτίβα μεταβλητού μήκους σε 2 ή περισσότερα χωριστά μοτίβα. Το MAST (Motif Alignment & Search Tool) σκανάρει τις ακολουθίες για την ύπαρξη μοτίβων που ανακαλύφθηκαν με το MEME. Επίσης, ταξινομεί τις αλληλουχίες με βάση το καλύτερο σκορ ταιριάσματος με όλα τα μοτίβα.

2.2.4 HMMER 3.1

Το HMMER χρησιμοποιείται για αναζήτηση πρωτεϊνικών επικρατειών σε ακολουθίες επερώτησης, χρησιμοποιώντας ως βάση δεδομένων μια σειρά από προφίλ Hidden Markov Models (Eddy, Wheeler, and the HMMER development team 2015). Αυτά τα προφίλ έχουν κατασκευασθεί από . πολλαπλές στοιχίσεις μίας οικογένειας ή μιας συγκεκριμένης επικράτειας.

Τα προγράμματα του HMMER που χρησιμοποιήσαμε ήταν το hmmbuild, το οποίο χτίζει ένα HMM προφίλ από μία εισερχόμενη πολλαπλή στοίχιση, το hmpress, που μορφοποιεί μία βάση δεδο-

μένων HMM σε binary format για χρήση αυτού στο hmmscan και το hmmsearch, το οποίο σκανάρει μία πρωτεϊνική ακολουθία με τη πρωτεϊνικού προφίλ HMM βάση δεδομένων.

2.2.5 SeaView 4.4.2

Το SeaView είναι ένα γραφικό περιβάλλον χρήστη για πολλαπλή στοίχιση αλληλουχιών και φυλογένεση (Gouy, Guindon, and Gascuel 2010). Η πολλαπλή στοίχιση μπορεί να πραγματοποιηθεί με το πρόγραμμα muscle, το οποίο χρησιμοποιήθηκε και στην παρούσα εργασία, ή το Clustal Omega. Επίσης, υπολογίζει τα φυλογενετικά δένδρα με χρήση είτε μέγιστης φειδωλότητας, είτε μέγιστης πιθανοφάνειας, είτε με μεθόδους αποστάσεων, όπως των Neighbor Joinning και BioNJ αλγορίθμων. Υπολογίζει τα bootstrap των κλάδων σε κάθε δένδρο, ενώ δίνει τη δυνατότητα αποθήκευσης των δένδρων σε Newick format, ή σε SVG, PDF ή PostScript files.

2.2.6 TreeDyn & ETE Toolkit

Το TreeDyn και το ETE Toolkit είναι προγράμματα επεξεργασίας φυλογενετικών δένδρων. Το Treedyn δέχεται εισερχόμενα αρχεία τύπου newick ή TGF, και οι εικόνες μπορούν να αποθηκευτούν σε τύπο TGF, Postscript, SVG, bitmap, GIF κ.α. Μερικά από τα εργαλεία που περιλαμβάνει είναι η μεγέθυνση-σμίκρυνση, το ladderize, το swarp, ο χρωματισμός και το collapse. Το ETE Toolkit (γραμμένο σε Python) παρέχει εργαλεία για οπτικοποίηση, ανάλυση και χειρισμό φυλογενετικών δέντρων (Huerta-Cepas, Serra, and Bork 2016). Μερικά από τα εργαλεία που διαθέτει είναι ο χαρακτηρισμός των κλαδιών, η αυτόματη ανίχνευση ορθολόγων και παραλόγων, η ανεξάρτητη επεξεργασία διαφορετικών κομματιών του ίδιου φυλογενετικού δένδρου.

2.2.8 USEARCH

Το USEARCH είναι ένα εργαλείο ανάλυσης ακολουθιών και παρέχει αλγορίθμους για ανάλυση και ομαδοποίηση με βάση την ομοιότητα (Edgar 2010).

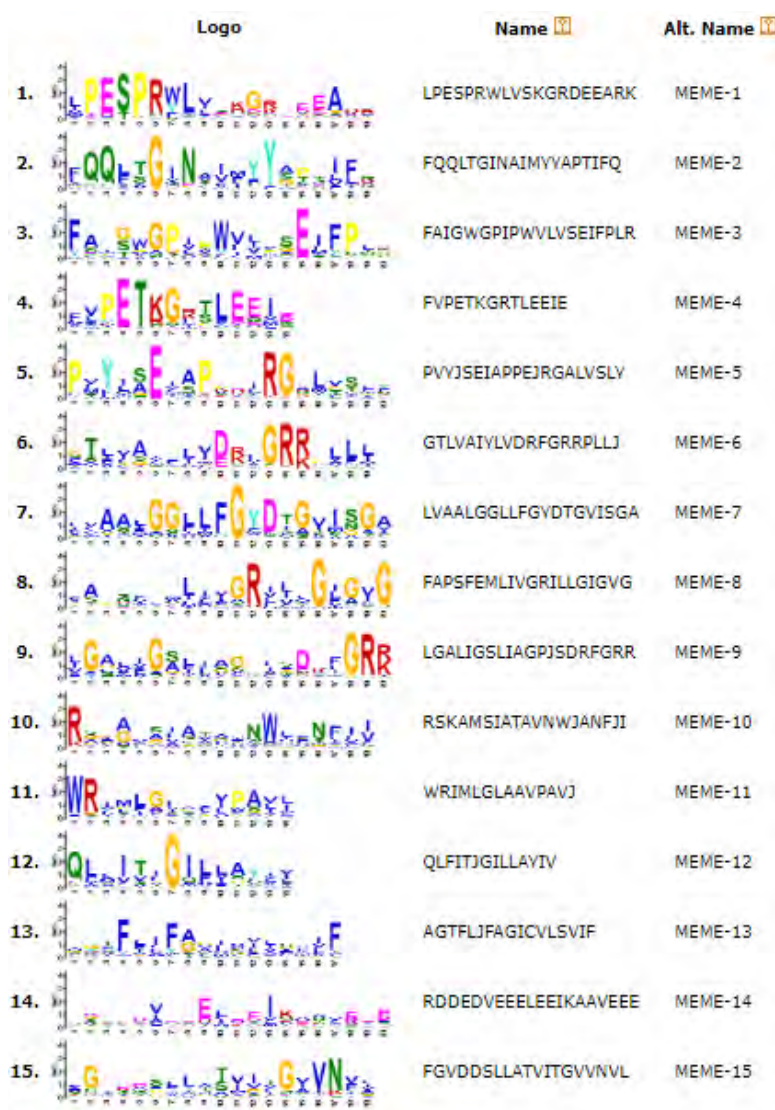
2.2.9 TIMETREE (THE TIMESCALE OF LIFE)

Το TimeTree είναι μία δημόσια βάση δεδομένων που δείχνει πότε περίπου έζησε ο κοινός πρόγονος δύο οργανισμών ή taxa (Kumar *et al.* 2017). Δίνει τη δυνατότητα αναζήτησης με τρία εργαλεία. Το NODE TIME που βρίσκει πότε πραγματοποιήθηκε ο διαχωρισμός δύο ειδών, το TIMELINE που δίνει το χρονοδιάγραμμα της εξέλιξης ενός taxa και το TIMETREE, το οποίο κατασκευάζει ένα «χρονοφυλογενετικό δένδρο» για ένα σύνολο ειδών ή μία επιθυμητή λίστα.

3. ΑΠΟΤΕΛΕΣΜΑΤΑ-ΣΥΖΗΤΗΣΗ

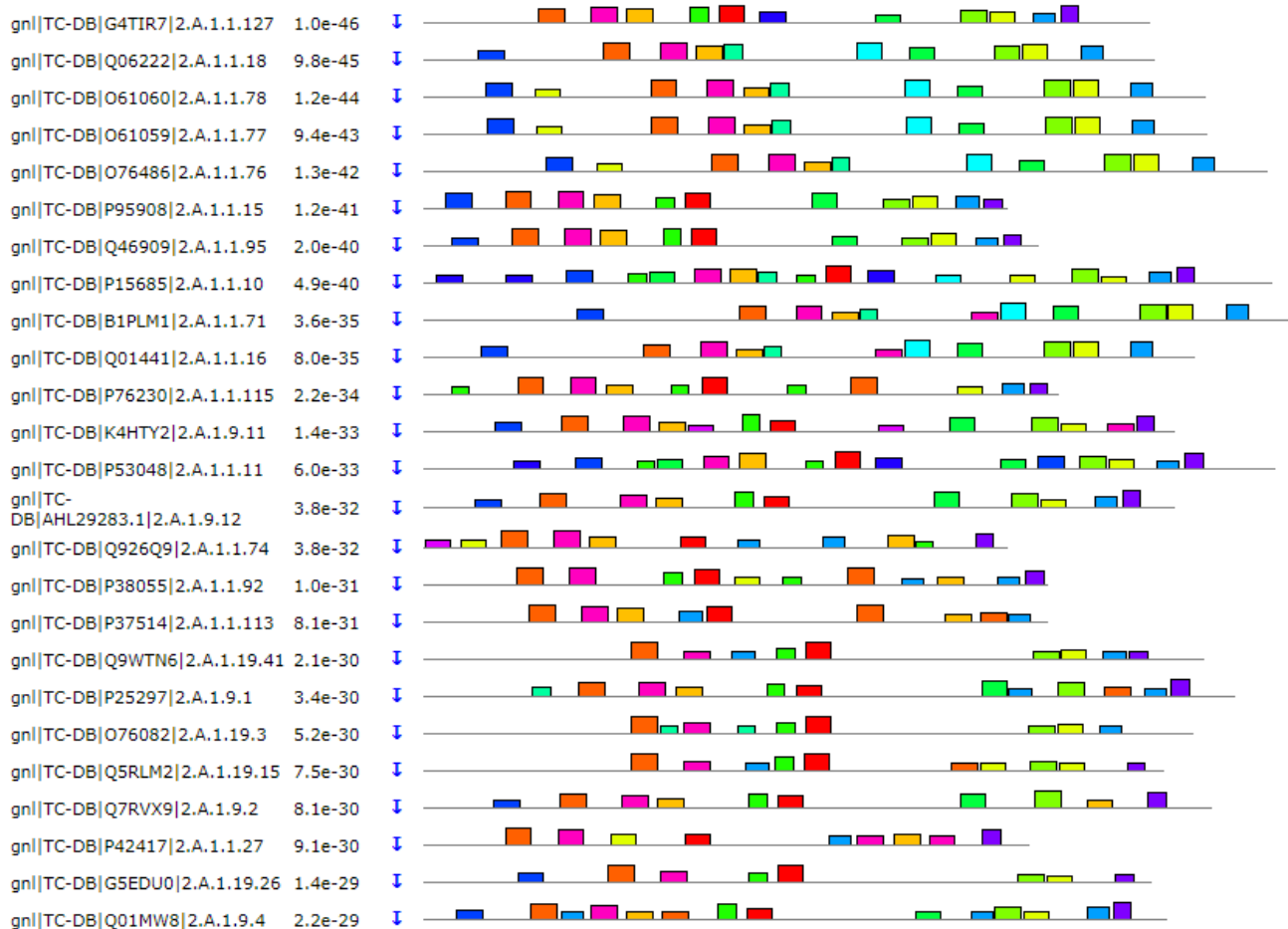
3.1 Αναγνώριση μοτίβων της οικογένειας 2.A.1.1

Αρχικά, κατεβάσαμε ολόκληρη τη βιβλιοθήκη της TCDB (tcdb_all_proteins_original_fasta.fa), η οποία περιείχε 16462 πρωτεΐνες. Από αυτές οι 876 ανήκουν στην οικογένεια 2.A.1 (MFS superfamily) (all_2_A_1_seqs_from_tcdb.fa), ενώ οι 133 στην υποοικογένεια 2.A.1.1 (MFS sugar transporters family) (2a11_from_tcdb.fa). Στη συνέχεια, θέλαμε να δούμε ποια κοινά δομικά στοιχεία μοιράζονται οι πρωτεΐνες της 2.A.1.1. Έτσι, με το πρόγραμμα USEARCH αφαιρέσαμε πολύ συγγενικές ακολουθίες (2_a_1_1_from_tcdb_70_percent_centroids) με όριο ταύτισης 70% και πάνω στο αμινοξικό επίπεδο. Οι ακολουθίες που απέμειναν χρησιμοποιήθηκαν για να βρούμε 15 κοινά μοτίβα της 2.A.1.1 (Εικόνα 12), με χρήση του MEME Suite 4.12.0 και παραμέτρους «-protein -nmotifs 15 -minw 10 -maxw 20 -maxsize 550000».



Εικόνα 12. Μοτίβα των 2.A.1.1 αντιπροσώπων

Στη συνέχεια, τα παραπάνω μοτίβα τα χρησιμοποιήσαμε για να σκανάρουμε όλες τις ακολουθίες της TCDB με το MAST (Εικόνα 13). Όπως ήταν αναμενόμενο, οι ακολουθίες με τα καλύτερα σκορ ανήκαν στην 2.A.1.1. Ωστόσο, τα MEMES για την 2.A.1.1. μπορούσαν επίσης να ανιχνεύσουν με αρκετά καλό σκορ και ακολουθίες των 2.A.1.9 (The Phosphate: H⁺ Symporter (PHS) Family) και 2.A.1.19 (The Organic Cation Transporter (OCT) Family), μαζί με άλλες. Έτσι, αποφασίσαμε να ελέγξουμε την μονοφυλία των 2.A.1.1, 2.A.1.9 και 2.A.1.19, που είναι αρκετά συγγενικές οικογένειες.

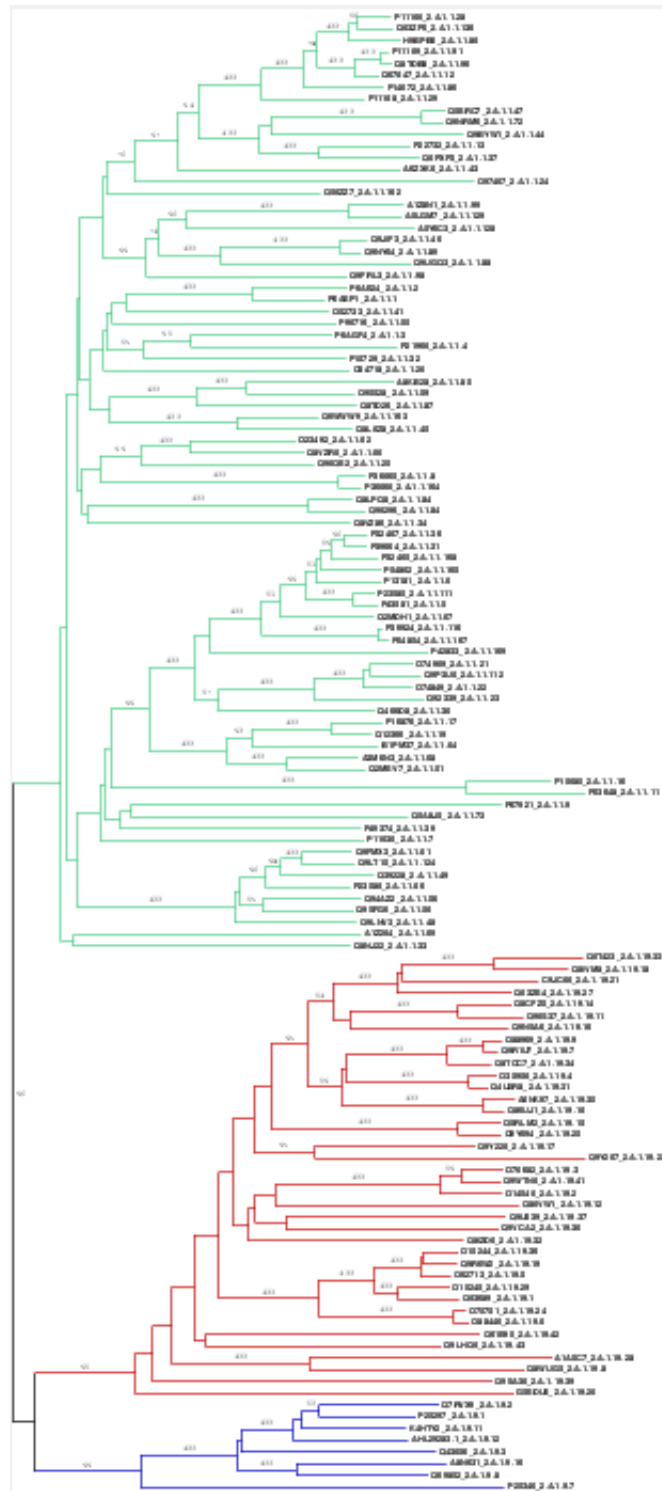


Εικόνα 13. Ενδεικτική παρουσίαση μερικών αποτελεσμάτων του MAST της 2.A.1.1 έναντι της βιβλιοθήκης της TCDB

3.2 Έλεγχος μονοφυλίας της οικογένειας 2.A.1.1

Έπειτα, ελέγξαμε και αφαιρέσαμε ακολουθίες των 2.A.1.1, 2.A.1.9 και 2.A.1.19 που είχαν χαρακτηριστεί ως πιθανά μέλη. Έτσι απέμειναν 85 πρωτεΐνες από την 2.A.1.1 (2a11_from_tcdb_edited.fa), 38 πρωτεΐνες από την 2.A.1.19 (2a119_from_tcdb_edited.fa) και 8 από την 2.A.1.9 (2a19_from_tcdb_edited.fa). Αυτές τις ακολουθίες τις κάναμε πολλαπλή στοίχιση,

διόρθωση και φυλογένεση (BioNJ, Poisson, 100 bootstraps) με το SeaView. Μετά την διόρθωση και την αφαίρεση προβληματικών αλληλουχιών, η 2.A.1.1 περιέχει 80 ακολουθίες (2a11_from_tcdb_edited_aln_edited.fst). Από την φυλογενετική ανάλυση των τριών οικογενειών φάνηκε ότι και οι 3 είναι μονοφυλετικές (Εικόνα 14).



Εικόνα 14. Φυλογενετικό δένδρο των οικογενειών 2.A.1.1, 2.A.1.9 και 2.A.1.19 (BioNJ, Poisson, 100 bootstraps).

3.3 Δημιουργία HMMs για τον εντοπισμό των 2.A.1.1

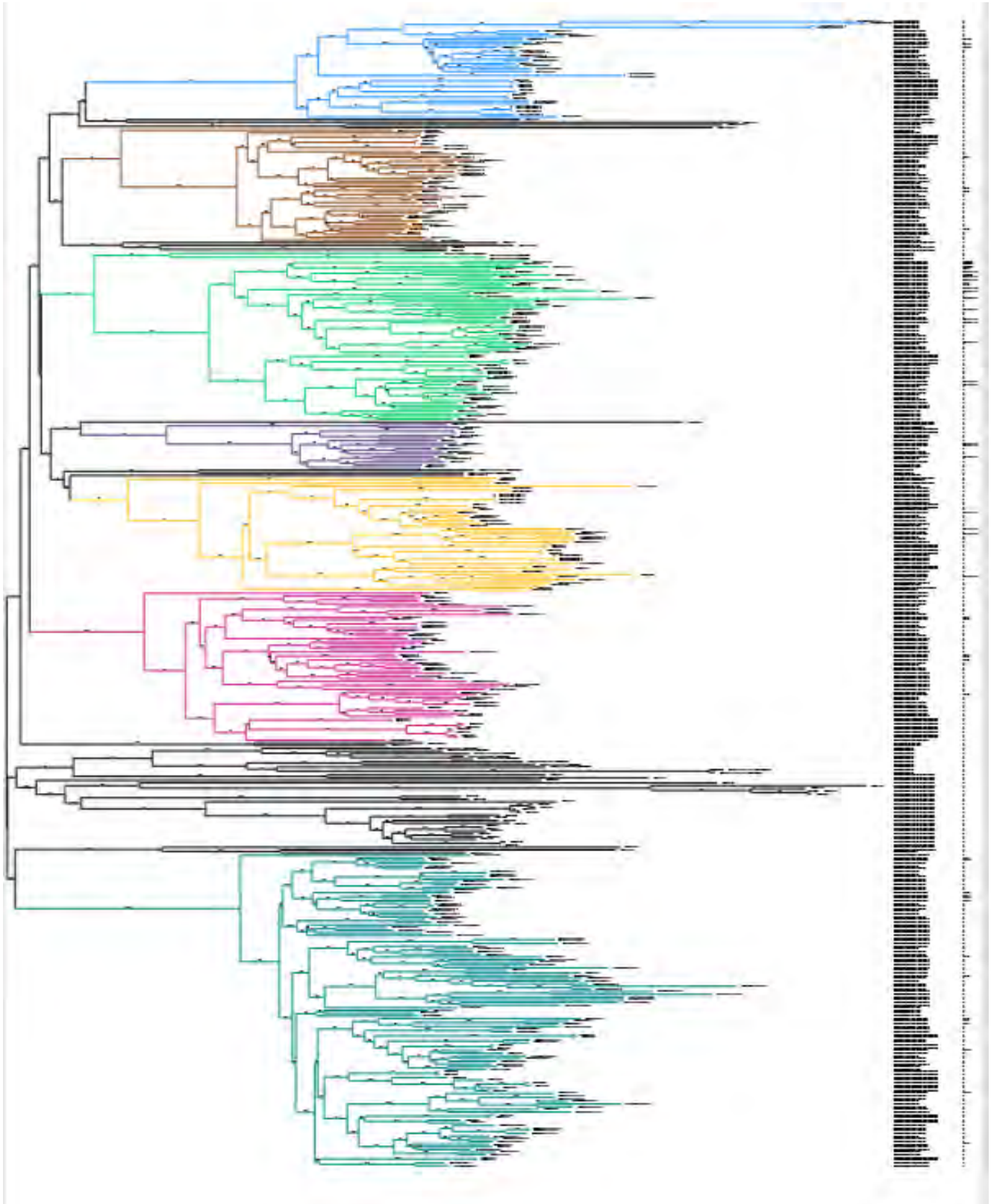
Οι 3 οικογένειες είναι αρκετά συγγενικές, επομένως, εάν δημιουργούσαμε ένα HMM μόνο από ακολουθίες 2.A.1.1, αυτό θα μπορούσε να εντοπίσει και 2.A.1.9 και 2.A.1.19. Επομένως, δεν θα μπορούσαμε να ξέρουμε αν οι ακολουθίες που εντοπίζονται ανήκουν στην 2.A.1.1 ή είναι μακρινά ομόλογα των 2.A.1.9 και 2.A.1.19. Για αυτό το λόγο, δημιουργήσαμε 3 Hidden Markov models για τις επεξεργασμένες 2.A.1.1, 2.A.1.9 και 2.A.1.19 χρησιμοποιώντας τις εντολές hmmbuild, hmmpress και hmmscan. Επομένως, τα 3 αυτά μοντέλα ανταγωνίζονται μεταξύ τους και έτσι μπορούμε να δούμε για μια ακολουθία ποιο μοντέλο την εντοπίζει με το καλύτερο σκορ. Στη συνέχεια, σκανάραμε όλες τις ακολουθίες της 2.A.1 με τα 3 μοντέλα και με βάση τα αποτελέσματα που έδωσε το hmmscan, ορίσαμε ως τιμή κατώφλι για το μοντέλο 2.A.1.1 $E\text{-value} \leq 1e^{-30}$.

Στο επόμενο βήμα κατεβάσαμε τα 15 πρωτεώματα που αναφέρθηκαν στην ενότητα 2.1, τα οποία αποτελούν εξελικτικούς αντιπροσώπους των φυτών με προσθήκη τα πρωτεώματα του ανθρώπου, του σαχαρομύκητα και του *E.coli*. Ενώσαμε όλα τα πρωτεώματα σε ένα fasta αρχείο (proteome.plants_plus_others.fa), το οποίο διαθέτει 390795 αλληλουχίες. Τα σκανάραμε με τα HMMs των τριών οικογενειών (αφού τα ενώσαμε στο file 2a11_2a19_2a119.hmm) χρησιμοποιώντας την εντολή:

```
hmmscan -E 1e-30 -o plants_plus_other_scanned --domtblout
plants_plus_other_scanned.domtblout
/home/alex/Desktop/2A11_2A119_2A19_new/hmmer/Hidden_Markov_Models/2a11_2a19_2a119.
hmm /home/alex/Desktop/proteomes/proteome.plants_plus_others.fa
```

3.4 Φυλογενετική ανάλυση των 2.A.1.1 από αντιπροσωπευτικά πρωτεώματα

Με το παραπάνω παραγόμενο αρχείο domtblout που περιέχει 2156 αποτελέσματα, βρήκαμε τις 562 ακολουθίες που χτύπησαν με καλύτερο σκορ την 2.A.1.1 και στη συνέχεια τις απομονώσαμε σε fasta format (plants_hits_2A11.fa). Πραγματοποιήσαμε πολλαπλή στοίχιση αυτών, με το SeaView, (plants_hits_2A11_aln.fst) και μετά από επεξεργασία αυτής της στοίχισης (plants_hits_2A11_aln_edited_gda.fst) φτιάξαμε το φυλογενετικό δέντρο (BioNJ, Poisson, 500 bootstraps) plants_hits_2A11_aln_edited_gda-BioNJ_tree (Εικόνα 15). Σύμφωνα με αυτό, υπάρχουν 7 υποοικογένειες στα φυτά και κάθε μία στο δέντρο παρουσιάζεται με διαφορετικό χρώμα.



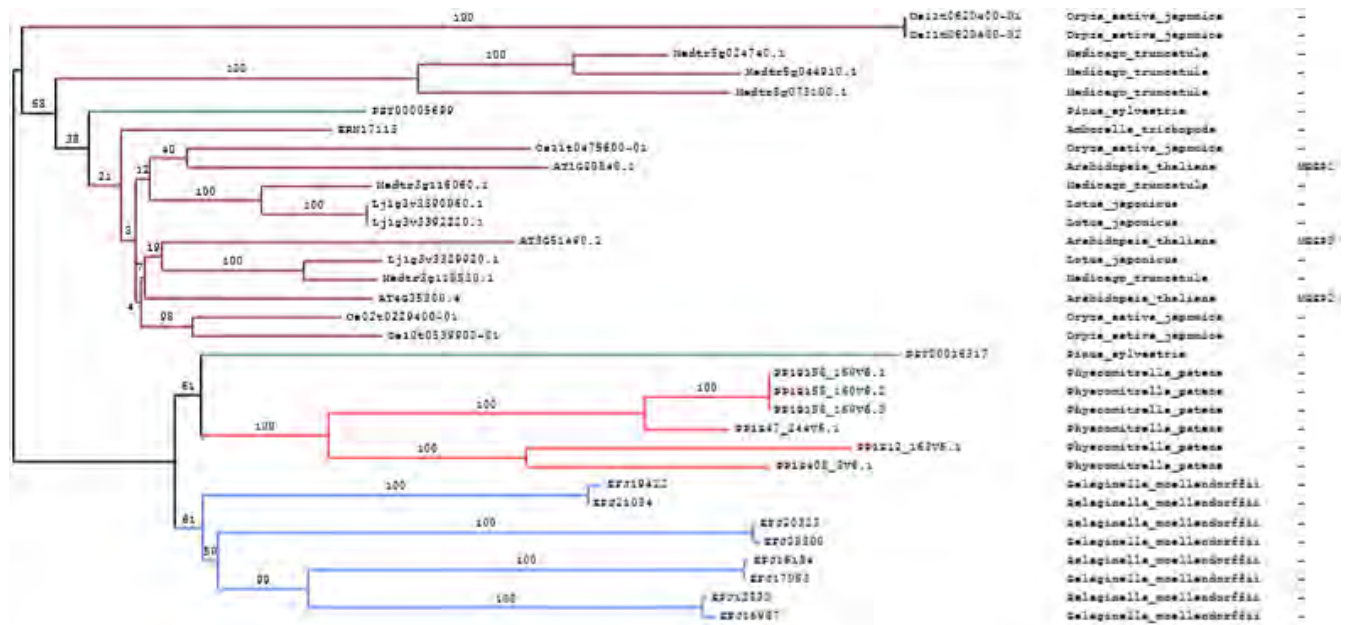
Εικόνα 15. Οι 7 υποοικογένειες της οικογένειας 2.A.1.1 στα φυτά (BioNJ, Poisson, 500 bootstraps).

Έπειτα, απομονώσαμε με το TreeDyn κάθε μία από τις 7 υποοικογένειες και με perl script τις πλήρεις ακολουθίες τους. Αυτές τις στοιχίσαμε, τις επεξεργαστήκαμε και φτιάξαμε 7 επιμέρους φυλογενετικά δέντρα, που παρατίθενται παρακάτω σε μορφή εικόνας. Στον Πίνακα 3 παρουσιάζονται οι υποοικογένειες καθώς και ο χαρακτηρισμός που τους δόθηκε και το σύνολο των πρωτεϊνών που απαρτίζουν κάθε μία.

Υποοικογένεια	Χαρακτηρισμός	Σύνολο ακολουθιών
Subfamily1	TMT	33
Subfamily2	INT	39
Subfamily3	ESL	62
Subfamily4	VGT	16
Subfamily5	pGLcT/SGB1	43
Subfamily6	PMT	56
Subfamily7	STP	117

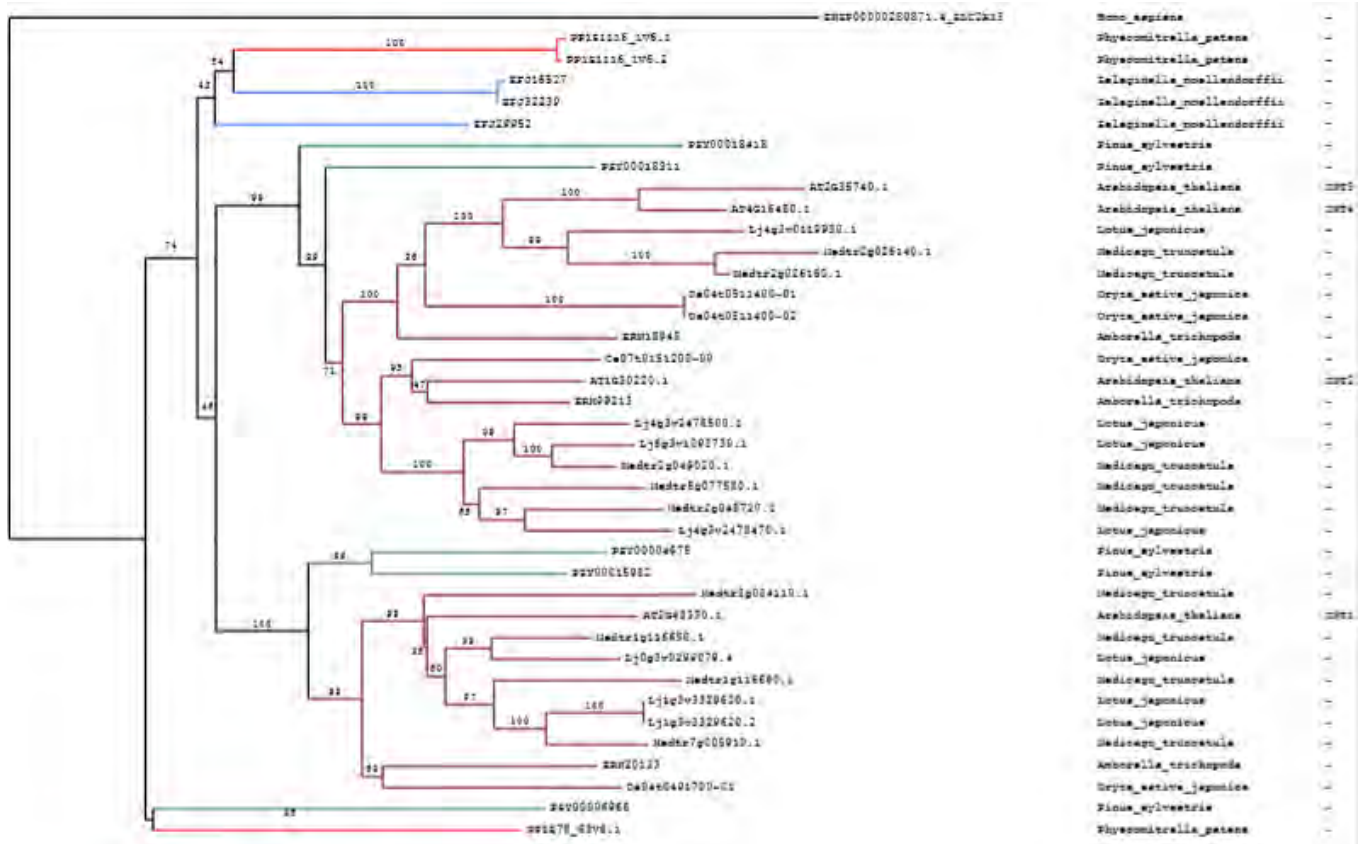
Πίνακας 3. Οι 7 φυτικές υποοικογένειες της 2.A.1.1 και ο αριθμός των ομολόγων που εντοπίστηκαν

Συγκεκριμένα, η TMT (Εικόνα 16) περιλαμβάνει 33 πρωτεΐνες και οι πιο απομακρισμένες εξελικτικά ανήκουν στο *Psyscomitrella patens*. Επομένως, η υποοικογένεια αυτή εμφανίστηκε στον κοινό πρόγονο των αγγειόσπερμων και βρυόφυτων, πριν 500 εκατομύρια χρόνια περίπου. Γενικά, είναι μια σχετικά μικρή υποοικογένεια, με 1 έως λίγα ομόλογα ανά οργανισμό.



Εικόνα 16. Φυλογενετικό δέντρο της Subfamily 1 (TMT) (BioNJ, Poisson, 500 bootstraps).

Η INT (Εικόνα 17) περιέχει 39 αλληλουχίες, από τις οποίες 3 ανήκουν στο *Psyscomitrella patens*. Επομένως, η υποοικογένεια αυτή εμφανίστηκε στον κοινό πρόγονο των αγγειόσπερων και βρυόφυτων, πριν 500 εκατομύρια χρόνια περίπου. Όπως και η TMT, ή INT είναι μια σχετικά μικρή υποοικογένεια, με 1 έως λίγα ομόλογα ανά οργανισμό.



Εικόνα 17. Φυλογενετικό δέντρο της Subfamily2 (INT) (BioNJ, Poisson, 500 bootstraps).

Η ESL (Εικόνα 18) είναι η δεύτερη μεγαλύτερη υποοικογένεια με 62 ακολουθίες. Τα περισσότερα ομόλογα που περιέχει ανήκουν στα αγγειόσπερμα φυτά, ενώ περιέχει και ομόλογα από το *Pinus sylvestris*, το *Selaginella moellendorffii* και το *Physcomitrella patens*. Επομένως, η υποοικογένεια αυτή εμφανίστηκε στον κοινό πρόγονο των αγγειόσπερμων και βρυόφυτων, πριν 500 εκατομύρια χρόνια περίπου. Χωρίζεται σε δύο κύριες ομάδες, την ομάδα 1 που περιλαμβάνει ομόλογα των αγγειόσπερμων φυτών, του *Pinus sylvestris*, του *Selaginella moellendorffii* και του *Physcomitrella patens* και την ομάδα 2 που αποτελείται εξολοκλήρου από ομόλογα αγγειόσπερμων φυτών.



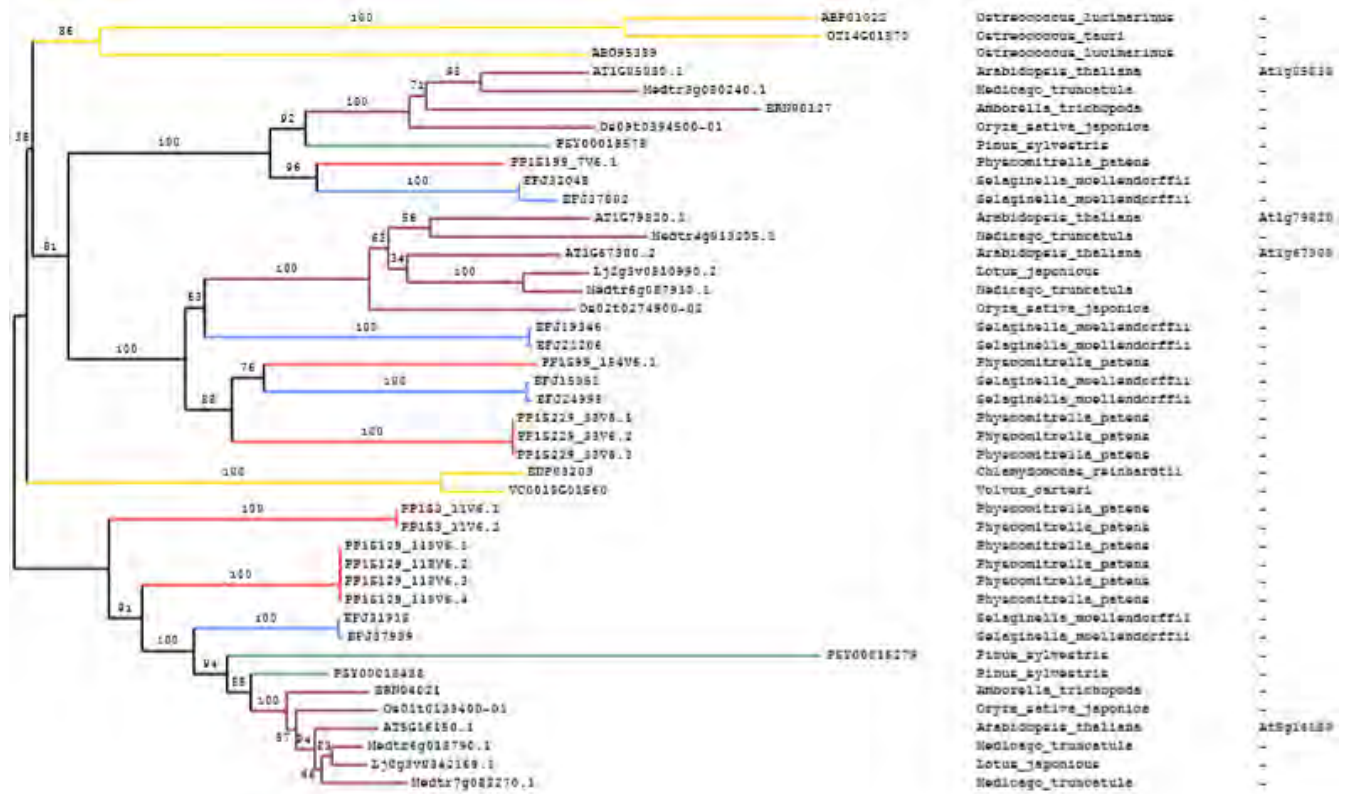
Εικόνα 18. Φυλογενετικό δέντρο της Subfamily3 (ESL) (BioNJ, Poisson, 500 bootstraps).

Η τέταρτη και μικρότερη υποοικογένεια, VGT (Εικόνα 19), περιλαμβάνει 16 πρωτεΐνες από διάφορες εξελικτικές γραμμές, όπως αγγειόσπερμα, γυμνόσπερμα, βρυόφυτα και το χλωρόφυτο *Ostreococcus lucimarinus*. Επομένως, η υποοικογένεια αυτή εμφανίστηκε στον κοινό πρόγονο των αγγειόσπερμων και χλωρόφυτων, πριν 950 εκατομύρια χρόνια περίπου. Φαίνεται ότι αν και πολύ αρχαία, αυτή η υποοικογένεια δεν ανέχεται γονιδιακούς διπλασιασμούς.



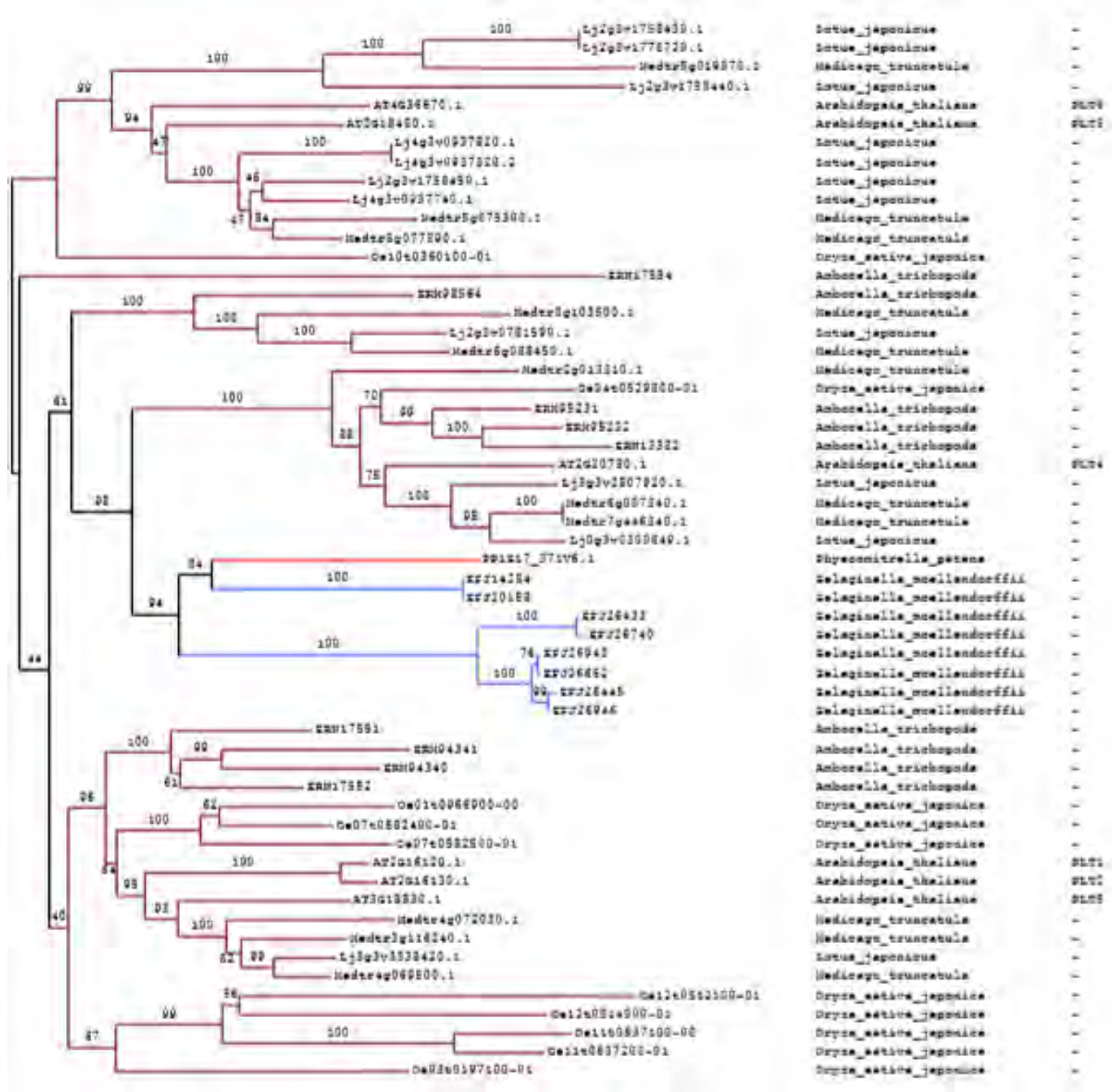
Εικόνα 19. Φυλογενετικό δέντρο της Subfamily4 (VGT) (BioNJ, Poisson, 500 bootstraps).

Η pGLcT/SGB1 (Εικόνα 20) είναι η μόνη οικογένεια που περιέχει ομόλογα από όλους τους αντιπροσώπους των φυτών. Επομένως, η υποοικογένεια αυτή εμφανίστηκε στον κοινό πρόγονο των αγγειόσπερμων και χλωρόφυτων, πριν 950 εκατομύρια χρόνια περίπου. Περιέχει 43 ακολουθίες, από τις οποίες οι 5 ανήκουν στα χλωρόφυτα. Και αυτή είναι γενικά μια μικρή αλλά αρχαία υποοικογένεια που φαίνεται να μην ανέχεται γονιδιακούς διπλασιασμούς.



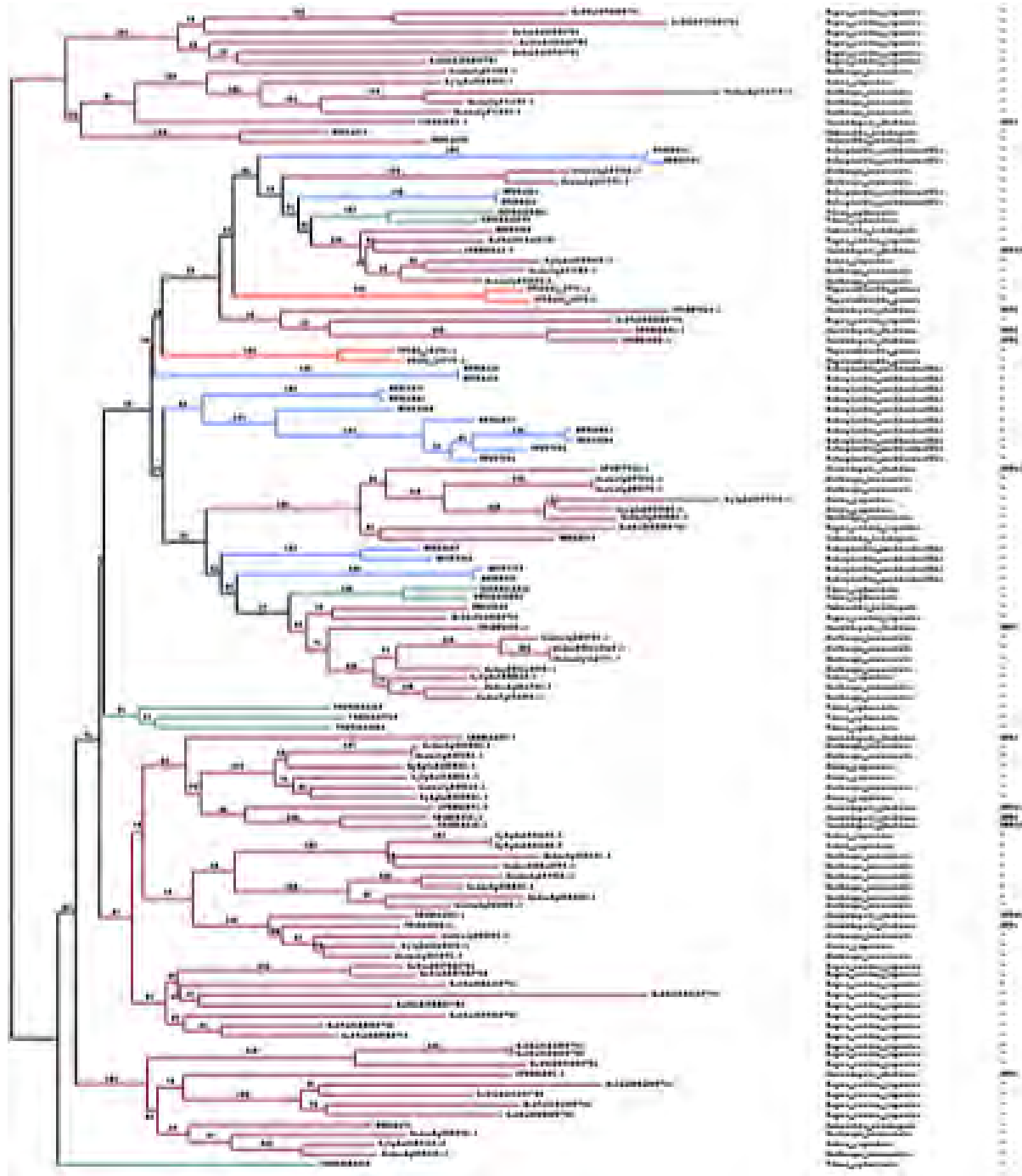
Εικόνα 20. Φυλογενετικό δέντρο της Subfamily5 (pGLcT/SGB1) (BioNJ, Poisson, 500 bootstraps).

Η PMT (Εικόνα 21) περιέχει 56 πρωτεΐνες, που ανήκουν σχεδόν εξολοκλήρου στα αγγειόσπερμα φυτά. Περιέχει μόλις μία ανεξάρτητη ομάδα με 8 πρωτεΐνες από το *Selaginella moellendorffii* και μία από το *Physcomitrella patens*. Επομένως, η υποοικογένεια αυτή εμφανίστηκε στον κοινό πρόγονο των αγγειόσπερμων και βρυόφυτων, πριν 500 εκατομύρια χρόνια περίπου. Γενικά, αυτή η υποοικογένεια υπέστη αρκετούς γονιδιακούς διπλασιασμούς κατά την εξέλιξη των αγγειόσπερμων.



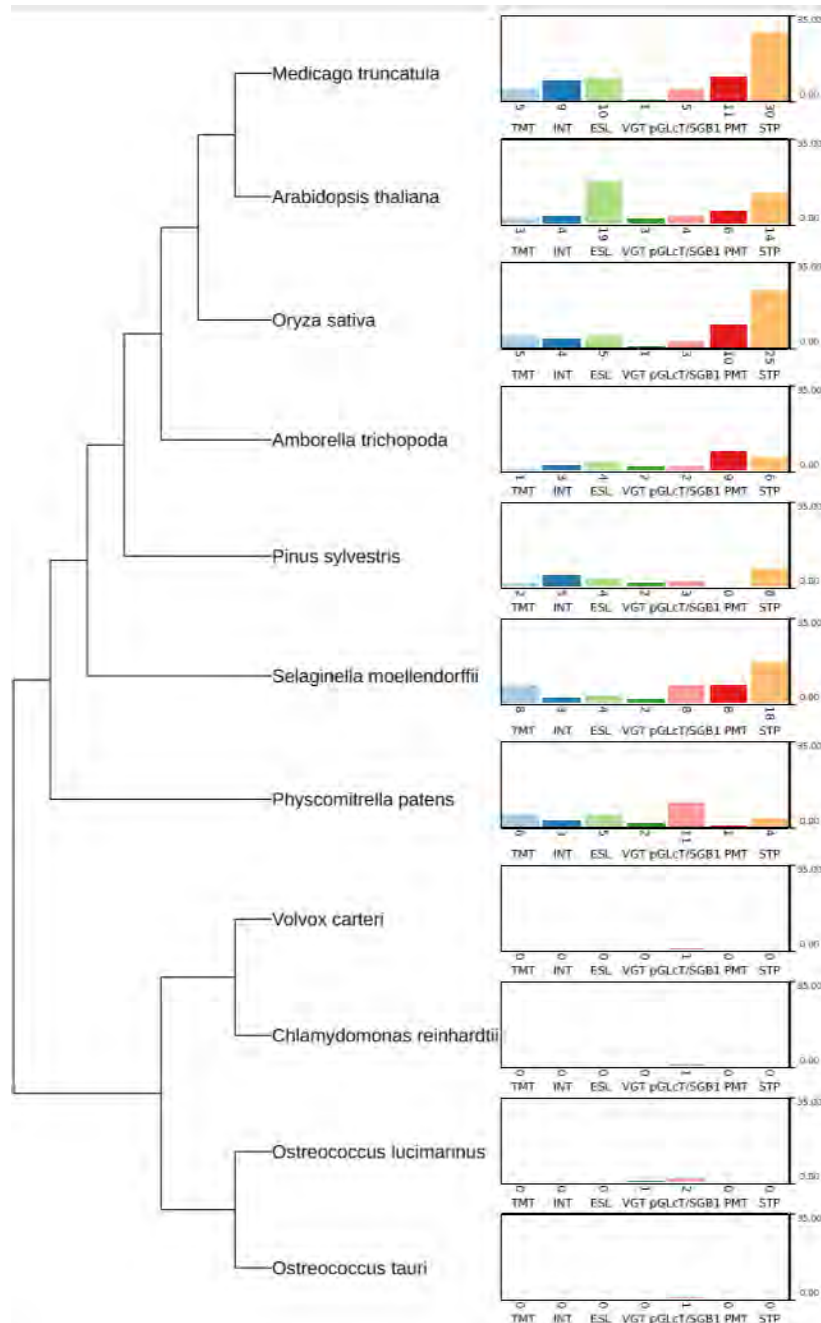
Εικόνα 21. Φυλογενετικό δέντρο της Subfamily6 (PMT) (BioNJ, Poisson, 500 bootstraps).

Η STP (Εικόνα 22) είναι η μεγαλύτερη υποοικογένεια με 117 ακολουθίες, από αγγειόσπερμα, γυμνόσπερμα και βρυόφυτα, Επομένως, η υποοικογένεια αυτή εμφανίστηκε στον κοινό πρόγονο των αγγειόσπερμων και βρυόφυτων, πριν 500 εκατομύρια χρόνια περίπου. Γενικά, αυτή η υποοικογένεια υπέστη αρκετούς γονιδιακούς διπλασιασμούς κατά την εξέλιξη των αγγειόσπερμων.



Εικόνα 22. Φυλογενετικό δέντρο της Subfamily7 (STP) (BioNJ, Poisson, 500 bootstraps).

Τέλος, κατασκευάσαμε το κλαδόγραμμα της Εικόνας 23, το οποίο δημιουργήθηκε σε newick format με βάση την γνωστή φυλογένεση των φυτών. Για κάθε έναν οργανισμό, υπολογίσαμε πόσες από τις ακολουθίες του εμφανίζονται σε κάθε υποοικογένεια και αυτές εμφανίζονται γραφικά δίπλα από κάθε κλαδί. Κάθε μπάρα απεικονίζει μία από τις επτά υποοικογένειες. Για να κατασκευαστεί, αυτό το δέντρο χρησιμοποιήθηκε Python script, με τη βοήθεια του Etetoolkit.



Εικόνα 23. Ο αριθμός ομολόγων των 7 υποοικογενειών σε 12 αντιπροσωπευτικά πρωτεώματα φυτικών οργανισμών.

4. ΣΥΜΠΕΡΑΣΜΑΤΑ

Συνοψίζοντας, η παρούσα ερευνητική εργασία είχε ως στόχο την εξελικτική ανάλυση των μεταφορέων σακχάρων της οικογένειας MFS (2.A.1.1 κατά ταξινόμηση TCDB) στα φυτά. Αρχικά αναπτύξαμε μια γενική βιοπληροφορική μέθοδο εντοπισμού τους σε όλους τους ευκαρυώτες, βασισμένου σε Hidden Markov Models και ικανού να τις διαχωρίσει από άλλες συγγενικές οικογένειες μεταφορέων. Η εξελικτική ανάλυση έδειξε ότι η 2.A.1.1 χωρίζεται σε επτά υποοικογένειες στα φυτά. Οι υποοικογένειες 1 (TMT), 2 (INT), 3 (ESL), 6 (PMT) και 7 (STP) περιλαμβάνουν ακολουθίες και του *Arabidopsis thaliana* και του *Physcomitrella patens*, άρα με βάση τον κοινό πρόγονο τους, υπολογίζουμε ότι οι υποοικογένειες αυτές εμφανίστηκαν περίπου πριν 500 εκατομύρια χρόνια. Οι υποοικογένειες 4 (VGT) και 5 (pGlcT/SGB1), εκτός από ακολουθίες αγγειόσπερμων διαθέτουν και ακολουθίες χλωρόφυτων (π.χ. από *Ostreococcus lucimarinus*, *Ostreococcus tauri*, *Volvox carteri*, *Chlamydomonas reinhardtii*). Άρα με βάση τον κοινό πρόγονο τους, υπολογίζουμε ότι οι υποοικογένειες αυτές εμφανίστηκαν περίπου πριν 950 εκατομύρια χρόνια. Ένα ενδιαφέρον εύρημα είναι ότι οι αρχαιότερες υποοικογένειες δεν είναι και οι μεγαλύτερες. Πιθανόν να υπάρχει εξελικτική πίεση για να μην επιβιώνουν τυχόν διπλασιασμένα γονίδια αυτών των υποοικογενειών στα πρωτεώματα των φυτών.

5. ΒΙΒΛΙΟΓΡΑΦΙΑ

1. “ATPase Pumps.” n.d. Accessed June 19, 2018. <http://www.zoology.ubc.ca/~gardner/ATPase%20pumps.htm>.
2. “BeginnersGuide/Overview - Python Wiki.” n.d. Accessed June 19, 2018. <https://wiki.python.org/moin/BeginnersGuide/Overview>.
3. Büttner, Michael. 2007. “The Monosaccharide Transporter(-like) Gene Family in Arabidopsis.” 2007. <https://febs.onlinelibrary.wiley.com/doi/abs/10.1016/j.febslet.2007.03.016>.
4. César-Razquin, Adrián, Berend Snijder, Tristan Frappier-Brinton, Ruth Isserlin, Gergely Gyimesi, Xiaoyun Bai, Reinhart A. Reithmeier, et al. 2015. “A Call for Systematic Research on Solute Carriers.” *Cell* 162 (3): 478–87. <https://doi.org/10.1016/j.cell.2015.07.022>.
5. Chen, Li-Qing, Lily S. Cheung, Liang Feng, Widmar Tanner, and Wolf B. Frommer. 2015. “Transport of Sugars.” *Annual Review of Biochemistry* 84 (1): 865–94. <https://doi.org/10.1146/annurev-biochem-060614-033904>.
6. Chen, Li-Qing, Bi-Huei Hou, Sylvie Lalonde, Hitomi Takanaga, Mara L. Hartung, Xiao-Qing Qu, Woei-Jiun Guo, et al. 2010. “Sugar Transporters for Intercellular Exchange and Nutrition of Pathogens.” *Nature* 468 (7323): 527–32. <https://doi.org/10.1038/nature09606>.
7. Chien, Sufan. 2013. “Figure 8, [Permeability of the Cell Membrane...].” Text. 2013. <https://www.ncbi.nlm.nih.gov/books/NBK6156/figure/A76034/>.
8. Cooper, Geoffrey M. 2000. “Figure 12.17, Model for the Facilitated Diffusion of Glucose.” Text. 2000. <https://www.ncbi.nlm.nih.gov/books/NBK9847/figure/A1991/>.
9. Cooper, Geoffrey M., and Robert E. Hausman. 2011. *ΤΟ ΚΥΤΤΑΡΟ Μια Μοριακή Προσέγγιση. ΠΡΩΤΗ ΕΛΛΗΝΙΚΗ ΕΚΔΟΣΗ*.
10. Deng, Dong, and Nieng Yan. 2016. “GLUT, SGLT, and SWEET: Structural and Mechanistic Investigations of the Glucose Transporters.” *Protein Science : A Publication of the Protein Society* 25 (3): 546–58. <https://doi.org/10.1002/pro.2858>.
11. Doidy, Joan, Emily Grace, Christina Khün, Frankçoise Simon-Plas, Leonardo Casieri, and Daniel Wipf. 2012. “Sugar Transporters in Plants and in Their Interactions with Fungi: Trends in Plant Science.” 2012. [https://www.cell.com/trends/plant-science/fulltext/S1360-1385\(12\)00061-1?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS1360138512000611%3Fshowall%3Dtrue](https://www.cell.com/trends/plant-science/fulltext/S1360-1385(12)00061-1?_returnURL=https%3A%2F%2Flinkinghub.elsevier.com%2Fretrieve%2Fpii%2FS1360138512000611%3Fshowall%3Dtrue).
12. Eddy, Sean R., Travis J. Wheeler, and the HMMER development team. 2015. *HMMER User’s Guide*. Version 3.1b2. <http://hmmer.org/>.
13. Edgar, Robert C. 2010. “Search and Clustering Orders of Magnitude Faster than BLAST.” *Bioinformatics (Oxford, England)* 26 (19): 2460–61. <https://doi.org/10.1093/bioinformatics/btq461>.
14. Frommer, Wolf B., Davide Sosso, and Li-Qing Chen. 2013. “SWEET Glucoside Transporter Superfamily.” In *Encyclopedia of Biophysics*, 2556–58. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-16712-6_685.
15. Gouy, Manolo, Stéphane Guindon, and Olivier Gascuel. 2010. “SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building.” *Molecular Biology and Evolution* 27 (2): 221–24. <https://doi.org/10.1093/molbev/msp259>.
16. Govers, Roland. 2014. “Cellular Regulation of Glucose Uptake by Glucose Transporter GLUT4.” *Advances in Clinical Chemistry* 66: 173–240.

17. Höglund, Pär J., Karl J.V. Nordström, Helgi B. Schiöth, and Robert Fredriksson. 2011. “The Solute Carrier Families Have a Remarkably Long Evolutionary History with the Majority of the Human Families Present before Divergence of Bilaterian Species.” *Molecular Biology and Evolution* 28 (4): 1531–41. <https://doi.org/10.1093/molbev/msq350>.
18. Huang, Shaohui, and Michael P. Czech. 2007. “The GLUT4 Glucose Transporter.” *Cell Metabolism* 5 (4): 237–52. <https://doi.org/10.1016/j.cmet.2007.03.006>.
19. Huerta-Cepas, Jaime, François Serra, and Peer Bork. 2016. “ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data.” *Molecular Biology and Evolution* 33 (6): 1635–38. <https://doi.org/10.1093/molbev/msw046>.
20. “Introduction - MEME Suite.” n.d. Accessed June 19, 2018. <http://meme-suite.org/>.
21. Kumar, Sudhir, Glen Stecher, Michael Suleski, and S. Blair Hedges. 2017. “TimeTree: A Resource for Timelines, Timetrees, and Divergence Times.” *Molecular Biology and Evolution* 34 (7): 1812–19. <https://doi.org/10.1093/molbev/msx116>.
22. Mueckler, M., C. Caruso, S. A. Baldwin, M. Panico, I. Blench, H. R. Morris, W. J. Allard, G. E. Lienhard, and H. F. Lodish. 1985. “Sequence and Structure of a Human Glucose Transporter.” *Science (New York, N.Y.)* 229 (4717): 941–45.
23. Mueckler, Mike, and Bernard Thorens. 2013. “The SLC2 (GLUT) Family of Membrane Transporters.” *Molecular Aspects of Medicine* 34 (0): 121–38. <https://doi.org/10.1016/j.mam.2012.07.001>.
24. Nagamatsu, S., J. M. Kornhauser, C. F. Burant, S. Seino, K. E. Mayo, and G. I. Bell. 1992. “Glucose Transporter Expression in Brain. CDNA Sequence of Mouse GLUT3, the Brain Facilitative Glucose Transporter Isoform, and Identification of Sites of Expression by in Situ Hybridization.” *The Journal of Biological Chemistry* 267 (1): 467–72.
25. “PLASMA MEMBRANE STRUCTURE AND FUNCTION TRANSPORT ACROSS MEMBRANE. - Ppt Download.” n.d. Accessed June 19, 2018. <http://slideplayer.com/slide/8411421/>.
26. Reddy, Vamsee S., Maksim A. Shlykov, Rostislav Castillo, Eric I. Sun, and Milton H. Saier. 2012. “The Major Facilitator Superfamily (MFS) Revisited.” *The FEBS Journal* 279 (11): 2022–35. <https://doi.org/10.1111/j.1742-4658.2012.08588.x>.
27. Shi, Yigong. 2013. “Common Folds and Transport Mechanisms of Secondary Active Transporters.” *Annual Review of Biophysics* 42: 51–72. <https://doi.org/10.1146/annurev-biophys-083012-130429>.
28. Thorens, Bernard, and Mike Mueckler. 2010. “Glucose Transporters in the 21st Century.” *American Journal of Physiology - Endocrinology and Metabolism* 298 (2): E141–45. <https://doi.org/10.1152/ajpendo.00712.2009>.
29. Wang, Wei, Hui Zhou, Baiquan Ma, Albert Owiti, Schuyler S. Korban, and Yuepeng Han. 2016. “Divergent Evolutionary Pattern of Sugar Transporter Genes Is Associated with the Difference in Sugar Accumulation between Grasses and Eudicots.” *Scientific Reports* 6 (June). <https://doi.org/10.1038/srep29153>.
30. Wright, Ernest M., Donald D.F. Loo, and Bruce A. Hirayama. 2011. “Biology of Human Sodium Glucose Transporters | Physiological Reviews.” 2011. https://www.physiology.org/doi/abs/10.1152/physrev.00055.2009?url_ver=Z39.88-2003&rfr_id=ori%3Arid%3Acrossref.org&rfr_dat=cr_pub%3Dpubmed.
31. Yan, Nieng. 2015. “Structural Biology of the Major Facilitator Superfamily Transporters.” *Annual Review of Biophysics* 44 (1): 257–83. <https://doi.org/10.1146/annurev-biophys-060414-033901>.