



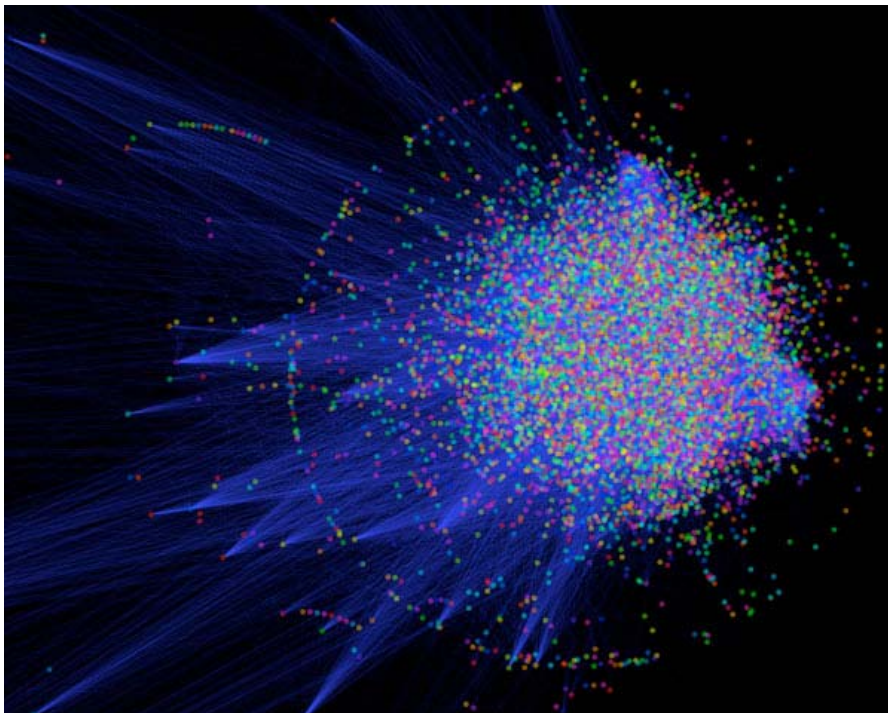
ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

**Πρόγραμμα Μεταπτυχιακών Σπουδών του Τμήματος Βιοχημείας
και Βιοτεχνολογίας «ΕΦΑΡΜΟΓΕΣ ΜΟΡΙΑΚΗΣ ΒΙΟΛΟΓΙΑΣ-
ΜΟΡΙΑΚΗ ΓΕΝΕΤΙΚΗ, ΔΙΑΓΝΩΣΤΙΚΟ ΔΕΙΚΤΕΣ»**

ΓΑΛΛΙΟΠΟΥΛΟΥ ΕΛΕΝΗ

ΤΙΤΛΟΣ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΔΙΑΤΡΙΒΗΣ

**«ΔΙΚΤΥΟ ΠΡΩΤΕΪΝΙΚΩΝ ΑΛΛΗΛΕΠΙΔΡΑΣΕΩΝ
ΤΗΣ ΠΡΩΤΕΪΝΗΣ FRA10AC1»**



ΣΕΠΤΕΜΒΡΙΟΣ 2011

ΤΙΤΛΟΣ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΔΙΑΤΡΙΒΗΣ
**«Δίκτυο πρωτεϊνικών αλληλεπιδράσεων της πρωτεΐνης
FRA10AC1»**

ΜΕΛΗ ΤΡΙΜΕΛΟΥΣ ΕΠΙΤΡΟΠΗΣ

- 1. Μαμούρης Ζήσης:** Καθηγητής Γενετικής Ζωικών Πληθυσμών, Τμήμα Βιοχημείας & Βιοτεχνολογίας, Πανεπιστήμιο Θεσσαλίας.

- 2. Μοσχονάς Νικόλαος:** Καθηγητής Ιατρικής Μοριακής Γενετικής, Ιατρική Σχολή, Πανεπιστήμιο Πάτρας.

- 3. Σαραφίδου Θεολογία:** Λέκτορας Μοριακής Γενετικής Ζωικών Οργανισμών, Τμήμα Βιοχημείας & Βιοτεχνολογίας, Πανεπιστήμιο Θεσσαλίας.

Η παρούσα εργασία πραγματοποιήθηκε στο εργαστήριο Γενετικής, Συγκριτικής και Εξελικτικής Βιολογίας του Τμήματος Βιοχημείας & Βιοτεχνολογίας του Πανεπιστημίου Θεσσαλίας υπό την επίβλεψη του Καθηγητή Γενετικής Ζωικών Πληθυσμών, Μαμούρη Ζήση τον οποίο και θα ήθελα να ευχαριστήσω για την άριστη συνεργασία μας. Επίσης, θα ήθελα να ευχαριστήσω την κ. Σαραφίδου, Λέκτορα Μοριακής Γενετικής Ζωικών Οργανισμών, για τη σωστή ερευνητική καθοδήγηση και τις πολύτιμες συμβουλές της οι οποίες ήταν καθοριστικές για την ολοκλήρωση της εργασίας αυτής καθώς και για τη βοήθεια και την κατανόησή της. Επίσης ευχαριστώ τον Καθηγητή Ιατρικής Μοριακής Γενετικής Νικόλαο Μοσχονά, για τη συμμετοχή του στην τριμελή επιτροπή. Ευχαριστώ θερμά τα μέλη του εργαστηρίου Σταμάτη Κώστα, Γιαννούλη Θέμη, Γεωργίου Στέλλα για την πολύτιμη βοήθεια και συμπαράστασή τους. Τέλος, ευχαριστώ θερμά την οικογένεια μου και τους στενούς φίλους μου για την συνεχή υποστήριξη και κατανόηση.

ΠΕΡΙΕΧΟΜΕΝΑ

1. ΕΙΣΑΓΩΓΗ	1
1.1 Εύθραυστες χρωμοσωματικές θέσεις	2
➤ 1.1.1 Συνήθεις εύθραυστες θέσεις	2
➤ 1.1.2 Σπάνιες εύθραυστες θέσεις	2
➤ 1.1.3 Κλινική σημασία των εύθραυστων θέσεων	3
1.2 Η σπάνια εύθραυστη χρωμοσωματική θέση <i>FRA10A</i>	4
➤ 1.2.1 Εύθραυστη θέση <i>FRA10A</i> και πιθανός παθολογικός φαινότυπος	4
➤ 1.2.2 Το γονίδιο <i>FRA10AC1</i> χαρτογραφείται στην εύθραυστη θέση <i>FRA10A</i>	5
<i>Γενωμική οργάνωση</i>	5
<i>Πρότυπο έκφρασης και εναλλακτικά μετάγραφα</i>	6
➤ 1.2.3 Η πρωτεΐνη <i>FRA10AC1</i>	6
<i>Πρωτεϊνικές αλληλεπιδράσεις της <i>FRA10AC1</i></i>	7
<i>Πληροφορίες από βάσεις δεδομένων</i>	7
<i>Πληροφορίες από πειραματικές προσεγγίσεις</i>	8
<i>Οι πρωτεϊνικές αλληλεπιδράσεις της <i>FRA10AC1</i> συνηγορούν υπέρ της συμμετοχής της στη διαδικασία της συναρμογής (splicing)</i>	9
1.3 Μηχανισμός συναρμογής (splicing)	11
➤ 1.3.1 Πρωτεΐνες-συστατικά του σωματίου συναρμογής	15
1.4 Πρωτεομική (proteomics) και δίκτυα πρωτεϊνικών αλληλεπιδράσεων (protein interactome)	16
➤ 1.4.1 Πρωτεομική	16
➤ 1.4.2 Πρωτεϊνικές αλληλεπιδράσεις	17
<i>Βάσεις δεδομένων πρωτεϊνικών αλληλεπιδράσεων</i>	19
1.5 Σκοπός της εργασίας	20
2. ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ	21
ΥΛΙΚΑ	22
2.1 Καλλιέργεια βακτηρίων	22
2.2 Καλλιέργεια κυττάρων θηλαστικών	22
2.3 Εκκινήτες	23
2.4 Βακτηριακά στελέχη	23
2.5 Κυτταρικές σειρές θηλαστικών	24
2.6 Πλασμιδιακοί φορείς	24
2.7 Πλασμιδιακές κατασκευές	26

ΜΕΘΟΔΟΙ	27
2.8 Απομόνωση νουκλεϊκών οξέων	27
2.8.1 Απομόνωση πλασμιδιακού DNA σε μικρή κλίμακα	27
2.8.2 Απομόνωση πλασμιδιακού DNA σε μεγάλη κλίμακα για αντίδραση αλληλούχησης ή παροδική διαμόλυνση	28
2.9 Τεμαχισμός DNA με περιοριστικές ενδονουκλεάσες	29
2.10 Αντίδραση σύνδεσης (ligation)	30
2.11 Ηλεκτροφόρηση DNA σε πήκτωμα αγαρόζης	30
2.12 Απομόνωση τεμαχίων DNA από πηκτή αγαρόζης και καθαρισμός προϊόντων PCR και πέψης	31
2.13 Αλυσιδωτή αντίδραση πολυμεράσης (PCR)	32
2.14 Μετασχηματισμός βακτηρίων με θερμικό σοκ	33
2.15 Ηλεκτροφόρηση πρωτεϊνών σε αποδιατακτικό πήκτωμα πολυακρυλαμίδης (SDS-PAGE)	34
2.16 Χρώση πηκτωμάτων	36
2.17 Παραγωγή ανασυνδυσασμένων πρωτεϊνών σε βακτήρια και απομόνωση με χρωματογραφία συγγένειας	36
2.18 Καλλιέργεια κυττάρων θηλαστικών (HeLa)	39
➤ Συνθήκες και βασικές αρχές κυτταρικής καλλιέργειας	39
➤ Διατήρηση κυτταροκαλλιιεργειών	39
➤ Κρυοδιατήρηση κυττάρων	40
➤ Επανακαλλιέργεια	41
2.19 Παροδική διαμόλυνση κυττάρων θηλαστικών με πλασμιδιακό DNA	41
➤ Μέθοδος συνκατακρήμνισης του DNA με Ca ₃ (PO ₄) ₂	42
2.20 Μικροσκοπία φθορισμού	43
3. ΑΠΟΤΕΛΕΣΜΑΤΑ – ΣΥΖΗΤΗΣΗ	44
3.1 <i>In silico</i> προσδιορισμός του δικτύου πρωτεϊνικών αλληλεπιδράσεων του σωματίου συναρμογής (spliceosome) με αφετηρία την πρωτεΐνη FRA10AC1	45
➤ <i>In silico</i> ταυτοποίηση των πρωτεϊνών του σωματίου συναρμογής	45
➤ Ανασύσταση του δικτύου πρωτεϊνικών αλληλεπιδράσεων του σωματίου συναρμογής με αφετηρία την πρωτεΐνη FRA10AC1	61
3.2 Κλωνοποίηση των cDNA πλήρους μήκους των υποψήφιων αλληλεπιδρώντων της πρωτεΐνης DGCR14, RED και RACK1 και παραγωγή της ανασυνδυσασμένης GST-DGCR14 σε βακτήρια	86
➤ 3.2.1 Κλωνοποίηση του cDNA του γονιδίου RED	86
➤ 3.2.2 Κλωνοποίηση του cDNA του γονιδίου RACK1	89
➤ 3.2.3 Παραγωγή ανασυνδυσασμένης πρωτεΐνης GST-DGCR14 σε βακτήρια και απομόνωση με χρωματογραφία συγγένειας	90
BIBΛΙΟΓΡΑΦΙΑ	91

ΠΕΡΙΛΗΨΗ

Το γονίδιο *FRA10AC1* έχει χαρτογραφηθεί στη σπάνια αυτοσωματική χρωμοσωματική θέση του ανθρώπου *FRA10A* στην περιοχή 10q23.3 η οποία έχει αναφερθεί ότι σχετίζεται με ήπια νοητική υστέρηση και αναπτυξιακές ανωμαλίες στα άτομα-φορείς. Η εμφάνιση της *FRA10A* οφείλεται στην δραματική επέκταση της τρινουκλεοτιδικής επανάληψης (CGG) η οποία βρίσκεται στην 5' μη μεταφραζόμενη περιοχή του γονιδίου *FRA10AC1* με αποτέλεσμα την υπερμεθυλίωση της περιοχής και τη μεταγραφική καταστολή του γονιδίου. Το γονίδιο *FRA10AC1* είναι μοναδικό στο γονιδίωμα του ανθρώπου και εκφράζεται σε όλους τους ιστούς, κωδικοποιώντας μια συντηρημένη πυρηνική πρωτεΐνη, άγνωστης λειτουργίας για την οποία υπάρχουν ορθόλογα αποκλειστικά στους πολυκύτταρους ευκαρυωτικούς οργανισμούς. Πρόσφατα δεδομένα υποδεικνύουν ότι η *FRA10AC1* αλληλεπιδρά με δυο πρωτεΐνες-συστατικά του σωματίου συναρμογής (DGCR14, SAP145), γεγονός που υποδηλώνει την πιθανή συμμετοχή και της ίδιας σε αυτήν τη διαδικασία, υπόθεση που ενισχύεται και από την ταυτοποίησή της ως υπομονάδα του συμπλόκου C του σωματίου σε δύο πρόσφατες πρωτεομικές μελέτες.

Σε αυτήν την εργασία, στο πλαίσιο της λειτουργικής μελέτης της πρωτεΐνης *FRA10AC1* και ειδικότερα στη διερεύνηση του ρόλου της στη λειτουργία και τη δυναμική του σωματίου συναρμογής, πραγματοποιήθηκε *in silico* η κατασκευή του δικτύου πρωτεϊνικών αλληλεπιδράσεων του σωματίου συναρμογής του ανθρώπου. Η ανασύσταση του δικτύου έγινε με αφετηρία την *FRA10AC1* και τις πειραματικά προσδιορισμένες αλληλεπιδράσεις της με τις πρωτεΐνες DGCR14 και SAP145. Αρχικά, έγινε η συλλογή ενός συνόλου μη επικαλυπτόμενων δεδομένων όλων των πρωτεϊνών που είχαν ταυτοποιηθεί ως συστατικά του σωματίου συναρμογής με βάση το χαρακτηριστικό μοναδικό κωδικό τους, Gene ID. Έτσι, συλλέχθηκαν 479 πρωτεΐνες, συμπεριλαμβανομένου της *FRA10AC1*. Στη συνέχεια, το δίκτυο επεκτάθηκε εξετάζοντας τις δυαδικές πρωτεϊνικές αλληλεπιδράσεις αντλώντας δεδομένα από πειράματα μεγάλης και μικρής κλίμακας στον άνθρωπο, στη *D. melanogaster* και στον *C. elegans* που υπάρχουν καταχωρημένα στις αντίστοιχες βάσεις δεδομένων UniHI, HPRD, Droid και Worm Interactome Database. Βασικό κριτήριο για την εισαγωγή μιας πρωτεΐνης στο δίκτυο ήταν η ταυτοποίησή της ως πρωτεΐνη-συστατικό του σωματίου συναρμογής. Έτσι, κατασκευάστηκε ένα αξιόπιστο δίκτυο το οποίο περιλαμβάνει 255 πρωτεΐνες, ποσοστό 93,5% των οποίων είναι πυρηνικές, και 471 αλληλεπιδράσεις. Τέλος, πραγματοποιήθηκε η κατασκευή των προκαταρκτικών εργαλείων (πλασμιδιακές κατασκευές) για το μελλοντικό πειραματικό έλεγχο συγκεκριμένων πρωτεϊνικών αλληλεπιδράσεων της πρωτεΐνης DGCR14, η οποία αλληλεπιδρά με την *FRA10AC1*.

ABSTRACT

FRA10AC1 gene maps at human chromosomal rare fragile site *FRA10A* at 10q23.3 region. *FRA10A* expression is associated with the expansion of a (CGG)_n repeat located in the 5' UTR region of the gene, resulting in hypermethylation of the region and gene silencing. *FRA10AC1* is unique in the human genome, highly and ubiquitously expressed, and encodes for a conserved nucleoplasmic protein found exclusively in multicellular eukaryotes. By experimental approaches, *FRA10AC1* has been shown to interact with two spliceosomal proteins (DGCR14, SAP145), implying its function in splicing or/and other related processes.

In this study, in the context of *FRA10AC1* functional analysis, and particular in order to evaluate its role in the spliceosomal function and dynamic, we have determined *in silico* the human spliceosomal complex interaction network, starting from the experimental data from *FRA10AC1* binary protein interactions with DGCR14 and SAP145. Firstly, we have created a set of non-overlapping data of all the proteins that have been identified, so far, as spliceosomal subunits by collecting data from all published proteomic studies and converted them at the gene level using the unique Gene ID code. The dataset contains 479 proteins, including *FRA10AC1*. Subsequently, we have retrieved experimental data concerning all binary protein interactions, either from large or small scale studies, in human, *D. melanogaster* and *C. elegans* by investigating UniHI, HPRD, DroID and Worm Interactome Database respectively. For the inclusion of a protein in the network, the main criterion was its identification as a spliceosomal component. Thus, a high-confident spliceosomal protein-protein interaction (PPI) network was constructed which includes 255 proteins (nodes), 93,5% of which are localized to the nucleus, and 471 interactions (edges). Furthermore, preliminary tools (plasmid constructs) have been developed for future experimental approaches of particular protein interactions of DGCR14 protein, which interacts with *FRA10AC1*.

1. ΕΙΣΑΓΩΓΗ

1.1 Εύθραυστες χρωμοσωματικές θέσεις

Οι εύθραυστες χρωμοσωματικές θέσεις είναι γενετικοί τόποι οι οποίοι εμφανίζουν «χάσματα» ή θραύσεις κατά τη μετάφραση της μίτωσης, όταν τα κύτταρα καλλιεργηθούν υπό συγκεκριμένες συνθήκες. Με βάση τη συχνότητά τους στον πληθυσμό, διακρίνονται σε δυο κατηγορίες, στις σπάνιες και στις συνήθεις. Περαιτέρω, με βάση τον τρόπο επαγωγής τους -που γενικά στοχεύει στην αναστολή της αντιγραφής του DNA- διακρίνονται στις ευαίσθητες στο φυλλικό οξύ, που αποτελούν την κύρια κατηγορία των σπάνιων εύθραυστων θέσεων και σε αυτές οι οποίες επάγονται με την προσθήκη δισταμυκίνης (distamycin A) ή βρωμο-δεοξουριδίνης (BrdU).

➤ 1.1.1 Συνήθεις εύθραυστες θέσεις

Οι συνήθεις εύθραυστες θέσεις αποτελούν εγγενές χαρακτηριστικό των χρωμοσωμάτων όλων των ανθρώπων (Lukusa & Fryns 2008). Έως τώρα, έχουν ταυτοποιηθεί κυτταρογενετικά 89 συνήθεις εύθραυστες θέσεις στο γονιδίωμα του ανθρώπου οι οποίες χαρτογραφούνται σε περιοχές που χαρακτηρίζονται από γονιδιωματική αστάθεια, ιδιαίτερα σε καρκινικά κύτταρα (Sutherland & Baker 2000). Από αυτές, οι 13 έχουν χαρακτηριστεί σε μοριακό επίπεδο (Lukusa & Fryns 2008). Ωστόσο, από την ανάλυση των αντίστοιχων αλληλουχιών δεν έχει προκύψει κάποιο αναγνωρίσιμο στοιχείο που να μπορεί να συσχετιστεί με την ευθραυστότητα σε αυτές τις περιοχές. Παρ' όλα αυτά, έχουν περιγραφεί συγκεκριμένα δομικά χαρακτηριστικά τα οποία σχηματίζουν ασυνήθεις δευτεροταγείς δομές οι οποίες είναι δυνατόν να διαταράξουν την αντιγραφή του DNA και την ακετυλίωση των ιστονών με συνέπεια να μην είναι δυνατή η συμπύκνωση της χρωματίνης (Zlotorynski *et al.* 2003, Pichiorri *et al.* 2008). Ως αποτέλεσμα, η χρωματίνη παραμένει σε αποσυσπειρωμένη μορφή, η οποία κυτταρογενετικά εμφανίζεται ως ασυνέχεια-εύθραυστη θέση στα χρωμοσώματα (Jiang *et al.* 2009).

➤ 1.1.2 Σπάνιες εύθραυστες θέσεις

Σε αντίθεση με τις συνήθεις εύθραυστες θέσεις, οι σπάνιες εμφανίζονται σε ένα μικρό μόνο ποσοστό του πληθυσμού (<5%). Έως τώρα έχουν περιγραφεί κυτταρογενετικά 31 σπάνιες θέσεις (Lukusa & Fryns 2008) με συχνότητα εμφάνισης στον πληθυσμό που ποικίλλει στις διάφορες μελέτες από 1/769 (Quack *et al.* 1978) έως 1/15 (Petit *et al.* 1986). Η πλειοψηφία των σπάνιων εύθραυστων θέσεων (24 από τις 31) εκδηλώνονται όταν τα κύτταρα καλλιεργηθούν απουσία φυλλικού οξέος. Έως τώρα έχουν κλωνοποιηθεί δέκα σπάνιες εύθραυστες θέσεις, οκτώ από τις οποίες είναι ευαίσθητες στο φυλλικό οξύ [*FRAXA* (Kremer *et al.* 1991, Verkerk *et al.*

1991, Fu *et al.* 1991), *FRAXE* (Knight *et al.* 1993), *FRAXF* (Parrish *et al.* 1994), *FRA16A* (Nancarrow *et al.* 1994), *FRA10A* (Sarafidou *et al.* 2004), *FRA11B* (Jones *et al.* 1995), *FRA11A* (Debacker *et al.* 2007) και *FRA12A* (Winnepenincx *et al.* 2007, Lukusa & Fryns 2008)] και δύο που εκδηλώνονται παρουσία δισταμυκίνης ή/και BrdU [*FRA10B* (Hewett *et al.* 1998) και *FRA16B* (Yu *et al.* 1997)]. Η πιο γνωστή, ευαίσθητη στο φυλλικό όξύ, σπάνια εύθραυστη θέση είναι η *FRAXA* που χαρτογραφείται στη ζώνη Χq27.3 και σχετίζεται με το σύνδρομο του ευθραύστου Χ (OMIM # 309550), την πιο συχνή αιτία οικογενούς νοητικής υστέρησης. Η μοριακή ανάλυση αυτής της θέσης έδειξε ότι η εκδήλωσή της οφείλεται στην επέκταση πολυμορφικών τρινουκλεοτιδικών επαναλήψεων του τύπου (CGG)*n* που εντοπίζονται στην 5' μη μεταφραζόμενη περιοχή του γονιδίου *FMR1*. Ακολούθως, προσδιορίστηκε ότι η μοριακή βάση των σπάνιων εύθραυστων θέσεων αυτής της κατηγορίας είναι παρόμοια, δηλαδή ότι όλες οφείλονται στην επέκταση επαναλήψεων του τύπου (CGG)*n*. Μια πιθανή εξήγηση για τη μοριακή βάση αυτών των σπάνιων εύθραυστων θέσεων είναι η αναστολή του σχηματισμού των νουκλεοσωμάτων λόγω της επαναλαμβανόμενης αλληλουχίας σε σύγκριση με τις μη επαναλαμβανόμενες (Wang *et al.* 1996). Όπως στην περίπτωση της *FRAXA*, έτσι και για τις θέσεις *FRAXE*, *FRAXF*, *FRA10A* και *FRA12A* προσδιορίστηκε ότι η τρινουκλεοτιδική επανάληψη εντοπίζεται εντός των 5' μη μεταφραζόμενων περιοχών των γονιδίων *FMR2*, *FAM11A*, *FRA10AC1* και *DIP2B* αντίστοιχα. Επίσης, σε ορισμένες περιπτώσεις, η επέκταση των επαναλήψεων πέραν ενός αριθμού βρέθηκε ότι συνδέεται με την υπερμεθυλίωση της περιοχής και την επακόλουθη μεταγραφική καταστολή του γονιδίου (Sutherland & Richards 1995). Η μοριακή ανάλυση των θέσεων *FRA10B* και *FRA16B* έδειξε ότι σχετίζονται με την επέκταση, έως και αρκετών κιλοβάσεων, μινι-δορυφορικών επαναλήψεων πλούσιων σε ΑΤ οι οποίες δύναται να σχηματίζουν δευτεροταγείς δομές που αναστέλλουν την αντιγραφή του DNA (Lukusa & Fryns 2008).

➤ 1.1.3 Κλινική σημασία των εύθραυστων θέσεων

Όσον αφορά στην κλινική σημασία της εμφάνισης των εύθραυστων χρωμοσωματικών θέσεων, οι συνήθεις θέσεις δεν είναι δυνατόν να συσχετιστούν με αυξημένη προδιάθεση για κάποια διαταραχή, καθώς υπάρχουν σε όλους τους ανθρώπους και πιθανότατα σε ομοζυγωτία. Στην περίπτωση των σπάνιων εύθραυστων θέσεων, έως τώρα, μόνο δύο από αυτές, η *FRAXA* και η *FRAXE*, που χαρτογραφούνται στο χρωμόσωμα Χ, έχουν συσχετιστεί σαφώς με παθολογικό φαινότυπο νοητικής υστέρησης. Για τις υπόλοιπες σπάνιες εύθραυστες θέσεις, υπάρχουν διάφορες μελέτες που αναφέρουν αυξημένη συχνότητα εμφάνισής τους σε

άτομα με διάφορες διαταραχές, κυρίως νοητική υστέρηση, χωρίς όμως να υπάρχει έως τώρα σαφής συσχέτιση γονοτύπου-φαινοτύπου για κάποια από αυτές. Εξάλλου, τρεις σπάνιες αυτοσωματικές εύθραυστες θέσεις, οι *FRA10B*, *FRA16B*, *FRA17A*, έχουν βρεθεί και σε ομοζυγωτία η οποία δεν προσδίδει κάποιο παθολογικό φαινότυπο. Αυτό υποδεικνύει ότι οι συγκεκριμένες εύθραυστες θέσεις δεν επηρεάζουν τη λειτουργία γονιδίων που είναι απαραίτητα για την ανάπτυξη (Sutherland & Baker 2000). Ωστόσο, ορισμένες σπάνιες αυτοσωματικές εύθραυστες θέσεις-ευαίσθητες στο φυλλικό οξύ δεν έχουν βρεθεί σε ομοζυγωτία, παρόλο που, με βάση τη συχνότητα των ετεροζυγωτών, τέτοια άτομα θα ήταν αναμενόμενο να ταυτοποιηθούν. Το γεγονός αυτό πιθανώς υποδηλώνει ότι μία τέτοια κατάσταση δεν είναι συμβατή με τη ζωή (Sutherland & Baker 2000).

1.2 Η σπάνια εύθραυστη χρωμοσωματική θέση *FRA10A*

Η *FRA10A* ανήκει στις σπάνιες αυτοσωματικές εύθραυστες θέσεις που εμφανίζονται όταν τα κύτταρα καλλιεργηθούν απουσία φυλλικού οξέος στο θρεπτικό μέσο. Αποτελεί την πιο συχνά εμφανιζόμενη σπάνια εύθραυστη θέση στο γονιδίωμα του ανθρώπου με συχνότητα που εκτιμάται στο 1/600, σε σύνολο 30.000 περίπου ατόμων που προέρχονται από διαφορετικές μελέτες και πληθυσμιακές ομάδες (Kahkonen *et al.* 1989). Η παρατηρούμενη συχνότητα ποικίλλει σημαντικά μεταξύ των διαφορετικών μελετών από 0% (Takahashi *et al.* 1988) έως 2% (Petit *et al.* 1986), γεγονός που μπορεί να οφείλεται στην πληθυσμιακή προέλευση των ατόμων ή/και στις πειραματικές διαδικασίες που εφαρμόστηκαν. Σε όσες περιπτώσεις εξετάστηκε η γονεϊκή προέλευση της *FRA10A*, διαπιστώθηκε ότι η κληρονόμησή της ήταν μητρικής προέλευσης (Sutherland 1982 & 1985, Kahkonen *et al.* 1989).

➤ 1.2.1 Εύθραυστη θέση *FRA10A* και πιθανός παθολογικός φαινότυπος

Μέχρι τώρα, η εμφάνιση της *FRA10A* δεν έχει συσχετιστεί με συγκεκριμένο παθολογικό φαινότυπο. Ωστόσο, διάφορες μελέτες δείχνουν ότι η έκφρασή της μπορεί να σχετίζεται με νοητική υστέρηση, νευρολογικές διαταραχές και αναπτυξιακές ανωμαλίες, όπως δυσμορφικά χαρακτηριστικά, μικρό ανάστημα, υποσπαδία και δυσχέρεια λόγου (Mavrou *et al.* 1991, Petit *et al.* 1986, Sutherland 1982 & 1985, Kahkonen *et al.* 1989, Sarafidou *et al.* 2004). Αυτός ο συσχετισμός προέρχεται από μελέτες που ελέγχουν την παρουσία εύθραυστων θέσεων στον καρυότυπο ατόμων με νοητική υστέρηση σε σχέση με υγιή άτομα ή, σε ορισμένες περιπτώσεις, με άτομα από το γενικό πληθυσμό που έχουν επιλεγεί τυχαία. Σε ορισμένες μελέτες, η *FRA10A* ανιχνεύθηκε αποκλειστικά στα προσβεβλημένα άτομα (Sutherland 1985, Petit *et al.* 1986, Mavrou *et al.* 1991) ενώ σε άλλες εντοπίστηκαν

και ορισμένα υγιή άτομα-φορείς (Sutherland 1982, Kahkonen *et al.* 1989, Sarafidou *et al.* 2004), γεγονός που μπορεί να υποδεικνύει ότι η ετεροζυγωτία αυτής της θέσης δεν συσχετίζεται με παθολογικό φαινότυπο. Εναλλακτικά, αυτό το γεγονός μπορεί να οφείλεται σε μειωμένη διεισδυτικότητα του παθολογικού φαινοτύπου στον πληθυσμό έτσι ώστε να μην εμφανίζονται σε ένα ποσοστό ατόμων-φορέων τα κλινικά συμπτώματα. Το φαινόμενο της μειωμένης διεισδυτικότητας έχει δειχθεί στην περίπτωση του συνδρόμου του εύθραυστου X που σχετίζεται με την εκδήλωση νοητικής υστέρησης και την εμφάνιση της εύθραυστης χρωμοσωματικής θέσης *FRAXA*. Έτσι, συγκεκριμένου στελέχους ποντίκια-κnockout για το γονίδιο *FMR1*, δεν εμφάνιζαν παθολογικό φαινότυπο ενώ ένα διαφορετικό στέλεχος ποντικών knockout εμφάνιζαν τα χαρακτηριστικά συμπτώματα του συνδρόμου (Dobkin *et al.* 2000). Τέλος, η συσχέτιση της *FRA10A* με το φαινότυπο της νοητικής υστέρησης μπορεί να οφείλεται σε μεροληψία κατά τη διακρίβωση (ascertainment bias) διότι στις περισσότερες μελέτες επιλέχθηκαν προς εξέταση άτομα προερχόμενα π.χ. από ιδρύματα περίθαλψης ατόμων με νοητική υστέρηση, σχολεία ειδικής εκπαίδευσης κ.λπ. και όχι τυχαίο δείγμα από το γενικό πληθυσμό.

➤ **1.2.2 Το γονίδιο *FRA10AC1* χαρτογραφείται στην εύθραυστη θέση *FRA10A***

Ο προσδιορισμός της μοριακής βάσης της εύθραυστης θέσης *FRA10A* οδήγησε στην κλωνοποίηση του γονιδίου *FRA10AC1*, βάσει της χρωμοσωματικής του θέσης και του εντοπισμού της πολυμορφικής επανάληψης (CGG)_n στην 5' μη μεταφραζόμενη περιοχή του. Όπως προαναφέρθηκε, όπως ισχύει και για τις άλλες εύθραυστες θέσεις αυτής της κατηγορίας, η κυτταρογενετική έκφραση της *FRA10A* οφείλεται στην επέκταση κατά ~200 φορές της πολυμορφικής επανάληψης (CGG)_n η οποία οδηγεί στην υπερμεθυλίωση της περιοχής και συνεπάγεται τη μεταγραφική καταστολή του αντίστοιχου αλληλομόρφου.

➤ Γενωμική οργάνωση

Ο γενετικός τόπος *FRA10AC1* καταλαμβάνει ~35 kb στη χρωμοσωματική περιοχή 10q23.3 του ανθρώπου και χαρτογραφείται ανάμεσα στο γονίδιο *PDE6C* το οποίο κωδικοποιεί τη φωσφοδιεστεράση 6C και στο γονίδιο *LG11* που σχετίζεται με μία κληρονομική μορφή επιληψίας. Το γονίδιο *FRA10AC1* περιλαμβάνει 19 εξόνια, 5 από τα οποία (13-17) υπόκεινται σε εναλλακτική συναρμογή. Οι τρινουκλεοτιδικές επαναλήψεις (CGG)_n, εντοπίζονται στο εξόνιο 1 στην 5' μη μεταφραζόμενη περιοχή του γονιδίου και είναι πολυμορφικές στο γενικό πληθυσμό, με αριθμό επαναλήψεων

που κυμαίνεται από 8 έως 14. Το κωδικόνιο έναρξης της μετάφρασης βρίσκεται στην αρχή του εξονίου 2 και το κωδικόνιο λήξης στο τελευταίο εξόνιο.

➤ Πρότυπο έκφρασης και εναλλακτικά μετάγραφα

Το cDNA του γονιδίου *FRA10AC1* έχει μήκος 1344 bp. Με ανάλυση κατά Northern, χρησιμοποιώντας δυο διαφορετικούς ανιχνευτές οι οποίοι ανιχνεύουν τις πρώτες 616 bp στο 5' άκρο του cDNA του γονιδίου *FRA10AC1* και ένα μέρος της αλληλουχίας του εξονίου 19, βρέθηκε ότι το γονίδιο *FRA10AC1* είναι μεταγραφικά ενεργό σε όλους τους ιστούς ενήλικων ατόμων παρουσιάζοντας υψηλότερα επίπεδα έκφρασης σε όργανα με υψηλή μεταγραφική δραστηριότητα (εγκέφαλος, καρδιά, σκελετικοί μυς, ήπαρ). Επίσης, με πειράματα RT-PCR, χρησιμοποιώντας εκκινητές για τα εξόνια 11 και 19, επιβεβαιώθηκε η μεταγραφική ενεργότητα του γονιδίου *FRA10AC1* σε όλους τους ιστούς. Με το τελευταίο πείραμα σε συνδυασμό με αλληλούχηση προσδιορίστηκε επίσης η ύπαρξη πέντε εναλλακτικών μεταγράφων. Ειδικότερα, το κύριο μετάγραφο περιλαμβάνει τα εξόνια 1-12, 18 και 19. Ένα από τα εναλλακτικά μετάγραφα εντοπίστηκε σχεδόν σε όλους τους ιστούς και περιλαμβάνει τα εξόνια 13-16. Επιπλέον, βρέθηκαν τρία εναλλακτικά μετάγραφα αποκλειστικά στις γονάδες, που περιλαμβάνουν το εξόνιο 16 σε συνδυασμό με τα εξόνια 13, 14, 15 αντίστοιχα. Όλα αυτά τα εναλλακτικά μετάγραφα αντιστοιχούν σε διαφορετική ακολουθία αμινοξέων στο καρβοξυ-τελικό άκρο, γεγονός που πιθανόν να συνεπάγεται αντίστοιχο αριθμό ισομορφών της πρωτεΐνης *FRA10AC1*, ετερογενών στο καρβοξυ-τελικό άκρο (Sarafidou *et al*, 2004).

➤ **1.2.3 Η πρωτεΐνη *FRA10AC1***

Το γονίδιο *FRA10AC1* κωδικοποιεί μια πρωτεΐνη 315 αμινοξέων που εμφανίζει ορθόλογα σε πλήθος ευκαρυωτικών, αποκλειστικά πολυκύτταρων, ζωικών και φυτικών οργανισμών, αλλά όχι σε προκαρυωτικούς οργανισμούς. Πιο συγκεκριμένα, η ακολουθία της πρωτεΐνης του ανθρώπου είναι ταυτόσημη κατά 86% με την ακολουθία της ορθόλογης πρωτεΐνης του ποντικού και παρουσιάζει ομολογία της τάξεως του 92%. Επιπλέον, παρουσιάζει ομοιότητα 35% με την ορθόλογη πρωτεΐνη του *C. elegans* και 58% με του *Anopheles gambiae*. Η κεντρική περιοχή της πρωτεΐνης (αμινοξέα 105-240) είναι αυτή που εμφανίζει τον υψηλότερο βαθμό ταυτότητας, που κυμαίνεται από 45 έως 67% με τα υπόλοιπα ορθόλογα μόρια.

Κανένα από τα ορθόλογα μόρια δεν έχει μελετηθεί λειτουργικά και δε φέρει κάποιο χαρακτηριστικό μοτίβο που να εντάσσει την πρωτεΐνη σε συγκεκριμένη οικογένεια πρωτεϊνών. Ωστόσο προσδιορίστηκε *in silico* ένα διμερές σήμα πυρηνικού εντοπισμού, στην περιοχή που περιλαμβάνει τα αμινοξέα 223-238 και 245-262, η οποία ανήκει σε μία ευρύτερη περιοχή-πλούσια σε λυσίνη (αμινοξέα 129-269)

(Sarafidou *et al*, 2004). Ο πυρηνικός εντοπισμός της πρωτεΐνης, και ειδικότερα ο εντοπισμός της στο πυρηνόπλασμα, επιβεβαιώθηκε και πειραματικά, τόσο για την ανασυνδυασμένη πρωτεΐνη μετά από πειράματα παροδικής διαμόλυνσης κυττάρων θηλαστικών, όσο και για την ενδογενή πρωτεΐνη, η οποία ανιχνεύθηκε με ειδικό πολυκλωνικό αντίσωμα.

➤ Πρωτεϊνικές αλληλεπιδράσεις της FRA10AC1

Μία προσέγγιση για να ενταχθεί μία πρωτεΐνη άγνωστης λειτουργίας, όπως η FRA10AC1, σε συγκεκριμένο βιολογικό μονοπάτι είναι να προσδιοριστούν οι πρωτεΐνες με τις οποίες αλληλεπιδρά. Συνήθως, για αυτό το σκοπό εφαρμόζεται το σύστημα δύο υβριδίων του σακχαρομύκητα με το οποίο εξετάζονται αποκλειστικά δυαδικές πρωτεϊνικές αλληλεπιδράσεις. Παρόλο που πλήθος δεδομένων από πειράματα μεγάλης κλίμακας πρωτεϊνικών αλληλεπιδράσεων υπάρχουν σε αντίστοιχες βάσεις δεδομένων, τα δεδομένα αυτά δεν μπορούν να θεωρηθούν πλήρως αξιόπιστα γιατί παρουσιάζουν υψηλό ποσοστό ψευδώς θετικών αποτελεσμάτων και αντιπροσωπεύουν ένα σχετικά μικρό ποσοστό του συνόλου των πρωτεϊνικών αλληλεπιδράσεων του ανθρώπου (Von Mering *et al.* 2002, Suthram *et al.* 2006). Κατά συνέπεια, θα πρέπει να εφαρμόζονται, συνδυαστικά, στοχευμένα πειράματα μικρής κλίμακας, τα αποτελέσματα των οποίων να επιβεβαιώνονται με συμπληρωματικές προσεγγίσεις.

Πληροφορίες από βάσεις δεδομένων

Από τη διερεύνηση των βάσεων δεδομένων που περιλαμβάνουν καταχωρήσεις σχετικά με τις αλληλεπιδράσεις των πρωτεϊνών του ανθρώπου (HPRD, <http://www.hprd.org/>), προκύπτει ότι η FRA10AC1 αλληλεπιδρά άμεσα (δυαδικές αλληλεπιδράσεις) με 10 πρωτεΐνες (Stelzl *et al.* 2005). Αυτές οι αλληλεπιδράσεις προέρχονται αποκλειστικά από πειράματα μεγάλης κλίμακας (high-throughput). Με βάση τον υποκυτταρικό εντοπισμό της FRA10AC1 αποκλειστικά στο πυρηνόπλασμα, οι αλληλεπιδράσεις με επτά από τις αναφερόμενες πρωτεΐνες (EEF1D, MAP3K12, GDPD2, LPPR1, C21orf33, VPS29, GLRX3) πιθανώς δεν είναι αξιόπιστες καθώς οι τελευταίες εντοπίζονται σε διαφορετικά κυτταρικά διαμερίσματα (βάσει του Gene Ontology Annotation, <http://www.ebi.ac.uk/GOA/>). Από τις υπόλοιπες τρεις πρωτεΐνες που αναφέρονται, η μία (c7orf25) είναι εντελώς άγνωστης λειτουργίας και τοπολογίας, η δεύτερη (ADPRHL2, Gene ID: 54936) ανήκει στην οικογένεια των ενζύμων ADP-ριβοζη-γλυκο-υδρολασών και, παρόλο που εντοπίζεται και στον πυρήνα, το κύριο οργανίδιο όπου δρα είναι τα μιτοχόνδρια και, τέλος, η τρίτη (HABP4, Ki-1/57) είναι πυρηνική. Ενδιαφέρον είναι ότι η τελευταία πρωτεΐνη έχει

συσχετιστεί με τη διαδικασία της συναρμογής (splicing) (Bressan GC *et al.* 2009), γεγονός που συνάδει με τα στοιχεία που συνηγορούν υπέρ της συμμετοχής της FRA10AC1 σε αυτήν τη διαδικασία (βλ. παρακάτω) και για αυτό πιθανώς χρειάζεται περαιτέρω διερεύνηση μέσω πειραματικών προσεγγίσεων.

Πληροφορίες από πειραματικές προσεγγίσεις

Η πειραματική προσέγγιση για την ταυτοποίηση των πρωτεϊνικών αλληλεπιδράσεων της FRA10AC1 με το σύστημα των δύο υβριδίων του σακχαρομύκητα οδήγησε στην εύρεση δύο πρωτεϊνών, της DGCR14 και της SAP145. Οι αλληλεπιδράσεις αυτές επιβεβαιώθηκαν με δοκιμή συν-κατακρήμνισης κατά την οποία η πρωτεΐνη FRA10AC1 είχε παραχθεί σε σύντηξη με την τρανσφεράση της S-γλουταθειόνης (GST) σε βακτήρια ενώ η άλλη πρωτεΐνη του ζεύγους είχε παραχθεί ραδιενεργά σημασμένη (μεθειονίνη, S³⁵) με μεταγραφή-μετάφραση *in vitro*. Επιπλέον, η αλληλεπίδραση της FRA10AC1 με την DGCR14 επιβεβαιώθηκε και με συν-ανοσοκατακρήμνιση, κατά την οποία πραγματοποιήθηκε κατακρήμνιση της πρωτεΐνης FRA10AC1 από πρωτεϊνικό εκχύλισμα κυττάρων θηλαστικών (*HeLa*) με το αντίσωμα αντι-FRA10AC1 και ακολούθως ανιχνεύθηκε στο κατακρήμνισμα η πρωτεΐνη DGCR14 με ανάλυση Western με το αντίστοιχο αντίσωμα. Η ίδια διαδικασία ακολουθήθηκε και αντίστροφα για τις δυο πρωτεΐνες γεγονός που έδειξε ότι η FRA10AC1 και η DGCR14 εντοπίζονται *in vivo* στο ίδιο σύμπλοκο. Η αντίστοιχη δοκιμή συν-ανοσοκατακρήμνισης δεν ήταν επιτυχής για την SAP145 πιθανώς διότι οι αλληλεπιδράσεις μεταξύ τους μπορεί να είναι είτε παροδικές, είτε ασθενείς ή να υπάρχουν λίγα πρωτεϊνικά μόρια που αλληλεπιδρούν με αποτέλεσμα να μην είναι ανιχνεύσιμα (Sarafidou *et al*, υπό συγγραφή).

Επίσης, με συν-κατακρήμνιση εντοπίστηκαν οι περιοχές αλληλεπίδρασης των πρωτεϊνών FRA10AC1 και DGCR14, κατά την οποία διαφορετικά τμήματα της πρωτεΐνης DGCR14 παρήχθησαν σε σύντηξη με την τρανσφεράση της S-γλουταθειόνης (GST) σε βακτήρια και διαφορετικά τμήματα της πρωτεΐνης FRA10AC1 παρήχθησαν σε σύντηξη με την EGFP σε κύτταρα θηλαστικών (*HeLa*) έπειτα από παροδική διαμόλυνση. Από τις δοκιμές συν-κατακρήμνισης βρέθηκε ότι η κεντρική περιοχή της FRA10AC1 (αμινοξέα 96-253), η οποία είναι και αυτή που εμφανίζει την υψηλότερη συντήρηση μεταξύ των ορθολόγων μορίων, αλληλεπιδρά με την καρβοξυτελική περιοχή της DGCR14 (αμινοξέα 294-472) η οποία περιέχει ένα μοτίβο coiled-coil, δομή που συμμετέχει σε πρωτεϊνικές αλληλεπιδράσεις.

Επιπλέον, τα αποτελέσματα των αλληλεπιδράσεων είναι σε συμφωνία με τον υποκυτταρικό εντοπισμό στο πυρηνόπλασμα των τριών πρωτεϊνών FRA10AC1,

DGCR14 και SAP145, ο οποίος εξετάστηκε σε κύτταρα *HeLa* (Sarafidou *et al*, υπό συγγραφή).

Οι πρωτεϊνικές αλληλεπιδράσεις της FRA10AC1 συνηγορούν υπέρ της συμμετοχής της στη διαδικασία της συναρμογής (splicing)

Το γονίδιο *DGCR14* έχει ταυτοποιηθεί μέσω κλωνοποίησης θέσης ως πιθανό υποψήφιο για το σύνδρομο DiGeorge, ένα σύνδρομο με κύριο χαρακτηριστικό τις καρδιακές ανωμαλίες οι οποίες συνοδεύονται από χαρακτηριστικό προσωπείο. Η κύρια μοριακή αιτία του συνδρόμου είναι ένα έλλειμμα μεγέθους 3 Mb στη χρωμοσωματική περιοχή 22q11.2 (Rizzu *et al.* 1996, Gong *et al.* 1997, OMIM #188400). Τελικά, βρέθηκε ότι το γονίδιο *DGCR14* δεν σχετίζεται, τουλάχιστον με τα κύρια χαρακτηριστικά του συνδρόμου, αλλά ότι το κύριο υπεύθυνο γονίδιο είναι το *Tbx*, το οποίο επίσης χαρτογραφείται στην περιοχή αυτή (Huh & Ornitz 2010). Μέχρι τώρα, δεν υπάρχουν μελέτες που να προσδιορίζουν λειτουργικά το ρόλο της πρωτεΐνης *DGCR14*. Ωστόσο, υπάρχουν πληροφορίες από βιοχημικά πειράματα απομόνωσης του σωματίου συναρμογής και χαρακτηρισμού των συστατικών του με φασματομετρία μάζας που υποδεικνύουν την πιθανή λειτουργία της πρωτεΐνης. Έτσι, από πέντε ανεξάρτητες μελέτες, η *DGCR14* προσδιορίστηκε κατ' επανάληψη ως πρωτεΐνη-συστατικό του σωματίου συναρμογής (Rappsilber *et al.* 2002, Zhou *et al.* 2002, Makarova *et al.* 2004, Bessonov *et al.* 2008, Bessonov *et al.* 2010) και μάλιστα ως συστατικό του συμπλόκου C (Bessonov *et al.* 2008, Bessonov *et al.* 2010).

Η πρωτεΐνη *SAP145* κωδικοποιείται από το γονίδιο *SF3B2* και αποτελεί την υπομονάδα 2 του παράγοντα συναρμογής 3b, ενός πρωτεϊνικού συμπλόκου του σωματίου συναρμογής. Η πρωτεΐνη *SAP145* αποτελείται από 872 αμινοξέα, είναι υδρόφιλη, το αμινοτελικό της άκρο είναι πλούσιο σε προλίνη και το καρβοξυτελικό της άκρο πλούσιο σε γλουταμινικό οξύ. Η πρωτεΐνη αυτή δεν διαθέτει γνωστά μοτίβα πρόσδεσης του RNA. Ο παράγοντας συναρμογής 3b μαζί με τον παράγοντα συναρμογής 3a και την υπομονάδα 12S του RNA σχηματίζουν το μικρό πυρηνικό ριβονουκλεοπρωτεϊνικό σύμπλοκο U2 snRNP. Το σύμπλοκο συναρμογής 3b/3a προσδένεται στο πρόδρομο mRNA σε μια αλληλουχία ανοδικά της θέσης διακλάδωσης του ιντρονίου (branch site) και βοηθάει το μικρό πυρηνικό ριβονουκλεοπρωτεϊνικό σύμπλοκο U2 snRNP να «αγκιστρωθεί» στο πρόδρομο mRNA (Gozani *et al.* 1996). Η πρωτεΐνη *SAP145* αλληλεπιδρά άμεσα με την υπομονάδα 4 του παράγοντα συναρμογής 3b, *SAP49* (Champion-Arnaud & Reed 1994). Η πρωτεΐνη *SAP145* χαρακτηρίστηκε κι αυτή ως πρωτεΐνη-συστατικό του σωματίου συναρμογής σε όλες τις μελέτες που περιλαμβάνουν βιοχημικά πειράματα απομόνωσης του σωματίου συναρμογής και χαρακτηρισμού των συστατικών του με

φασματομετρία μάζας (Rappsilber *et al.* 2002, Zhou *et al.* 2002, Jurica *et al.* 2002, Hartmuth *et al.* 2002, Makarov *et al.* 2002, Makarova *et al.* 2004, Chen *et al.* 2007, Herold *et al.* 2008, Bessonov *et al.* 2008, Bessonov *et al.* 2010) και μάλιστα επίσης ως συστατικό του συμπλόκου C (Jurica *et al.* 2002, Herold *et al.* 2008, Bessonov *et al.* 2008, Bessonov *et al.* 2010)

Επομένως, εξετάζοντας τις πρωτεϊνικές αλληλεπιδράσεις της FRA10AC1 εξάγεται το συμπέρασμα ότι πιθανώς και η FRA10AC1 σχετίζεται με τη διαδικασία της συναρμογής. Μάλιστα, στις πιο πρόσφατες και εκτενείς εργασίες απομόνωσης και προσδιορισμού των συστατικών του σωματίου συναρμογής με βιοχημικές μεθόδους βρέθηκε ότι και αυτή αποτελεί συστατικό του σωματίου συναρμογής και ειδικά του συμπλόκου C (Bessonov *et al.* 2008, Bessonov *et al.* 2010) όπως και οι δύο άλλες. Το γεγονός αυτό συνάδει με τον υποκυτταρικό εντοπισμό της και με το ότι η πρωτεΐνη είναι συντηρημένη στους ευκαρυωτικούς οργανισμούς οι οποίοι, σε αντίθεση με τους προκαρυωτικούς, χρησιμοποιούν τη διαδικασία της συναρμογής για την ωρίμανση των πρόδρομων μορίων mRNA.

Η συσχέτιση της FRA10AC1 με το μηχανισμό συναρμογής και, με την ευρύτερη έννοια, με το μεταβολισμό του RNA είναι ενδιαφέρουσα επειδή πλέον αρκετές πρωτεΐνες-συστατικά ριβονουκλεοπρωτεϊνικών συμπλόκων φαίνεται να σχετίζονται με νευρολογικές διαταραχές, όπως για παράδειγμα η πρωτεΐνη TDP-43 που σχετίζεται με την νευρομυική διαταραχή πλάγια μωσατροφική σκλήρυνση (amyotrophic lateral sclerosis, ALS), η SMN η οποία είναι απαραίτητη για την επιβίωση των νευρώνων και η FMRP που σχετίζεται με το σύνδρομο του ευθραύστου X (Ule 2008). Αυτά τα δεδομένα ενισχύουν την πιθανή συσχέτιση της FRA10AC1 με το φαινότυπο της νοητικής υστέρησης που παρατηρείται στα άτομα-φορείς της εύθραυστης θέσης *FRA10A*.

Εναλλακτικά, η πρωτεΐνη FRA10AC1 θα μπορούσε να σχετίζεται με άλλες συναφείς διαδικασίες καθώς στα βιοχημικά πειράματα απομόνωσης του σωματίου συναρμογής ταυτοποιήθηκαν, εκτός από παράγοντες συναρμογής, και διάφορες άλλες πρωτεΐνες που δεν συμμετέχουν σε αυτήν καθ' αυτήν τη συναρμογή αλλά διαθέτουν μοτίβα ή περιοχές (domains) σύνδεσης στο RNA, πρωτεΐνες οι οποίες συμμετέχουν σε άλλα στάδια της γονιδιακής έκφρασης ή σε άλλα στάδια ωρίμανσης του πρόδρομου mRNA (προσθήκη δομής καλύπτρας στο 5' άκρο, πολυαδενυλίωση) καθώς και πρωτεΐνες (νουκλεοπορίνες) που συμμετέχουν στην έξοδο του ώριμου mRNA από τον πυρήνα (Reed 2003).

1.3 Μηχανισμός συναρμογής (splicing)

Τα γονίδια των ευκαρυωτικών οργανισμών εμφανίζουν μεγάλη ετερογένεια ως προς τον αριθμό και τα μεγέθη των ιντρονίων που φέρουν. Τα εξόνια είναι σχετικά μικρά σε μήκος (~100-200 bp), ενώ τα ιντρόνια είναι σχετικά μεγάλα (>1 kb). Το πρόδρομο μετάγραφο έχει την ίδια οργάνωση με το γονίδιο και ονομάζεται πρόδρομο mRNA (pre-mRNA). Η διαδικασία απομάκρυνσης των ιντρονίων ονομάζεται συναρμογή του RNA (splicing). Η συναρμογή λαμβάνει χώρα στον πυρήνα παράλληλα με τις άλλες τροποποιήσεις του νεοσυντιθέμενου RNA. Το μετάγραφο καλύπτεται στο 5' άκρο, πολυαδενολιώνεται στο 3' άκρο και μεταφέρεται στο κυτταρόπλασμα διαμέσου των πυρηνικών πόρων. Η διαδικασία της συναρμογής επιτελείται από το σωματίο συναρμογής, ένα μακρομοριακό σύμπλοκο το οποίο συναρμολογείται διαδοχικά πάνω στο πρόδρομο mRNA (Lewin B. 2004). Τα επιμέρους συστατικά του σωματίου μπορούν να απομονωθούν κατά τα διαδοχικά στάδια της δημιουργίας του διαχωρίζοντας με βιοχημικές μεθόδους κλάσματα, που περιέχουν σύμπλοκα διαφορετικών μεγεθών. Η συσκευή συναρμογής περιλαμβάνει τόσο πρωτεΐνες όσο και μικρά πυρηνικά RNAs (U snRNAs, uridine-rich small nuclear RNAs) τα οποία απαντώνται σε ριβονουκλεοπρωτεϊνικά σύμπλοκα που ονομάζονται U snRNPs. Στα θηλαστικά έχουμε δύο κατηγορίες πρωτεϊνών στα U snRNPs, τις μικρές πυρηνικές ριβονουκλεοπρωτεΐνες (snRNPs) και τους παράγοντες συναρμογής (μη-snRNPs) (Will & Luhrmann, 2001).

Τα ριβονουκλεοπρωτεϊνικά σύμπλοκα snRNP που συμμετέχουν στη συναρμογή είναι τα U1, U2, U4, U5 και U6 με βάση τα αντίστοιχα U snRNAs που περιέχουν. Αυτά αποτελούν το μείζον σωματίο συναρμογής, το οποίο είναι υπεύθυνο για την απομάκρυνση των περισσότερων ιντρονίων τύπου U2 από τα πρόδρομα mRNAs (ιντρόνια που διαθέτουν τις συντηρημένες θέσεις αναγνώρισης GT και AG στις 5' και 3' θέσεις συναρμογής τα οποία απομακρύνονται με τον κοινό και καλά χαρακτηρισμένο μηχανισμό συναρμογής, (Sharp & Burge 1997) (Εικ. 1). Επίσης, υπάρχει και το ελάσσον σωματίο συναρμογής το οποίο περιέχει τις υπομονάδες snRNP U11, U12, U4atac, U6atac που απομακρύνει τα σπάνια ιντρόνια τύπου U12 [ιντρόνια τα οποία διαφέρουν με τα σύνηθη ιντρόνια στη 5' θέση συναρμογής και στη θέση σχηματισμού θηλιάς, αποτελούν το 0,15-0,34% των ιντρονίων του ανθρώπου και συναντώνται κυρίως σε σπονδυλωτά, φυτά, έντομα και σε κάποιους μύκητες, (Sharp & Burge 1997, Burge *et al.* 1998)] (Εικ. 1). Τα ιντρόνια τύπου U12 βρίσκονται σε γονίδια ή σε οικογένειες γονιδίων που κωδικοποιούν πρωτεΐνες που είναι απαραίτητες για την αντιγραφή και επιδιόρθωση του DNA, τη μεταγραφή και τη μετάφραση καθώς και τη μεταγωγή σήματος (Konig *et al.* 2007).

Κάθε snRNP περιέχει ένα μόνο snRNA και αρκετές (<20) πρωτεΐνες. Μια κεντρική δομή για όλα τα snRNP απαρτίζεται από μια ομάδα 8 πρωτεϊνών οι οποίες ονομάζονται Sm καθώς και επιπλέον πρωτεΐνες ειδικές για κάθε snRNP. Τα snRNAs σχηματίζονται αρχικά ως πρόδρομα μόρια, εξέρχονται στο κυτταρόπλασμα με τη βοήθεια ενός πολυπρωτεϊνικού συμπλόκου εξόδου και συνδέονται με τις πρωτεΐνες Sm μέσω μιας συντηρημένης αλληλουχίας που είναι παρούσα σε όλα τα snRNA εκτός του U6 snRNA το οποίο συνδέεται με μια ομάδα πρωτεϊνών παρόμοιων με τις Sm (Sm-like, LSm). Στο κυτταρόπλασμα γίνεται η ωρίμανση των πρόδρομων snRNAs και ακολουθεί η είσοδος του ενδιάμεσου συμπλόκου στον πυρήνα με τη βοήθεια των πρωτεϊνών Snurportin-1 και Importin β. Τέλος, σε κάθε snRNP συνδέονται οι ειδικές πρωτεΐνες (Lewin B. 2004).

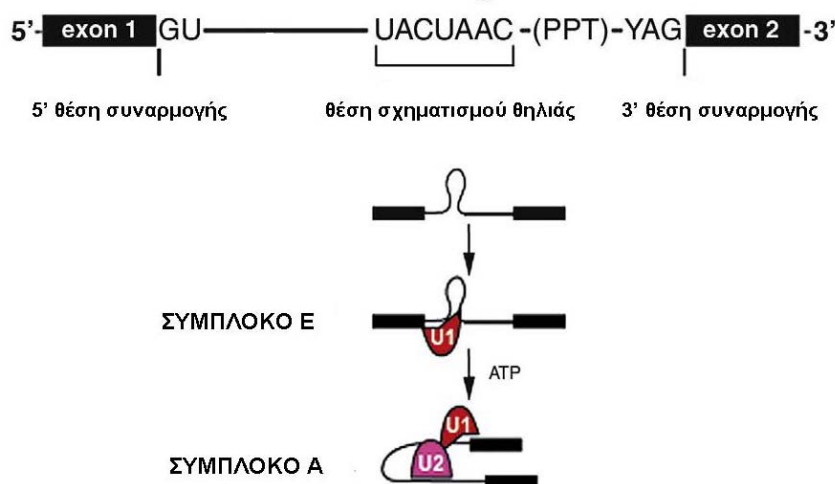


Εικ. 1: Διαφορές ιντρονίων τύπου U2 και ιντρονίων τύπου U12 (Προσαρμοσμένη από Will & Luhrmann 2005).

Το σωματίο συναρμογής σχηματίζεται αρχικά πάνω στο πρόδρομο mRNA και οι υπομονάδες U snRNP αναγνωρίζουν τα όρια εξονίων-ιντρονίων και στρατολογούν τους παράγοντες που καταλύουν τη συναρμογή του πρόδρομου mRNA. Αυτή η στρατολόγηση γίνεται με συγκεκριμένη σειρά με αποτέλεσμα ο σχηματισμός του σωματίου συναρμογής να επιτυγχάνεται με την αλληλεπίδραση των συστατικών του και την διαδοχική δημιουργία πολυπρωτεϊνικών συμπλόκων με την ακόλουθη σειρά: E, A, B και C (Lewin B. 2004). Ο σχηματισμός του σωματίου συναρμογής εξαρτάται από αλληλεπιδράσεις RNA-RNA, RNA-πρωτεΐνης και πρωτεΐνης-πρωτεΐνης (Green 1991, Guthrie 1991).

Το πρώτο βήμα για το μείζον σωματίο συναρμογής είναι ο σχηματισμός του συμπλόκου E (πρώιμο σύμπλοκο προ-συναρμογής). Το σύμπλοκο αυτό περιλαμβάνει την υπομονάδα U1 snRNP η οποία ζευγαρώνει με την 5' θέση συναρμογής μέσω μιας μονόκλωνης περιοχής (αλληλουχία 4-6 bp, συμπληρωματική με τη θέση συναρμογής) στο 5' άκρο του, τον παράγοντα συναρμογής U2AF ο οποίος είναι προσδεδμένος στην περιοχή ανάμεσα στη θέση σχηματισμού θηλιάς (branch site) και στη 3' θέση συναρμογής, τον παράγοντα συναρμογής SF1 (συνδέει τον παράγοντα U2AF με την υπομονάδα U1 snRNP) και μέλη της οικογένειας πρωτεϊνών SR. Οι πρωτεΐνες SR φέρουν μια περιοχή πλούσια σε κατάλοιπα σερίνης

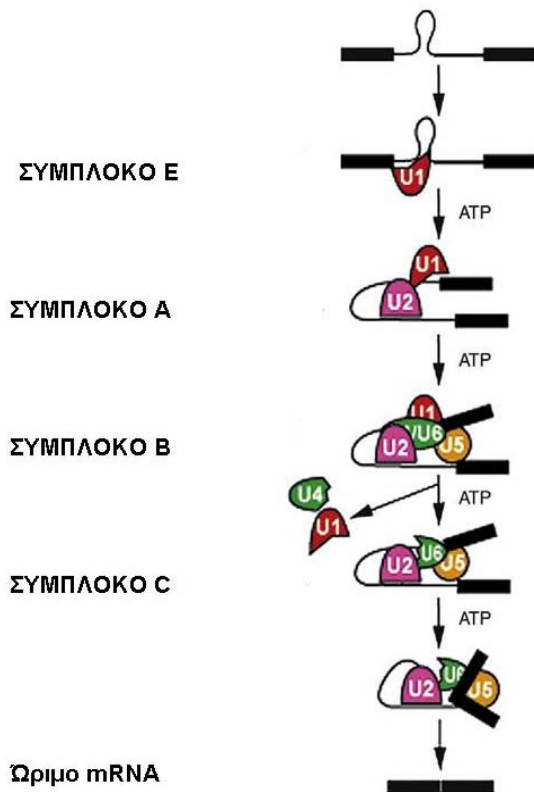
και αργινίνης η οποία είναι μεταβλητή σε μέγεθος και είναι απαραίτητη για την αλληλεπίδραση μεταξύ τους. Επίσης, οι πρωτεΐνες SR προσδένονται στο RNA και αποτελούν απαραίτητο συστατικό του σωματίου συναρμογής δημιουργώντας ένα σκελετό πάνω στο υπόστρωμα-RNA και συνδέουν τον παράγοντα U2AF με την υπομονάδα U1 snRNP. Το σύμπλοκο E μετατρέπεται σε πρόδρομο σύμπλοκο A με την πρόσδεση της υπομονάδας U2 snRNP, λόγω συμπληρωματικότητας του U2 snRNA στη θέση σχηματισμού της θηλιάς η οποία απαιτεί ενέργεια μέσω υδρόλυσης ATP (Lewin B. 2004, Green 1991, Guthrie 1991) (Εικ. 2).



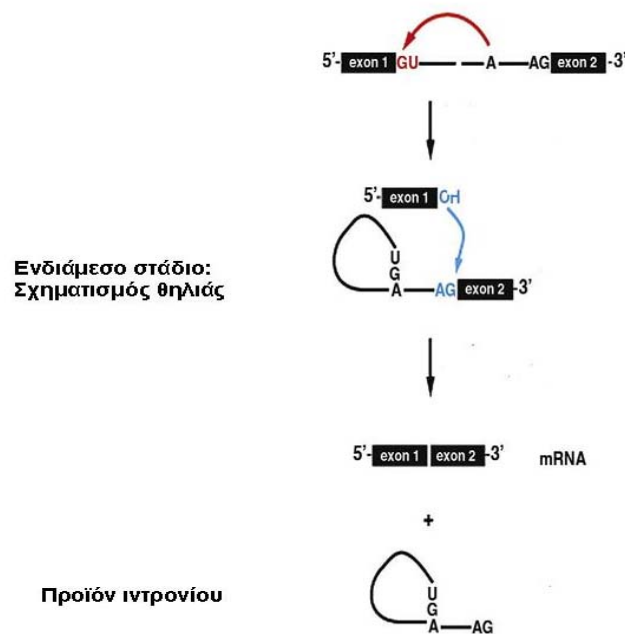
Εικ. 2: Στάδια σχηματισμού πρόδρομου σωματίου συναρμογής. Στο πάνω μέρος της εικόνας φαίνεται οι συντηρημένες θέσεις στα όρια των ιντρονίων και στο κάτω μέρος ο σχηματισμός του πρόδρομου σωματίου συναρμογής με αφετηρία το σύμπλοκο E και ακολούθως το σύμπλοκο A (προσαρμοσμένη από Ritchie *et al.* 2009).

Στο πρόδρομο αυτό σωματίο συναρμογής συνδέεται ένα τριμερές σύμπλοκο (tri-snRNP) που αποτελείται από τις υπομονάδες U5 και U4/U6 snRNP, με αποτέλεσμα το σχηματισμό του συμπλόκου B1. Το σύμπλοκο B1 μετατρέπεται στο σύμπλοκο B2 με την αποδέσμευση της υπομονάδας U1 snRNP με αποτέλεσμα την προσέγγιση του U6 snRNA στην 5' θέση συναρμογής. Η μετατόπιση της θέσης του U5 snRNA από την 5' θέση συναρμογής κοντά στις αλληλουχίες του ιντρονίου και η αποδέσμευση του U4 snRNA από το U6 snRNA οδηγεί στην αλληλεπίδραση U6-U2 snRNA καθώς και στο σχηματισμό ενδομοριακής φουρκέτας στο εσωτερικό του U6 snRNA. Οι δομικές αυτές αλλαγές οδηγούν στο σχηματισμό του συμπλόκου C κατά το οποίο πραγματοποιείται το πρώτο καταλυτικό βήμα της συναρμογής για το πρόδρομο mRNA (Εικ. 3). Όπως προαναφέρθηκε, η πρωτεΐνη FRA10AC1 καθώς και οι πρωτεΐνες με τις οποίες αλληλεπιδρά (DGCR14, SAP145) έχουν ταυτοποιηθεί ως συστατικά του συμπλόκου C του σωματίου συναρμογής, γεγονός που πιθανώς υποδηλώνει ότι σχετίζονται με το πρώτο καταλυτικό βήμα της διαδικασίας. Στη

συνέχεια, το σωματίο συναρμογής αποσυναρμολογείται και απελευθερώνονται τα ιντρόνια και το ώριμο mRNA με αποτέλεσμα την ολοκλήρωση του δεύτερου καταλυτικού βήματος (Lewin B. 2004, Green 1991, Guthrie 1991) (Εικ. 4).



Εικ. 3: Στάδια σχηματισμού σωματίου συναρμογής με το διαδοχικό σχηματισμό των συμπλόκων E, A, B, C στο πρόδρομο mRNA (προσαρμοσμένη από Ritchie *et al.* 2009).



Εικ. 4: Σχηματική αναπαράσταση του μηχανισμού συναρμογής που έχει ως αποτέλεσμα την παραγωγή του ώριμου mRNA και την απελευθέρωση των ιντρονίων (προσαρμοσμένη από Ritchie *et al.* 2009).

➤ 1.3.1 Πρωτεΐνες-συστατικά του σωματίου συναρμογής

Οι εκτιμήσεις για τον αριθμό των πρωτεϊνών που συμμετέχουν στη διαδικασία της συναρμογής στον άνθρωπο ποικίλλουν αλλά φαίνεται πως πρόκειται για εκατοντάδες πρωτεΐνες. Ένα εύλογο ερώτημα είναι ποιος είναι ο ρόλος όλων αυτών των πρωτεϊνών και σε τι οφείλεται αυτή η πολυπλοκότητα της μηχανής της συναρμογής ενώ ουσιαστικά απαιτούνται συγκεκριμένα ενζυμικά βήματα. Μία παράμετρος που συνεισφέρει στην πολυπλοκότητα του συστήματος είναι η ακρίβεια με την οποία επιτελείται η συναρμογή. Έτσι, τα συστατικά του σωματίου συναρμογής θα πρέπει να αναγνωρίζουν συγκεκριμένες, μικρές αλληλουχίες οι οποίες εντοπίζονται σε πολύ μεγάλη έκταση και να επιτελέσουν τη συναρμογή σωστά. Επιπλέον, θα πρέπει να ληφθεί υπ' όψιν η εναλλακτική συναρμογή. Πλέον είναι γνωστό ότι η πλειοψηφία των γονιδίων του ανθρώπου (92-94%) υφίστανται εναλλακτική συναρμογή (Wang *et al.* 2008). Τέλος, σε αυτόν τον πολύ μεγάλο αριθμό πρωτεϊνών του σωματίου συναρμογής συμπεριλαμβάνονται και πρωτεΐνες που συμμετέχουν σε άλλες, λειτουργικά συνδεδεμένες, διαδικασίες όπως η πολυαδενυλίωση, η έξοδος από τον πυρήνα, κ.λπ.

Οι πρωτεΐνες που αποτελούν υπομονάδες του σωματίου συναρμογής, ταυτοποιήθηκαν ως επί το πλείστον πρωτεωμικές μελέτες που περιλαμβάνουν βιοχημικά πειράματα απομόνωσης του σωματίου και χαρακτηρισμού των συστατικών του με φασματομετρία μάζας (Bennett *et al.* 1992). Πιο συγκεκριμένα, κατασκευάστηκαν τεχνητά πρόδρομα μόρια mRNA, συνήθως ραδιοσημασμένα με βιοτίνη (Rappsilber *et al.* 2002), π.χ. πρόδρομα μόρια mRNA που προέρχονται από γονίδια του αδενοϊού (AdML, adenovirus major late) ή από γονίδια της *Drosophila melanogaster* όπως το *Ftz* (Fushi tarazu), τα οποία χρησιμοποιήθηκαν για την *in vitro* ανασύσταση, απομόνωση και ταυτοποίηση των επιμέρους συμπλόκων του σωματίου, όπως το σύμπλοκο B και C (Bessonov *et al.* 2008, Makarova *et al.* 2004, Jurica *et al.* 2002) και το σύμπλοκο A (Behzдания *et al.* 2007). Σε άλλες περιπτώσεις, απομονώθηκε ολόκληρη η μηχανή συναρμογής *in vivo*, συνήθως από εκχύλισμα κυττάρων θηλαστικών *HeLa* (Chen *et al.* 2007). Τα παραπάνω πειράματα αφορούν κυρίως το σωματίο συναρμογής του ανθρώπου αλλά υπάρχουν και αντίστοιχα δεδομένα από τη *Drosophila melanogaster* (Herold *et al.* 2008) και από τον *Saccharomyces cerevisiae* (Neubauer G. 2005). Στα αρχικά πειράματα απομονώθηκαν δεκάδες πρωτεΐνες ως συστατικά του σωματίου συναρμογής αλλά με την πρόοδο των πειραματικών προσεγγίσεων, έχουν πλέον ταυτοποιηθεί εκατοντάδες πρωτεΐνες.

Πλέον, υπάρχουν πολυάριθμες μελέτες που αφορούν τη δομή του σωματίου συναρμογής και των επιμέρους συμπλόκων του (κυρίως μέσω της ηλεκτρονικής

μικροσκοπίας, της κρυσταλλογραφίας και της φασματοσκοπίας NMR) καθώς και τη σύστασή του (τουλάχιστον 10 πρωτεομικές μελέτες ταυτοποίησης των επιμέρους συστατικών του). Ωστόσο, οι γνώσεις σχετικά με τις αλληλεπιδράσεις των διαφορετικών υπομονάδων και τη δυναμική τους, οι οποίες είναι απαραίτητες για τον προσδιορισμό της ολοκληρωμένης εικόνας της λειτουργίας του σωματίου είναι σχετικά περιορισμένες. Έτσι, αξιόπιστα δεδομένα από στοχευμένα πειράματα μικρής κλίμακας υπάρχουν για ένα μικρό σχετικά ποσοστό των πρωτεϊνών-συστατικών του σωματίου. Επειδή αυτή η εργασία εστιάζεται στον προσδιορισμό των πρωτεϊνικών αλληλεπιδράσεων του σωματίου συναρμογής εκκινώντας από την πρωτεΐνη FRA10AC1, θα ακολουθήσει μία σύντομη ανασκόπηση σχετικά με τις πρωτεομικές μελέτες που εστιάζονται στον προσδιορισμό των αλληλεπιδράσεων των πρωτεϊνών.

1.4 Πρωτεομική (proteomics) και δίκτυα πρωτεϊνικών αλληλεπιδράσεων (protein interactome)

➤ 1.4.1 Πρωτεομική

Η αποτελεσματική ταυτοποίηση πεπτιδίων και πρωτεϊνών έχει αναπτυχθεί τα τελευταία χρόνια (αναφορά για πρώτη φορά από τον Maurice Wilkins το 1994) οδηγώντας στην ανακάλυψη ενός καινούργιου τομέα στην επιστήμη της Βιολογίας, την Πρωτεομική (proteomics). Η Πρωτεομική είναι η ανάλυση μεγάλης κλίμακας του συνόλου ή των υποσυνόλων των εκφραζόμενων πρωτεϊνών ενός οργανισμού, κυτταρικού τύπου ή ιστού, σε ποιοτικό αλλά και σε ποσοτικό επίπεδο, η οποία συνεισφέρει σε μεγάλο βαθμό στην κατανόηση της λειτουργίας των γονιδίων στη μετα-γονιδιωματική εποχή. Κάθε γονιδίωμα είναι πανομοιότυπο σε όλα τα κύτταρα ενός οργανισμού και αλλάζει κατά μικρό ποσοστό στους διάφορους οργανισμούς ενός είδους. Αντίθετα, η έκφραση των πρωτεϊνών διαφέρει ανάμεσα στους διάφορους τύπους κυττάρων. Επίσης διαφέρει και η ποσότητα των πρωτεϊνών που παράγονται κάθε φορά ανάλογα με τις ανάγκες του οργανισμού και τις εξωτερικές συνθήκες. Το γεγονός αυτό καθιστά σημαντικό το διαχωρισμό, τον χαρακτηρισμό και την ταυτοποίηση των πρωτεϊνών (Alderton 2004).

Το αντικείμενο μελέτης της Πρωτεομικής διακρίνεται σε τρεις κύριους τομείς. Πρώτον, το χαρακτηρισμό και την ταυτοποίηση των πρωτεϊνών καθώς και τις τροποποιήσεις αυτών σε μετα-μεταφραστικό επίπεδο, δεύτερον, τα διαφορετικά επίπεδα έκφρασης των πρωτεϊνών και την πιθανή συσχέτισή τους με ένα σύνολο ασθενειών και τέλος την μελέτη των πρωτεϊνικών αλληλεπιδράσεων. Όμως, επειδή είναι δύσκολο να προβλεφθεί η λειτουργία μιας πρωτεΐνης με βάση τη δομή της ή την

ομολογία της με άλλες πρωτεΐνες, κεντρική σημασία για την λειτουργική της ανάλυση έχει ο καθορισμός των συστατικών ενός συμπλόκου ή μιας κυτταρική δομής στα οποία ανήκει (Pandey & Mann 2000).

➤ 1.4.2 Πρωτεϊνικές αλληλεπιδράσεις

Οι πρωτεϊνικές αλληλεπιδράσεις αποτελούν μια σημαντική παράμετρο για την κατανόηση της λειτουργίας των βιομορίων διότι οι πρωτεΐνες δεν μπορούν να επιτελέσουν τις λειτουργίες τους χωρίς να αλληλεπιδράσουν με άλλα μακρομόρια. Έτσι, η ταυτοποίηση των φυσικών πρωτεϊνικών αλληλεπιδράσεων, των συμπλόκων καθώς και η δυναμική των πρωτεϊνών σε συγκεκριμένο χώρο και χρόνο, σε επίπεδο κυττάρου ή οργανισμού, είναι σημαντική για την κατανόηση της πρωτεϊνικής λειτουργίας, των βιοχημικών μονοπατιών και δικτύων στα οποία ανήκουν και μακροπρόθεσμα στην κατανόηση της συντονισμένης λειτουργίας των «πρωτεϊνικών μηχανών» και της αποτελεσματικής κυτταρικής λειτουργίας (Figeys 2008). Επίσης, αυτές οι πληροφορίες είναι περισσότερο απαραίτητες κυρίως όταν δεν υπάρχουν στοιχεία για την υποθετική πρωτεϊνική λειτουργία όπως στις περιπτώσεις στις οποίες η πρωτεΐνη δεν περιέχει ένα καλά χαρακτηρισμένο μοτίβο ή στις περιπτώσεις όπου δεν υπάρχουν πληροφορίες για την λειτουργία της από τα ορθόλογα μόριά της. Συνεπώς, οι αλληλεπιδράσεις μεταξύ των βιομορίων παίζουν σημαντικό ρόλο στην πλειοψηφία των κυτταρικών διαδικασιών. Αυτή η γνώση, δηλαδή του ρόλου και των συνεπειών των πρωτεϊνικών αλληλεπιδράσεων αξιοποιείται στη λειτουργική γονιδιωματική και στη συστημική βιολογία. Επιπλέον, αξιοποιείται στην κατανόηση των μηχανισμών που σχετίζονται με κάποια ασθένεια, στην ανακάλυψη νέων ειδικών δεικτών ή πρωτεϊνικών προτύπων, στην διαγνωστική ιστολογική απεικόνιση των πρωτεϊνών και στην ανακάλυψη νέων φαρμάκων και καινοτόμων θεραπειών (Figeys 2008).

Από τις αρχές της δεκαετίας του 2000, οι περισσότερες πληροφορίες σε σχέση με τα δίκτυα πρωτεϊνικών αλληλεπιδράσεων προέρχονται από μεγάλης κλίμακας πειράματα υψηλής απόδοσης, αρχικά σε γονιδιώματα οργανισμών-μοντέλων όπως στους *S.cerevisiae*, *D.melanogaster* και *C.elegans* (Gavin *et al* 2002, Giot *et al* 2003, Ito *et al* 2001, Li *et al* 2004, Uetz *et al* 2000) και αργότερα στον άνθρωπο (Stelzl *et al* 2005, Rual *et al* 2005). Η διερεύνηση των πρωτεϊνικών αλληλεπιδράσεων σε μεγάλη κλίμακα πραγματοποιείται με τη χρήση δύο πειραματικών μεθόδων (Figeys 2008). Η πρώτη αφορά την αναζήτηση πρωτεϊνικών αλληλεπιδράσεων με το σύστημα των «δύο υβριδίων» στο ζωικό μοντέλο του σακχαρομύκητα με τη βοήθεια της οποίας ταυτοποιούνται αποκλειστικά δυαδικές αλληλεπιδράσεις. Η δεύτερη αφορά την απομόνωση πολυπρωτεϊνικών συμπλόκων

που συγκροτούνται αυθόρμητα σε κυτταροκαλλιέργειες και την επακόλουθη μοριακή ταυτοποίηση των πρωτεϊνών, που αποτελούν συστατικά του συμπλόκου, με φασματομετρία μάζας. Με τη βοήθεια της δεύτερης πειραματικής μεθόδου προσδιορίζονται όχι μόνο οι δυαδικές αλληλεπιδράσεις αλλά και έμμεσες αλληλεπιδράσεις πρωτεϊνών που συμμετέχουν στο ίδιο πολυπρωτεϊνικό σύμπλοκο. Η τεχνική του συστήματος των «δύο υβριδίων» είναι η πιο ευρέως χρησιμοποιούμενη μέθοδος χαρτογράφησης των πρωτεϊνικών αλληλεπιδράσεων (Fields & Song 1989). Πιο συγκεκριμένα, τα δύο πρωτεϊνικά μόρια τα οποία ελέγχονται για πιθανή αλληλεπίδραση, συντήκονται το ένα με την περιοχή (domain) δέσμευσης στο DNA των γονιδίων αναφοράς του σακχαρομύκητα και το άλλο με την περιοχή (domain) ενεργοποίησης της μεταγραφής των γονιδίων αναφοράς του σακχαρομύκητα. Η πιθανή αλληλεπίδραση εξετάζεται με τη σύνδεση αυτών των δύο περιοχών και την επαγόμενη μεταγραφική ενεργοποίηση των γονιδίων αναφοράς (Fields & Song 1989).

Η απομόνωση των πρωτεϊνών που συνιστούν ένα πολυπρωτεϊνικό σύμπλοκο πραγματοποιείται με χρωματογραφικές τεχνικές [υγρή χρωματογραφία υψηλής πίεσης (HPLC), χρωματογραφία συγγένειας]. Πριν το διαχωρισμό των πεπτιδίων με υγρή χρωματογραφία υψηλής πίεσης, προηγείται ενζυμική πέψη των πρωτεϊνών (κυρίως με θρυψίνη, ένζυμο που παρουσιάζει το χαμηλότερο ποσοστό αυτοπρωτεόλυσης). Ο διαχωρισμός των πεπτιδίων είναι αποτέλεσμα της συνδυαστικής δράσης μιας στατικής και μιας κινητής φάσης. Στην αρχή, το δείγμα εισάγεται στη κορυφή της στήλης και με τη βοήθεια της κινητής φάσης, τα συστατικά του μετακινούνται με τη μορφή ζωνών και τελικά εκκλύονται το ένα μετά το άλλο. Τα πεπτίδια κατανέμονται μεταξύ της στατικής και της κινητής φάσης, με αποτέλεσμα να μετακινούνται με διαφορετικές ταχύτητες κατά μήκος της στήλης. Με τη χρωματογραφία συγγένειας ο διαχωρισμός των πρωτεϊνών πραγματοποιείται με τη χρήση ενός ειδικού αντισώματος ή μιας ειδικής ετικέτας (tag). Στη συνέχεια, πραγματοποιείται ηλεκτροφόρηση μιας ή δυο διαστάσεων με επακόλουθη ενζυμική πέψη για το διαχωρισμό των πεπτιδίων. Οι τεχνικές αυτές σε συνδυασμό με τη φασματομετρία μάζας αυξάνουν την ταχύτητα ταυτοποίησης των πρωτεϊνών που αλληλεπιδρούν μέσα σε ένα σύμπλοκο. Η φασματομετρία μάζας είναι μια μέθοδος ανάλυσης απαραίτητη για τον χαρακτηρισμό και την ταυτοποίηση των πρωτεϊνών και ιδιαίτερα χρήσιμη για την αξιολόγηση των μετα-μεταφραστικών τροποποιήσεων. Η μέθοδος στηρίζεται σε τρία βασικά βήματα. Αρχικά, με τη μέθοδο ιονισμού (ESI, MALDI) τα πεπτίδια μετατρέπονται σε ιόντα στην αέρια φάση επειδή είναι μεγάλα πολικά βιομόρια, μη πτητικά. Ακολουθεί διαχωρισμός και ανίχνευση των πεπτιδίων με βάση την αναλογία της μάζας τους προς το φορτίο τους (m/z) προσδίδοντας

υψηλό επίπεδο εξειδίκευσης με μεγάλη ευαισθησία. Τέλος, ο ανιχνευτής ανιχνεύει τα ιόντα που προκύπτουν από τον αναλυτή μάζας με αποτέλεσμα τη δημιουργία ενός χαρακτηριστικού φάσματος για κάθε πεπτιδίο το οποίο χρησιμοποιείται για τον προσδιορισμό της μοριακής μάζας του πεπτιδίου (Kicman *et al.* 2007).

Βάσεις δεδομένων πρωτεϊνικών αλληλεπιδράσεων

Οι πρωτεϊνικές αλληλεπιδράσεις που έχουν ταυτοποιηθεί με πειράματα τόσο μεγάλης όσο και μικρής κλίμακας υπάρχουν πλέον καταχωρημένες σε βάσεις δεδομένων, πολλές από τις οποίες είναι ελεύθερα προσβάσιμες. Οι κύριες βάσεις δεδομένων που περιλαμβάνουν δεδομένα από τη βιβλιογραφία και αφορούν αλληλεπιδράσεις πρωτεϊνών του ανθρώπου είναι η HPRD (Human Protein Reference Database, <http://www.hprd.org/>), η MINT (the Molecular Interaction Database, <http://160.80.34.4/mint/Welcome.do>), η INTACT (The IntAct molecular interaction database, <http://www.ebi.ac.uk/intact/>), η DIP (The Database of Interacting Proteins, <http://dip.doe-mbi.ucla.edu/dip/Main.cgi>) και η BIOGRID (Biological General Repository for Interaction Datasets, <http://thebiogrid.org/>). Η HPRD είναι μια βάση δεδομένων που δημιουργήθηκε το 2003 και περιλαμβάνει πληροφορίες για πρωτεϊνικές αλληλεπιδράσεις, μετα-μεταφραστικές τροποποιήσεις και για την έκφραση των πρωτεϊνών σε διάφορους ιστούς. Η HPRD περιλαμβάνει ένα σύνολο 38167 πρωτεϊνικών αλληλεπιδράσεων από τις οποίες οι 8958 έχουν ταυτοποιηθεί με το σύστημα των δυο υβριδίων του σακχαρομύκητα και έχει το πλεονέκτημα ότι ενημερώνεται αρκετά συχνά. Από την ενημέρωσή της κατά το διάστημα 2006-2009 έχουν προστεθεί επιπλέον 5000 πρωτεϊνικές αλληλεπιδράσεις (Prasad *et al.* 2008). Στην MINT κάθε αλληλεπίδραση βαθμολογείται με ένα σκορ από 0 έως 1 ανάλογα με την ποιότητα και ποσότητα των πληροφοριών που την υποστηρίζουν πειραματικά (Ceol *et al.* 2010). Η INTACT περιλαμβάνει δεδομένα από τη βιβλιογραφία τα οποία έχουν επεξεργαστεί από ειδικούς (curators) και απεικονίζονται σε μορφή γράφου περιλαμβάνοντας λεπτομερείς πληροφορίες για τις πρωτεϊνικές αλληλεπιδράσεις (Aranda *et al.* 2009). Η DIP εκτός από τις πληροφορίες για τις αλληλεπιδράσεις πρωτεϊνών που έχουν ταυτοποιηθεί πειραματικά περιλαμβάνει και δίκτυα αλληλεπιδράσεων που σχηματίζονται στις βιολογικές διαδικασίες (Xenarios *et al.* 2002). Η BIOGRID περιλαμβάνει πληροφορίες για πρωτεϊνικές και γενετικές αλληλεπιδράσεις από έξι διαφορετικούς οργανισμούς-μοντέλα (Breitkreutz *et al.* 2007).

Επιπλέον, υπάρχουν βάσεις δεδομένων που περιλαμβάνουν πρωτεϊνικές αλληλεπιδράσεις που παρατηρούνται μεταξύ ορθόλογων πρωτεϊνών στους οργανισμούς-μοντέλα *S.cerevisiae*, *C.elegans*, και *D.melanogaster*. Ενδεικτικά,

αναφέρεται η DroID (The Comprehensive Drosophila Interactions Database) που περιλαμβάνει φυσικές και γενετικές πρωτεϊνικές αλληλεπιδράσεις στη *D. melanogaster* (Yu *et al.* 2008) και η Worm Interactome Database που περιλαμβάνει 5534 πρωτεϊνικές αλληλεπιδράσεις που έχουν ταυτοποιηθεί με το σύστημα δύο υβριδίων του σακχαρομύκητα οι οποίες εκτιμάται ότι καλύπτουν το 15% του πρωτεϊνώματος του *C. elegans* (Chen *et al.* 2005).

1.5 Σκοπός της εργασίας

Η εργασία εντάσσεται στην ευρύτερη προσπάθεια λειτουργικής ανάλυσης της πρωτεΐνης FRA10AC1 του ανθρώπου. Η FRA10C1 κωδικοποιείται από το γονίδιο *FRA10AC1* που χαρτογραφείται στην εύθραυστη χρωμοσωματική θέση *FRA10A* στην περιοχή 10q23.3. η οποία έχει συσχετιστεί με φαινότυπο ήπιας νοητικής υστέρησης στα άτομα-φορείς. Οι ενδείξεις από προηγούμενες μελέτες συνηγορούν υπέρ της συμμετοχής της FRA10AC1 στη διαδικασία της συναρμογής ή/και σε άλλες στενά συσχετιζόμενες διαδικασίες του μεταβολισμού του RNA.

Ειδικότερα, η εργασία στοχεύει στην *in silico* ένταξη της FRA10AC1 στο δίκτυο των πρωτεϊνικών αλληλεπιδράσεων του σωματίου συναρμογής (spliceosome), αποτέλεσμα που αφ' ενός θα παρέχει χρήσιμες πληροφορίες για το ρόλο της FRA10AC1 στη δυναμική και λειτουργία του σωματίου και αφ' ετέρου θα παρέχει το πλαίσιο για το σχεδιασμό στοχευμένων πειραματικών προσεγγίσεων ταυτοποίησης πρωτεϊνικών αλληλεπιδράσεων. Για την ανασύσταση του δικτύου με αφετηρία την FRA10AC1 αξιοποιούνται δεδομένα από ελεύθερα προσβάσιμες βάσεις δεδομένων που περιλαμβάνουν καταχωρήσεις για δυαδικές πρωτεϊνικές αλληλεπιδράσεις του ανθρώπου και οργανισμών-μοντέλων (*D. melanogaster* και *C. elegans*) και ως βασικό κριτήριο για έλεγχο της αξιοπιστίας τους ορίζεται η βιοχημική ταυτοποίηση των πρωτεϊνών ως συστατικά του σωματίου συναρμογής. Επιπλέον, η εργασία στοχεύει στην κατασκευή των προκαταρκτικών εργαλείων (πλασμιδιακές κατασκευές) για το μελλοντικό πειραματικό έλεγχο συγκεκριμένων πρωτεϊνικών αλληλεπιδράσεων της πρωτεΐνης DGCR14, η οποία αλληλεπιδρά με την FRA10AC1.

2. ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ

ΥΛΙΚΑ

2.1 Καλλιέργεια βακτηρίων

Όλα τα υλικά έγιναν με δις-απιονισμένο νερό και αποστειρώθηκαν σε κλίβανο ατμού. Σε κάθε περίπτωση προστίθενται μετά την αποστείρωση το κατάλληλο αντιβιοτικό.

LB (Luria Broth): Υγρό μέσο ανάπτυξης των βακτηρίων.

Η σύστασή του είναι:

- ❖ 10 gr ανά λίτρο casein-tryptone
- ❖ 5 gr ανά λίτρο εκχύλισμα ζύμης
- ❖ 10 gr NaCl ανά λίτρο.

Το pH ρυθμίζεται στο 7-7,5 με διάλυμα NaOH.

Για να παρασκευαστεί στερεό θρεπτικό μέσο πρέπει να προστεθούν στα παραπάνω συστατικά επιπλέον 15 gr άγαρ/λίτρο. Το μίγμα αποστειρώνεται και τοποθετείται σε τρυβλία Petri όπου στερεοποιείται σε θερμοκρασία δωματίου (LB agar Cat. No 01-384, Scharlau Microbiology, LB υγρό Cat. No 02-384, Scharlau Microbiology).

Διαλύματα αντιβιοτικών: Τα αντιβιοτικά που χρησιμοποιήθηκαν ήταν η αμπικιλίνη, η καναμυκίνη και η χλωραμφαινικόλη. Τα δύο πρώτα ήταν διαλυμένα σε δις-απιονισμένο νερό και το τελευταίο σε απόλυτη αιθανόλη. Πιο συγκεκριμένα, η αμπικιλίνη χρησιμοποιήθηκε σε τελική συγκέντρωση διαλύματος 100 µg/ml από αρχικό διάλυμα συγκέντρωσης 100 mg/ml, η καναμυκίνη χρησιμοποιήθηκε σε τελική συγκέντρωση διαλύματος 30 µg/ml από αρχικό διάλυμα συγκέντρωσης 10 mg/ml και τέλος η χλωραμφαινικόλη χρησιμοποιήθηκε σε τελική συγκέντρωση διαλύματος 34 µg/ml από αρχικό διάλυμα συγκέντρωσης 34 mg/ml.

2.2 Καλλιέργεια κυττάρων θηλαστικών

- Θρεπτικό μέσο DMEM (Dulbecco's Modified Eagle's Medium): Υγρό μέσο ανάπτυξης κυττάρων θηλαστικών (GIBCO, 41966).

Η σύσταση του είναι:

- 4,5 g/L glucose
- L- glutamine
- Pyruvate

- Ρυθμιστικό διάλυμα PBS (Phosphate Buffered Saline 1X, pH=7.4): Το συγκεκριμένο διάλυμα δεν περιέχει CaCl₂ και MgCl₂ (GIBCO, Cat no 10010).
- Παγκρεατικό ένζυμο θρυψίνη (GIBCO, Cat no 15400-054): Το ένζυμο βρίσκεται σε μορφή διαλύματος με EDTA 10% και χρησιμοποιείται για την

θρυψινοποίηση των κυττάρων σε τελική αραιώση 1:10 σε ρυθμιστικό διάλυμα PBS.

- Κρυοδιατήρηση κυτταροκαλλιιεργειών: Για την κρυοδιατήρηση των κυτταροκαλλιιεργειών χρησιμοποιήθηκε κρυοπροστατευτικό μέσο Dimethyl sulphoxide (DMSO) της εταιρίας SIGMA (Cat. No D2650).

2.3 Εκκινητές

Οι εκκινητές που χρησιμοποιήθηκαν ήταν της εταιρίας Metabion International AG. Σχεδιάστηκαν με τέτοιο τρόπο ώστε με την PCR να σχηματιστεί cDNA πλήρους μήκους των γονιδίων *RED* και *RACK* που περιείχε στα άκρα του θέσεις αναγνώρισης περιοριστικών ενδονουκλεασών (Πιν. 1), οι θέσεις αναγνώρισης των περιοριστικών ενδονουκλεασών είναι υπογραμμισμένες) για επόμενη κλωνοποίηση σε πλασμιδιακούς φορείς έκφρασης σε κύτταρα θηλαστικών. Επίσης φαίνονται και οι εκκινητές των πλασμιδιακών φορέων pEGFP-C1 και pGEM-T Easy.

Πιν. 1: Αλληλουχίες εκκινητών. Οι θέσεις αναγνώρισης των περιοριστικών ενδονουκλεασών φαίνονται υπογραμμισμένες.

ΟΝΟΜΑ	ΑΛΛΗΛΟΥΧΙΑ ΕΚΚΙΝΗΤΩΝ	Tm
RED_F	5'- <u>AGATCT</u> CGA GAT CTA TGC CGG AGC GAG ATA GTG-3'	Tm=71°C
RED_R -new	5'- <u>GAATTC</u> CGG AAT TCG TAT TTT GGT CTT TTG AC-3'	Tm=63°C
RACK_F_new	5'- <u>GAGCTC</u> CGG AGC TCC CAT GAC TGA GCA GAT GAC C-3'	Tm=75°C
RACK1_R	5'- <u>CTGCAG</u> CGG CTG CAG GCG TGT GCC AAT GGT CAC-3'	Tm=76°C
pEGFP.Fw	5'- CAT GGT CCT GCT GGA GTT CGT G-3'	Tm=66°C
T7	5'- TAA TAC GAC TCA CTA TAG GG-3'	Tm=54°C
SP6	5'- TAT TTA GGT GAC ACT ATA G-3'	Tm=49°C

2.4 Βακτηριακά στελέχη

DH5a: Στελέχη βακτηρίων που χρησιμοποιούνται για μετασχηματισμό κατά την κλωνοποίηση τμημάτων DNA σε πλασμιδιακούς φορείς καθώς η απόδοσή τους είναι υψηλή (>2x10⁶ μετασχηματισμένα βακτήρια/μg πλασμιδιακού DNA). Επιπλέον, υπάρχει δυνατότητα επιλογής των ανασυνδυσασμένων πλασμιδίων λόγω της μεταλλαγής που φέρουν στο γονίδιο *LacZ* (στο καρβοξυτελικό τμήμα της β-γαλακτοσιδάσης) που εμποδίζει την σύνδεση των υπομονάδων και συνεπώς την ενεργότητα της β-γαλακτοσιδάσης, η οποία όμως συμπληρώνεται από φορείς που παράγουν το α-πολυπεπτίδιο. Έτσι, τα μη ανασυνδυσασμένα πλασμίδια αποκαθιστούν στα μετασχηματισμένα βακτήρια την ενεργότητα του ενζύμου. Αντίθετα, στα μετασχηματισμένα βακτήρια που έχουν δεχθεί τα ανασυνδυσασμένα πλασμίδια, το γονίδιο της β-γαλακτοσιδάσης διακόπτεται και δεν μπορεί να

μεταφραστεί σε β-γαλακτοσιδάση. Αυτό έχει ως αποτέλεσμα οι βακτηριακές αποικίες να είναι λευκές, παρουσία X-gal (αδρανές χρωμογόνο υπόστρωμα για τη β-γαλακτοσιδάση) (Miller J.H. 1972).

BL21 (DE3) Codon Plus: Στελέχη βακτηρίων τα οποία χρησιμοποιήθηκαν για την έκφραση χιμαιρικών πρωτεϊνών σε βακτηριακά κύτταρα. Τα στελέχη αυτά διαθέτουν ένα παράγωγο του βακτηριοφάγου λ, τον βακτηριοφάγο DE3 που βρίσκεται σε λυσιγονία. Ο βακτηριοφάγος αυτός φέρει το γονίδιο της RNA πολυμεράσης T7 υπό τον έλεγχο του υποκινητή lacUV5, ο οποίος επάγεται με IPTG (isopropyl-beta-D-thiogalactopyranoside, σταθερό συνθετικό ανάλογο της λακτόζης). Η προσθήκη IPTG απενεργοποιεί τον καταστολέα της λακτόζης και επάγει τη σύνθεση της β-γαλακτοζιδάσης, επάγοντας την έκφραση των κλωνοποιημένων γονιδίων που είναι υπό τον έλεγχο του οπερονίου της λακτόζης. Έτσι, προσθέτοντας IPTG σε κύτταρα μετασχηματισμένα με ένα πλασμίδιο που διαθέτει τον υποκινητή T7, παράγεται η RNA πολυμεράση T7 που με τη σειρά της μεταγράφει το γονίδιο που έχει κλωνοποιηθεί στο πλασμίδιο. Επιπλέον, έχουν το πλασμίδιο RIL που περιέχει γονίδια για τα tRNAs argU, ileW, leuY και proL (που είναι σπανιότερα στην *E.coli*) το οποίο διαθέτει γονίδιο ανθεκτικότητας σε χλωραμφαινικόλη.

2.5 Κυτταρικές σειρές θηλαστικών

Χρησιμοποιήθηκε η κυτταρική σειρά *HeLa* η οποία προέρχεται από επιθηλιακά κύτταρα ανθρώπου από αδενοκαρκίνωμα του τραχήλου της μήτρας. Η κυτταρική σειρά καλλιεργήθηκε σε θρεπτικό υλικό DMEM εμπλουτισμένο με εμβρυικό ορό βοοειδούς (10% FBS) και αντιβιοτικά στρεπτομυκίνη και πενικιλίνη (1% streptomycin-penicillin) για αποφυγή μολύνσεων από μικροοργανισμούς. Τα κύτταρα *HeLa* αναπτύσσονται σε κατάλληλη θερμοκρασία (37°C) και κατάλληλα επίπεδα CO₂ (5%) σχηματίζοντας μονοστοιβάδες.

2.6 Πλασμιδιακοί φορείς

Οι πλασμιδιακοί φορείς που χρησιμοποιήθηκαν έχουν την ικανότητα να πολλαπλασιάζονται σε πολλά αντίγραφα ανά κύτταρο. Περιέχουν ένα τουλάχιστον γονίδιο ανθεκτικότητας σε αντιβιοτικό (αμπικιλίνη, τετρακυκλίνη, καναμυκίνη και χλωραμφαινικόλη είναι τα πιο κοινά αντιβιοτικά), γεγονός που είναι απαραίτητο για την επιλογή των μετασχηματισμένων βακτηρίων που περιέχουν τα ανασυνδυασμένα πλασμίδια. Εκτός από την ανθεκτικότητα σε αντιβιοτικά, έχουν κι άλλες ιδιότητες, όπως είναι η δυνατότητα να μεταβολίζουν τη λακτόζη (γονίδιο *lacZ* της β-γαλακτοσιδάσης) γεγονός που επιτρέπει την επιλογή των βακτηριακών αποικιών που

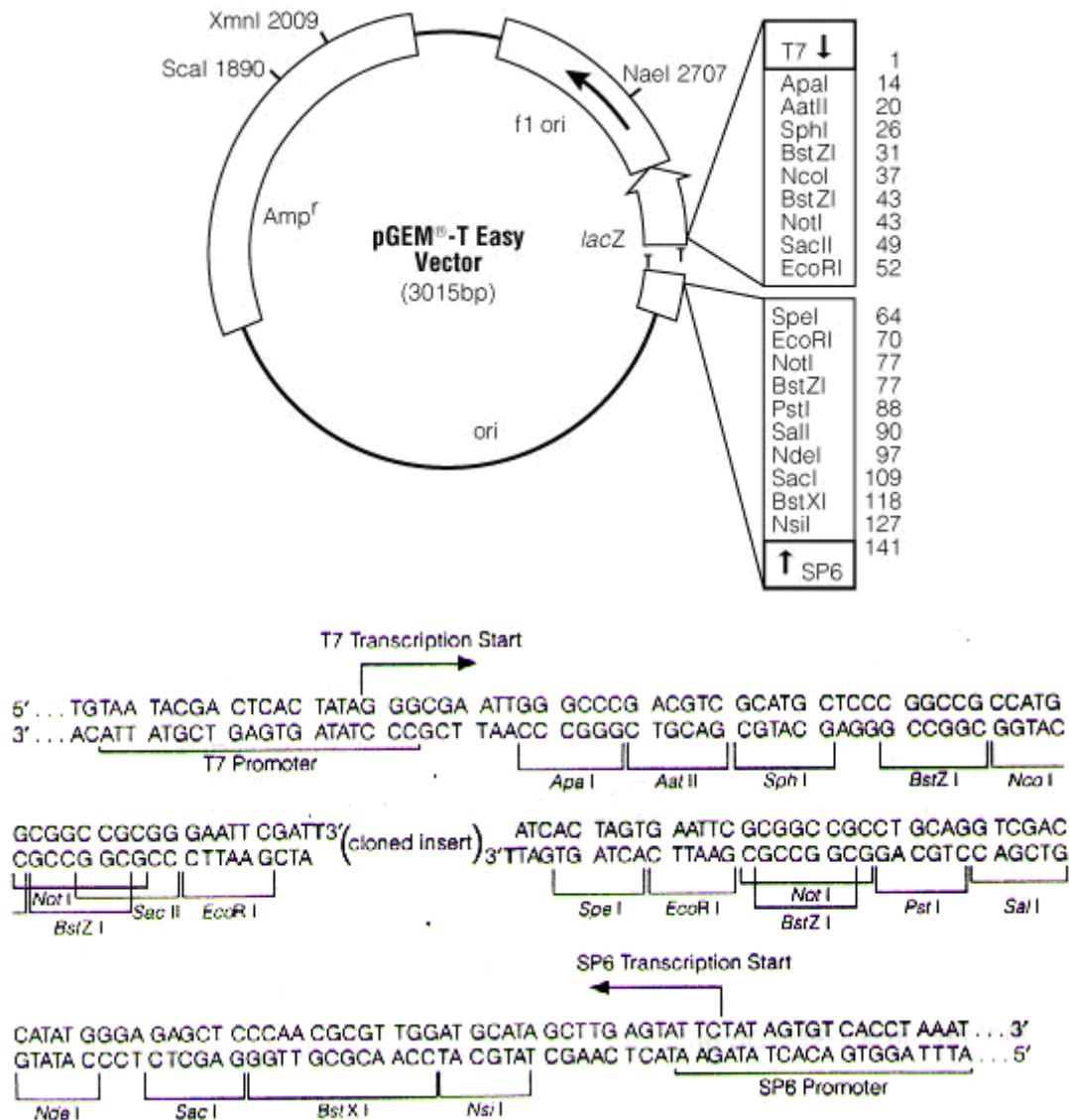
περιέχουν τα ανασυνδυασμένα πλασμίδια υπό τη δράση κατάλληλων υποστρωμάτων (Watson *et al.* 2007). Οι πλασμιδιακοί φορείς χρησιμοποιήθηκαν για την κλωνοποίηση τμημάτων DNA και για την έκφραση γονιδίων σε βακτηριακά κύτταρα και κύτταρα θηλαστικών.

➤ Φορείς κλωνοποίησης

pGEM-T Easy (Promega, #A1360): Αυτός ο πλασμιδιακός φορέας (Εικ. 5) χρησιμοποιήθηκε για την αλληλούχηση ενισχυμένων με PCR τμημάτων. Έχει μέγεθος 3015 bp και παρέχεται ως γραμμικός φορέας που είναι κατάλληλος για την κλωνοποίηση προϊόντων PCR καθώς στα 3' άκρα του φέρει προεξέχοντα νουκλεοτίδια θυμιδίνης που υβριδοποιούνται με τα προεξέχοντα νουκλεοτίδια αδενίνης που προστίθενται από πολλές Taq DNA πολυμεράσες στα 5' άκρα των προϊόντων PCR. Εκατέρωθεν της κασέτας κλωνοποίησης βρίσκονται οι υποκινητές T7 και SP6 που επιτρέπουν την αλληλούχηση και τις αντιδράσεις μεταγραφής *in vitro* με τις αντίστοιχες πολυμεράσες RNA. Επιπλέον, φέρει γονίδιο ανθεκτικότητας στην αμπικιλίνη και διαθέτει το γονίδιο *LacZ* που κωδικοποιεί το ένζυμο β-γαλακτοσιδάση παρέχοντας τη δυνατότητα επιλογής κυανών-λευκών αποικιών. Το σταθερό συνθετικό ανάλογο της λακτόζης, IPTG (isopropyl-beta-D-thiogalactopyranoside) αποτελεί μεταγραφικό ενισχυτή του γονιδίου *LacZ*.

➤ Φορείς έκφρασης σε κύτταρα θηλαστικών

pEGFP-C1 (Clontech #6084-1, Genbank ID #U55763): Ο φορέας αυτός (4,7 kb, Εικ. 6) κωδικοποιεί μια παραλλαγή του άγριου τύπου GFP (GFPmut1) με αποτέλεσμα ο φθορισμός της να διατηρείται για μεγάλα χρονικά διαστήματα. Η κασέτα κλωνοποίησης βρίσκεται μεταξύ της αλληλουχίας που κωδικοποιεί την πρωτεΐνη EGFP και του σήματος πολυαδενυλίωσης του ιού SV40. Οι χιμαιρικές πρωτεΐνες φέρουν την EGFP είτε στο καρβοξυτελικό τους άκρο (φορείς C1-3) είτε στο αμινοτελικό τους άκρο (φορείς N1-3). Έτσι επιτρέπουν τον άμεσο υποκυτταρικό εντοπισμό τους ή την ανίχνευσή τους με το αντίστοιχο αντίσωμα. Ο φορέας έχει τον υποκινητή CMV που παρέχει υψηλά επίπεδα έκφρασης και την θέση έναρξης της αντιγραφής του ιού SV40 που του επιτρέπει να πολλαπλασιάζεται σε κύτταρα θηλαστικών τα οποία εκφράζουν το αντιγόνο T. Επιπλέον, ο φορέας διαθέτει μια θέση έναρξης της αντιγραφής για τον πολλαπλασιασμό του σε βακτηριακά κύτταρα (κυρίως *E.coli*) και γονίδια ανθεκτικότητας στη νεομυκίνη και στην καναμυκίνη.

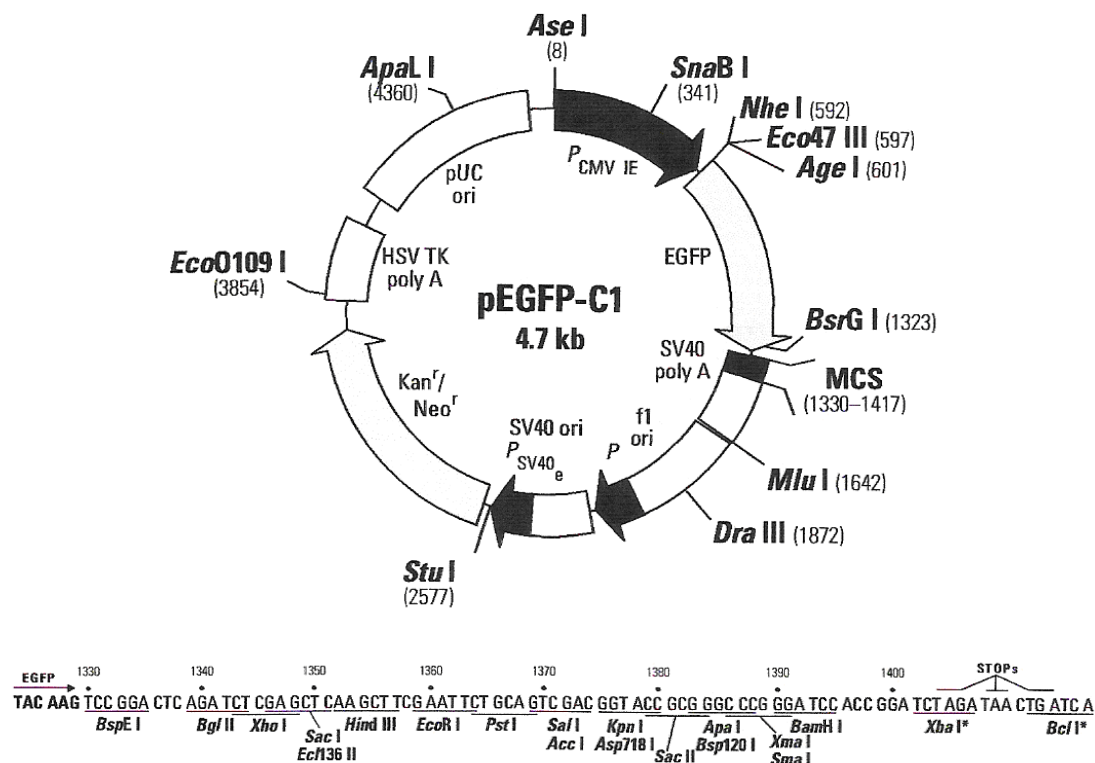


Εικ. 5: Ο πλασμιδιακός φορέας pGEM-T Easy (πάνω) και η αλληλουχία της κασέτας κλωνοποίησης (κάτω). Απεικονίζονται τα γονίδια *Amp^r* και *LacZ*, οι αλληλουχίες των εκκινήτων T7 και SP6 και οι θέσεις αναγνώρισης των διάφορων περιοριστικών ενζύμων (<http://www.promega.com/>).

2.7 Πλασμιδιακές κατασκευές

Χρησιμοποιήθηκαν οι εξής πλασμιδιακές κατασκευές: 1) η *pGEX-3X/DGCR14* (Sarafidou *et al*, υπό συγγραφή) που περιλαμβάνει το cDNA σχεδόν πλήρους μήκους του γονιδίου *DGCR14* (περιλαμβάνει τα νουκλεοτίδια που αντιστοιχούν στα αμινοξέα 25-472, δηλαδή λείπουν τα νουκλεοτίδια που κωδικοποιούν τα 24 αμινοτελικά αμινοξέα) στον πλασμιδιακό φορέα έκφρασης σε βακτηριακά κύτταρα pGEX-3X και 2) η *pcDNA/RACK1* (προσφορά από τον Prof. Jean-Luc Parent, Hospital Fleuriment, CHUS, Quebec) η οποία περιλαμβάνει το cDNA πλήρους μήκους του γονιδίου *RACK1* στον πλασμιδιακό φορέα έκφρασης σε κύτταρα

θηλαστικών pcDNA. Αυτή η πλασμιδιακή κατασκευή δεν χρησιμοποιήθηκε άμεσα επειδή δεν έφερε κάποιον επίτοπο σε σύντηξη με το cDNA ώστε να είναι δυνατή η ανίχνευση μετά την υπερέκφραση της αντίστοιχης πρωτεΐνης.



Εικ. 6: Ο πλασμιδιακός φορέας pEGFP-C1 (πάνω) και η αλληλουχία της κασέτας κλωνοποίησης (κάτω). Απεικονίζονται τα γονίδια *Kan^r* και *Neo^r*, ο υποκινητής CMV, η θέση έναρξης της αντιγραφής του ιού SV40 και οι θέσεις αναγνώρισης των διάφορων περιοριστικών ενζύμων (<http://www.promega.com/>).

ΜΕΘΟΔΟΙ

2.8 Απομόνωση νουκλεϊκών οξέων

2.8.1 Απομόνωση πλασμιδιακού DNA σε μικρή κλίμακα

Η απομόνωση πλασμιδιακού DNA σε μικρή κλίμακα πραγματοποιείται με αλκαλική λύση. Η διαδικασία είναι η ακόλουθη:

- Ανάπτυξη μοναδιαίων αποικιών των μετασχηματισμένων βακτηρίων σε 2-3 ml LB υγρό θρεπτικό υλικό με το κατάλληλο αντιβιοτικό (εξαρτάται από το πλασμίδιο) στους 37°C ολονύκτια υπό ανάδευση.
- Μεταφορά 1,5 ml της υγρής καλλιέργειας σε σωλήνα erpendorf.

- Συλλογή των κυττάρων με φυγοκέντρηση στις 3000 rpm για 15min σε θερμοκρασία δωματίου και άμεση απομάκρυνση του υπερκείμενου θρεπτικού υλικού.
- Επαναιώρηση των κυττάρων σε 300 μl διαλύματος P1 το οποίο έχει σύσταση:
 - 15mM Tris-Cl, pH=8
 - 10mM EDTA, pH=8

Η επαναιώρηση γίνεται με vortex μέχρι να μην φαίνονται συσσωματώματα κυττάρων.

- Προσθήκη 300 μl διαλύματος αποδιάταξης P2 το οποίο έχει σύσταση:
 - 200mM NaOH
 - 1% SDS

Ακολουθεί ήπια ανάδευση και επώαση για 5-10 min μέχρι το διάλυμα να γίνει διαυγές

- Προσθήκη 300 μl ρυθμιστικού διαλύματος P3 η σύσταση του οποίου είναι 3M οξικό κάλιο, pH 5,5 για την αντίστοιχη εξουδετέρωση. Ακολουθεί ισχυρή ανάδευση και επώαση στον πάγο για 15 min.
- Φυγοκέντρηση του δείγματος στις 13000 rpm στους 4°C για 15 min.
- Συλλογή του υπερκείμενου προσπαθώντας να μην συλλεχθούν άσπρα συσσωματώματα και επαναληπτική φυγοκέντρηση στις ίδιες συνθήκες.
- Μεταφορά του υπερκείμενου σε καθαρά σωληνάκια eppendorfs
- Κατακρήμνιση του πλασμιδιακού DNA με την προσθήκη ισοπροπανόλης (0,8 του όγκου του υπερκείμενου), ανάδευση και φυγοκέντρηση στις 12000 rpm στους 4°C για 15 min.
- Έκπλυση του DNA με 70% αιθανόλη.
- Φυγοκέντρηση στις 12000 rpm για 10 min στους 4°C.
- Απομάκρυνση της αιθανόλης με γυάλινη πιπέτα pasteur.
- Επαναδιάλυση του DNA σε 30 μl δις απιονισμένο νερό. Η ποσότητα του πλασμιδιακού DNA που απομονώνεται εξαρτάται από τον αριθμό αντιγράφων του πλασμιδίου και είναι της τάξεως των αρκετών μg (Sambrook *et al* 1989).

2.8.2 Απομόνωση πλασμιδιακού DNA σε μεγάλη κλίμακα για αντίδραση αλληλούχησης ή παροδική διαμόλυνση

Η διαδικασία που ακολουθείται είναι παρόμοια με την παραπάνω με τη διαφορά ότι ο όγκος της βακτηριακής καλλιέργειας αυξάνεται στα 2 lt. Οι όγκοι των διαλυμάτων P1, P2, P3 αυξάνονται στα 3 mL ενώ οι φυγοκεντρήσεις πραγματοποιούνται στις 3000 rpm για 20 min στους 4°C. Την κατακρήμνιση και την επαναδιάλυση του πλασμιδιακού DNA ακολουθεί επιπλέον επεξεργασία διότι το δείγμα πρέπει να είναι καθαρό για τη διαδικασία της αλληλούχησης ή της παροδικής διαμόλυνσης η οποία περιλαμβάνει:

- Διαδοχικές εκχυλίσεις με φαινόλη-χλωροφόρμιο μέχρι να είναι εντελώς καθαρή η μεσόφαση.
- Συλλογή του υπερκείμενου το οποίο περιέχει το DNA.
- Επανακατακρήμνιση με οξικό νάτριο (CH₃COONa, 1/10 του υπερκείμενου διαλύματος) συγκέντρωσης 3M και pH=5.2 και απόλυτη αιθανόλη (2,5 όγκοι του συνολικού όγκου του υπερκείμενου και του οξικού νατρίου).
- Ήπια ανάδευση και επώαση στους -80°C για 20 min.
- Φυγοκέντρηση στις 12000 rpm για 15 min στους 4°C με αποτέλεσμα τη συλλογή του DNA.
- Έκπλυση της πελλέτας με 70% αιθανόλη.
- Επαναδιάλυση σε ~1 ml δις-απιονισμένο νερό.
- Επώαση με RNAση σε τελική συγκέντρωση 100 µg/ml για 1h στους 37°C για την απομάκρυνση του RNA.
- Επώαση με πρωτεϊνάση K σε τελική συγκέντρωση 100 µg/ml για 2h στους 55°C σε κατάλληλο ρυθμιστικό διάλυμα 1x το οποίο περιέχει:
 - 50mM Tris.Cl, pH 7.5
 - 100mM NaCl
 - 0,02% SDS
 - 2mM EDTA
- Επακόλουθη εκχύλιση με φαινόλη/χλωροφόρμιο.
- Απομόνωση του DNA με προσθήκη απόλυτης αιθανόλης.
- Έκπλυση με 70% αιθανόλη και επαναδιάλυση σε ~2 ml δις- απιονισμένο νερό. Η ποσότητα του DNA είναι της τάξεως των λίγων mg (Sambrook *et al*, 1989).

Εναλλακτικά, για την απομόνωση πλασμιδιακού DNA που επρόκειτο να αλληλουχηθεί ή να χρησιμοποιηθεί σε διαμόλυνση των κυττάρων *HeLa*, εφαρμόστηκε το πρωτόκολλο της εταιρίας Invitrogen, PureLink™ Quick Plasmid Miniprep Kit (Catalog No K2100-10).

2.9 Τεμαχισμός DNA με περιοριστικές ενδονουκλεάσες

Η κλωνοποίηση DNA επιτυγχάνεται με την τεχνολογία του ανασυνδυασμένου DNA. Οι τεχνικές αυτές βασίζονται σε χρήση ειδικών ενζύμων που ονομάζονται περιοριστικές ενδονουκλεάσες. Τα ένζυμα αυτά απομονώθηκαν από διάφορους προκαρυωτικούς οργανισμούς, κυρίως βακτήρια, και η φυσιολογική τους λειτουργία είναι να προστατεύουν τους μικροοργανισμούς αυτούς από την εισβολή ξένου DNA μεθυλιώνοντας τις θέσεις στόχους των περιοριστικών ενδονουκλεασών. Έχουν την ιδιότητα να αναγνωρίζουν συγκεκριμένες αλληλουχίες DNA και να τεμαχίζουν το DNA σε συγκεκριμένες θέσεις (Watson *et al*. 2007).

Οι πέψεις πραγματοποιήθηκαν με βάση τις άριστες συνθήκες στις οποίες λειτουργεί ένα ένζυμο (θερμοκρασία, pH, μοριακότητα ρυθμιστικού διαλύματος σε άλατα). Ο τεμαχισμός του DNA πραγματοποιήθηκε με περίσσεια ενζύμου (3-5units/μg DNA) σε κατάλληλο ρυθμιστικό διάλυμα. Τα ρυθμιστικά διαλύματα που χρησιμοποιήθηκαν ήταν: NEB buffer 1 ή 3 και buffer H των εταιριών Biolabs και Takara αντίστοιχα σε κατάλληλη αραιώση. Με βάση τις οδηγίες της εταιρίας, χρησιμοποιήθηκε η κατάλληλη θερμοκρασία (συνήθως 37°C) και σε ορισμένες περιπτώσεις BSA (bovine serum albumin, αλβουμίνη ορού βοοειδούς). Η επώαση με την περιοριστική ενδονουκλεάση πραγματοποιείται για τουλάχιστον 3h ή και ολονύκτια για καλύτερα αποτελέσματα.

2.10 Αντίδραση σύνδεσης (ligation)

Για την αντίδραση σύνδεσης χρησιμοποιούνται ένζυμα που ονομάζονται λιγάσες. Οι λιγάσες έχουν την ιδιότητα να επανενώνουν τα τμήματα DNA που έχουν προέλθει από τον τεμαχισμό με ένζυμο περιορισμού. Έτσι δημιουργούνται νέα μόρια DNA που περιέχουν τμήματα από διαφορετικές πηγές (Watson *et al.* 2007). Για τη σύνδεση γραμμικών μορίων DNA χρησιμοποιείται η λιγάση του DNA του βακτηριοφάγου T4, η οποία καταλύει το σχηματισμό φωσφοδιεστερικών δεσμών μεταξύ των νουκλεοτιδίων στο DNA (Weiss *et al.* 1968).

Στην αντίδραση σύνδεσης, το ένθεμα και ο φορέας είναι σε μοριακή αναλογία 10:1 αντίστοιχα. Η συνολική ποσότητα του φορέα και του ενθέματος ήταν τουλάχιστον 200 ng σε τελικό όγκο 20μl. Για την αντίδραση σύνδεσης χρησιμοποιήθηκε η T4 DNA λιγάση της εταιρίας Promega (Cat. No. M180A) και ρυθμιστικό διάλυμα 10x (Promega, Cat. No. C126A 21542410) με την ακόλουθη σύσταση:

- 300mM Tris-Cl pH 7,8
- 100mM MgCl₂
- 100mM DTT
- 10mM ATP

Το διάλυμα τοποθετείται σε πάγο για να ξεπαγώσει και αναμιγνύεται με vortex μέχρι να γίνει τελείως διαυγές. Το ένζυμο προστίθεται τελευταίο και η αντίδραση σύνδεσης επωάζεται ολονύκτια στους 4°C.

2.11 Ηλεκτροφόρηση DNA σε πήκτωμα αгарόζης

Η ανίχνευση, ο έλεγχος ποιότητας και ο αδρός ποσοτικός προσδιορισμός μορίων DNA γίνεται με τη μέθοδο της ηλεκτροφόρησης. Η ηλεκτροφόρηση είναι μια μέθοδος που στηρίζεται στο γεγονός ότι τα περισσότερα μακρομόρια όπως το DNA,

το RNA και οι πρωτεΐνες, διαχωρίζονται ανάλογα με το μέγεθός τους εφόσον είναι ηλεκτρικά φορτισμένα. Τα νουκλεϊκά οξέα είναι αρνητικά φορτισμένα, λόγω των φωσφορικών τους ομάδων, με αποτέλεσμα να μετακινούνται προς τον θετικό πόλο κατά την διάρκεια της ηλεκτροφόρησης σε πηκτή αγαρόζης (πολυσακχαρίτης που παράγεται από φύκη). Η πηκτή αγαρόζης φέρει στο εσωτερικό της πόρους, που το μέγεθός τους εξαρτάται άμεσα από την συγκέντρωση της αγαρόζης. Όσο μικρότερη είναι η συγκέντρωση της χρησιμοποιούμενης αγαρόζης τόσο μεγαλύτερο είναι το μέγεθος των πόρων του σχηματιζόμενου πλέγματος. Έτσι, η ταχύτητα με την οποία το DNA μετακινείται κατά την ηλεκτροφόρησή του εξαρτάται:

- α) από το μέγεθός του
- β) την διαμόρφωσή του (υπερελικωμένο, γραμμικό)
- γ) τη συγκέντρωση της αγαρόζης
- δ) την τάση που εφαρμόζεται
- ε) την σύσταση και την ιοντική ισχύ του διαλύματος ηλεκτροφόρησης.

Για την ανίχνευση του DNA χρησιμοποιείται βρωμιούχο αιθίδιο (EtBr) το οποίο έχει την ιδιότητα να παρεμβάλλεται μεταξύ των βάσεων και να καθιστά το DNA ορατό κάτω από υπεριώδη ακτινοβολία. Η συγκέντρωση των πηκτωμάτων που χρησιμοποιήθηκαν κυμαινόταν από 1 έως 1,5% ανάλογα με το μέγεθος του DNA.

Για την πηκτή αγαρόζης χρησιμοποιήθηκαν:

- ❖ Αγαρόζη (σε κατάλληλες ποσότητες ανάλογα με τη σύσταση του πηκτώματος).
- ❖ Ρυθμιστικό διάλυμα 1xTAE. Το πυκνό διάλυμα 50xTAE έχει την εξής σύσταση:
 - 200 mM Tris
 - 200 mM οξικό οξύ
 - 1 mM EDTA pH 8.0.

Το διάλυμα ζεσταίνεται σε φούρνο μικροκυμάτων μέχρι βρασμού ώστε να λιώσει η αγαρόζη και στη συνέχεια προστίθεται EtBr σε συγκέντρωση 0,1 µg/ml. Μετά τη στερεοποίηση, η πηκτή τοποθετείται σε συσκευή ηλεκτροφόρησης που περιέχει ρυθμιστικό διάλυμα 1xTAE. Η διάρκεια και η τάση που εφαρμόστηκε ήταν 1-2h και 80-100V αντίστοιχα (Reed R.J. 2004).

2.12 Απομόνωση τεμαχίων DNA από πηκτή αγαρόζης και καθαρισμός προϊόντων PCR και πέψης

Για την απομόνωση τεμαχίων DNA από πηκτή αγαρόζης, διαχωρίστηκε το αντίστοιχο τμήμα της πηκτής με τη χρήση μιας καθαρής λεπίδας και το DNA απομονώθηκε με το πρωτόκολλο της εταιρίας QIAGEN, QIAquick gel extraction kit (Cat. No 28704).

Για τον καθαρισμό των προϊόντων PCR και των προϊόντων πέψης με περιοριστικές ενδονουκλεάσες χρησιμοποιήθηκε το πρωτόκολλο της εταιρίας Invitrogen, PureLink™ PCR Purification Kit (Cat. No K3100-02).

2.13 Αλυσιδωτή αντίδραση πολυμεράσης (PCR)

Η αλυσιδωτή αντίδραση πολυμεράσης (PCR) χρησιμοποιείται συνήθως για την αύξηση της ποσότητας DNA μιας συγκεκριμένης αλληλουχίας, η οποία βρίσκεται ανάμεσα σε δύο περιοχές γνωστής αλληλουχίας (Mullis KB 1990). Στην αντίδραση χρησιμοποιούνται δύο ολιγονουκλεοτίδια ως εναρκτήρια μόρια (εκκινητές) για μια σειρά από συνθετικές αντιδράσεις, οι οποίες καταλύονται από μία πολυμεράση DNA (π.χ. την Taq πολυμεράση) ανθεκτική σε υψηλές θερμοκρασίες. Έτσι, επιτυγχάνεται η σύνθεση της περιοχής του DNA που πλαισιώνεται από τους δύο εκκινητές. Το κάθε ολιγονουκλεοτίδιο-εκκινητής είναι συμπληρωματικό με μια από τις δύο διαφορετικές αλυσίδες του DNA-στόχου. Αρχικά, πραγματοποιείται αποδιάταξη του DNA με θέρμανση. Το μίγμα της αντίδρασης στη συνέχεια ψύχεται σε θερμοκρασία, η οποία επιτρέπει την υβριδοποίηση των εκκινητών με τις συμπληρωματικές αλυσίδες και τέλος, πραγματοποιείται επιμήκυνση των αναδιαταγμένων εναρκτήριων μορίων με την DNA πολυμεράση. Ο κύκλος της αποδιάταξης, υβριδοποίησης και επιμήκυνσης επαναλαμβάνεται πολλές φορές (25-35 κύκλοι) και τα προϊόντα κάθε κύκλου χρησιμοποιούνται ως υπόστρωμα DNA στους επόμενους κύκλους.

Στην παρούσα εργασία, η PCR χρησιμοποιήθηκε για την ενίσχυση αλληλουχιών DNA, ώστε να κλωνοποιηθούν σε κατάλληλο πλασμιδιακό φορέα. Τα ολιγονουκλεοτίδια που χρησιμοποιήθηκαν, σχεδιάστηκαν έτσι ώστε να διαθέτουν αλληλουχίες αναγνώρισης για κατάλληλα περιοριστικά ένζυμα στο άκρο 5' τους, προκειμένου το τμήμα DNA που θα προκύψει να μπορεί κατόπιν να κλωνοποιηθεί σε κατάλληλο πλασμιδιακό φορέα έκφρασης.

Για την κάθε αντίδραση χρησιμοποιούνται 100ng (για DNA από βιβλιοθήκη cDNA) ή 1ng (για πλασμιδιακό DNA) DNA, εκκινητές (forward και reverse) σε συγκέντρωση 50 pmol ο καθένας, ρυθμιστικό διάλυμα αντίδρασης 10x, μίγμα δεοξυνουκλεοτιδίων dNTPs σε τελική συγκέντρωση 2 mM το καθένα (αρχική συγκέντρωση 50 mM), διάλυμα ιόντων Mg²⁺ σε τελική συγκέντρωση 1 mM (αρχική συγκέντρωση 50 mM) και θερμοσταθερή πολυμεράση Taq (1unit/αντίδραση).

Ακολουθούν 35 κύκλοι ενίσχυσης στις συνθήκες που αναφέρονται παρακάτω:

1. αποδιάταξη του DNA με θέρμανση στους 95°C για 45sec
2. υβριδοποίηση των εκκινητών στο DNA στους 57-62°C για 45sec
3. επιμήκυνση του DNA από την Taq πολυμεράση στους 72°C για 1min

Για να βρεθεί η κατάλληλη θερμοκρασία υβριδοποίησης για τους εκκινητές πραγματοποιήθηκε gradient PCR (PCR κλίσης θερμότητας) κατά την οποία δοκιμάστηκαν διαφορετικές θερμοκρασίες υβριδοποίησης των εκκινητών στην ίδια αντίδραση.

Η αποτελεσματικότητα της PCR υπολογίζεται από την εξειδίκευση, την απόδοση και την ακρίβεια με την οποία πραγματοποιείται η αύξηση του DNA. Μία PCR υψηλής εξειδίκευσης ενισχύει μόνο ένα προϊόν, το επιθυμητό. Μία αντίδραση υψηλής απόδοσης οδηγεί σε υψηλό ποσό προϊόντος, με σχετικά λίγους κύκλους. Μια ιδανική PCR έχει υψηλή εξειδίκευση, απόδοση και ακρίβεια. Οι παράμετροι αυτοί εξαρτώνται από διάφορους παράγοντες, όπως το ρυθμιστικό διάλυμα του ενζύμου, η συγκέντρωση ιόντων Mg^{+2} , η θερμοκρασία, η διάρκεια των κύκλων και η πολυμεράση (Watson *et al.* 2007).

2.14 Μετασχηματισμός βακτηρίων με θερμικό σοκ

Μετασχηματισμός είναι ο μηχανισμός εκείνος κατά τον οποίον το βακτήριο δέχεται τμήματα ελεύθερου DNA προερχόμενα από άλλη πηγή (πλασμίδια). Ήταν ο πρώτος μηχανισμός μεταφοράς γενετικού υλικού που ανακαλύφθηκε από τον Griffith F. σε στελέχη *S. pneumoniae* και για πολλά χρόνια είχε θεωρηθεί ως ο μοναδικός τρόπος μεταφοράς. Λίγα βακτήρια είναι ικανά πρόσληψης εξωγενούς DNA *in vivo*, είναι δηλαδή επιδεκτικά κύτταρα (competent cells). Τα περισσότερα είδη βακτηρίων έχουν την ιδιότητα να προσλαμβάνουν και να ενσωματώνουν DNA από το θρεπτικό υλικό τους. Τις περισσότερες όμως φορές το «ξένο» DNA αποδομείται στην προσπάθεια ενσωμάτωσής του στο βακτηριακό γονιδίωμα. Σε άλλες περιπτώσεις και υπό συγκεκριμένες συνθήκες, ιδιαίτερα όταν το DNA είναι πλασμιδιακό, μπορεί να ενσωματωθεί πλήρως στο βακτηριακό γονιδίωμα. Το 1972 ο Cohen και οι συνεργάτες του παρατήρησαν ότι τα ιόντα Ca^{++} έχουν την ικανότητα να βελτιώνουν την ικανότητα των κυττάρων *E.coli* να προσλαμβάνουν πλασμιδιακό DNA πιθανώς λόγω της κατακρήμνισης και συσσώρευσης του DNA στην εξωτερική επιφάνεια του βακτηρίου (Cohen, Chang *et al.* 1972). Η πλήρης ενσωμάτωσή του πλασμιδιακού DNA στο βακτηριακό γονιδίωμα επιτυγχάνεται με την προσωρινή αλλαγή της θερμοκρασίας (θερμικό σοκ) από τους 37°C στους 42°C.

Στην παρούσα εργασία μετασχηματίστηκαν δεκτικά βακτήρια *E.coli* των στελεχών DH5a και BL21 codon plus. Ακολουθήθηκε η παρακάτω διαδικασία (Sambrook *et al.* 1989):

- ❖ Μεταφορά ενός δείγματος δεκτικών κυττάρων από τους -80°C σε φρέσκο πάγο.
- ❖ Επώση για 10 min για να πραγματοποιηθεί η τήξη του δείγματος.

- ❖ Προσθήκη του πλασμιδιακού DNA και ήπια ανάδευση.
- ❖ Επώαση σε πάγο για 20 min, κατά την οποία το DNA έρχεται σε επαφή με τα βακτήρια.
- ❖ Τοποθέτηση του δείγματος σε υδατόλουτρο στους 42°C για 90 sec και άμεση μεταφορά στον πάγο όπου παραμένει για 1 min με αποτέλεσμα τα βακτήρια να υφίστανται θερμικό σοκ προκειμένου να δεχθούν το πλασμίδιο στο εσωτερικό τους.
- ❖ Προσθήκη 1 mL υγρού LB και επώαση για 1h στους 37°C ώστε να εκφραστεί το γονίδιο ανθεκτικότητας του πλασμιδίου ώστε να είναι δυνατή η επιλογή των μετασχηματισμένων βακτηρίων.
- ❖ Συλλογή των κυττάρων με φυγοκέντρηση στις 2500 rpm για 10 min σε θερμοκρασία δωματίου.
- ❖ Απομάκρυνση της μεγαλύτερης ποσότητας του υπερκειμένου θρεπτικού υλικού αφήνοντας ελάχιστη ποσότητα για την επαναιώρηση της πελλέτας των κυττάρων.
- ❖ Επίστρωση των κυττάρων σε αποστειρωμένο και στεγνό τρυβλίο Petri με το κατάλληλο αντιβιοτικό (εξαρτάται από το πλασμίδιο).
- ❖ Επώαση στους 37°C για την ανάπτυξη των βακτηριακών κυττάρων.

2.15 Ηλεκτροφόρηση πρωτεϊνών σε αποδιατακτικό πήκτωμα πολυακρυλαμίδης (SDS-PAGE)

Για την ανάλυση και το διαχωρισμό των πρωτεϊνών με κριτήριο το μέγεθός τους χρησιμοποιήθηκε η μέθοδος της ηλεκτροφόρησης πρωτεϊνών σε αποδιατακτικό πήκτωμα πολυακρυλαμίδης (Raymond & Weintraub 1959). Η παρουσία του ανιονικού απορρυπαντικού SDS (δωδεκακυλοθειικού νατρίου) στο πήκτωμα και στο διάλυμα ηλεκτροφόρησης καθώς και το διάλυμα Laemmli που προστίθεται στα δείγματα έχει ως αποτέλεσμα οι πρωτεΐνες να αποδιατάσσονται και να φορτίζονται αρνητικά έτσι ώστε να διαχωρίζονται σύμφωνα με το μοριακό τους βάρος.

Τα πολυπεπτίδια προσδένουν SDS, σε ποσότητα ανάλογη με το μέγεθός τους, φορτίζονται αρνητικά και παρουσία ηλεκτρικού πεδίου μετακινούνται προς την άνοδο. Πιο συγκεκριμένα, το SDS διασπά τους υδρογονικούς δεσμούς και τις υδρόφοβες αλληλεπιδράσεις αποδιατάσσοντας με αυτόν τον τρόπο τις πρωτεΐνες. Η πλήρης αποδιάταξη πραγματοποιείται με την επίδραση της μερκαπτοαιθανόλης η οποία ανάγει τους δισουλφιδικούς δεσμούς που αναπτύσσονται ανάμεσα σε δύο ομάδες -κυστεινών της ίδιας ή δύο διαφορετικών πολυπεπτιδικών αλυσίδων καθώς και με τη θέρμανση των δειγμάτων στους 100°C. Ο σχηματισμός του πηκτώματος γίνεται σε θερμοκρασία δωματίου με πολυμερισμό των μονομερών ακρυλαμίδης με

τη βοήθεια δύο πολυμεριστικών παραγόντων: του υπερθειικού αμμωνίου (ammonium persulfate, APS) και του TEMED (N,N,N,N-τετραμεθυλο-1,2-διαμινοαιθάνιο), το οποίο καταλύει το σχηματισμό ελεύθερων ριζών από το APS. Έτσι, σχηματίζεται ένα πλέγμα που αποτελείται από πόρους μέσα στους οποίους κινούνται οι πρωτεΐνες. Γενικά, πηκτώματα με μικρή συγκέντρωση πολυακρυλαμίδης έχουν μεγαλύτερους πόρους και το αντίστροφο.

Το «καλούπι» στο οποίο σχηματίστηκε το πήκτωμα ήταν ένα «σάντουιτς» επίπεδων τζαμιών. Το πήκτωμα αποτελείται από δύο φάσεις:

- την ανώτερη (stacking gel) στην οποία πραγματοποιείται η ευθυγράμμιση των πρωτεϊνών σε ενιαίο μέτωπο
- και την υποκείμενη, δηλαδή το πήκτωμα διαχωρισμού (separating gel).

Σύμφωνα με το πρωτόκολλο ηλεκτροφορητικού διαχωρισμού του Laemmli, η ανώτερη φάση του πηκτώματος δεν διαχωρίζει τις πρωτεΐνες αλλά μάλλον τις «ευθυγραμμίζει» σε μία μικρή ζώνη, γεγονός που εξασφαλίζει ότι οι πρωτεΐνες κάθε δείγματος θα φθάσουν ταυτόχρονα στο κυρίως πήκτωμα διαχωρισμού (Laemmli 1970).

Οι συγκεντρώσεις των συστατικών του πηκτώματος ευθυγράμμισης είναι:

- ❖ 4,5% ακρυλαμίδη/δισακρυλαμίδη (29:1)
- ❖ 1,25 M Tris-Cl pH 6,8
- ❖ 20% SDS
- ❖ 20% APS (Ammonium Persulfate)
- ❖ 0,2% TEMED (N,N,N',N'- tetramethylethylene diamine)

Οι συγκεντρώσεις των συστατικών του πηκτώματος διαχωρισμού είναι:

- ❖ 10% ακρυλαμίδη/δισακρυλαμίδη (29:1)
- ❖ 1.25 M Tris.Cl pH 8,8
- ❖ 20% SDS
- ❖ 20% APS
- ❖ 0,05% TEMED

Ακολουθεί αποδιάταξη των πρωτεϊνικών δειγμάτων με διάλυμα 5X Laemmli, το οποίο έχει την ακόλουθη σύσταση:

- 200 mM Tris-Cl pH 6,8
- 8% SDS
- 40% γλυκερόλη
- 20% μερκαπτοαιθανόλη
- 0,4% bromophenol blue

και θέρμανση στους 100°C για 10 min.

Η τοποθέτηση των δειγμάτων γίνεται σε κατάλληλες θέσεις υποδοχής (slots) που δημιουργούνται στο πήκτωμα ευθυγράμμισης με την τοποθέτηση ειδικής «χτένας», όσο το πήκτωμα είναι ακόμα υγρό. Η ηλεκτροφόρηση πραγματοποιήθηκε σε ρυθμιστικό διάλυμα 1x από αρχικό διάλυμα 10x με την εξής σύσταση:

- 250mM Tris
- 1,92M γλυκίνη
- 1% SDS

με pH ρυθμισμένο στο 8,3 με διάλυμα HCl.

Η ηλεκτροφόρηση έγινε με σταθερή ένταση πεδίου 35-45 mA μέχρι η χρωστική μπλε της βρωμοφαινόλης (συστατικό του διαλύματος Laemmli) να φθάσει λίγο πριν το άκρο του πήκτωματος. Μετά την ολοκλήρωση της ηλεκτροφόρησης οι πρωτεΐνες στο πήκτωμα εμφανίζονται με χρώση κυανού Coomassie ή ακολουθεί ηλεκτρομεταφορά των πολυπεπτιδίων σε μεμβράνη για ανάλυση Western (Reed R.J. 2004).

2.16 Χρώση πήκτωμάτων

Η χρώση και η μονιμοποίηση των πολυπεπτιδίων στο πήκτωμα πραγματοποιήθηκε με διάλυμα Coomassie brilliant blue με την ακόλουθη σύσταση:

- 45% v/v μεθανόλη
- 10% v/v οξικό οξύ
- 0,25% χρωστικής Coomassie Brilliant blue R250 (διαλυτή σε μεθανόλη)

Η διακριτική ικανότητα της μεθόδου είναι 30-100 ng πρωτεΐνης. Για τη χρώση, το πήκτωμα τοποθετείται στο διάλυμα της χρωστικής με συνεχή ήπια ανάδευση από 2h έως ολονύκτια σε θερμοκρασία δωματίου. Το πήκτωμα ξεπλένεται με νερό και μεταφέρεται σε διάλυμα αποχρωματισμού, ώστε να απομακρυνθεί η χρωστική που δεν έχει δεσμευτεί στα πολυπεπτίδια, με την εξής σύσταση:

- 30% v/v μεθανόλη
- 10% v/v οξικό οξύ
- 3% v/v γλυκερόλη, εάν είναι επιθυμητή η αποξήρανση

Ακολουθεί επώαση υπό ανάδευση σε θερμοκρασία δωματίου μέχρι το πήκτωμα να αποχρωματιστεί πλήρως (2h έως ολονύκτια).

2.17 Παραγωγή ανασυνδυασμένων πρωτεϊνών σε βακτήρια και απομόνωση με χρωματογραφία συγγένειας

➤ Παραγωγή ανασυνδυασμένης πρωτεΐνης σε βακτήρια

Οι προκαρυωτικοί φορείς έκφρασης φέρουν κατάλληλους επαγωγίμους υποκινητές οι οποίοι ρυθμίζουν την έκφραση του επιθυμητού γονιδίου που έχει κλωνοποιηθεί σε αυτούς. Για την παραγωγή των ανασυνδυασμένων πρωτεϊνών, πραγματοποιείται

μετασχηματισμός βακτηριακών δεκτικών κυττάρων με τις αντίστοιχες πλασμιδιακές κατασκευές και επαγωγή της έκφρασης με την προσθήκη κατάλληλων ουσιών (IPTG) (Reece R.J. 2004).

Στην παρούσα εργασία χρησιμοποιήθηκε η πλασμιδιακή κατασκευή *pGEX-DGCR14* και το στέλεχος *E. coli* BL21 (DE3) Codon plus. Η διαδικασία που ακολουθήθηκε είναι η εξής (Sambrook *et al.* 1989):

- Μετασχηματισμός και καλλιέργεια των βακτηρίων σε στερεό θρεπτικό υλικό παρουσία αμπικιλίνης και χλωραμφαινικόλης
- Απομόνωση μοναδιαίας αποικίας και εμβολιασμός σε υγρό θρεπτικό υλικό (3ml LB υγρό με αμπικιλίνη και χλωραμφαινικόλη)
- Ολονύκτια επώαση στους 37°C υπό ανάδευση.
- Εμβολιασμός φρέσκου υγρού θρεπτικού υλικού (200 ml) με την ολονύκτια καλλιέργεια (αραίωση 1:100) παρουσία των κατάλληλων αντιβιοτικών χρησιμοποιώντας φλάσκες των 2 lt για να υπάρχει κατάλληλος αερισμός για την ανάπτυξη των βακτηρίων.
- Επώαση της καλλιέργειας μέχρι την εκθετική φάση ανάπτυξης (οπτική πυκνότητα, $O.D_{600nm} = 0,6-0,7$).
- Συλλογή μικρής ποσότητας της καλλιέργειας (~10 ml) η οποία αποτελεί τον αρνητικό έλεγχο του πειράματος.
- Προσθήκη IPTG (ισοπρότυλο-β-D-θειογαλακτοσίδιο) σε τελική συγκέντρωση 0,1-1mM [αρχική συγκέντρωση (στοκ) 0,5M] και επώαση για διάφορους χρόνους (2h έως ολονύκτια) και σε διάφορες θερμοκρασίες (20-37°C) για την ανεύρεση των βέλτιστων συνθηκών παραγωγής της ανασυνδυασμένης πρωτεΐνης.
- Συλλογή των κυττάρων με φυγοκέντρηση στις 3000 rpm για 20 min.
- Έλεγχος ενός κλάσματος της πελλέτας των βακτηρίων για την έκφραση της πρωτεΐνης και αποθήκευση της υπόλοιπης βακτηριακής πελέτας στους -80° C ή άμεση χρήση της για την απομόνωση της πρωτεΐνης.

➤ Χρωματογραφία συγγένειας με υπόστρωμα σεφαρόζη-γλουταθειόνη

Η μέθοδος αυτή εφαρμόστηκε για την απομόνωση της χιμαιρικής πρωτεΐνης GST-DGCR14 που περιλαμβάνει στο αμινοτελικό άκρο της την τρανσφεράση της S-γλουταθειόνης και στο καρβοξυτελικό της άκρο το προϊόν του προς μελέτη γονιδίου. Οι τρανσφεράσες της S-γλουταθειόνης είναι μια οικογένεια ενζύμων που συμμετέχουν στην προστασία του κυττάρου από ξενοβιοτικές χημικές ενώσεις. Καταλύουν την προσθήκη της γλουταθειόνης σε αυτά τα υποστρώματα με αποτέλεσμα να αυξάνουν τη διαλυτότητά τους στο νερό και να προάγουν την

περαιτέρω ενζυματική αποδόμησή τους (Strange, Jones & Fryer 2000). Η γλουταθειόνη είναι ένα τριπεπτίδιο που αποτελείται από γλουταμικό οξύ, κυστεΐνη και γλυκίνη και συνδέεται με την τρανσφεράση της γλουταθειόνης με υψηλό βαθμό συγγένειας (Walker *et al.* 1993).

Η απομόνωση της χιμαιρικής πρωτεΐνης πραγματοποιήθηκε με την παρακάτω διαδικασία (Frangioni *et al.* 1993).

- Επαναδιάλυση του βακτηριακού ιζήματος σε διάλυμα δέσμευσης το οποίο περιέχει:
 - 50 mM Tris pH 8
 - 100 mM NaCl
 - 1 mM EDTA
 - 0,1% NP-40
- Επώαση με λυσοζύμη σε πάγο για 1h (ένζυμο το οποίο διασπά το κυτταρικό τοίχωμα των βακτηρίων) σε τελική συγκέντρωση 300μg/ml (αρχική συγκέντρωση 10 mg/ml).
- Προσθήκη DTT (διθειοθρεϊτόλη) σε τελική συγκέντρωση 5 mM (αρχική συγκέντρωση 0,1M). Η διαδικασία αυτή ακολουθείται διότι η πρωτεΐνη εντοπίζεται στο αδιάλυτο κλάσμα.
- Επώαση για 10 min σε πάγο.
- Προσθήκη απορρυπαντικού sarcosyl (άλας νατρίου, υπεύθυνο για τη λύση των κυττάρων) σε τελική συγκέντρωση 1,5% (από αρχικό διάλυμα 10%).
- Επώαση για 30 min σε πάγο.
- Επεξεργασία με υπέρηχους για 2-3 min ώστε να σπάσουν μηχανικά τα νουκλεϊκά οξέα και να μειωθεί το ιξώδες του δείγματος (λόγω υπερθέρμανσης, κάθε 20 sec το δείγμα τοποθετείται σε πάγο για 1min).
- Φυγοκέντρηση στις 12000 rpm στους 4°C για 20 min.
- Συλλογή του υπερκλειμένου, επαναληπτική φυγοκέντρηση και συλλογή του υπερκλειμένου.
- Προσθήκη μη ιοντικού απορρυπαντικού Triton σε τελική συγκέντρωση 2% (από αρχικό διάλυμα 20%) για να μειωθούν τα επίπεδα υπεροξειδίων, καρβονυλικών ενώσεων και αλάτων.
- Φυγοκέντρηση και συλλογή του υπερκλειμένου.

Για την απομόνωση της ανασυνδυασμένης πρωτεΐνης χρησιμοποιήθηκαν σφαιρίδια γλουταθειόνης. Η συγγένεια των σφαιριδίων γλουταθειόνης με την τρανσφεράση της S-γλουταθειόνης που βρίσκεται στο αμινοτελικό άκρο της πρωτεΐνης επιτρέπει την πρόσδεσή της στα σφαιρίδια με μεγάλο βαθμό συγγένειας (Reece R.J. 2004). Στην παρούσα εργασία χρησιμοποιήθηκαν σφαιρίδια Sepharose 4B της εταιρίας

Pharmacia. Πριν τη χρήση των σφαιριδίων προηγήθηκε η εξισορρόπησή τους με προσθήκη διαλύματος δέσμευσης, ανάδευση και φυγοκέντρηση σε χαμηλές στροφές (800 rpm στους 4°C) για 3-4 min (δύο φορές). Στη συνέχεια, το πρωτεϊνικό εκχύλισμα προστέθηκε στα σφαιρίδια και το δείγμα επωάστηκε ολονύκτια στους 4°C με περιστροφική ανάδευση. Ακολούθησαν τρεις διαδοχικές εκπλύσεις της στήλης με διάλυμα δέσμευσης, κατά τις οποίες τα σφαιρίδια συλλέγονται με φυγοκέντρηση στις 800 rpm για 1 min στους 4°C. Το δείγμα διατηρείται παρουσία διαλύματος δέσμευσης στους 4°C.

2.18 Καλλιέργεια κυττάρων θηλαστικών (*HeLa*)

➤ Συνθήκες και βασικές αρχές κυτταρικής καλλιέργειας

Ο χειρισμός και η ανάπτυξη όλων των κυτταρικών σειρών πραγματοποιείται σε ειδικά διαμορφωμένο χώρο (cell culture room) και η ανάπτυξη των κυττάρων γίνεται σε επωαστικό κλίβανο με τις ακόλουθες συνθήκες:

- σταθερή θερμοκρασία 37°C
- κατάλληλες συνθήκες υγρασίας
- ατμόσφαιρα εμπλουτισμένη με 5% CO₂ για την διατήρηση του pH στο θρεπτικό υλικό της καλλιέργειας.

Σε αυτήν την εργασία χρησιμοποιήθηκε η κυτταρική σειρά *HeLa* για την μελέτη του υποκυτταρικού εντοπισμού της ανασυνδυασμένης πρωτεΐνης RED-GFP. Τα κύτταρα *HeLa* αναπτύσσονται σε φλάσκες ή τρυβλία σχηματίζοντας μονοστοιβάδες. Η παρατήρηση των κυττάρων γίνεται σε οπτικό μικροσκόπιο και η μακροχρόνια συντήρησή τους σε δεξαμενή υγρού αζώτου (-196 °C ως -156 °C) (Sambrook *et al.* 1989).

➤ Διατήρηση κυτταροκαλλιιεργειών

Η κυτταροκαλλιέργεια διατηρείται σε φλάσκες (T-75 ή T-25) ενώ οι πειραματικές διαδικασίες πραγματοποιούνται σε τρυβλία διαμέτρου 3 ή 10 cm. Τα κύτταρα, όταν καλλιεργούνται σε στερεό υπόστρωμα, διαχωρίζονται και συλλέγονται σε εκθετική φάση ανάπτυξης, δηλαδή όταν έχουν καλύψει ~80% της διαθέσιμης επιφάνειας, όταν η ανάπτυξή τους δεν παρεμποδίζεται από την έλλειψη χώρου. Σε πολλά είδη κυττάρων, όταν παρατηρείται έλλειψη χώρου, αναστέλλεται ο κυτταρικός πολλαπλασιασμός και τροποποιούνται οι μεταβολικές τους διαδικασίες. Ορισμένα μάλιστα είδη κυττάρων όταν δεν υπάρχει διαθέσιμη επιφάνεια για να προσκολληθούν, πεθαίνουν, αποκολλώνται και αιωρούνται στο θρεπτικό μέσο της καλλιέργειας. Επομένως, η ανακαλλιέργειά τους θα πρέπει να πραγματοποιείται σε

τακτά χρονικά διαστήματα (3-4 μέρες), έτσι ώστε να διασφαλίζεται ο κατάλληλος αριθμός των κυττάρων για πειράματα αλλά και η ομαλή τους ανάπτυξη. Η αποκόλληση των κυττάρων από το υπόστρωμα είναι δυνατή:

- α) με μηχανικό τρόπο
- β) με χημικά μέσα
- γ) με ενζυμική δράση

Συγκεκριμένα, στην παρούσα εργασία, για την αποκόλληση των κυττάρων χρησιμοποιήθηκε το παγκρεατικό ένζυμο θρυψίνη (διαλυμένο σε EDTA που είναι χηλικό αντιδραστήριο), το οποίο διασπά τις συνδέσεις των κυττάρων μεταξύ τους αλλά και με το στερεό τους υπόστρωμα. Η αποκόλληση γίνεται ως εξής:

- Απομάκρυνση του θρεπτικού μέσου.
- Έκπλυση των κυττάρων με ρυθμιστικό διάλυμα φωσφορικών αλάτων 1xPBS, pH 7,4.
- Επώαση με το παγκρεατικό ένζυμο θρυψίνη σε ρυθμιστικό διάλυμα 1xPBS σε κατάλληλη αραίωση 1:10.
- Προσθήκη θρεπτικού μέσου DMEM (προθερμασμένου στους 37°C) εμπλουτισμένου με FBS το οποίο αναστέλλει τη δράση της θρυψίνης. Η προσθήκη πραγματοποιείται μόλις γίνει η αποκόλληση όλων των κυττάρων από την επιφάνεια της φλάσκας (2-10 min).

Με το τέλος της διαδικασίας, το ένζυμο δεν δρα πλέον έναντι των πρωτεϊνών των κυττάρων, αντιθέτως δρα έναντι των πρωτεϊνών του θρεπτικού υλικού.

➤ Κρυοδιατήρηση κυττάρων

Κατά την κρυοδιατήρηση των κυττάρων πρέπει να πραγματοποιείται σταδιακή μετάβασή τους στην τελική θερμοκρασία κατάψυξης, ώστε να αποφευχθεί ο σχηματισμός πάγου στο εσωτερικό τους και να ελαχιστοποιηθεί η απώλεια νερού. Όταν τα κύτταρα βρίσκονται σε εκθετική φάση ανάπτυξης (~80% κάλυψη της διαθέσιμης επιφάνειας, για κύτταρα που αναπτύσσονται σχηματίζοντας μονοστοιβάδα) γίνεται η κατάψυξή τους με την εξής διαδικασία:

- Διαδοχικές πλύσεις με 1x PBS
- Διαδικασία της θρυψινοποίησης
- Συλλογή των κυττάρων με φυγοκέντρηση για 7 min στις 1200 rpm, σε θερμοκρασία δωματίου.
- Επαναίωση της πελλέτας των κυττάρων σε θρεπτικό μέσο εμπλουτισμένο με FBS (πενικιλίνη και στρεπτομυκίνη). Προστίθεται αρχικά μικρή ποσότητα θρεπτικού μέσου για την επαναδιάλυση της πελλέτας των κυττάρων και έπειτα προστίθεται και η υπόλοιπη ποσότητα.

- Προσθήκη ίσου όγκου διαλύματος 20% DMSO (σε DMEM με 10% FBS, πενικιλίνη και στρεπτομυκίνη). Η προσθήκη του διαλύματος αυτού γίνεται σταγόνα-σταγόνα με συνεχή ανάδευση έτσι ώστε να αποφευχθεί το οσμωτικό σοκ και να ελαχιστοποιηθεί η αύξηση της θερμοκρασίας που προκαλείται από την διάλυση του DMSO.
- Μεταφορά του δείγματος σε κατάλληλο φιαλίδιο (cryovial) το οποίο, για λόγους μόνωσης, τυλίγεται με χαρτί.
- Επώαση για 20-30 min στους -20°C .
- Παραμονή του δείγματος στους -80°C ολονύκτια και αποθήκευση σε δεξαμενή υγρού αζώτου (-196°C ως -156°C).

➤ Επανακαλλιέργεια

Κατά τη διαδικασία αυτή, τα κύτταρα που έχουν παραμείνει παγωμένα ακόμα και για μεγάλο χρονικό διάστημα σε δεξαμενή υγρού αζώτου και πρόκειται να επανακαλλιεργηθούν, μεταφέρονται σε υδατόλουτρο στους 37°C όπου και αναδεύονται ήπια μέχρι την πλήρη τήξη του θρεπτικού μέσου (1-2 min). Η διαδικασία γίνεται υπό άσηπτες συνθήκες και όσο το δυνατόν ταχύτερα, προκειμένου τα κύτταρα να μην εκτεθούν στο DMSO, το οποίο μπορεί να δράσει τοξικά καθώς και για να αποφευχθεί ο σχηματισμός ενδοκυτταρικών κρυστάλλων κατά τη διαδικασία της ενυδάτωσής τους. Ακολουθεί μεταφορά στη φιάλη καλλιέργειας στην οποία έχει ήδη προστεθεί DMEM (με 10% FBS, πενικιλίνη και στρεπτομυκίνη).

2.19 Παροδική διαμόλυνση κυττάρων θηλαστικών με πλασμιδιακό DNA

Υπό κατάλληλες συνθήκες, μόρια ανασυνδυασμένου πλασμιδιακού DNA μπορεί να εισαχθούν και να ενσωματωθούν στον πυρήνα ευκαρυωτικών κυττάρων όπου είναι δυνατόν να εκφραστούν. Αυτή η δυνατότητα, επιτρέπει τη μελέτη της λειτουργίας και της ρύθμισης της έκφρασης των γονιδίων. Η έκφραση του πλασμιδιακού DNA μπορεί να είναι μόνιμη (stable) εφόσον το ξένο DNA ενσωματώνεται στο χρωμόσωμα ή παροδική (transient) εφόσον δεν ενσωματώνεται στο χρωμόσωμα αλλά αποδομείται μετά την παρέλευση ενός μιτωτικού κύκλου. Στην παρούσα εργασία η έκφραση του πλασμιδιακού DNA ήταν παροδική, δηλαδή τα κύτταρα που ενσωμάτωσαν το πλασμιδιακό DNA το εξέφρασαν για 12-72 h. Ωστόσο, τα επίπεδα των πρωτεϊνών είναι ικανοποιητικά ώστε να χρησιμοποιηθούν σε βιοχημικές μελέτες, σε πειράματα υποκυτταρικού εντοπισμού και σε ανίχνευση πρωτεϊνικών αλληλεπιδράσεων (Reece R.J. 2004). Το πλασμιδιακό DNA που χρησιμοποιείται πρέπει να είναι ελεύθερο πρωτεϊνών, RNA ή άλλων χημικών ουσιών. Για αυτό το λόγο, το DNA πρέπει να ανασυσταθεί σε δις-απιονισμένο νερό και να

ελεγχθεί ο λόγος απορρόφησης του στα 260/280 nm. Αποδεκτές είναι τιμές μεγαλύτερες του 1,5.

Η κυτταρική σειρά *HeLa* που χρησιμοποιήθηκε αναπτύσσεται προσκολλημένη στον πυθμένα των πλακών καλλιέργειας (adherent cells) σχηματίζοντας μονοστοιβάδες. Η παροδική διαμόλυνση των κυττάρων *HeLa* πραγματοποιήθηκε με τη μέθοδο συνκατακρήμνισης του DNA με $\text{Ca}_3(\text{PO}_4)_2$.

➤ Μέθοδος συνκατακρήμνισης του DNA με $\text{Ca}_3(\text{PO}_4)_2$

Είναι μια μέθοδος παροδικής διαμόλυνσης σε κύτταρα θηλαστικών η οποία επιτρέπει την είσοδο κλώνων cDNA που έχουν εισαχθεί σε πλασμιδιακούς φορείς έκφρασης και την παροδική έκφρασή τους σε αυτά. Οι πλασμιδιακοί αυτοί φορείς έκφρασης διαθέτουν υποκινητές που έχουν δυνατότητα ισχυρής ενεργοποίησης των γονιδίων που φέρουν (π.χ. CMV, SV40). Η παροδική διαμόλυνση επιτυγχάνεται με την συνκατακρήμνιση του DNA με $\text{Ca}_3(\text{PO}_4)_2$ (Graham FL & Van der Eb AJ, 1973). Με την προσθήκη του $\text{Ca}_3(\text{PO}_4)_2$ στο διάλυμα του DNA δημιουργούνται σύμπλοκα (με τη μορφή λεπτών κρυστάλλων) μεταξύ του DNA και θετικών φορισμένων μορίων-φορέων, τα οποία προσκολλώνται στην κυτταρική μεμβράνη και εισέρχονται στο κύτταρο με ενδοκύτωση (Orrantia & Chang 1990). Το ποσοστό των διαμολυσμένων κυττάρων κυμαίνεται μεταξύ 5% και 20% ενώ η απόδοση εξαρτάται κυρίως από την καθαρότητα του πλασμιδιακού DNA και την καλή κατάσταση των κυττάρων.

Η μέθοδος περιλαμβάνει τα εξής στάδια (Sambrook, *et al.* 1989):

1. Επίστρωση των κυττάρων σε τρυβλία 10 cm, 24 ώρες πριν τη διαμόλυνση. Τα κύτταρα πρέπει να καλύπτουν το ~50% της διαθέσιμης επιφάνειας καλλιέργειας.
2. Αλλαγή του θρεπτικού μέσου 1-2 h πριν τη διαμόλυνση.
3. Η ποσότητα του δείγματος DNA πρέπει να είναι 20 μg σε τελικό όγκο H_2O 450 μl.
4. Προσθήκη στο διάλυμα DNA, 50 μl διαλύματος CaCl_2 2,5 M με καλή ανάμιξη του μίγματος.
5. Σταδιακή προσθήκη με ταυτόχρονη ανάμιξη 500 μl διαλύματος 2xHEBS με την ακόλουθη σύσταση:
 - 280mM NaCl
 - 50mM Hepes
 - 5mM Na_2HPO_4 , pH 7,12

6. Επώαση 5-20 min σε θερμοκρασία δωματίου για να σχηματιστούν τα σύμπλοκα DNA-Ca₃(PO₄)₂. Ο χρόνος επώασης εξαρτάται από την καθαρότητα του DNA για να αποφευχθεί η δημιουργία ιζήματος.
7. Προσθήκη του μίγματος σταγόνα-σταγόνα στα κύτταρα.
8. Επώαση των κυττάρων για 24-48 h σε επωαστικό κλίβανο.
9. Έκπλυση των κυττάρων δυο φορές με διάλυμα φωσφορικών αλάτων 1XPBS και αλλαγή θρεπτικού μέσου.

2.20 Μικροσκοπία φθορισμού

Η ανίχνευση και ο υποκυτταρικός εντοπισμός της ανασυνδυασμένης πρωτεΐνης RED-EGFP στα κύτταρα *HeLa*, βασίστηκε στην έκφραση της πράσινης φθορίζουσας πρωτεΐνης EGFP. Για την ανίχνευση χρησιμοποιήθηκε μικροσκοπία φθορισμού. Τα προς εξέταση κύτταρα καλλιεργήθηκαν σε αποστειρωμένες γυάλινες καλυπτρίδες σε τρυβλίο διαμέτρου 10 cm όπου και πραγματοποιήθηκε η διαδικασία της παροδικής διαμόλυνσης όπως αναφέρθηκε παραπάνω. Στη συνέχεια, πραγματοποιήθηκαν δυο διαδοχικές πλύσεις με ρυθμιστικό διάλυμα 1XPBS. Τα κύτταρα μονιμοποιήθηκαν με διάλυμα φορμαλδεΐδης 4% (αρχικό διάλυμα 37%) για 5 min και ξεπλύθηκαν με ρυθμιστικό διάλυμα 1XPBS. Για την παρατήρηση των κυττάρων στο μικροσκόπιο φθορισμού, η καλυπτρίδα μεταφέρθηκε σε αντικειμενοφόρο πλάκα με τη χρήση διαλύματος επικάλυψης και παρεμπόδισης αποχρωματισμού (DPX mountant for microscopy, BDH Laboratory Supplies England, D2650). Τέλος, ακολούθησε στεγανοποίηση του δείγματος με βερνίκι στα άκρα της καλυπτρίδας, παρατήρηση και φωτογράφιση του δείγματος.

3. ΑΠΟΤΕΛΕΣΜΑΤΑ - ΣΥΖΗΤΗΣΗ

3.1 *In silico* προσδιορισμός του δικτύου πρωτεϊνικών αλληλεπιδράσεων του σωματίου συναρμογής (spliceosome) με αφετηρία την πρωτεΐνη FRA10AC1

➤ *In silico* ταυτοποίηση των πρωτεϊνών του σωματίου συναρμογής

Για να προσδιοριστεί η ολοκληρωμένη λειτουργία του σωματίου συναρμογής είναι απαραίτητος ο προσδιορισμός των συστατικών του, της δομής και λειτουργίας των υπομονάδων του καθώς και των δυναμικών αλληλεπιδράσεων μεταξύ τους. Η εργασία αυτή στοχεύει στην ανασύσταση του δικτύου πρωτεϊνικών αλληλεπιδράσεων του σωματίου συναρμογής με αφετηρία την πρωτεΐνη FRA10AC1. Για αυτό το σκοπό είναι απαραίτητη αρχικά η συλλογή όλων των πληροφοριών σχετικά με τη σύσταση του σωματίου συναρμογής. Τα τελευταία χρόνια, με σειρά πρωτεομικών μελετών, έχει επιτευχθεί η βιοχημική απομόνωση και η ταυτοποίηση των υπομονάδων είτε ολόκληρου του σωματίου είτε των επιμέρους συμπλόκων του από διαφορετικές ερευνητικές ομάδες μέσω χρωματογραφίας συγγένειας σε συνδυασμό με φασματομετρία μάζας και πλέον ο αριθμός των πρωτεϊνών-συστατικών του σωματίου ανέρχεται σε εκατοντάδες (Πίν. 2).

Πιν. 2: Αριθμός πρωτεϊνών που απομονώθηκαν είτε από ολόκληρο το σωματίο συναρμογής είτε από τα επιμέρους σύμπλοκα στις διαφορετικές ερευνητικές εργασίες.

Ερευνητική εργασία	Αριθμός πρωτεϊνών	Σύμπλοκα του σωματίου συναρμογής
Bessonov <i>et al</i> 2008	258	B,C
Herold <i>et al</i> 2008	166	B,C
Behzadnia <i>et al</i> 2007	113	A
Chen <i>et al</i> 2007	160	ολόκληρο το σωματίο συναρμογής
Makarova <i>et al</i> 2004	93	B
Jurica <i>et al</i> 2002	71	C
Rappsilber <i>et al</i> 2002	272	ολόκληρο το σωματίο συναρμογής
Zhou <i>et al</i> 2002	137	ολόκληρο το σωματίο συναρμογής
Makarov <i>et al</i> 2002	83	ολόκληρο το σωματίο συναρμογής
Hartmuth <i>et al</i> 2002	70	ολόκληρο το σωματίο συναρμογής

Οι πρωτεΐνες αυτές, που απομονώθηκαν είτε από ολόκληρο το σωματίο συναρμογής είτε από τα επιμέρους σύμπλοκα, πρέπει να αποτελούν ένα σύνολο μη επικαλυπτόμενων δεδομένων διαφορετικών πρωτεϊνών το οποίο στη συνέχεια θα χρησιμοποιηθεί για την κατασκευή του δικτύου των πρωτεϊνικών αλληλεπιδράσεων του σωματίου συναρμογής. Η συλλογή αυτών των δεδομένων δεν γίνεται άμεσα διότι στις διαφορετικές ερευνητικές εργασίες καταγράφονται διαφορετικές ονομασίες για

την ίδια πρωτεΐνη, για διαφορετικές ισομορφές μιας πρωτεΐνης, για εναλλακτικά μετάγραφα που προέρχονται από το ίδιο γονίδιο ή για το ίδιο γονίδιο. Για αυτό το λόγο πραγματοποιήθηκε *in silico* η αναγωγή όλων των κωδικών των πρωτεϊνών ή γονιδίων-μεταγράφων που καταγράφονται ως συστατικά του σωματίου συναρμογής, σε επίπεδο γονιδίου, χρησιμοποιώντας ένα μοναδικό χαρακτηριστικό κωδικό (Gene ID, Genbank, <http://www.ncbi.nlm.nih.gov/gene>). Με αυτό τον τρόπο πραγματοποιήθηκε η σύγκριση των δεδομένων των διαφορετικών εργασιών και εξασφαλίστηκε η μοναδικότητά τους.

Έτσι, δημιουργήθηκε ένα σύνολο 479 Gene IDs (Πιν. 3). Αυτό το σύνολο περιλαμβάνει IDs που αντιστοιχούν σε μικρές πυρηνικές ριβονουκλεοπρωτεΐνες (snRNPs), παράγοντες συναρμογής (μη-snRNPs), πρωτεΐνες Sm, οι οποίες αποτελούν μια κεντρική δομή για όλα τα snRNPs, επιπλέον πρωτεΐνες ειδικές για κάθε snRNP, μια ομάδα πρωτεϊνών παρόμοιων με τις Sm (Sm-like, Lsm) οι οποίες συνδέονται με το U6 snRNA, πρωτεΐνες SR που αποτελούν απαραίτητο συστατικό του σωματίου συναρμογής και κυρίως του συμπλόκου E (πρώιμο σύμπλοκο συναρμογής), τριμερή πρωτεϊνών (tri-snRNP) που αποτελούνται από τις υπομονάδες U5 και U4/U6 snRNP, κυκλοφιλλίνες (cyclophilins) που διευκολύνουν την αναδίπλωση των πρωτεϊνών, ελικάσες, πρωτεΐνες που δε σχετίζονται με τη διαδικασία συναρμογής αλλά διαθέτουν μοτίβα ή περιοχές (domains) σύνδεσης στο RNA, πρωτεΐνες οι οποίες εμπλέκονται και σε άλλα στάδια της γονιδιακής έκφρασης ή σε άλλα στάδια ωρίμανσης του πρόδρομου mRNA (προσθήκη δομής καλύπτρας στο 5' άκρο, πολυαδενυλίωση), πρωτεΐνες (νουκλεοπορίνες) που συμμετέχουν στην έξοδο του ώριμου mRNA από τον πυρήνα και πρωτεΐνες με άγνωστη λειτουργία που έχουν συσχετιστεί με τη διαδικασία της συναρμογής λόγω των αλληλεπιδράσεων με πρωτεΐνες που αποτελούν συστατικά του σωματίου συναρμογής (Πιν. 3).

Πιν. 3: Τα συστατικά του σωματίου συναρμογής (spliceosome) και οι αντίστοιχες εργασίες που αναφέρονται

Gene ID*	Γονίδιο	Bessonov <i>et al 2008</i> [@]	Herold <i>et al 2008</i> [@]	Chen <i>et al 2007</i> [@]	Behzadnia <i>et al 2007</i> [@]	Makarova <i>et al 2004</i> [@]	Makarov <i>et al 2002</i> [@]	Rappsilber <i>et al 2002</i> [@]	Zhou <i>et al 2002</i> [@]	Jurica <i>et al 2002</i> [@]	Hartmuth <i>et al 2002</i> [@]
6626	U1-A										
6627	U2A'										
6629	U2B''										
10291	SF3a120										
8175	SF3a66										
10946	SF3a60										
23451	SF3b155										
10992	SF3b145										
23450	SF3b130										
10262	SF3b49										
51639	SF3b14a										
84844	SF3b14b										
11338	U2AF65										
7307	U2AF35										
1665	hPRP43										
84991	SPF45										
23350	SR140										
10523	CHERP										
11325	SF3b125										
10594	U5 220K										
23020	U5 200K										
9343	U5 116K										
9410	U5 40K										
24148	U5 102K										
10907	U5 15K										
9416	U5 100K										
10421	U5 52K										
9129	U4/U6 90K										
9128	U4/U6 60K										
10465	U4/U6 20K										
26121	U4/U6 61K										
4809	U4/U6 15.5K										
9092	U4/U6.U5 110K										
10713	U4/U6.U5 65K										

84950	PRPF38A										
24144	TFIP11										
57819	LSm2										
27258	LSm3										
25804	LSm4										
11157	LSm6										
51690	LSm7										
51691	LSm8										
6628	Sm protein B										
6632	Sm protein D1										
6633	Sm protein D2										
6634	Sm protein D3										
6635	Sm protein E										
6636	Sm protein F										
6637	Sm protein G										
27339	hPRP19										
988	CDC5L										
10286	SPF27										
5356	PRL1										
3312	CCAP1 (HSPA8)										
51503	CCAP2										
56259	catenin b-like 1										
51729	Npw38BP										
56949	hSYF1										
51340	CRNKL1										
57461	hlsy1										
22938	SKIP										
55696	RBM22										
10450	Cyp-E										
51645	PPIL1										
9716	KIAA0560										
8896	G10										
79753	SNIP1										
84811	MGC13125										
51634	CGI-79										
9584	RBM39										

1655	DDX5										
1994	ELAV-like 1										
10521	DDX17										
4236	MFAP1										
3550	RED										
55234	SMU1										
23759	PPIL2										
8449	DDX16										
8899	hPRP4-Kinase										
9967	THRAP3										
26986	PABP1										
23517	SKIV2L2										
8106	PABPN1										
7737	RNF113A										
10283	NY-CO-10										
57703	KIAA1604										
3315	hsp27										
94104	GCFC										
59286	UBL5										
91603	CCDC16										
151903	CCDC12										
22944	HsKin17										
1659	hPRP22										
8559	hPRP18										
51362	hPRP17										
9785	hPRP16										
10569	hSLU7										
51428	Abstrakt										
23582	GCIP										
60625	DDX35										
55094	Q9BRR8										
58509	c19orf29										
53938	PPIase-like 3b										
23398	PPWD1										
84292	MORG1										
2483	FRG1										

51070	NOSIP										
27238	GPKOW										
163859	C1orf55										
26017	DKFZP586O0120										
10399	RACK1										
8607	TIP-49										
9360	PPIG										
9130	FAM50A										
26240	FAM50B										
51759	C9orf78										
118924	FRA10AC1										
63932	CXorf56										
8220	DGCR14										
81576	CCDC130										
114034	TOE1										
79576	NKAP										
54819	ZCCHC10										
8558	CDK10										
151613	TTC14										
55100	WDR70										
4795	NFKBIL1										
3728	JUP										
9775	eIF4A3										
55110	Magoh										
9939	Y14										
5411	Pinin										
10921	RNPS1										
22985	Acinus										
10284	SAP18										
10189	Aly										
7919	UAP56										
9984	THOC1										
57187	THOC2										
84321	THOC3										
8563	THOC5										
79228	THOC6										

22916	CBP20											
4686	CBP80											
3608	NF45											
124245	ZC3H18											
4904	YB-1											
55421	ELG											
1654	DDX3											
51593	ASR2											
9774	BCLAF1											
8531	DBPA											
10179	RBM7											
3308	HSP70											
9532	BAG2											
5930	RBBP6											
79171	RBM42											
10250	SRm160											
23524	SRm300											
6426	SF2											
6432	9G8											
6428	SRp20											
8683	SRp30c											
10772	SRp38											
6430	SRp40											
10929	SRp46											
6431	SRp55											
6429	SRp75											
6427	SC35											
27429	hTra-2											
29896	hTra-2 alpha											
6434	hTra-2 beta											
3178	hnRNP A1											
220988	hnRNP A3											
3182	hnRNP AB											
3181	hnRNP A2/B1											
3183	hnRNP C											
3184	hnRNP D											

3185	hnRNP F										
27316	hnRNP G										
27288	hnRNP G-T										
3187	hnRNP H1										
3188	hnRNP H2										
3189	hnRNP H3										
3190	hnRNP K										
4670	hnRNP M										
10492	hnRNP Q										
10236	hnRNP R										
3192	hnRNP U										
5093	PCBP1										
5094	PCBP2										
22913	RALY										
9782	matrin 3										
10949	hnRNP A0										
8761	PABPC4										
1660	DDX9										
83714	NRIP2										
9829	DNaJ homolog										
8241	RBM10										
10181	RBM5										
10915	TCERG1										
29105	TAFIIB										
11017	RY1										
153527	ZMAT2										
6631	U1-C										
6625	U1-70K										
58517	RBM25										
10285	SPF30										
22827	PUF60										
55660	FBP11										
83443	SF3b10										
11193	FBP21										
55082	FLJ10154										
140890	SFRS12										

11100	E1B-AP5										
27154	BRPF3										
606634	Dermcidin										
23013	SHARP										
3306	HSPA2										
64783	RBM15										
11269	DDX19B										
80145	THOC7										
8243	SMC1										
79718	TBLR1										
1832	Desmoplakin										
10643	IGF2BP3										
3301	DNAJA1										
25929	GEMIN5										
10016	PDCD6										
6294	SAFB										
79811	SAFB like										
23435	TARDBP										
301	ANX1										
28974	C19Orf53										
90957	DHX57										
2091	FBL										
26354	GNL3										
63906	GPATC3										
80789	INTS5										
9667	SAFB2										
7030	TFE3										
91746	YTHDC1										
85437	ZCCHC19										
55596	ZCCHC8										
1387	CREBBP										
4926	NUMA1										
22872	SEC31L1										
8148	TAF15										
9419	CRIP1										
3310	HSPA6										

135295	SRp35										
23360	FNBP4										
10432	RBM14										
9188	DDX21										
29894	CPSF1										
2079	ERH										
5514	PPP1R10										
84132	USP42										
22826	SPF31										
9879	hPRP5										
10084	Npw38										
5546	PRCC										
80312	FLJ10839										
57794	SF4										
2521	FUS										
3609	NFAR										
7536	SF1										
27336	tat SF1										
728642	CDC2L2										
79002	MGC2803										
7756	ZNF207										
9184	BUB3										
5511	NIPP1										
29115	HCNGP										
5499	PPP1CA										
57805	p30 DBC										
25956	SEC31-like 2										
55109	AGGF1										
1153	CIRP										
6729	SRp54										
10642	IGF2BP1										
8570	KHSRP										
8880	FUBP1										
5037	PEBP1										
9987	hnRNP D-like										
5725	hnRNP I										

3191	hnRNP L										
2130	EWS										
10482	NXF1										
103	ADAR										
11051	CPSF 25 kDa										
10898	CPSF 30 kDa										
11052	CPSF6										
51692	CPSF 73 kDa										
53981	CPSF 100 kDa										
1479	CSTF3										
6741	Lupus La protein										
5393	EXOSC9										
5394	EXOSC10										
22894	RRP44										
54512	RRP41										
22803	Dhm1-like protein										
55280	FLJ10998										
10212	DDX39										
1653	DDX1										
26512	DDX26										
55342	STRBP										
8939	FUBP3										
64062	RBM26										
23543	RBM9										
54502	RBM47										
84271	KIAA1649										
8227	SFRS17A										
125950	RAVER1										
55192	DNAJC17										
51663	ZFR										
11340	EXOSC8										
55339	WDR33										
541565	ENSP00000290008										
23196	C9orf10										
84081	CCDC55										
25949	SYF2										

85395	C21orf70										
10856	CGI-46										
55777	KIAA1461										
79074	C2orf49										
51637	C14orf166										
51493	C22orf28										
55119	PRPF38B										
11083	C20orf158										
26173	KIAA1440										
65123	FLJ21919										
81608	Q8WVN3										
4800	NFYA										
6936	GCF										
2969	P78347										
3054	HCFC1										
6878	TAF6										
2547	XRCC6										
8467	SMARCA5										
5431	POLR2B										
5430	POLR2A										
8334	HIST1H2AC										
8337	HIST2H2AA3										
3146	HMGB1										
3149	HMGB3										
5931	RBBP7										
6188	RPS3										
6189	RPS3A										
6191	RPS4X										
6192	RPS4Y1										
6193	RPS5										
6201	RPS7										
6202	RPS8										
6203	RPS9										
6204	RPS10										
6206	RPS12										
6207	RPS13										

6209	RPS15											
6210	RPS15A											
6217	RPS16											
6218	RPS17											
6222	RPS18											
6223	RPS19											
6230	RPS25											
6235	RPS29											
2197	FAU											
6205	RPS11											
6125	RPL5											
6146	RPL22											
6147	RPL23A											
6160	RPL31											
6728	SRP19											
6730	SRP68											
6731	SRP72											
1915	EEF1A1											
7175	TPR											
8826	IQGAP1											
3838	KPNA2											
8812	CCNK											
8900	CCNA1											
142	PARP1											
10270	AKAP8											
8554	PIAS1											
8575	PRKRA											
4691	NCL											
7431	VIM											
10376	ENSP00000243115											
203068	TUBB											
4863	NPAT											
23091	KIAA0853											
22974	TPX2											
9715	KIAA0773											
6279	S100A8											

6280	S100A9										
6421	PSF										
9295	SFRS11										
5496	PP2Cγ										
4343	fSAP113										
51747	LUC7A										
51755	CrkRS										
9589	WTAP										
25962	fSAP121										
81550	FLJ21007										
10592	SMC2										
221092	hnRPUL2										
85313	PPIL4										
23217	ZFR2										
29890	RBM15B										
11218	GEMIN3										
6651	SON3										
22999	RIMS1										
23567	ZNF346										
25992	SNED1										
10444	Catenin b-like 1										
9703	LOC9703/KIAA0100										
84433	CARD11										
1294	Collagen type VII										
23258	DENND5A										
1938	EEF2										
5317	PKP1										
2997	GYS1										
64718	UNKL										
23404	EXOSC2										
23016	EXOSC7										
118460	EXOSC6										
56915	EXOSC5										
51013	EXOSC1										
29072	SETD2										
10142	AKAP9										

50628	GEMIN4										
1828	DSG1										
4841	p54nrb										
79009	DDX50										
22907	DHX30										
170506	DHX36										
5935	RBM3										
112846	RBM8B										
22828	RBM16										
54439	RBM27										
284695	ZFP326										
104	ADAR2										
10657	SAM68										
26135	SERBP1										
5901	Ran										
2733	GLE1L										
8480	GLE2										
10605	PAIP1										
8021	NUP214										
23225	NUP210										
9688	NUP93										
4927	NUP88										
57122	NUP107										
4928	NUP98										
53371	NUP54										
129401	NUP35										
55746	NUP133										
23165	NUP205										
6749	FACT80										
27044	SND1										
10902	BRD8										
57003	CCDC47										
55210	ATAD3A										
83858	ATAD3B										
9611	NCoR1										
7153	TopIIa										

1842	hECM2										
4627	Myosin										
71	Actin										
4869	NUMATRIN										
4001	Lamin B										
23345	Nuspectrin										
3320	HSP90α										
3324	HSP90β										
7184	HSP108										
7520	ku80										
55614	Tri-snRNP 110K										
8503	KIAA1590										
5836	PYGL										

* Κάθε συστατικό-υπομονάδα συμβολίζεται με το μοναδιαίο κωδικό Gene ID
 © Η γκρι σκίαση υποδηλώνει ότι το αντίστοιχο μόριο έχει ταυτοποιηθεί στη συγκεκριμένη εργασία

Από το σύνολο των διαφορετικών πρωτεϊνών, ποσοστό της τάξεως του 50% έχουν ταυτοποιηθεί ως πρωτεΐνες του σωματίου συναρμογής σε περισσότερες από μια ερευνητικές εργασίες ενώ το 25% του συνόλου έχει ταυτοποιηθεί σε τουλάχιστον πέντε εργασίες, γεγονός που εντάσσει αυτές τις πρωτεΐνες αξιόπιστα σε αυτό το σύμπλοκο. Το υπόλοιπο 50% των πρωτεϊνών έχει ταυτοποιηθεί σε μία μόνο από τις δέκα ερευνητικές εργασίες, γεγονός που μπορεί να υποδεικνύει ότι αυτές οι πρωτεΐνες έχουν χαρακτηριστεί ψευδώς ως συστατικά του σωματίου συναρμογής. Εναλλακτικά, αυτό το γεγονός μπορεί να οφείλεται στη φύση των πρωτεϊνικών αλληλεπιδράσεων οι οποίες μπορεί να είναι ασθενείς ή παροδικές, στις διαφορετικές συνθήκες που εφαρμόζονται (π.χ. απομόνωση του σωματίου συναρμογής *in vivo* και *in vitro*, συγκέντρωση αλάτων) ή στο ότι σε ορισμένες ερευνητικές εργασίες πραγματοποιήθηκε η απομόνωση ολόκληρου του σωματίου συναρμογής ενώ στις άλλες έγινε η απομόνωση των επιμέρους συμπλόκων του που δημιουργήθηκαν κατά το σχηματισμό του. Επιπλέον, σε ορισμένες περιπτώσεις διέφερε το συνθετικό πρόδρομο μόριο mRNA που χρησιμοποιείται για την ανασύσταση του σωματίου συναρμογής και την ακόλουθη απομόνωσή του.

➤ **Ανασύσταση του δικτύου πρωτεϊνικών αλληλεπιδράσεων του σωματίου συναρμογής με αφετηρία την πρωτεΐνη FRA10AC1**

Μία προσέγγιση για την κατανόηση της λειτουργίας της πρωτεΐνης FRA10AC1 στο πλαίσιο του ευρύτερου σωματίου συναρμογής είναι ο προσδιορισμός της τοπολογίας της στο δίκτυο πρωτεϊνικών αλληλεπιδράσεων. Καταρχάς, η δημιουργία αυτού του δικτύου ξεκίνησε από την πρωτεΐνη FRA10AC1 και τις αλληλεπιδράσεις της με τις πρωτεΐνες DGCR14 και SAP145 (SF3b2) που είχαν προσδιοριστεί παλαιότερα με το σύστημα δύο υβριδίων στον σακχαρομύκητα και είχαν επιβεβαιωθεί με δοκιμή συνκατακρήμνισης (Sarafidou *et al.* υπό συγγραφή). Ακολούθως, πραγματοποιήθηκε η συλλογή και η αξιολόγηση δεδομένων από μία σειρά βάσεων δεδομένων σχετικά με τις άμεσες αλληλεπιδράσεις αυτής και των αλληλεπιδρώντων της. Στη συνέχεια, το δίκτυο επεκτάθηκε με την αξιολόγηση περαιτέρω πληροφοριών που αφορούν τα μόρια που ταυτοποιήθηκαν και η διαδικασία αυτή επαναλήφθηκε πολλαπλές φορές, οδηγώντας στην κατασκευή του γράφου των πρωτεϊνικών αλληλεπιδράσεων της FRA10AC1. Το γράφημα αυτό επεκτάθηκε μέχρι εκείνο το σημείο κατά το οποίο καμμία από τις αναφερόμενες αλληλεπιδρούσες πρωτεΐνες δεν περιλαμβανόταν στο σύνολο των δεδομένων (βλ. Πιν. 3) που αποτελούν το βασικό κριτήριο για την κατασκευή αυτού του δικτύου.

Ειδικότερα, για να κατασκευαστεί αυτό το δίκτυο αναζητήθηκαν αποκλειστικά δυαδικές πρωτεϊνικές αλληλεπιδράσεις, οι οποίες ταυτοποιήθηκαν: α) με το σύστημα δύο υβριδίων, με πειράματα μεγάλης κλίμακας (high-throughput), στον άνθρωπο (Stelzl *et al* 2005, Rual *et al* 2005), στη *D. melanogaster* (Giot *et al* 2003) και στον *C. elegans* (Li *et al* 2004, Simonis *et al* 2009). Για τον εντοπισμό αυτών των αλληλεπιδράσεων διερευνήθηκαν *in silico* οι βάσεις δεδομένων UniHI (<http://theoderich.fb3.mdc-berlin.de:8080/unih/>, σύμβολα CCSB-Y2H, MDC-Y2H), Droid (<http://www.droidb.org/IMBrowser.jsp>) και Worm Interactome Database (http://interactome.dfci.harvard.edu/C_elegans/) από τις οποίες εξήχθησαν οι δυαδικές αλληλεπιδράσεις για τον άνθρωπο, τη *D. melanogaster* και τον *C. elegans* αντίστοιχα και β) με στοχευμένα πειράματα μικρής κλίμακας που αφορούν πρωτεΐνες του ανθρώπου. Η εξαγωγή των δεδομένων για αυτές τις αλληλεπιδράσεις έγινε από την HPRD (www.hprd.org) η οποία επιλέχθηκε επειδή οι πληροφορίες που περιλαμβάνει ανανεώνονται συχνά και παράλληλα παρέχει τη δυνατότητα μετάβασης στις αντίστοιχες ερευνητικές εργασίες.

Οι παραπάνω βάσεις δεδομένων ελέχθησαν αρχικά για τις αλληλεπιδράσεις της πρωτεΐνης FRA10AC1. Ξεκινώντας, έγινε ο έλεγχος των αλληλεπιδράσεων για τον άνθρωπο στις βάσεις δεδομένων UniHI και HPRD. Για τη διερεύνηση των αλληλεπιδράσεων στη Droid και στη Worm Interactome Database χρησιμοποιήθηκαν οι κωδικοί των ορθόλογων πρωτεϊνών που εξήχθησαν από την HomoloGene (<http://www.ncbi.nlm.nih.gov/homologene/>). Για τις πρωτεΐνες που αναφέρονται ότι αλληλεπιδρούν προσδιορίστηκε το αντίστοιχο ορθόλογο μόριο στον άνθρωπο (εφόσον υπήρχε) χρησιμοποιώντας είτε τη βάση δεδομένων HomoloGene είτε διαθέσιμες πληροφορίες από την ίδια την Droid. Έτσι δημιουργήθηκε μία λίστα με υποψήφια αλληλεπιδρώντα μόρια με τη FRA10AC1. Για όλα αυτά τα μόρια βρέθηκε ο χαρακτηριστικός μοναδικός κωδικός (Gene ID) από τη βάση δεδομένων Genbank (<http://www.ncbi.nlm.nih.gov/gene>) ώστε να είναι δυνατός στη συνέχεια ο έλεγχος της λίστας των 479 πρωτεϊνών-συστατικών του σωματίου συναρμογής (βλ. Πιν. 3). Για όσα μόρια ταυτοποιήθηκαν ως συστατικά του σωματίου συναρμογής, θεωρήθηκε αξιόπιστη η αλληλεπίδραση οπότε και προσθέτονταν στο γράφο των αλληλεπιδράσεων. Με αυτό τον τρόπο η μόνη αλληλεπίδραση που ταυτοποιήθηκε για την πρωτεΐνη FRA10AC1 (CG16890) ήταν στη *D. melanogaster* με την πρωτεΐνη DGCR14 (CG1474). Η αλληλεπίδραση αυτή έχει ήδη ταυτοποιηθεί και πειραματικά στον άνθρωπο (Sarafidou *et al.* υπό συγγραφή), γεγονός που υποδεικνύει ότι είναι εξελικτικά συντηρημένη και πιθανώς λειτουργικά σημαντική. Συνεπώς, για την FRA10AC1 δεν προέκυψε κάποια επιπλέον πληροφορία εκτός από την επιβεβαίωση της αλληλεπίδρασής της με την DGCR14 στη *D. melanogaster*. Έτσι, η αρχή του

γράφου περιλαμβάνει τις άμεσες αλληλεπιδράσεις της FRA10AC1 με την DGCR14 και την SAP145 που είχαν ήδη προσδιοριστεί πειραματικά (Εικ. 7)

Στο δίκτυο αυτό απεικονίζονται μόνο πρωτεΐνες του ανθρώπου παρόλο που οι πληροφορίες προέρχονται και από άλλους οργανισμούς, υποθέτωντας ότι οι πρωτεϊνικές αλληλεπιδράσεις είναι πιθανώς εξελικτικά συντηρημένες μεταξύ του ανθρώπου, της *D. melanogaster* και του *C. elegans*.



Εικ. 7: Απεικόνιση των αλληλεπιδράσεων της πρωτεΐνης FRA10AC1. Στην παρένθεση αναγράφεται ο χαρακτηριστικός μοναδικός κωδικός του γονιδίου (Gene ID). Η κόκκινη γραμμή υποδηλώνει ότι η αντίστοιχη αλληλεπίδραση ανιχνεύτηκε σε πρωτεΐνες της *D. melanogaster* και η διακεκομμένη γραμμή ότι η αλληλεπίδραση ταυτοποιήθηκε πειραματικά.

Ακολούθως, χρησιμοποιήθηκε η βάση δεδομένων UniHI για τον εντοπισμό δυαδικών αλληλεπιδράσεων της DGCR14 χωρίς όμως να ταυτοποιηθεί κάποια πρωτεΐνη. Επίσης, χρησιμοποιήθηκε η HPRD στην οποία αναφέρεται η πρωτεΐνη LYST (Gene ID=1130) που αλληλεπιδρά με την DGCR14 χωρίς όμως να αποτελεί συστατικό του σωματίου συναρμογής (Πιν. 3). Ο κωδικός του ορθόλογου γονιδίου στη *D. melanogaster* (CG1474) χρησιμοποιήθηκε για τη διερεύνηση της DroiD ώστε να εντοπιστούν τα μόρια με τα οποία αλληλεπιδρά και στη συνέχεια τα αντίστοιχα ορθόλογα μόρια στον άνθρωπο εντοπίστηκαν μέσω της HomoloGene. (Πιν. 4).

Με βάση το Gene ID αυτών των 30 αλληλεπιδρώντων μορίων, έγινε έλεγχος της λίστας με τα συστατικά του σωματίου συναρμογής (Πιν. 3) και έτσι ταυτοποιήθηκαν δύο πρωτεΐνες, η IK cytokine (RED) και η NUP107 με κωδικούς 3550 και 57122 για τα αντίστοιχα γονίδια του ανθρώπου.

Με παρόμοιο τρόπο, χρησιμοποιήθηκε ο κωδικός της ορθόλογης πρωτεΐνης στον *C. elegans* (F42H10.7) και εντοπίστηκαν 12 πρωτεΐνες που αναφέρονται ως αλληλεπιδρώντα (Πιν. 5).

Πιν. 4: Αλληλεπιδρώντα μόρια της DGCR14 (CG1474) στη *D. melanogaster* (στοιχεία από τη βάση δεδομένων Droid)

Πρωτεϊνικές αλληλεπιδράσεις της DGCR14 (CG1474) στη <i>D.melanogaster</i>		
Κωδικός γονιδίου στη <i>D. melanogaster</i>	Ορθόλογα γονίδια στον άνθρωπο	Gene ID
CG16890	<i>FRA10AC1</i>	118924
CG9638	<i>TADA2B</i>	93624
CG9113	<i>AP1G1</i>	164
CG4303	<i>SMARCD3</i>	6604
CG10604	<i>BSX</i>	390259
CG7937	<i>TLX2</i>	3196
CG7134	<i>CDC14A</i>	8556
CG10576	<i>PA2G4</i>	5036
CG11164	<i>RNASEH2B</i>	79621
CG11984	<i>KCMF1</i>	56888
CG12010	<i>SPATA5L1</i>	79029
CG12985	<i>DIDO1</i>	11083
CG15071	*	
CG18005	<i>IK cytokine (RED)</i>	3550
CG18012	<i>ALG1</i>	56052
CG2701	<i>ASCC2</i>	84164
CG32406	<i>TNS1</i>	7145
CG6410	<i>SNX16</i>	64089
CG8097	<i>DALRD3</i>	55152
CG9012	<i>CLTC</i>	1213
CG11061	<i>GOLGA2</i>	2801
CG3052	<i>NHLH1</i>	4807
CG3654	<i>JARID2</i>	3720
CG9423	<i>KPNA4</i>	3840
CG18584	*	
CG9762	<i>NDUFB5</i>	4711
CG2072	<i>MAD1L1</i>	8379
CG5140	<i>TRAIP</i>	10293
CG6743	<i>NUP107</i>	57122
CG1422	<i>USO1</i>	8615

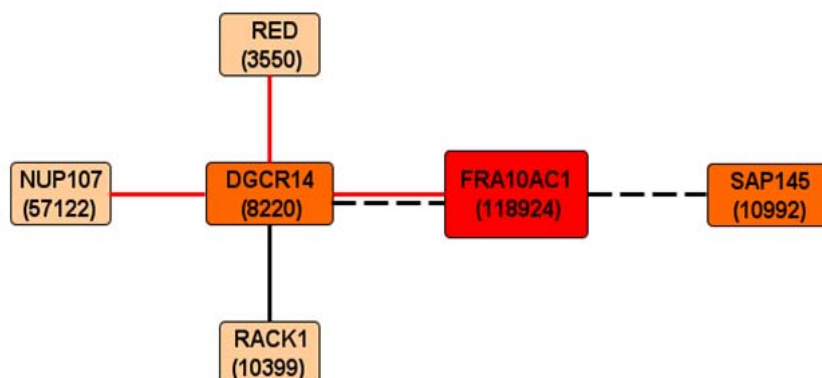
* Δεν ταυτοποιήθηκε ορθόλογο μόριο στον άνθρωπο

Από τον έλεγχο των πρωτεϊνών-συστατικών του σωματίου συναρμογής (Πιν. 3) με βάση το Gene ID, ταυτοποιήθηκε η πρωτεΐνη GNB2L1/RACK1 (K04D7.1 στον *C. elegans*). Οι αλληλεπιδράσεις της DGCR14 που ταυτοποιήθηκαν με τις πρωτεΐνες RACK1/GNB2L1, RED, NUP107 προστέθηκαν στον προκαταρκτικό γράφο (Εικ. 8).

Πιν. 5: Αλληλεπιδρώντα μόρια της DGCR14 (F42H10.7) στον *C. elegans* (στοιχεία από την Worm Interactome Database)

Πρωτεϊνικές αλληλεπιδράσεις της DGCR14 (F42H10.7) στον <i>C. elegans</i>		
Κωδικός γονιδίου στον <i>C. elegans</i>	Ορθόλογα γονίδια στον άνθρωπο	Gene ID
C27A2.6	<i>DVL2</i>	1856
F43C1.2	<i>MAPK1</i>	5594
H20J04.5	<i>PFDN2</i>	5202
K09B11.9	<i>USO1</i>	8615
T04D1.3	<i>SH3GL1</i>	6455
T04H1.2	<i>GTPBP2</i>	54676
W04D2.1	<i>ACTN2</i>	88
W05H7.4	*	
Y57G11C.24	<i>EPS8</i>	2059
ZK1098.4	<i>EIF2B1</i>	1967
K04D7.1	<i>GNB2L1</i>	10399
F26F12.1	<i>COL9A1</i>	1297

* Δεν ταυτοποιήθηκε ορθόλογο μόριο στον άνθρωπο



Εικ. 8: Επέκταση του προκαταρκτικού γράφου με την αντίστοιχη απεικόνιση των πρωτεϊνικών αλληλεπιδράσεων της DGCR14. Όπως και στην Εικ. 7, στις παρενθέσεις αναγράφεται το Gene ID του αντίστοιχου γονιδίου. Επίσης, οι κόκκινες, μαύρες και διακεκομμένες γραμμές συμβολίζουν αντίστοιχα ότι η αλληλεπίδραση ανιχνεύτηκε σε πρωτεΐνες της *D. melanogaster*, του *C. elegans* ή πειραματικά, αντίστοιχα.

Η πρωτεΐνη RACK1/GNB2L1 (~36 kD) αποτελεί υποδοχέα της πρωτεϊνικής κινάσης C. Περιλαμβάνει 7 μοτίβα (WD repeats), 40 αμινοξέων το καθένα, που συμμετέχουν σε πρωτεϊνικές αλληλεπιδράσεις. Το αντίστοιχο γονίδιο χαρτογραφείται στη χρωμοσωματική περιοχή 5q35.3, εκφράζεται σε όλους τους ιστούς και είναι συντηρημένο σε πολλούς, αποκλειστικά ευκαρυωτικούς οργανισμούς, συμπεριλαμβανομένου και του μονοκύτταρου σακχαρομύκητα *S. cerevisiae*. Η RACK1 συμμετέχει σε ποικίλες κυτταρικές διαδικασίες όπως στη ρύθμιση της

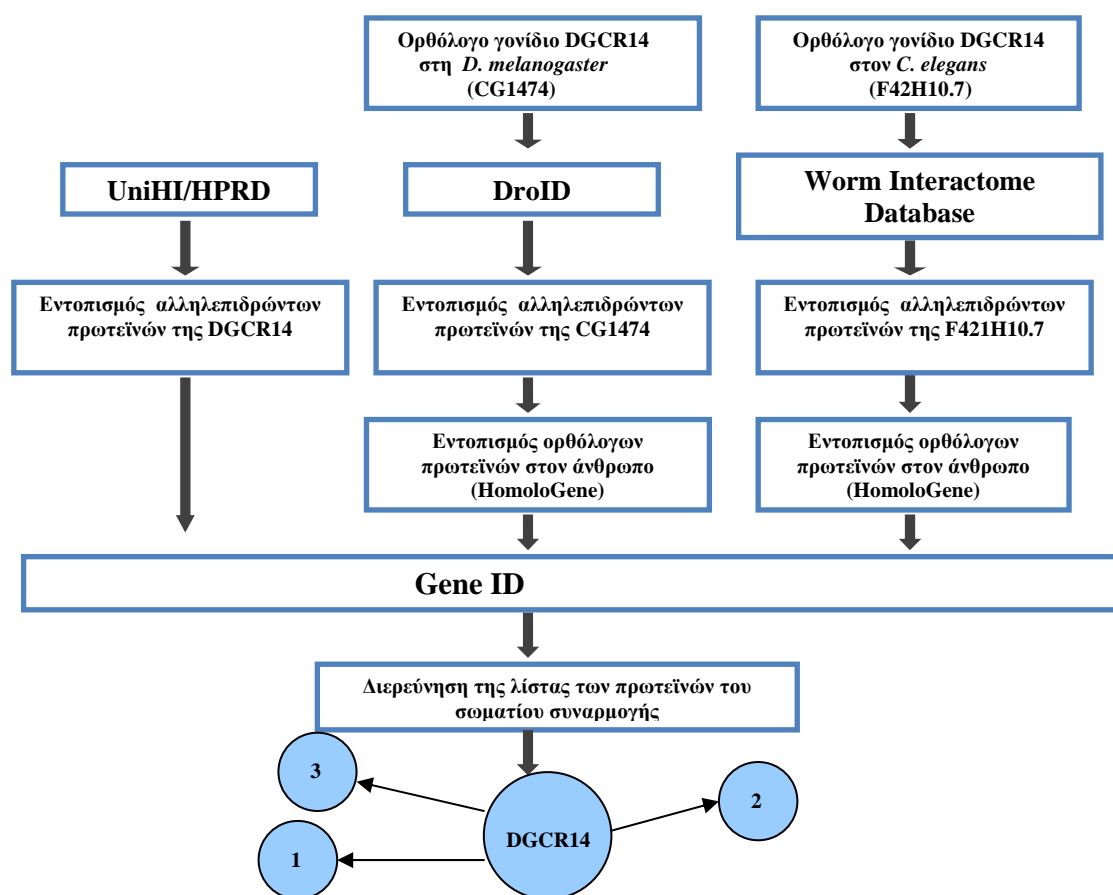
μετάφρασης των πρωτεϊνών, ενεργοποιώντας τη συναρμολόγηση των ριβοσωμάτων, στο κυτταρικό στρες και στην ανάπτυξη των ιστών. Εντοπίζεται τόσο στο κυτταρόπλασμα όσο και στον πυρήνα όπου αλληλεπιδρά με πρωτεΐνες που συμμετέχουν στη ρύθμιση του μηχανισμού συναρμογής (Wang *et al.* 2003, Goncalves *et al.* 2010). Καθώς, λοιπόν, η πρωτεΐνη RACK1 σχετίζεται με πρωτεΐνες που σχετίζονται με τον μηχανισμό συναρμογής και έχει ταυτοποιηθεί ως πρωτεΐνη-συστατικό του σωματίου συναρμογής σε πειράματα μεγάλης κλίμακας (Makarova *et al.* 2002, Makarova *et al.* 2004, Bessonov *et al.* 2008, Herold *et al.* 2008) είναι λογικό να συμπεριλαμβάνεται στο δίκτυο του ευρύτερου συμπλόκου συναρμογής.

Η πρωτεΐνη RED (~70 kDa) αποτελεί μια κυτοκίνη που αναστέλλει την επαγωγή του αντιγόνου HLA τάξης II (HLA Class II) από την ιντερφερόνη γ (IFNγ). Αποτελείται από 557 αμινοξέα και περιλαμβάνει επαναλαμβανόμενα κατάλοιπα αργινίνης, γλουταμικού και ασπαρτικού οξέος (RED repeats) (Krief *et al.* 1994). Το αντίστοιχο γονίδιο χαρτογραφείται στο χρωμόσωμα 5 (5q22-24), εκφράζεται σε πληθώρα ιστών και είναι συντηρημένο σε ευκαρυωτικούς οργανισμούς χωρίς να συμπεριλαμβάνεται ο μονοκύτταρος σακχαρομύκητας *S. cerevisiae*. Η πρωτεΐνη RED εντοπίζεται στον πυρήνα και πιθανώς συμμετέχει στη ρύθμιση της μεταγραφής (Assier *et al.* 1999). Η ορθόλογη πρωτεΐνη στον *C. elegans* (*smu-2*), συμμετέχει στην εναλλακτική συναρμογή διαφοροποιώντας την επιλογή των θέσεων συναρμογής των πρόδρομων mRNA (Spartz *et al.* 2004). Κατά συνέπεια, λαμβάνοντας υπόψη αυτά τα δεδομένα καθώς και ότι η πρωτεΐνη RED απομονώθηκε σε πειράματα μεγάλης κλίμακας ως συστατικό του σωματίου συναρμογής (Bessonov *et al.* 2008, Herold *et al.* 2008, Chen *et al.* 2007, Makarova *et al.* 2004, Makarov *et al.* 2002, Rappsilber *et al.* 2002, Zhou *et al.* 2002) είναι λογικό να σχετίζεται με τον μηχανισμό της συναρμογής και να συμπεριλαμβάνεται σε αυτό το δίκτυο των πρωτεϊνικών αλληλεπιδράσεων.

Η πρωτεΐνη NUP107 είναι μία νουκλεοπορίνη, που, μαζί με τις νουκλεοπορίνες NUP160, NUP133 και NUP96, αποτελεί υπομονάδα ενός ετερο-ολιγομερούς συμπλόκου, το οποίο συμμετέχει στο σχηματισμό των πυρηνικών πόρων (nucleus pore complex) (Boehmer *et al.* 2003). Κατά τη διάρκεια της μεσόφασης, η NUP107 είναι στενά συνδεδεμένη με τη NUP133 στις δύο πλευρές του συμπλόκου των πυρηνικών πόρων και κατά τη μίτωση εντοπίζεται στους κινητοχώρους γεγονός που δηλώνει μια σύνδεση μεταξύ των δομικών συστατικών του συμπλόκου των πυρηνικών πόρων και των κινητοχώρων (Belgareh *et al.* 2001). Η αλληλεπίδραση της DGCR14 με πρωτεΐνες όπως η NUP107 υποδηλώνει ότι ο μηχανισμός της συναρμογής σχετίζεται και με άλλες διαδικασίες όπως η έξοδος του mRNA από τον πυρήνα (Reed 2003).

Συνεπώς, οι αλληλεπιδράσεις της πρωτεΐνης DGCR14 με τις πρωτεΐνες RACK1 και RED οι οποίες σχετίζονται με τη διαδικασία της συναρμογής, είναι λογικό να συμπεριλαμβάνονται στο δίκτυο των πρωτεϊνικών αλληλεπιδράσεων του ευρύτερου συμπλόκου συναρμογής.

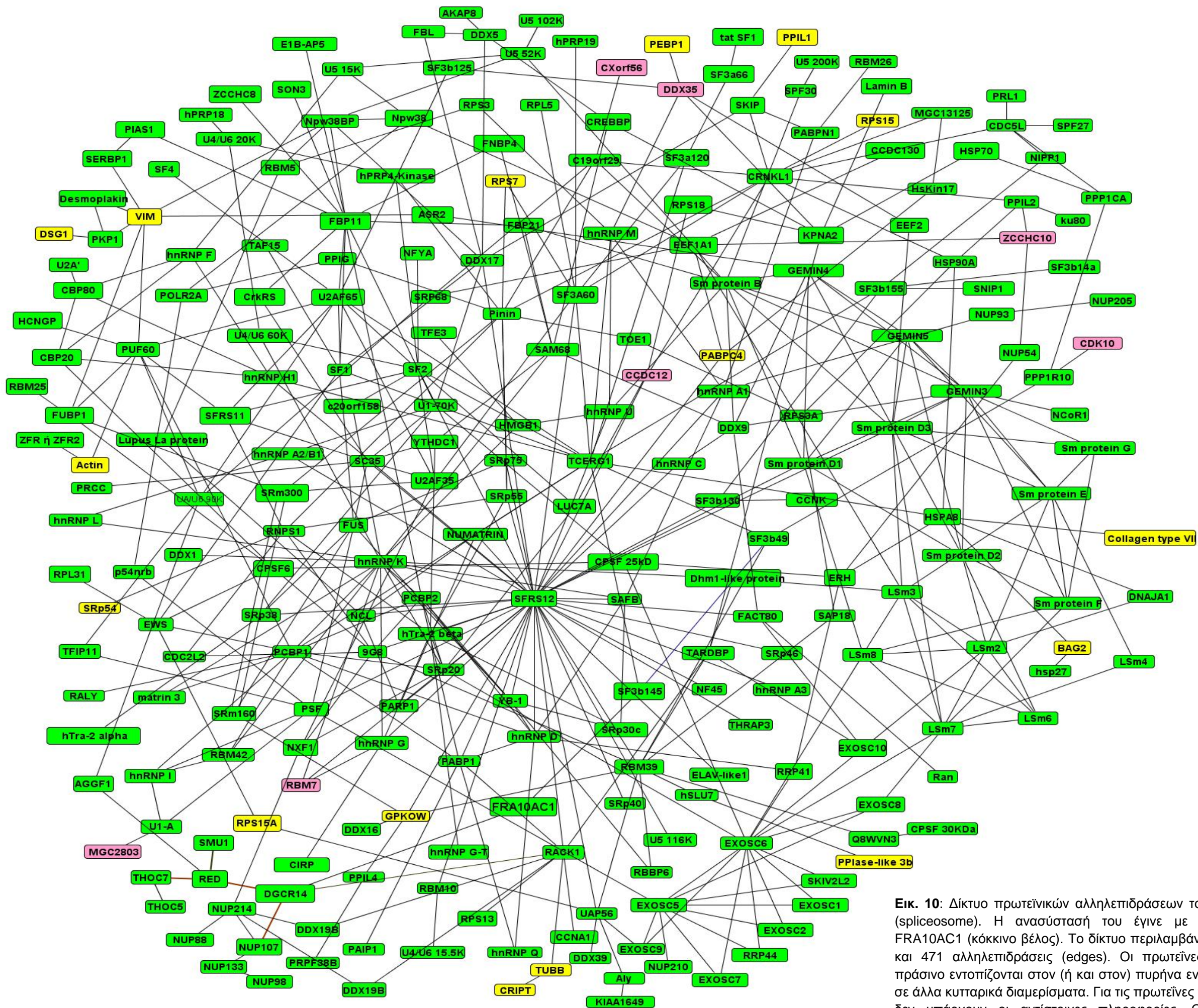
Σχηματικά τα στάδια που ακολουθήθηκαν για την κατασκευή του δικτύου των πρωτεϊνικών αλληλεπιδράσεων, παρουσιάζονται στην Εικ. 9.



Εικ. 9: Σχηματική απεικόνιση της διαδικασίας ανασύστασης του δικτύου πρωτεϊνικών αλληλεπιδράσεων του σωματίου συναρμογής.

Με την ίδια διαδικασία που αναφέρθηκε παραπάνω ταυτοποιήθηκαν και οι υπόλοιπες πρωτεϊνικές αλληλεπιδράσεις του σωματίου συναρμογής (Πιν. 6) και το προκαταρκτικό αυτό δίκτυο επεκτάθηκε με αποτέλεσμα την κατασκευή ενός γράφου (Εικ. 10) που περιλαμβάνει 255 πρωτεΐνες (nodes) και 471 αλληλεπιδράσεις (edges). Έτσι, με αυτήν την προσέγγιση, περισσότερες από τις μισές πρωτεΐνες που είχαν ταυτοποιηθεί ως συστατικά του σωματίου συναρμογής έχουν ενσωματωθεί στο δίκτυο των πρωτεϊνικών αλληλεπιδράσεων. Κατά μέσο όρο, κάθε πρωτεΐνη αλληλεπιδρά περίπου με άλλες δύο αλλά ο αριθμός αυτός κυμαίνεται από μία έως σαράντα αλληλεπιδράσεις. Ιδιαίτερα ενδιαφέρον είναι το γεγονός πως από τις 248

πρωτεΐνες του δικτύου για τις οποίες υπάρχουν πληροφορίες για τον υποκυτταρικό τους εντοπισμό, το 93,5 % των πρωτεϊνών αυτών εντοπίζονται στον (ή και στον) πυρήνα (πράσινα χρωματισμένες πρωτεΐνες του γράφου), δεδομένο συμβατό με την προβλεπόμενη λειτουργία τους στη διαδικασία της συναρμογής. Η υποκυτταρική τοπολογία προσδιορίστηκε είτε βάσει ερευνητικών εργασιών είτε από βιο-υπολογιστικά προγράμματα πρόβλεψης και οι αντίστοιχες πληροφορίες αντλήθηκαν από τη βάση δεδομένων HPRD (Gene Ontology Annotation, <http://www.ebi.ac.uk/GOA/>). Η αξιοπιστία αυτού του δικτύου (Εικ. 10) θα ελεγχθεί μελλοντικά με σύγκριση του προτύπου έκφρασης των αλληλεπιδρώντων πρωτεϊνών από δεδομένα που υπάρχουν κατατεθειμένα σε αντίστοιχες βάσεις δεδομένων, με κατάλληλα εργαλεία βιοπληροφορικής, με μελέτες συν-εξέλιξης των αλληλεπιδρώντων πρωτεϊνών κ.λπ. Ωστόσο, η τελική απόδειξη της αξιοπιστίας του θα είναι ο πειραματικός έλεγχος των προβλεπόμενων αλληλεπιδράσεων.



Εικ. 10: Δίκτυο πρωτεϊνικών αλληλεπιδράσεων του σωματίου συναρμογής (spliceosome). Η ανασύστασή του έγινε με αφετηρία την πρωτεΐνη FRA10AC1 (κόκκινο βέλος). Το δίκτυο περιλαμβάνει 255 πρωτεΐνες (nodes) και 471 αλληλεπιδράσεις (edges). Οι πρωτεΐνες που συμβολίζονται με πράσινο εντοπίζονται στον (ή και στον) πυρήνα ενώ οι κίτρινες εντοπίζονται σε άλλα κυτταρικά διαμερίσματα. Για τις πρωτεΐνες που συμβολίζονται με ροζ δεν υπάρχουν οι αντίστοιχες πληροφορίες. Οι αλληλεπιδράσεις που απεικονίζονται στο δίκτυο, αναγράφονται αναλυτικά στον Πιν. 6.

Πιν. 6: Οι πρωτεϊνικές αλληλεπιδράσεις από τις βάσεις δεδομένων UniHI, HPRD, DroiD και C. elegans interactome που απεικονίζονται στο γράφο (Εικ. 10).

Gene ID	Γονίδιο	UniHI database	HPRD database	ορθόλογο στη <i>D.melanogaster</i>	<i>DroiD database</i>	ορθόλογο στον <i>C.elegans</i>	<i>C. elegans interactome database</i>
626	U1-A	hnRNP I (5725) MGC2803 (79002)	MGC2803 (79002) hnRNP I (5725) PSF (6421)	CG4528	X	K08D10.4	X
10291	SF3a120	X	TCERG1 (10915) SF3A3 (10946) SF3a66 (8175)	CG16941	CG2925 (SF3A3, 10946)	ZK930.3	F11A10.2 (SF3a66,8175) T13H5.4 (SF3A3, 10946)
8175	SF3a66	X	SF3A1 (10291) tat SF1 (27336)	CG10754	X	F11A10.2	W07E6.4 (SF3A1, 10291)
10946	SF3a60	X	SF3a120 (10291) TCERG1 (10915) U2AF35 (7307) SFRS12 (140890)	CG2925	CG16941 (SF3a120, 10291) CG3193 (CRNKL1, 51340) CG5519 (hPRP19, 27339) CG17489 (RPL5, 6125) CG1883 (RPS7, 6201)	T13H5.4	W07E6.4 (SF3a120, 10291)
23451	SF3b155	X	NIPP1 (5511) SNIP1 (79753) SF3b14a (51639) TCERG1 (10915) SF3b130 (23450)	CG2807	X	T08A11.2	X
10992	SF3b145	SF3b49 (10262)	TCERG1(10915) SF3b49 (10262)	CG3605	CG3780 (SF3b49, 10262)	W03F9.10	X
23450	SF3b130	X	SF3b155 (23451) SFRS12 (140890)	CG13900	CG15218 (CCNK,8812)	K02F2.3	X
10262	SF3b49	X	TCERG1(10915)	CG3780	CG11266 (RBM39, 9584) CG33526 (PPP1R10, 5514)	Δεν υπάρχει ορθόλογο	
11338	U2AF65	SF2 (6426) SF1 (7536) PUF60 (22827)	PUF60 (22827) SF2 (6426) SF1 (7536) U2AF35 (7307) SFRS11 (9295) TCERG1 (10915) FBP11 (55660)	CG9998	CG6418 (SF3b125,11325) CG31550 (SF4, 57794) CG3582 (U2AF35,7307) CG7757 (U4/U6 90K, 9129)	Y92C3B.2	Y116A8C.35 (U2AF35,7307)
7307	U2AF35	U2AF65 (11338)	U2AF65 (11338) SF2 (6426) SC35 (6427) NXF1 (10482) SF3A3 (10946)	CG3582	CG7971 (SRw300, 23524) CG9998 (U2AF65,11338)	Y116A8C35	Y92C3B.2 (U2AF65, 11338)
11325	SF3b125	X	X	CG6418	CG3225 (DDX35, 60625) CG9998 (U2AF65, 11338)	C46F11.4	X
9343	U5 116K	X	SFRS12 (140890)	CG4849	X	ZK328.2	X
24148	U5 102K	X	U5 52K (10421)	CG6841	X	Y59A8B.6	X
10907	U5 15K	X	hnRNP F (3185)	CG3058	X	δεν υπάρχει ορθόλογο	

			Npw38 (10084) U5 52K (10421)				
9121	U5 52K	X	U5 102K (24148) Sm protein B (6628) U5 15K (10907) Npw38BP (51729)	CG5198	X	Y47D3A.27	X
9129	U4/U6 90K	X	X	CG7757	CG9998 (U2AF65,11338) CG5064 (SRP68,6730) CG12085 (PUF60, 22827) CG7238 (TFIP11,24144)	M03C11.7	X
9128	U4/U6 60K	X	SF2 (6426) U4/U6 20K (10465)	CG6322	X	C36B1.5	F26B1.2 (hnRNPK, 3190)
91465	U4/U6 20K	X	hPRP18 (8559) U4/U6 60K (9128)	CG17266	hPRP18 (8559)	T01B7.4	X
914144	TFIP11	X	PSF (6421) FUS (2521) U1-A (6626) p54nrb (4841) hnRNP I (5725) SFRS12 (140890) EXOSC5 (56915)	CG7238	CG7757 (U4/U6 90K, 9129) CG4211 (PSF, 6421)	C07E3.1	X
917819	LSm2	LSm3 (27258)	DNAJA1 (3301) Sm protein D2 (6633) Sm protein E (6635) LSm3 (27258) LSm6 (11157) LSm7 (51690) LSm8 (51691) EXOSC10 (5394)	CG10418	CG2021 (LSm8, 51691)	T10G3.6	X
27258	LSm3	LSm2 (57819) EEF1A1 (1915)	LSm2 (57819) EEF1A1 (1915) Sm protein D3 (6634) Sm protein E (6635) LSm6 (11157) LSm7 (51690) LSm8 (51691) Dhm1-like protein (22803)	CG31184	X	Y62E10A.12	X
25804	LSm4	X	Sm protein E (6635)	CG17768	X	F32A5.7	ZK593.7 (LSm7, 51690)
11157	LSm6	X	LSm2 (57819) LSm3 (27258) Sm protein D2 (6633) LSm7 (51690) LSm8 (51691)	CG9344	X	Y71G12B.14	X

51690	LSm7	X	LSm2 (57819) LSm3 (27258) LSm6 (11157) Sm protein D3 (6634) LSm8 (51691) EXOSC8 (11340) EXOSC6 (118460)	CG13277	CG8427 (SmD3, 6634) CG18591 (SmE, 6635) CG16792 (SmF, 6636)	ZK593.7	F32A5.7 (LSm4, 25804)
51691	LSm8	X	LSm2 (57819) LSm3 (27258) LSm6 (11157) LSm7 (51690)	CG2021	CG10418 (LSm2, 57819) CG16792 (Sm protein F, 6636)	Y73B6BL.32	X
6628	Sm protein B	X	GEMIN4 (50628) GEMIN3 (11218) U5 52K (10421) SPF30 (10285) Sm protein D3 (6634) GEMIN5 (25929) DDX9 (1660) Sm protein D1 (6632) FBP21 (11193)	CG5352	CG1676 (c19orf29, 58509)	δεν υπάρχει ορθόλογο	
6632	Sm protein D1	X	Sm protein B (6628) GEMIN3 (11218) GEMIN5 (25929) LSm8 (51691) GEMIN4 (50628) SFRS12 (140890) Sm protein D2 (6633)	CG10753	X	T28D9.10	X
6633	Sm protein D2	X	Sm protein D1 (6632) Sm protein F (6636) GEMIN3 (11218) GEMIN5 (25929) LSm2 (57819) LSm6 (11157) GEMIN4 (50628)	CG1249	CG16792 (Sm protein F, 6636)	C52E4.3	ZK652.1 (Sm protein F, 6636)
6634	Sm protein D3	X	Sm protein E (6635) Sm protein G (6637) Sm protein B (6628) GEMIN3 (11218) GEMIN5 (25929) LSm3 (27258) LSm7 (51690)	CG8427	CG13277 (LSm7, 51690) CG2238 (EEF2, 1938)	Y116A8C.42	X

6635	Sm protein E	X	SFRS12 (140890) Sm protein F (6636) Sm protein G (6637) GEMIN3 (11218) GEMIN5 (25929) LSm2 (57819) LSm3 (27258) LSm4 (25804) Sm protein D3 (6634)	δεν υπάρχει ορθόλογο		Y49E10.15	ZK652.1 (Sm protein F, 6636)
6636	Sm protein F	X	Sm protein D2 (6633) Sm protein E (6635) Sm protein G (6637) GEMIN3 (11218)	CG16792	CG2021 (LSm8, 51691) CG13277 (LSm7, 51690) CG1249 (Sm protein D2, 6633) CG18591 (Sm protein E, 6635)	ZK652.1	C52E4.3 (Sm protein D2,6633) Y49E10.15 (Sm protein E, 6635)
6637	Sm protein G	X	Sm protein E (6635) GEMIN3 (11218) Sm protein D3 (6634) Sm protein F (6636)	CG9742	X	Y71F9B.4	X
27339	hPRP19	X	X	CG5519	CG2925 (SF3A3, 10946)	T10F2.4	X
988	CDC5L	X	NIPP1 (5511) PPP1CA (5499) PRL1 (5356)	CG6905	CG4980 (SPF27, 10286) CG3193 (CRNKL1,51340)	D1081.8	X
10286	SPF27	X	X	CG4980	X	T12A2.7	X
5356	PRL1	X	CDC5L (988)	CG1796	X	D1054.15	X
3312	HSPA8	Collagen type VII (1294) ERH (2079)	BAG2 (9532) ERH (2079) DNAJA1 (3301) HSP90A (3320) TCERG1 (10915) Collagen type VII (1294)	CG8937	X	F26D10.3	X
51729	Npw38BP	X	VIM (7431) FBP11 (55660) Npw38 (10084) U5 52K (10421) DDX17 (10521)	CG2685	CG4887 (RBM5, 10181)	δεν υπάρχει ορθόλογο	
51340	CRNKL1	X	HSP90α (3320) PPIL2 (23759)	CG3193	CG13625 (BUD13, 84811) CG15525 (CCDC12, 151903) CG6905 (CDC5L, 988) CG1676 (c19orf29, 58509) CG8332 (RPS15, 6209) CG2925 (SF3A3, 10946)	M03F8.3	X

					CG8264 (SKIP, 22938) CG8383 (Pinin, 5411)		
22938	SKIP	X	PPIL1 (51645) PABPN1 (8106)	CG8264	CG10279 (DDX17, 10521) CG3193 (CRNKL1, 51340)	T27F2.1	C34D4.12 (PPIL1, 51645)
51645	PPIL1	X	SKIP (22938)	CG13892	X	C34D4.12	X
84811	BUD13	X	X	CG13625	CG5649 (HsKin17, 22944)	R08D7.1	X
9584	RBM39	X	UAP56 (7919)	CG11266	CG6046 (SAP18, 10284) CG3231 (RBBP6,5930) CG1078 (Q8WVN3, 81608) CG10327 (TARDBP, 23435) CG1420 (SLU7, 10569)	Y55F3AM.3	C33H5.12 (9G8, 6432)
1655	DDX5	X	FBL (2091) CREBBP (1387) AKAP8 (10270) DDX17 (10521)	CG10077	X	F58E10.3	X
1994	ELAV-like 1	X	SFRS12 (140890)	δεν υπάρχει ορθόλογο		F35H8.5	X
40521	DDX17	SF1 (7536) Npw38BP (51729)	SFRS12 (140890) SF1 (7536) DDX5 (1655) FBL (2091) Npw38BP (51729) U1-70K (6625)	CG10279	X	δεν υπάρχει ορθόλογο	
3550	RED	X	X	CG18005	CG1474 (DGCR14,8220) CG17143 (THOC7, 80145) CG8079 (AGGF1, 55109)	Y49F6B.4	CC4.3 (SMU1,55234)
55234	SMU1	X	X	CG5451	X	CC4.3	Y49F6B.4 (RED,3550)
23759	PPIL2	X	CRNKL1 (51340) HSP90A (3320)	CG7747	CG4622 (ZCCHC8, 55596) CG18801 (ku80, 7520)	F59E10.2	X
8449	DDX16	X	X	CG10689	X	C04H5.6	X
8899	hPRP4-Kinase	X	SF2 (6426) U4/U6 20K (10465)	CG7028	X	F22D6.5	X
9967	THRAP3	X	SFRS12 (140890)	δεν υπάρχει ορθόλογο		δεν υπάρχει ορθόλογο	
26986	PABP1	X	hnRNP D (3184) PCBP1 (5093) PCBP2 (5094) RACK1 (10399) PAIP1 (10605) hnRNP Q (10492) SFRS12 (140890)	δεν υπάρχει ορθόλογο		δεν υπάρχει ορθόλογο	
8106	PABPN1	X	SKIP (22938) hnRNP A1 (3178)	CG2163	CG10084 (RBM26, 64062)	C17E4.5	B0336.3 (RBM26, 64062)
151903	CCDC12	X	X	CG15525	CG3198 (LUC7A, 51747) CG8900 (RPS18,6222)	Δεν υπάρχει ορθόλογο	

					CG3193 (CRNKL1,51340) CG4602 (SFRS12, 140890)		
22944	HsKin17	X	X	CG5649	CG5119 (PABPC4, 8761)	Y52B11A.9	X
10569	hSLU7	PPlase-like 3b (53938)	PPlase-like 3b (53938)	CG1420	X	K07C5.6	X
60625	DDX35	X	X	CG3225	CG17919 (PEBP1, 5037) CG2238 (EEF2, 1938) CG6418 (SF3b125, 11325) CG4799 (KPNA2,3838)	Y67D2.6	X
58509	c19orf29	X	X	CG1676	CG16865 (cXorf56, 63932) CG8383 (Pinin, 5411) CG3193 (CRNKL1,51340) CG5064 (SRP68,6730)	W03H9.4	X
33938	PPlase-like 3b	SLU7 (10569)	SLU7 (10569)	CG11777	X	B0252.4	X
27238	GPKOW	X	X	CG10324	CG11266 (RBM39, 9584) CG5808 (PPIL4, 85313) CG10689 (DDX16, 8449) CG11274 (SRm160, 10250) CG4602 (SFRS12, 140890)	R11A8.2	X
10399	RACK1	DDX19B (11269)	DDX19B (11269) SFRS12 (140890) PABPC1 (26986) CCNA1 (8900) TUBB (203068)	CG7111	CG10327 (TARDBP, 23435)	K04D7.1	X
9360	PPIG	X	Pinin (5411) POLR2A (5430)	Δεν υπάρχει ορθόλογο		D1009.2	X
118924	FRA10AC1	X	X	CG16890	CG1474 (DGCR14,8220)		
63932	CXorf56	X	X	CG16865	CG1676 (c10orf29, 58509)	Y66D12A.8	X
8220	DGCR14	X	X	CG1474	CG18005 (RED,3550) CG6743 (NUP107, 57122)	F42H10.7	K04D7.1 (RACK1, 10399)
81576	CCDC130	X	EEF1A1 (1915)	CG15084	X	ZK1307.9	X
114034	TOE1	RPS3A (6189)	RPS3A (6189)	Δεν υπάρχει ορθόλογο		Y57A10A.25	X
54819	ZCCHC10	EEF1A1 (1915)	EEF1A1 (1915)	CG5516	CG8831 (NUP54, 53371)	Δεν υπάρχει ορθόλογο	
8558	CDK10	X	X	CG1362	X	Δεν υπάρχει ορθόλογο	
5411	Pinin	TOE1 (114034)	SRm300 (23524) SRp75 (6429) RNPS1 (10921) PPIC (9360) hPRP4-kinase (8899)	CG8383	CG1676 (c19orf29, 58509) CG10851 (SRp55, 6431) CG1883 (RPS7, 6201)	R186.7	X
10921	RNPS1	X	CDC2L2 (728642) SRp55 (6431) Pinin (5411) SRp54 (6729)	CG16788	X	K02F3.11	C33H5.12 (9G8, 6432)

			SFRS11 (9295) hTra-2 beta (6434)				
20284	SAP18	RPS3A (6189)	RPS3A (6189)	CG6046	X	Δεν υπάρχει ορθόλογο	
20189	Aly	X	UAP56 (7919)	CG1101	CG18259 (KIAA1649, 84271)	Δεν υπάρχει ορθόλογο	
2919	UAP56	X	RPS15A (6210) EXOSC9 (5393) Aly (10189) RBM39 (9584) DDX39 (10212) SFRS12 (140890)	CG7269	X	Δεν υπάρχει ορθόλογο	
2563	THOC5	X	THOC7 (80145)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
22916	CBP20	X	CBP80 (4686) hnRNP H1 (3187) hnRNP F (3185)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
4686	CBP80	X	hnRNP H1 (3187) hnRNP F (3185) U2A' (6627) CBP20 (22916)	CG7035	X	F37E3.1	X
2608	NF45	X	SFRS12 (140890)	CG5641	X	R11H6.5	X
2904	YB-1	X	FUS (2521) SRp30c (8683) PCBP2 (5094) SFRS12 (140890) hnRNP K (3190) hnRNP D (3184) RBBP6 (5930)	CG5654	X	Δεν υπάρχει ορθόλογο	
1654	DDX3	X	SFRS12 (140890)	CG9748	X		
51593	ASR2	X	VIM (7431) GEMIN4 (50628)	CG7843	X	E01A2.2	X
10179	RBM7	hnRNP K (3190) hnRNP G (27316)	SRp20 (6428) hnRNP G (27316) hnRNP K (3190)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
3308	HSP70	X	KPNA2 (3838) PPP1CA (5499)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
9532	BAG2	X	HSPA8 (3312)	CG7945	CG4264 (HSPA8, 3312)	H14N18.1	C09B8.6 (hsp27, 3315)
5930	RBBP6	X	YB-1 (4904)	CG3231	X	F36F2.3	X
10250	SRm160	X	SRm300 (23524) FUS (2521)	CG11274	X	F28D9.1	X
23524	SRm300	X	Pinin (5411) SRm160 (10250)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
6426	SF2	U2AF65 (11338)	U2AF65 (11338)	CG6987	X	Y111B2A.18	X

		c20orf158 (11083)	c20orf158 (11083)				
			NFYA (4800)				
			NXF1 (10482)				
			U1-70K (6625)				
			U2AF35 (7307)				
			SFRS11 (9295)				
			TCERG1 (10915)				
			U4/U6 60K (9128)				
			CrKRS (51755)				
			YTHDC1 (91746)				
			SFRS12 (140890)				
			SFRS17A (8227)				
			LUC7A (51747)				
			hPRP4-kinase (8899)				
6432	9G8	X	SFRS12 (140890)	CG10203	X	C33H5.12	X
			CDC2L2 (728642)				
			CPSF6 (11052)				
			NXF1 (10482)				
			PCBP2 (5094)				
6428	SRp20	X	hnRNP K (3190)	CG17136	CG10128 (h Tra-2 beta, 6434)	Δεν υπάρχει ορθόλογο	
			SFRS12 (140890)				
			PCBP2 (5094)				
			PCBP1 (5093)				
			NXF1 (10482)				
			RBM7 (10179)				
			EXOSC6 (118460)				
8683	SRp30c	X	hTra-2 beta (6434)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
			SAFB (6294)				
			YB-1 (4904)				
10772	SRp38	X	FUS (2521)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
			U1-70K (6625)				
			hTra-2 alpha (29896)				
			SFRS12 (140890)				
6430	SRp40	X	SFRS12 (140890)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
10929	SRp46	X	SFRS12 (140890)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
6431	SRp55	SRp75 (6429)	SRp75 (6429)	CG10851	CG10128 (h Tra-2 beta, 6434)	W02B12.3	X
		RNPS1 (10921)	RNPS1 (10921)				
			SFRS12 (140890)				
6429	SRp75	X	FUS (2521)	Δεν υπάρχει ορθόλογο		Y106G6D.3 και W09C3.2	X
			SFRS12 (140890)				

			Pinin (5411) SRp55 (6431)				
427	SC35	X	FUS (2521)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
			U1-70K (6625) U2AF35 (7307) SFRS11 (9295) CrKRS (51755) YTHDC1 (91746) SFRS12 (140890) PRCC (5546)				
6434	hTra-2 beta	x	SFRS12 (140890) hnRNP G (27316) hnRNP G-T (27288) RNPS1 (10921) SRp30c (8683) YTHDC1 (91746)	CG10128	CG17136 (SRp20, 6428)	C18D11.4	X
178	hnRNP A1	X	PABPN1 (8106) SAFB (6294) SFRS12 (140890)	CG9983	CG11092 (NUP93, 9688)	Δεν υπάρχει ορθόλογο	
20988	hnRNP A3	X	SFRS12 (140890)	Δεν υπάρχει ορθόλογο		F42A6.7	X
181	hnRNP A2/B1	X	U1-70K (6625) hnRNP L (3191) SFRS12 (140890)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
183	hnRNP C	X	SFRS12 (140890) DDX9 (1660)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
3184	hnRNP D	X	PCBP1 (5093) PCBP2 (5094) YB-1 (4904) SAFB (6294) SFRS12 (140890) PABP1 (26986) hnRNP Q (10492) RRP41 (54512)	Δεν υπάρχει ορθόλογο		Y73B6BL.6	X
27316	hnRNP G	hnRNP K (3190) CIRP (1153) RBM7 (10179)	CIRP (1153) RBM7 (10179) hTra-2 beta (6434) SFRS12 (140890) hnRNP K (3190)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
27288	hnRNP G-T	X	hTra-2 beta (6434)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
3187	hnRNP H1	X	CBP80 (4686)	Δεν υπάρχει		Y73B6BL.33	X

			hnRNP F (3185) CBP20 (22916) TCERG1 (10915) SFRS12 (140890)	ορθόλογο			
190	hnRNP K	hnRNP G (27316)	hnRNP G (27316)	Δεν υπάρχει ορθόλογο		F26B1.2	C36B1.5 (U4/U6 60 K, 9128)
		matrin 3 (9782)	matrin 3 (9782)				Y48B6A.3 (Dhm1-like protein, 22803)
		RBM7 (10179)	RBM7 (10179)				
		RBM42 (79171)	RBM42 (79171)				
		c20orf158 (11083)	c20orf158 (11083)				
			HMGB1 (3146)				
			SAFB (6294)				
			SRp20 (6428)				
			PCBP2 (5094)				
			SFRS12 (140890)				
			YB-1 (4904)				
			DDX1 (1653)				
			SAM68 (10657)				
			hnRNP L (3191)				
			PCBP1 (5093)				
670	hnRNP M	X	TCERG1 (10915)	CG9373	CG30122 (hnRNP U, 3192) CG9983 (hnRNP A1, 3178) CG32225 (DDX35, 60625) CG6779 (RPS3, 6188)	Δεν υπάρχει ορθόλογο	
192	hnRNP U	X	SFRS12 (140890)	CG30122	CG12223 (HMGB1, 3146) CG5352 (Sm protein B, 6628)	Y41E3.11	X
5093	PCBP1	X	PUF60 (22827)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
			RBM42 (79171)				
			matrin 3 (9782)				
			PCBP2 (5094)				
			hnRNP I (5725)				
			RALY (22913)				
			SRp20 (6428)				
			EWS (2130)				
			hnRNP K (3190)				
			hnRNP D (3184)				
			TCERG1 (10915)				
5094	PCBP2	X		Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
9782	matrin 3	hnRNP K (3190)	PCBP1 (5093)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
			hnRNP K (3190)				
8761	PABPC4	X	EEF1A1 (1915) SFRS12 (140890)	CG5119	X	Y106G6H.2	X

6660	DDX9	X		CG11680	X	T07D4.3	X
8241	RBM10	U4/U6 15.5K (4809)	U4/U6 15.5K (4809)	Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
			SFRS12 (140890)				
0181	RBM5	X	X	CG4896	CG10922 (Lupus La protein, 6741)	T08B2.5	X
					CG6779 (RPS3, 6188)		
10915	TCERG1	X	hnRNP M (4670)	Δεν υπάρχει ορθόλογο		ZK1127.9 και ZK1127.6	X
			SF3A3 (10946)				
			SF3B2 (10992)				
6625	U1-70K	X		CG8749	X	K04G7.10	X
8517	RBM25	X	X	CG4119	X	W04D2.6	X
0285	SPF30	X	Sm protein B (6628)	CG17454	X	Y71D11A.2	X
			U5 200K (23020)				
2827	PUF60	U2AF65 (11338)	VIM (7431)	CG12085	CG4027 (Actin, 71)	Y47G6A.20	X
		HCNGP (29115)	U2AF65 (11338)		CG7757 (U4/U6 90K,9129)		
		VIM (7431)	SFRS11 (9295)				
			FUBP1 (8880)				
			PCBP1 (5093)				
			HCNGP (29115)				
5660	FBP11	E1B-AP5 (11100)	U1-70K (6625)	CG3542	X	ZK1098.1	X
			SF1 (7536)				
			U2AF65 (11338)				
			FBP21 (11193)				
			PIAS1 (8554)				
			FNBP4 (23360)				
			SON3 (6651)				
			ZCCHC8 (55596)				
			CrkRS (51755)				
			PCBP2 (5094)				
			Npw38BP (51729)				
			E1B-AP5 (11100)				
11193	FBP21	X		CG4291	X	Δεν υπάρχει ορθόλογο	
140890	SFRS12	X	PABPC4 (8761)	CG4602	CG10324 (GPKOW, 27238)	D2089.1	X
			SRp40 (6430)		CG11266 (RBM39, 9584)		
			SF2 (6426)		CG15525 (CCDC12, 151903)		
			NF45 (3608)				
			NCL (4691)				
			UAP56 (7919)				
			SC35 (6427)				
			FACT80 (6749)				
			hnRNP H1 (3187)				
			Sm protein D1 (6632)				
			Sm protein D3				

			(6634)				
			SF3b130 (23450)				
			hnRNP A1 (3178)				
			YB-1 (4904)				
			hnRNP D (3184)				
			THRAP3 (9967)				
			U5 116K (9343)				
			hnRNP U (3192)				
			hnRNP G (27316)				
			RB10 (8241)				
			SAFB (6294)				
			PABP1 (26986)				
			hnRNP A3 (220968)				
			hnRNP K (3190)				
			NUMATRIN (4869)				
			ELAV-like1 (1994)				
			SRp75 (6429)				
			DDX17 (10521)				
			SRp55 (6431)				
			hnRNP A2/B1 (3181)				
			hnRNP C (3183)				
			hTra-2 beta (6434)				
			PARP1 (142)				
			9G8 (6432)				
			DDX3 (1654)				
			SRp20 (6428)				
			PSF (6421)				
			SRp38 (10772)				
			RACK1 (10399)				
			SRp46 (10929)				
11269	DDX19B	X	NUP214 (8021)	δεν υπάρχουν ορθόλογα		δεν υπάρχουν ορθόλογα	
			RACK1 (10399)				
80145	THOC7	X	THOC5 (8563)	CG17143	CG18005 (RED,3550)	Δεν υπάρχει ορθόλογο	
23435	TARDBP	X	Dhm1-like protein (22803)	CG10327	CG4799 (KPNA2, 3838)	δεν υπάρχουν ορθόλογα	
			EXOSC10 (5394)		CG11266 (RBM39, 9584)		
					CG7111 (RACK1, 10399)		
2091	FBL	X		CG9888	X	T01C3.7	X
7030	TFE3	X		Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
91746	YTHDC1	X		Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
1387	CREBBP	X		Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
8148	TAF15	X		Δεν υπάρχει		Δεν υπάρχει ορθόλογο	

				ορθόλογο			
2079	ERH	HSPA8 (3312)		CG1871	CG3689 (CPSF 25KD,11051) CG8831 (NUP54,53371) CG15481 (RRP41, 54512)	Δεν υπάρχει ορθόλογο	
5514	PPP1R10	X	PPP1CA (5499)	CG33526	CG1362 (CDK10, 8558)	Δεν υπάρχει ορθόλογο	
40084	Npw38	EEF1A1 (1915)	Npw38BP (51729)	Δεν υπάρχει ορθόλογο		T21D12.3 και T07A9.1	X
			POLR2A (5430) U5 15K (10907)				
57794	SF4	X	X	CG31550	CG9998 (U2AF65, 11338)	Δεν υπάρχει ορθόλογο	
2521	FUS	X		CG3606	X	C27H5.3	X
7536	SF1	U2AF65 (11338) DDX17 (10521)	U2AF65 (11338) DDX17 (10521)	CG5836	X	Y116A8C.32	X
			FUS (2521) TAF15 (8148) EWS (2130) FBP11 (55660) TCERG1 (10915)				
27336	tat SF1	X		Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
728642	CDC2L2	X		CG4268	X	ZC504.3 και B0495.2	X
6511	NIPP1	X				B0511.7	X
29115	HCNGP	X	PUF60 (22827)	CG2063	X	Δεν υπάρχει ορθόλογο	
5499	PPP1CA	X				F56C9.1	X
25109	AGGF1	X	X	CG8079	CG18005 (RED,3550) CG7757 (U4/U6 90K, 9129)	Y55D9A.2	X
8880	FUBP1	VIM (7431)	VIM (7431) PUF60 (22827)	Δεν υπάρχει ορθόλογο		C12D8.1	D1046.1 (CPSF6, 11052) F34D10.4 (SRm300, 23524) ZK849.2 (RBM25, 58517)
5037	PEBP1	X	X	CG10298	X	Δεν υπάρχει ορθόλογο	
5725	hnRNP I	U1-A (6626)				D2089.4	Y54H5A.3 (RBM42, 79171)
2130	EWS	RPL31 (6160)		Δεν υπάρχει ορθόλογο		Δεν υπάρχει ορθόλογο	
		RPS15A (6210) CPSF6 (11052)					
11051	CPSF 25 kDa	CPSF6 (11052)	CPSF6 (11052) EXOSC6 (118460) U1-70K (6625)	CG3689	CG7185 (CPSF6, 11052) CG1871 (ERH, 2079)	F43G9.5	D1046.1 (CPSF6, 11052)
10898	CPSF 30 kDa	X	X	CG3642	CG1078 (Q8WVN3, 81608)	F11A10.8	X
11052	CPSF6	CPSF 25KD (11051) EWS (2130)	CPSF 25KD (11051) EWS (2130)	CG7185	CG12357 (CBP20, 22916) CG3689 (CPSF 25 KD, 10051)	D1046.1	C12D8.1 (FUBP1, 8880) F43G9.5 (CPSF 25KD, 11051)
			9G8 (6432)				

6741	Lupus La protein	X	NCL (4691)	CG10922	CG4887 (RBM5, 10181)	C44E4.4	X
22803	Dhm1-like protein	X				Y48B6A.3	F26B1.2 (hnRNPk, 3190)
6653	DDX1	X	HNRNPK (3190)	CG9054	X	Y55F3BR.1	X
55119	PRPF38B	X	X	CG1622	X	Y108G3A1.2	X
41083	C20orf158	hnRNP K (3190)	hnRNP K (3190)	δεν υπάρχει ορθόλογο		δεν υπάρχει ορθόλογο	
		SF2 (6426)	SF2 (6426)				
1608	Q8WVN3	CPSF 30KD (10898)	CPSF 30KD (10898)	CG1078	CG3642 (CPSF 30KD, 10898)	δεν υπάρχει ορθόλογο	
3146	HMGB1	X	TFE3 (7030) hnRNP K (3190)	CG12223	X	F47D12.4	X
6730	SRP68	X		CG5064	CG1676 (c19orf29, 58509) CG7757 (U4/U6 90K,9129) CG8280 (EEF1A1, 1915)	F55C5.8	X
4915	EEF1A1	CCDC130 (81576) ZCCHC10 (54819) LSm3 (27258) Npw38 (10084)		CG8280	CG5065 (SRP68, 6730)	F31E3.5	X
3838	KPNA2	X	CREBBP (1387) HSP70 (3308)	CG4799	CG3225 (DDX35, 60625) CG6944 (Lamin B, 4001) CG8900 (RPS18,6222)	δεν υπάρχει ορθόλογο	
4691	NCL	X	Lupus La protein (6741) SFRS12 (140890) PARP1 (142) NUMATRIN (4869)	δεν υπάρχει ορθόλογο		δεν υπάρχει ορθόλογο	
7431	VIM	FUBP1 (8880)	FUBP1 (8880)	δεν υπάρχει ορθόλογο		δεν υπάρχει ορθόλογο	
		SERBP1 (26135) ASR2 (51593)	ASR2 (51593) Desmoplakin (1832) PUF60 (22827) SERBP1 (26135) Npw38BP (51729) PKD1 (5317)				
203068	TUBB	X	RACK1 (10399)	δεν υπάρχει ορθόλογο		K01G5.7	X
			CRIP1 (9419)				
6421	PSF	X	FUS (2521) hnRNP I (5725) U1-A (6626) EXOSC5 (56915) SFRS12 (140890) p54nrb (4841)	CG4211	CG7238 (TFIP11, 24144)	δεν υπάρχει ορθόλογο	
9295	SFRS11	X	RNPS1 (10921)	δεν υπάρχει ορθόλογο		δεν υπάρχει ορθόλογο	

			SF2 (6426)				
			SC35 (6427)				
			U2AF65 (11338)				
			PUF60 (22827)				
11747	LUC7A	X	SF2 (6426)	CG3198	CG17136 (SRp20, 6428)	C50D2.8	X
5313	PPIL4	X	X	CG5808	CG1622 (PRPF38B, 55119)	F39H2.2	X
1218	GEMIN3	X	DDX9 (1660)	δεν υπάρχει ορθόλογο		δεν υπάρχει ορθόλογο	
			NCoR1 (9611)				
			Sm protein B (6628)				
			Sm protein D1(6632)				
			Sm protein D2 (6633)				
			Sm protein D3 (6634)				
			Sm protein E (6635)				
			Sm protein F (6636)				
			Sm protein G (6637)				
			GEMIN4 (50628)				
			GEMIN5 (25929)				
1938	EEF2	X	X	δεν υπάρχει ορθόλογο		δεν υπάρχει ορθόλογο	
5317	PKP1	X	DSG1 (1828)	δεν υπάρχει ορθόλογο		δεν υπάρχει ορθόλογο	
			Desmoplakin (1832)				
			VIM (7431)				
118460	EXOSC6	X	CPSF 25 kDa (11051)	CG8025	CG6249 (EXOSC1, 51013)	δεν υπάρχει ορθόλογο	
			RRP41 (54512)		CG8395 (EXOSC7, 23016)		
			RRP44 (22894)				
			EXOSC7 (23016)				
			LSm7 (51690)				
			LSm8 (51691)				
			EXOSC8 (11340)				
			EXOSC9 (5393)				
			EXOSC5 (56915)				
			EXOSC2 (23404)				
			SRp20 (6428)				
			SKIV2L2 (23517)				
			EXOSC10 (5394)				
56915	EXOSC5	X	NUP210 (23225)	CG4043	X	C14A4.5	X
			PSF (6421)				
			EXOSC2 (23404)				
			RRP44 (22894)				
			EXOSC9 (5393)				
			EXOSC8 (11340)				

			EXOSC7 (23016)				
			EXOSC6 (118460)				
			SKIV2L2 (23517)				
			EXOSC1 (51013)				
50628	GEMIN4	X	GEMIN3 (11218)	δεν υπάρχει ορθόλογο		δεν υπάρχει ορθόλογο	
			GEMIN5 (25929)				
			Sm protein D1 (6632)				
			Sm protein D2 (6633)				
			Sm protein B (6628)				
			ASR2 (51593)				
			DDX9 (1660)				
8841	p54nrb	X	PSF (6421)	δεν υπάρχει ορθόλογο		δεν υπάρχει ορθόλογο	
			POLR2A (5430)				
0657	SAM68	X	CREBBP (1387)	δεν υπάρχει ορθόλογο		δεν υπάρχει ορθόλογο	
			DDX9 (1660)				
			hnRNP K (3190)				
			FNBP4 (23360)				
			FBP21 (11193)				
			YTHDC1 (91746)				
26135	SERBP1	VIM (7431)	VIM (7431)	CG4170 και CG41503	X	F56D12.5	X
			PIAS1 (8554)				
8021	NUP214	X	NUP88 (4927)	CG3820	CG3647 (NXF1, 10482)	C03D6.4	X
			NXF1 (10482)				
			DDX19B (11269)				
			NUP107 (57122)				
9688	NUP93	X	NUP205 (23165)	CG7262	X	Y37E3.15	X
57122	NUP107	X	NUP133 (55746)	CG6743	CG1474 (DGCR14,8220)	Δεν υπάρχει ορθόλογο	
			NUP214 (8021)				
53371	NUP54	X	X	CG8831	CG1871 (ERH, 2079)	K07F5.13	X
55746	NUP133	X	NUP98 (4928)	CG6958	X	δεν υπάρχει ορθόλογο	
			NUP107 (57122)				
23165	NUP205	X	NUP93 (9688)	CG11943	X	K12D12.2	X
6749	FACT80	X	Ran (5901)	CG4817	X	C32F10.5 και T20B12.8	X
			EXOSC10 (5394)				
			SFRS12 (140890)				
			DDX9 (1660)				
71	Actin	X	X	CG4027	CG12085 (PUF60,22827)	T25C8.2	X
					CG5215 (ZFR, 51663 και ZFR2, 23217)		
4869	NUMATRIN	X	NCL (4691)	δεν υπάρχει		δεν υπάρχει ορθόλογο	

				ορθόλογο			
			TCERG1 (10915)				
			SFRS12 (140890)				
4001	Lamin B	X	X	CG6944	CG4799 (KPNA2, 3838)	DY3.2	X
320	HSP90A	RPS3A (6189)	RPS3A (6189)	δεν υπάρχει ορθόλογο		δεν υπάρχει ορθόλογο	
			HSPA8 (3312)				
			CRNKL1 (51340)				
			PPIL2 (23759)				
			PARP1 (142)				

3.2 Κλωνοποίηση των cDNA πλήρους μήκους των υποψήφιων αλληλεπιδρώντων της πρωτεΐνης DGCR14, RED και RACK1 και παραγωγή της ανασυνδυασμένης GST-DGCR14 σε βακτήρια

Η αξιοπιστία του δικτύου πρωτεϊνικών αλληλεπιδράσεων που κατασκευάστηκε θα πρέπει να μελετηθεί περαιτέρω με προσεγγίσεις που περιλαμβάνουν συλλογή δεδομένων σχετικά με το πρότυπο έκφρασης των συμμετεχόντων πρωτεϊνών ή με την εφαρμογή κατάλληλων βιοϋπολογιστικών προγραμμάτων. Ωστόσο, τελική απόδειξη της αξιοπιστίας του δικτύου παραμένει η απόδειξη των προβλεπόμενων αλληλεπιδράσεων, με συνδυαστικές πειραματικές προσεγγίσεις. Σε αυτήν την εργασία δημιουργήθηκαν τα εργαλεία για τον έλεγχο των προβλεπόμενων αλληλεπιδράσεων της πρωτεΐνης DGCR14 με τις πρωτεΐνες RED και RACK1. Καθώς αυτές οι αλληλεπιδράσεις εντοπίστηκαν μεταξύ των ορθολόγων μορίων στη *D. melanogaster* και τον *C. elegans*, η πειραματική διερεύνησή τους θα συνεισφέρει και στον έλεγχο της υπόθεσης της εξελικτικής συντήρησης των αλληλεπιδράσεων, στην οποία βασίσθηκε ένα σημαντικό ποσοστό των αλληλεπιδράσεων του δικτύου.

Ειδικότερα πραγματοποιήθηκε η κλωνοποίηση των πλήρους μήκους μορίων cDNA των γονιδίων RED και RACK καθώς και η παραγωγή ανασυνδυασμένης πρωτεΐνης GST-DGCR14 σε βακτήρια και η απομόνωσή της με χρωματογραφία συγγένειας.

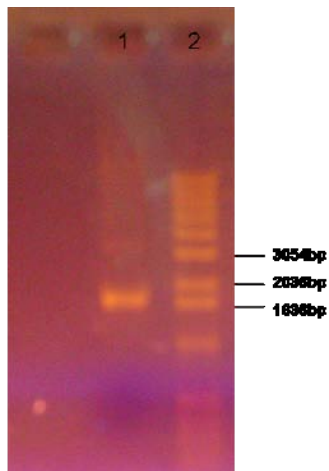
➤ 3.2.1 Κλωνοποίηση του cDNA του γονιδίου RED

Αρχικά, το cDNA του γονιδίου RED υποκλωνοποιήθηκε στο φορέα pGEM-T Easy, ο οποίος είναι κατάλληλος για την κλωνοποίηση προϊόντων PCR. Ακολούθησε έλεγχος της πλασμιδιακής κατασκευής με πέψη του ενθέματος με τα κατάλληλα περιοριστικά ένζυμα και επακόλουθη αλληλούχηση και στη συνέχεια απομόνωση του ενθέματος και κλωνοποίησή του στον ευκαρυωτικό φορέα έκφρασης pEGFP-C1, έτσι ώστε η πρωτεΐνη RED να παράγεται σε σύντηξη με την EGFP.

Υποκλωνοποίηση του cDNA του γονιδίου RED στο φορέα pGEM-T Easy

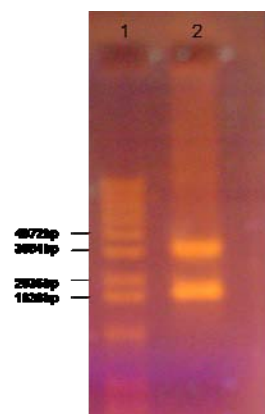
Το cDNA του γονιδίου RED απομονώθηκε με PCR που πραγματοποιήθηκε σε DNA προερχόμενο από βιβλιοθήκη cDNA εγκεφάλου του ανθρώπου. Οι εκκινητές ήταν σχεδιασμένοι με τέτοιο τρόπο ώστε να σχηματιστεί cDNA πλήρους μήκους που έφερε στα άκρα του θέσεις αναγνώρισης περιοριστικών ενδονουκλεασών για την επόμενη κλωνοποίηση. Στην PCR χρησιμοποιήθηκαν ως υπόστρωμα 100 ng DNA από τη βιβλιοθήκη ενώ η βέλτιστη θερμοκρασία υβριδοποίησης των εκκινητών

προσδιορίστηκε με PCR κλίσης θερμοκρασίας (gradient PCR) στους 57°C, κατά την οποία προκύπτει μία μόνο ζώνη του αναμενόμενου μεγέθους 1690 bp (Εικ. 11).



Εικ. 11: Προϊόν PCR του cDNA του γονιδίου RED. Η ηλεκτροφόρηση έγινε σε πήκτωμα αγαρόζης συγκέντρωσης 1,5%. Σε θερμοκρασία υβριδοποίησης εκκινητών 57°C, είναι εμφανής η αναμενόμενη ζώνη μεγέθους 1690bp (1). Το δείγμα 2 αντιστοιχεί σε δείκτη γνωστών μοριακών μεγεθών.

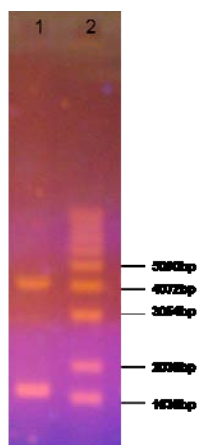
Στη συνέχεια, τεμαχίστηκε το τμήμα της πηκτής που περιλάμβανε το προϊόν της PCR και απομονώθηκε το DNA που περιελάμβανε. Ακολούθως, το καθαρό προϊόν PCR (~100 ng) χρησιμοποιήθηκε σε αντίδραση σύνδεσης με το φορέα pGEM-T Easy. Ακολούθησε μετασχηματισμός δεκτικών κυττάρων *E. coli* και έλεγχος των ανασυνδυασμένων μορίων με πέψη με την περιοριστική ενδονουκλεάση *EcoRI*. Η *EcoRI* υπάρχει στο φορέα σε δύο θέσεις εκατέρωθεν του ενθέματος (βλ. Υλικά και Μέθοδοι). Η πέψη ελέγχθηκε με ηλεκτροφόρηση σε πηκτή αγαρόζης. Από τους θετικούς κλώνους, επιλέχθηκε ένας (Εικ. 12) για τον οποίο έγινε και προσδιορισμός της αλληλουχίας του ενθέματος, χρησιμοποιώντας τους εκκινητές T7 και SP6 του φορέα.



Εικ. 12: Έλεγχος των ανασυνδυασμένων μορίων με πέψη με την περιοριστική ενδονουκλεάση EcoRI. Η ηλεκτροφόρηση έγινε σε πηκτή αγαρόζης συγκέντρωσης 1,5%. Φαίνεται η αναμενόμενη ζώνη του ενθέματος μεγέθους 1690 bp (2) καθώς και η ζώνη του φορέα pGEM-T Easy μεγέθους ~3kb (2). Το δείγμα 1 αντιστοιχεί σε δείκτη γνωστών μοριακών μεγεθών.

Κλωνοποίηση του cDNA του γονιδίου RED στο φορέα pEGFP-C1

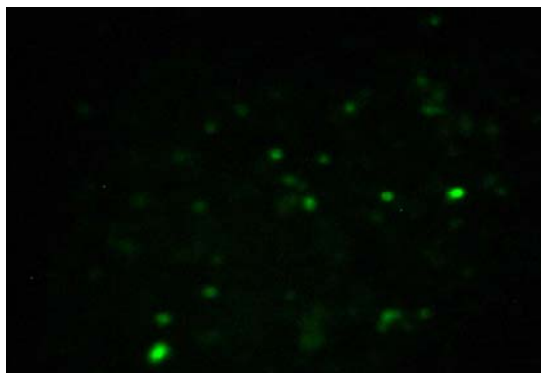
Καθώς οι αρχικοί εκκινητές έφεραν στα άκρα 5' τους θέσεις αναγνώρισης των περιοριστικών ενδονουκλεασών *Bgl*III και *Eco*RI, για την απομόνωση του ενθέματος από το φορέα pGEM-T Easy έγιναν δυο διαδοχικές πέψεις. Αρχικά, έγινε πέψη με την περιοριστική ενδονουκλεάση *Bgl*III και καθαρισμός του προϊόντος και στη συνέχεια η δεύτερη πέψη με την περιοριστική ενδονουκλεάση *Eco*RI και καθαρισμός του προϊόντος. Το προϊόν των πέψεων (~200 ng) χρησιμοποιήθηκε σε αντίδραση σύνδεσης με τον πλασμιδιακό φορέα pEGFP-C1 (ο οποίος είχε τεμαχιστεί με τα ίδια ένζυμα). Μετά το μετασχηματισμό και την καλλιέργεια των βακτηρίων, έγινε έλεγχος των ανασυνδυασμένων κλώνων με πέψη: α) με την περιοριστική ενδονουκλεάση *Pst*I η οποία έχει δυο θέσεις αναγνώρισης, μια στο ένθεμα στις 420 bp και μια στην κασέτα κλωνοποίησης (polylinker) του φορέα και β) με *Bgl*III και *Eco*RI, που είναι οι θέσεις κλωνοποίησης. Από τους θετικούς κλώνους επιλέχθηκε ένας (Εικ. 13), για τον οποίο προσδιορίστηκε η αλληλουχία του ενθέματος (εστιάζοντας στη θέση κλωνοποίησης 5') με τον εκκινητή pEGFP_Fw. Η αλληλουχία που προέκυψε ήταν ίδια με την κατατεθμεμένη αλληλουχία NM_006083.3.



Εικ. 13: Έλεγχος των ανασυνδυασμένων μορίων με πέψη με τις περιοριστικές ενδονουκλεάσες *Eco*RI και *Bgl*III. Η ηλεκτροφόρηση έγινε σε πηκτική αгарόζης συγκέντρωσης 1,5%. Φαίνεται η αναμενόμενη ζώνη του ενθέματος μεγέθους 1690 bp (2) καθώς και η ζώνη του φορέα pEGFP-C1 μεγέθους 4,7kb (1). Το δείγμα 2 αντιστοιχεί σε δείκτη γνωστών μοριακών μεγεθών.

Ο έλεγχος της λειτουργικότητας της πλασμιδιακής κατασκευής, δηλαδή η ικανότητά της να παράγει ανασυνδυασμένη πρωτεΐνη σε κύτταρα θηλαστικών, η οποία εντοπίζεται στον πυρήνα (Assier *et al.* 1999), πραγματοποιήθηκε με παροδική διαμόλυνση της κυτταρικής σειράς *HeLa*. Η καλλιέργεια των κυττάρων έγινε σε γυάλινη καλυπτήριδα, η οποία είχε προηγουμένως τοποθετηθεί σε τρυβλίο διαμέτρου 10 cm. Για τη διαμόλυνση χρησιμοποιήθηκαν 20 µg DNA από την πλασμιδιακή κατασκευή pEGFP-C1/RED. Η ποσοτικοποίηση και ο έλεγχος καθαρότητας του DNA

(260/280nm=1,6) έγινε με φωτομέτρηση. Η μονιμοποίηση των κυττάρων στην καλυπτρίδα και η παρατήρησή τους σε μικροσκόπιο φθορισμού πραγματοποιήθηκε 48 h μετά τη διαμόλυνσή τους (Εικ. 14).



Εικ. 14: Επιθηλιακά κύτταρα *HeLa* μετά από παροδική διαμόλυνση με την πλασμιδιακή κατασκευή pEGFP-C1/RED. Η ανασυνδυασμένη πρωτεΐνη RED-EGFP εντοπίζεται στον πυρήνα (μεγέθυνση 20X).

➤ 3.2.2 Κλωνοποίηση του cDNA του γονιδίου *RACK1*

Το πρώτο βήμα για την κλωνοποίηση του cDNA του γονιδίου *RACK1* είναι η υποκλωνοποίηση του στο φορέα pGEM-T Easy ο οποίος είναι κατάλληλος για την κλωνοποίηση προϊόντων PCR. Το βήμα αυτό πραγματοποιείται για να γίνει πλήρης έλεγχος της αλληλουχίας του ενισχυμένου τμήματος.

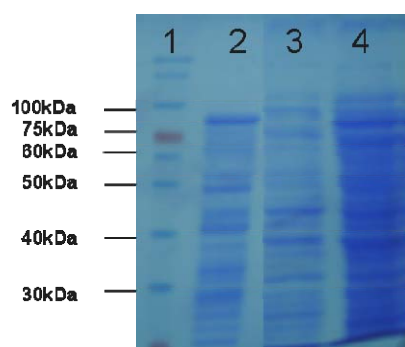
Υποκλωνοποίηση του cDNA του γονιδίου *RACK1* στο φορέα pGEM-T Easy

Το cDNA του γονιδίου *RACK1* απομονώθηκε με PCR από την πλασμιδιακή κατασκευή pcDNA3/*RACK1* (Jean-Luc Parent, CHUS, Quebec). Οι εκκινητές ήταν σχεδιασμένοι με τέτοιο τρόπο ώστε να σχηματιστεί cDNA πλήρους μήκους που έφερε στα άκρα του θέσεις αναγνώρισης περιοριστικών ενδονουκλεασών (*SacI* και *PstI*) για την επόμενη κλωνοποίηση. Στην PCR χρησιμοποιήθηκαν ως υπόστρωμα 10 ng DNA προερχόμενο από την πλασμιδιακή κατασκευή ενώ η βέλτιστη θερμοκρασία υβριδοποίησης των εκκινητών προσδιορίστηκε με PCR κλίσης θερμοκρασίας (gradient PCR) στους 57°C, κατά την οποία προκύπτει μία μόνο ζώνη του αναμενόμενου μεγέθους 970 bp. Ακολούθως, το τμήμα της πηκτής που περιλάμβανε το προϊόν της PCR τεμαχίστηκε και απομονώθηκε το DNA που περιελάμβανε. Το καθαρό προϊόν PCR (~100 ng) χρησιμοποιήθηκε σε αντίδραση σύνδεσης με το φορέα pGEM-T Easy. Ακολούθησε μετασχηματισμός κυττάρων *E. coli* και έλεγχος των ανασυνδυασμένων μορίων με πέψη με την περιοριστική ενδονουκλεάση *EcoRI*. Η *EcoRI* υπάρχει στο φορέα σε δύο θέσεις εκατέρωθεν του ενθέματος καθώς και στην αλληλουχία του ενθέματος στη θέση 158. Μετά την ηλεκτροφόρηση επιλέχθηκε ένας κλώνος για προσδιορισμό της αλληλουχίας του

ενθέματος, με τους εκκινητές T7 και SP6 του φορέα. Η αλληλουχία που προέκυψε ήταν ίδια με την κατατεθειμένη αλληλουχία NM_006098.4.

➤ 3.2.3 Παραγωγή ανασυνδυασμένης πρωτεΐνης GST-DGCR14 σε βακτήρια και απομόνωση με χρωματογραφία συγγένειας

Για την παραγωγή της ανασυνδυασμένης πρωτεΐνης GST-DGCR14 χρησιμοποιήθηκαν βακτήρια *E. coli* στελέχους BL21 (DE3) Codon plus που μετασχηματίστηκαν με πλασμιδιακή κατασκευή που έφερε ως ένθεμα το cDNA σχεδόν πλήρους μήκους του γονιδίου *DGCR14* (περιλαμβάνει τα αμινοξέα 25-472) στο φορέα κλωνοποίησης pGEX-3X). Ως βέλτιστες συνθήκες επαγωγής και έκφρασης της ανασυνδυασμένης πρωτεΐνης επιλέχθηκε η θερμοκρασία των 20°C, με 0,1 mM IPTG, ολονύκτια. Στη συνέχεια, τα κύτταρα συλλέχθηκαν και ακολούθησε απομόνωση του πρωτεϊνικού εκχυλίσματος, εφαρμόζοντας πρωτόκολλο απομόνωσης από το αδιάλυτο κλάσμα, όπου και εντοπίζεται η GST-DGCR14 (βλ. Υλικά και Μέθοδοι). Η απομόνωση της ανασυνδυασμένης πρωτεΐνης έγινε με χρωματογραφία συγγένειας με σφαιρίδια σεφαρόζης 4B συζευγμένα με γλουταθειόνη. Ο έλεγχος της υπερέκφρασης και της απομόνωσης της ανασυνδυασμένης πρωτεΐνης πραγματοποιήθηκε με ηλεκτροφόρηση σε αποδιατακτικό πήκτωμα πολυακρυλαμίδης (SDS- PAGE) και ακόλουθη χρώση του πηκτώματος με Coomassie Blue (Εικ. 14). Το αναμενόμενο μέγεθος της ανασυνδυασμένης πρωτεΐνης είναι ~80 kD, σε συμφωνία με τη ζώνη που εμφανίζεται στο αντίστοιχο δείγμα (2). Σε συμφωνία είναι επίσης το ότι αυτή η ζώνη δεν εμφανίζεται στο ολικό πρωτεϊνικό εκχύλισμα πριν την επαγωγή της έκφρασης, που χρησιμοποιήθηκε ως αρνητικός έλεγχος του πειράματος, αλλά εμφανίζεται στο ολικό πρωτεϊνικό εκχύλισμα μετά την ολονύκτια επαγωγή της έκφρασης της ανασυνδυασμένης πρωτεΐνης (Εικ. 15).



Εικ. 15: Παραγωγή και απομόνωση της ανασυνδυασμένης πρωτεΐνης GST-DGCR14. Η ανασυνδυασμένη πρωτεΐνη μετά την απομόνωσή της με χρωματογραφία συγγένειας εμφανίζει το αναμενόμενο μέγεθος ~80 kD (2). Τα δείγματα (3) και (4) αντιστοιχούν στο ολικό πρωτεϊνικό εκχύλισμα πριν και μετά την επαγωγή της έκφρασης της πρωτεΐνης, αντίστοιχα. Στο πρώτο δείγμα φαίνεται ο δείκτης γνωστών μοριακών μεγεθών (1).

BIBΛΙΟΓΡΑΦΙΑ

1. Alderton A.L, University of Kentucky, Lexington, KY, USA 2004 Elsevier Ltd.
2. Aranda B, Achuthan P, Alam-Faruque Y, Armean I, Bridge A, Derow C, Feuermann M, Ghanbarian AT, Kerrien S, Khadake J, Kerssemakers J, Leroy C, Menden M, Michaut M, Montecchi-Palazzi L, Neuhauser SN, Orchard S, Perreau V, Roechert B, van Eijk K, Hermjakob H. "The IntAct molecular interaction database in 2010." *Nucleic Acids Res.* 2010 Jan; 38(Database issue):D525-31. Epub 2009 Oct 22.
3. Assier E, Bouzinba-Segard H, Stolzenberg MC, Stephens R, Bardos J, Freemont P, Charron D, Trowsdale J, Rich T. "Isolation, sequencing and expression of RED, a novel human gene encoding an acidic-basic dipeptide repeat." *Gene.* 1999 Apr 16; 230(2):145-54.
4. Behzadnia N, Golas MM, Hartmuth K, Sander B, Kastner B, Deckert J, Dube P, Will CL, Urlaub H, Stark H, Lührmann R. "Composition and three-dimensional EM structure of double affinity-purified, human prespliceosomal A complexes." *EMBO J.* 2007 Mar 21; 26(6):1737-48. Epub 2007 Mar 1.
5. Belgareh N, Rabut G, Bai SW, van Overbeek M, Beaudouin J, Daigle N, Zatssepina OV, Pasteau F, Labas V, Fromont-Racine M, Ellenberg J, Doye V. "An evolutionarily conserved NPC subcomplex, which redistributes in part to kinetochores in mammalian cells." *J Cell Biol.* 2001 Sep 17; 154(6):1147-60.
6. Bennett M, Michaud S, Kingston J, Reed R. "Protein components specifically associated with prespliceosome and spliceosome complexes." *Genes Dev.* 1992 Oct; 6(10):1986-2000.
7. Bessonov S, Anokhina M, Krasauskas A, Golas MM, Sander B, Will CL, Urlaub H, Stark H, Lührmann R. "Characterization of purified human Bact spliceosomal complexes reveals compositional and morphological changes during spliceosome activation and first step catalysis." *RNA.* 2010 Dec; 16(12):2384-403. Epub 2010 Oct 27.
8. Bessonov S, Anokhina M, Will CL, Urlaub H, Lührmann R. "Isolation of an active step I spliceosome and composition of its RNP core." *Nature.* 2008 Apr 17; 452(7189):846-50. Epub 2008 Mar 5.
9. Boehmer T, Enninga J, Dales S, Blobel G, Zhong H. "Depletion of a single nucleoporin, Nup107, prevents the assembly of a subset of nucleoporins into the nuclear pore complex." *Proc Natl Acad Sci U S A.* 2003 Feb 4; 100(3):981-5. Epub 2003 Jan 27.
10. Breitkreutz BJ, Stark C, Reguly T, Boucher L, Breitkreutz A, Livstone M, Oughtred R, Lackner DH, Bähler J, Wood V, Dolinski K, Tyers M. "The BioGRID Interaction Database: 2008 update." *Nucleic Acids Res.* 2008 Jan; 36(Database issue):D637-40. Epub 2007 Nov 13.
11. Bressan GC, Quaresma AJ, Moraes EC, Manfiolli AO, Passos DO, Gomes MD, Kobarg J. Functional association of human Ki-1/57 with pre-mRNA splicing events. *FEBS J.* 2009 Jul; 276(14):3770-83. Epub 2009 Jun 11.
12. Burge CB, Padgett RA, Sharp PA. "Evolutionary fates and origins of U12-type introns." *Mol Cell.* 1998 Dec; 2(6):773-85.
13. Ceol A, Chatr Aryamontri A, Licata L, Peluso D, Briganti L, Perfetto L, Castagnoli L, Cesareni G. "MINT, the molecular interaction database: 2009 update." *Nucleic Acids Res.* 2010 Jan; 38(Database issue):D532-9. Epub 2009 Nov 6.
14. Champion-Arnaud P, Reed R. "The prespliceosome components SAP 49 and SAP 145 interact in a complex implicated in tethering U2 snRNP to the branch site." *Genes Dev.* 1994 Aug 15; 8(16):1974-83.
15. Chen N, Harris TW, Antoshechkin I, Bastiani C, Bieri T, Blasiar D, Bradnam K, Canaran P, Chan J, Chen CK, Chen WJ, Cunningham F, Davis P, Kenny E, Kishore R, Lawson D, Lee R, Muller HM, Nakamura C, Pai S, Ozersky P, Petcherski A, Rogers A, Sabo A, Schwarz EM, Van Auken K, Wang Q, Durbin R, Spieth J, Sternberg PW, Stein LD. "WormBase: a comprehensive data resource for *Caenorhabditis* biology and genomics." *Nucleic Acids Res.* 2005 Jan 1; 33(Database issue):D383-9.
16. Chen YI, Moore RE, Ge HY, Young MK, Lee TD, Stevens SW. "Proteomic analysis of in vivo-assembled pre-mRNA splicing complexes expands the catalog of participating factors." *Nucleic Acids Res.* 2007; 35(12):3928-44. Epub 2007 May 30.

17. Cohen SN, Chang AC, Hsu L. "Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA." *Proc Natl Acad Sci U S A*. 1972 Aug; 69(8):2110-4.
18. Debacker K, Winnepenninckx B, Longman C, Colgan J, Tolmie J, Murray R, van Luijk R, Scheers S, Fitzpatrick D, Kooy F. "The molecular basis of the folate-sensitive fragile site FRA11A at 11q13." *Cytogenet Genome Res*. 2007; 119(1-2):9-14. Epub 2007 Dec 14.
19. Dobkin C, Rabe A, Dumas R, El Idrissi A, Haubenstock H, Brown WT. "Fmr1 knockout mouse has a distinctive strain-specific learning impairment". *Neuroscience*. 2000; 100(2):423-9.
20. Fields S, Song O. "A novel genetic system to detect protein-protein interactions." *Nature*. 1989 Jul 20; 340(6230):245-6.
21. Figeys D. "Mapping the human protein interactome." *Cell Res*. 2008 Jul; 18(7):716-24. Review.
22. Frangioni JV, Neel BG. "Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins." *Anal Biochem*. 1993 Apr; 210(1):179-87.
23. Fu YH, Kuhl DP, Pizzuti A, Pieretti M, Sutcliffe JS, Richards S, Verkerk AJ, Holden JJ, Fenwick RG Jr, Warren ST, et al. "Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox." *Cell*. 1991 Dec 20; 67(6):1047-58.
24. Gavin AC, Börsche M, Krause R, Grandi P, Marzioch M, Bauer A, Schultz J, Rick JM, Michon AM, Cruciat CM, Remor M, Höfert C, Schelder M, Brajenovic M, Ruffner H, Merino A, Klein K, Hudak M, Dickson D, Rudi T, Gnau V, Bauch A, Bastuck S, Huhse B, Leutwein C, Heurtier MA, Copley RR, Edlmann A, Querfurth E, Rybin V, Drewes G, Raida M, Bouwmeester T, Bork P, Seraphin B, Kuster B, Neubauer G, Superti-Furga G. "Functional organization of the yeast proteome by systematic analysis of protein complexes." *Nature*. 2002 Jan 10; 415(6868):141-7.
25. Giot L, Bader JS, Brouwer C, Chaudhuri A, Kuang B, Li Y, Hao YL, Ooi CE, Godwin B, Vitols E, Vijayadamodar G, Pochart P, Machineni H, Welsh M, Kong Y, Zerhusen B, Malcolm R, Varrone Z, Collis A, Minto M, Burgess S, McDaniel L, Stimpson E, Spriggs F, Williams J, Neurath K, Ioime N, Agee M, Voss E, Furtak K, Renzulli R, Aanensen N, Carrola S, Bickelhaupt E, Lazovatsky Y, DaSilva A, Zhong J, Stanyon CA, Finley RL Jr, White KP, Braverman M, Jarvie T, Gold S, Leach M, Knight J, Shimkets RA, McKenna MP, Chant J, Rothberg JM. "A protein interaction map of *Drosophila melanogaster*." *Science*. 2003 Dec 5;302(5651):1727-36. Epub 2003 Nov 6.
26. Gonçalves KA, Borges JC, Silva JC, Papa PF, Bressan GC, Torriani IL, Kobarg J. "Solution structure of the human signaling protein RACK1." *BMC Struct Biol*. 2010 Jun 8; 10:15.
27. Gong W, Emanuel BS, Galili N, Kim DH, Roe B, Driscoll DA, Budarf ML. "Structural and mutational analysis of a conserved gene (DGSI) from the minimal DiGeorge syndrome critical region." *Hum Mol Genet*. 1997 Feb; 6(2):267-76.
28. Gozani O, Feld R, Reed R. "Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A." *Genes Dev*. 1996 Jan 15; 10(2):233-43.
29. Graham FL, van der Eb AJ. "A new technique for the assay of infectivity of human adenovirus 5 DNA." *Virology*. 1973 Apr;52(2):456-67
30. Graham FL, van der Eb AJ. "Transformation of rat cells by DNA of human adenovirus 5." *Virology*. 1973 Aug; 54(2):536-9.
31. Green MR. "Biochemical mechanisms of constitutive and regulated pre-mRNA splicing." *Annu Rev Cell Biol*. 1991; 7:559-99.
32. Guthrie C. "Messenger RNA splicing in yeast: clues to why the spliceosome is a ribonucleoprotein." *Science*. 1991 Jul 12; 253(5016):157-63.
33. Hartmuth K, Urlaub H, Vornlocher HP, Will CL, Gentzel M, Wilm M, Lührmann R. "Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method." *Proc Natl Acad Sci U S A*. 2002 Dec 24;99(26):16719-24. Epub 2002 Dec 11.
34. Herold N, Will CL, Wolf E, Kastner B, Urlaub H, Lührmann R. "Conservation of the protein composition and electron microscopy structure of *Drosophila melanogaster*

- and human spliceosomal complexes." *Mol Cell Biol.* 2009 Jan; 29(1):281-301. Epub 2008 Nov 3.
35. Hewett DR, Handt O, Hobson L, Mangelsdorf M, Eyre HJ, Baker E, Sutherland GR, Schuffenhauer S, Mao JI, Richards RI. "FRA10B structure reveals common elements in repeat expansion and chromosomal fragile site genesis." *Mol Cell.* 1998 May; 1(6):773-81.
 36. Huh SH, Ornitz DM. "Beta-catenin deficiency causes DiGeorge syndrome-like phenotypes through regulation of Tbx1." *Development.* 2010 Apr;137(7):1137-47.
 37. Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y. "A comprehensive two-hybrid analysis to explore the yeast protein interactome." *Proc Natl Acad Sci U S A.* 2001 Apr 10;98(8):4569-74. Epub 2001 Mar 13.
 38. Jiang Y, Lucas I, Young DJ, Davis EM, Karrison T, Rest JS, Le Beau MM. "Common fragile sites are characterized by histone hypoacetylation." *Hum Mol Genet.* 2009 Dec 1; 18(23):4501-12. Epub 2009 Aug 28.
 39. Jones C, Slijepcevic P, Marsh S, Baker E, Langdon WY, Richards RI, Tunnacliffe A. Physical linkage of the fragile site FRA11B and a Jacobsen syndrome chromosome deletion breakpoint in 11q23.3. *Hum Mol Genet.* 1994 Dec; 3(12):2123-30.
 40. Jurica MS, Licklider LJ, Gygi SR, Grigorieff N, Moore MJ. "Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis." *RNA.* 2002 Apr; 8(4):426-39.
 41. Kähkönen M, Tengström C, Alitalo T, Matilainen R, Kaski M, Airaksinen E. "Population cytogenetics of folate-sensitive fragile sites. II. Autosomal rare fragile sites." *Hum Genet.* 1989 Apr; 82(1):3-8.
 42. Keshava Prasad TS, Goel R, Kandasamy K, Keerthikumar S, Kumar S, Mathivanan S, Telikicherla D, Raju R, Shafreen B, Venugopal A, Balakrishnan L, Marimuthu A, Banerjee S, Somanathan DS, Sebastian A, Rani S, Ray S, Harrys Kishore CJ, Kanth S, Ahmed M, Kashyap MK, Mohmood R, Ramachandra YL, Krishna V, Rahiman BA, Mohan S, Ranganathan P, Ramabadrana S, Chaerkady R, Pandey A. "Human Protein Reference Database--2009 update." *Nucleic Acids Res.* 2009 Jan; 37(Database issue):D767-72. Epub 2008 Nov 6.
 43. Kicman AT, Parkin MC, Iles RK. "An introduction to mass spectrometry based proteomics-detection and characterization of gonadotropins and related molecules." *Mol Cell Endocrinol.* 2007 Jan 2; 260-262:212-27. Epub 2006 Nov 13. Review.
 44. Knight SJ, Flannery AV, Hirst MC, Campbell L, Christodoulou Z, Phelps SR, Pointon J, Middleton-Price HR, Barnicoat A, Pembrey ME, et al. "Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation." *Cell.* 1993 Jul 16; 74(1):127-34.
 45. König H, Matter N, Bader R, Thiele W, Müller F. "Splicing segregation: the minor spliceosome acts outside the nucleus and controls cell proliferation." *Cell.* 2007 Nov 16; 131(4):718-29.
 46. Kremer EJ, Pritchard M, Lynch M, Yu S, Holman K, Baker E, Warren ST, Schlessinger D, Sutherland GR, Richards RI. "Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n." *Science.* 1991 Jun 21; 252(5013):1711-4.
 47. Kremer EJ, Yu S, Pritchard M, Nagaraja R, Heitz D, Lynch M, Baker E, Hyland VJ, Little RD, Wada M, et al. "Isolation of a human DNA sequence which spans the fragile X." *Am J Hum Genet.* 1991 Sep; 49(3):656-61.
 48. Krief P, Augery-Bourget Y, Plaisance S, Merck MF, Assier E, Tanchou V, Billard M, Boucheix C, Jasmin C, Azzarone B. "A new cytokine (IK) down-regulating HLA class II: monoclonal antibodies, cloning and chromosome localization." *Oncogene.* 1994 Dec; 9(12):3449-56.
 49. Li S, Armstrong CM, Bertin N, Ge H, Milstein S, Boxem M, Vidalain PO, Han JD, Chesneau A, Hao T, Goldberg DS, Li N, Martinez M, Rual JF, Lamesch P, Xu L, Tewari M, Wong SL, Zhang LV, Berriz GF, Jacotot L, Vaglio P, Reboul J, Hirozane-Kishikawa T, Li Q, Gabel HW, Elewa A, Baumgartner B, Rose DJ, Yu H, Bosak S, Sequerra R, Fraser A, Mango SE, Saxton WM, Strome S, Van Den Heuvel S, Piano F, Vandenhaute J, Sardet C, Gerstein M, Doucette-Stamm L, Gunsalus KC, Harper JW, Cusick ME, Roth FP, Hill DE, Vidal M. "A map of the interactome network of the metazoan *C. elegans*." *Science.* 2004 Jan 23; 303(5657):540-3. Epub 2004 Jan 2.

50. Lukusa T, Fryns JP. "Human chromosome fragility." *Biochim Biophys Acta*. 2008 Jan; 1779(1):3-16. Epub 2007 Dec 3.
51. Makarov EM, Makarova OV, Urlaub H, Gentzel M, Will CL, Wilm M, Lührmann R. "Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome." *Science*. 2002 Dec 13; 298(5601):2205-8. Epub 2002 Oct 31.
52. Makarova OV, Makarov EM, Urlaub H, Will CL, Gentzel M, Wilm M, Lührmann R. "A subset of human 35S U5 proteins, including Prp19, function prior to catalytic step 1 of splicing." *EMBO J*. 2004 Jun 16; 23(12):2381-91. Epub 2004 Jun 3.
53. Mavrou A, Syrrou M, Tsenghi C, Metaxotou C. "Autosomal folate sensitive fragile sites in normal and mentally retarded individuals in Greece." *Am J Med Genet*. 1991 Feb-Mar; 38(2-3):437-9.
54. Mullis KB. "Target amplification for DNA analysis by the polymerase chain reaction." *Ann Biol Clin (Paris)*. 1990; 48(8):579-82.
55. Nancarrow JK, Kremer E, Holman K, Eyre H, Doggett NA, Le Paslier D, Callen DF, Sutherland GR, Richards RI. "Implications of FRA16A structure for the mechanism of chromosomal fragile site genesis." *Science*. 1994 Jun 24; 264(5167):1938-41.
56. Neubauer G. "The analysis of multiprotein complexes: the yeast and the human spliceosome as case studies." *Methods Enzymol*. 2005;405:236-63. Review.
57. Orrantia E, Chang PL. "Intracellular distribution of DNA internalized through calcium phosphate precipitation." *Exp Cell Res*. 1990 Oct;190(2):170-4.
58. Pandey A, Mann M. "Proteomics to study genes and genomes." *Nature*. 2000 Jun 15; 405(6788):837-46. Review.
59. Parrish JE, Oostra BA, Verkerk AJ, Richards CS, Reynolds J, Spikes AS, Shaffer LG, Nelson DL. "Isolation of a GCC repeat showing expansion in FRAXF, a fragile site distal to FRAXA and FRAXE." *Nat Genet*. 1994 Nov;8(3):229-35.
60. Petit P, Fryns JP, van den Berghe H, Hecht F. "Population cytogenetics of autosomal fragile sites." *Clin Genet*. 1986 Feb; 29(2):96-100.
61. Pichiorri F, Ishii H, Okumura H, Trapasso F, Wang Y, Huebner K. "Molecular parameters of genome instability: roles of fragile genes at common fragile sites." *JCell Biochem*. 2008 Aug 1; 104(5):1525-33.
62. Quack B, Nantois Y, Mottet J, Noël B. "Constitutional stereotyped gap in human chromosomes." *J Genet Hum*. 1978 Mar; 26(1):55-67.
63. Rappsilber J, Ryder U, Lamond AI, Mann M. "Large-scale proteomic analysis of the human spliceosome." *Genome Res*. 2002 Aug; 12(8):1231-45.
64. RAYMOND S, WEINTRAUB L. "Acrylamide gel as a supporting medium for zone electrophoresis." *Science*. 1959 Sep 18;130:711.
65. Reed R. "Coupling transcription, splicing and mRNA export." *Curr Opin Cell Biol*. 2003 Jun;15(3):326-31.
66. Ritchie DB, Schellenberg MJ, MacMillan AM. "Spliceosome structure: piece by piece" *Biochim Biophys Acta*. 2009 Sep-Oct;1789(9-10):624-33. Epub 2009 Sep 3.
67. Rizzu P, Lindsay EA, Taylor C, O'Donnell H, Levy A, Scambler P, Baldini A. "Cloning and comparative mapping of a gene from the commonly deleted region of DiGeorge and Velocardiofacial syndromes conserved in *C. elegans*." *Mamm Genome*. 1996 Sep;7(9):639-43.
68. Rual JF, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, Berriz GF, Gibbons FD, Dreze M, Ayivi-Guedehoussou N, Klitgord N, Simon C, Boxem M, Milstein S, Rosenberg J, Goldberg DS, Zhang LV, Wong SL, Franklin G, Li S, Albala JS, Lim J, Fraughton C, Llamosas E, Cevik S, Bex C, Lamesch P, Sikorski RS, Vandenhaute J, Zoghbi HY, Smolyar A, Bosak S, Sequerra R, Doucette-Stamm L, Cusick ME, Hill DE, Roth FP, Vidal M. "Towards a proteome-scale map of the human protein-protein interaction network." *Nature*. 2005 Oct 20;437(7062):1173-8. Epub 2005 Sep 28.
69. Sarafidou T, Apostolopoulou D, Katidou M, Moschonas N.K. (υπό συγγραφή). FRA10AC1 interacts with DGCR14, a spliceosomal component
70. Sarafidou T, Kahl C, Martinez-Garay I, Mangelsdorf M, Gesk S, Baker E, Kokkinaki M, Talley P, Maltby EL, French L, Harder L, Hinzmänn B, Nobile C, Richkind K, Finnis M, Deloukas P, Sutherland GR, Kutsche K, Moschonas NK, Siebert R, Géczy J; European Collaborative Consortium for the Study of ADLTE. "Folate-sensitive fragile site FRA10A is due to an expansion of a CGG repeat in a novel gene, FRA10AC1, encoding a nuclear protein." *Genomics*. 2004 Jul;84(1):69-81.

71. Sharp PA, Burge CB. "Classification of introns: U2-type or U12-type." *Cell*. 1997 Dec 26;91(7):875-9.
72. Simonis N, Rual JF, Carvunis AR, Tasan M, Lemmens I, Hirozane-Kishikawa T, Hao T, Sahalie JM, Venkatesan K, Gebreab F, Cevik S, Klitgord N, Fan C, Braun P, Li N, Ayivi-Guedehoussou N, Dann E, Bertin N, Szeto D, Dricot A, Yildirim MA, Lin C, de Smet AS, Kao HL, Simon C, Smolyar A, Ahn JS, Tewari M, Boxem M, Milstein S, Yu H, Dreze M, Vandenhoute J, Gunsalus KC, Cusick ME, Hill DE, Tavernier J, Roth FP, Vidal M. "Empirically controlled mapping of the *Caenorhabditis elegans* protein-protein interactome network." *Nat Methods*. 2009 Jan; 6(1):47-54.
73. Spartz AK, Herman RK, Shaw JE. "SMU-2 and SMU-1, *Caenorhabditis elegans* homologs of mammalian spliceosome-associated proteins RED and fSAP57, work together to affect splice site choice." *Mol Cell Biol*. 2004 Aug;24 (15):6811-23.
74. Stelzl U, Worm U, Lalowski M, Haenig C, Brembeck FH, Goehler H, Stroedicke M, Zenkner M, Schoenherr A, Koeppen S, Timm J, Mintzlaff S, Abraham C, Bock N, Kietzmann S, Goedde A, Toksöz E, Droege A, Krobitsch S, Korn B, Birchmeier W, Lehrach H, Wanker EE. "A human protein-protein interaction network: a resource for annotating the proteome." *Cell*. 2005 Sep 23;122(6):957-68.
75. Sutherland GR, Baker E. "The clinical significance of fragile sites on human chromosomes." *Clin Genet*. 2000 Sep;58(3):157-61.
76. Sutherland GR, Richards RI. "The molecular basis of fragile sites in human chromosomes." *Curr Opin Genet Dev*. 1995 Jun;5(3):323-7. Review.
77. Sutherland GR. "Heritable fragile sites on human chromosomes. VIII. Preliminary population cytogenetic data on the folic-acid-sensitive fragile sites." *Am J Hum Genet*. 1982 May;34(3):452-8.
78. Sutherland GR. "Heritable fragile sites on human chromosomes. XII. Population cytogenetics." *Ann Hum Genet*. 1985 May;49(Pt 2):153-61.
79. Suthram S, Shlomi T, Ruppin E, Sharan R, Ideker T. "A direct comparison of protein interaction confidence assignment schemes." *BMC Bioinformatics*. 2006 Jul 26; 7:360.
80. Takahashi E, Hori T, Murata M. "Population cytogenetics of rare fragile sites in Japan." *Hum Genet*. 1988 Feb; 78(2):121-6.
81. Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, Qureshi-Emili A, Li Y, Godwin B, Conover D, Kalbfleisch T, Vijayadamodar G, Yang M, Johnston M, Fields S, Rothberg JM. "A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*." *Nature*. 2000 Feb 10; 403(6770):623-7.
82. Ule J. "Ribonucleoprotein complexes in neurologic diseases." *Curr Opin Neurobiol*. 2008 Oct;18(5):516-23. Epub 2008 Oct 27. Review.
83. Verkerk AJ, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DP, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang FP, et al. "Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome." *Cell*. 1991 May 31;65(5):905-14.
84. Von Mering C, Krause R, Snel B, Cornell M, Oliver SG, Fields S, Bork P. "Comparative assessment of large-scale data sets of protein-protein interactions." *Nature*. 2002 May 23; 417(6887):399-403. Epub 2002 May 8.
85. Wang ET, Sandberg R, Luo S, Khrebtkova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. "Alternative isoform regulation in human tissue transcriptomes." *Nature*. 2008 Nov 27; 456(7221):470-6.
86. Wang S, Chen JZ, Zhang Z, Gu S, Ji C, Tang R, Ying K, Xie Y, Mao Y. "Cloning, expression and genomic structure of a novel human GNB2L1 gene, which encodes a receptor of activated protein kinase C (RACK)." *Mol Biol Rep*. 2003 Mar; 30(1):53-60.
87. Wang YH, Gellibolian R, Shimizu M, Wells RD, Griffith J. "Long CCG triplet repeat blocks exclude nucleosomes: a possible mechanism for the nature of fragile sites in chromosomes." *J Mol Biol*. 1996 Nov 8; 263(4):511-6.
88. Wang YH, Griffith J. "Methylation of expanded CCG triplet repeat DNA from fragile X syndrome patients enhances nucleosome exclusion." *J Biol Chem*. 1996 Sep 20; 271(38):22937-40.
89. Will CL, Lührmann R. "Spliceosomal UsnRNP biogenesis, structure and function." *Curr Opin Cell Biol*. 2001 Jun;13(3):290-301. Review.

90. Will CL, Lührmann R. "Splicing of a rare class of introns by the U12-dependent spliceosome". *Biol Chem*. 2005 Aug; 386(8):713-24.
91. Winnepeninckx B, Debacker K, Ramsay J, Smeets D, Smits A, FitzPatrick DR, Kooy RF. "CGG-repeat expansion in the DIP2B gene is associated with the fragile site FRA12A on chromosome 12q13.1." *Am J Hum Genet*. 2007 Feb;80(2):221-31. Epub 2006 Dec 12.
92. Xenarios I, Salwinski L, Duan XJ, Higney P, Kim SM, Eisenberg D. "DIP, the Database of Interacting Proteins: a research tool for studying cellular networks of protein interactions." *Nucleic Acids Res*. 2002 Jan 1; 30(1):303-5.
93. Yu J, Pacifico S, Liu G, Finley RL Jr. "DroID: the Drosophila Interactions Database, a comprehensive resource for annotated gene and protein interactions." *BMC Genomics*. 2008 Oct 7; 9:461.
94. Yu S, Mangelsdorf M, Hewett D, Hobson L, Baker E, Eyre HJ, Lapsys N, Le Paslier D, Doggett NA, Sutherland GR, Richards RI. "Human chromosomal fragile site FRA16B is an amplified AT-rich minisatellite repeat." *Cell*. 1997 Feb 7; 88(3):367-74.
95. Zhou Z, Licklider LJ, Gygi SP, Reed R. "Comprehensive proteomic analysis of the human spliceosome." *Nature*. 2002 Sep 12; 419(6903):182-5.
96. Zlotorynski E, Rahat A, Skaug J, Ben-Porat N, Ozeri E, Hershberg R, Levi A, Scherer SW, Margalit H, Kerem B. "Molecular basis for expression of common and rare fragile sites." *Mol Cell Biol*. 2003 Oct; 23(20):7143-51.

BIBΛΙΑ

- 📖 "Experiments in Molecular Genetics" J.H Miller, 1972
- 📖 "Ανασυνδυσσμένο DNA" (Watson, Caudy, Myers, Witkowski 2007)
- 📖 "Molecular Cloning" (Sambrook, Fritsch, Maniatis 1989)
- 📖 "Analysis of Genes and Genomes", Richard J Reed 2004)
- 📖 "Genes VII", Benjamin Lewin 2004

ΔΙΑΔΙΚΤΥΟ

- 🌐 <http://www.ncbi.nlm.nih.gov/gene>
- 🌐 <http://theoderich.fb3.mdc-berlin.de:8080/unihi/>
- 🌐 <http://www.droidb.org/IMBrowser.jsp>
- 🌐 http://interactome.dfci.harvard.edu/C_elegans/
- 🌐 www.hprd.org
- 🌐 <http://www.ncbi.nlm.nih.gov/homologene/>
- 🌐 <http://160.80.34.4/mint/Welcome.do>
- 🌐 <http://www.ebi.ac.uk/intact/>
- 🌐 <http://dip.doe-mbi.ucla.edu/dip/Main.cgi>
- 🌐 <http://thebiogrid.org/>