

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ –ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΚΑΙ
ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΠΙΣΤΗΜΩΝ**

**ΤΜΗΜΑ ΒΙΟΛΟΓΙΑΣ
&
ΙΝΣΤΙΤΟΥΤΟ ΜΟΡΙΑΚΗΣ ΒΙΟΛΟΓΙΑΣ ΚΑΙ
ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ**

Διδακτορική Διατριβή

**“Φυσική χαρτογράφηση υψηλής ευκρίνειας στη
χρωμοσωμική περιοχή 10q23 του ανθρώπου και
συμβολή στη λειτουργική ανάλυση του γονιδίου
Neuralized”**

Μαρία Κοκκινάκη

**Εργαστήριο Μοριακής Γενετικής Ανθρώπου
Επιβλέπων Καθηγητής: Ν. Κ. Μοσχονάς**

Ηράκλειο, Ιανουάριος 2005

**UNIVERSITY OF CRETE –SCHOOL OF SCIENCES
AND ENGINEERING**

**DEPARTMENT OF BIOLOGY
&
INSTITUTE OF MOLECULAR BIOLOGY AND
BIOTECHNOLOGY**

Ph. D Thesis

**“High resolution physical mapping of human
chromosomal region 10q23 and contribution to the
functional analysis of *Neuralized* gene”**

Maria Kokkinaki

**Human Molecular Genetics Lab
Advisor: Prof. N. K. Moschonas**

Heraklion, January 2005

Στους γονείς μου

«Και τον κόσμο αυτόν ανάγκη να τον βλέπεις και να τον λαβαίνεις»
Οδυσσέας Ελύτης, Άξιον Εστί

Επιβλέπων

Ν.Κ. Μοσχονάς (Καθηγητής)

Τριμελής Συμβουλευτική Επιτροπή

Ν.Κ. Μοσχονάς (Καθηγητής)

Δ. Αλεξανδράκη (Αν. Καθηγήτρια)

Χ. Λούης (Καθηγητής)

Επταμελής Εξεταστική Επιτροπή

Ν.Κ. Μοσχονάς (Καθηγητής)

Δ. Αλεξανδράκη (Αν. Καθηγήτρια)

Χ. Λούης (Καθηγητής)

Ι. Παπαματθαϊάκης (Καθηγητής)

Χ. Δελιδάκης (Αν. Καθηγητής)

Γ. Χαλεπάκης (Αν. Καθηγητής)

Δ. Καρδάσης (Επικ. Καθηγητής, Παν/μιο Κρήτης, Ιατρική Σχολή)

Τελειώνοντας το διδακτορικό μου, θέλω να ευχαριστήσω ειλικρινά τους ανθρώπους με τους οποίους μοιράστηκα αρκετές από τις εμπειρίες, ιδέες και γνώσεις, που περιγράφω στις σελίδες που ακολουθούν.

Ευχαριστώ τον επιβλέποντα καθηγητή μου Νίκο Μοσχονά για την απεριόριστη υπομονή του, τις συμβουλές και το ενδιαφέρον του σε όλη τη διάρκεια του διδακτορικού, από τα πρωταρχικά πειράματα μέχρι και την εκτύπωση της τελευταίας σελίδας της διατριβής. Περισσότερο τον ευχαριστώ για την ευκαιρία που μου έδωσε να γίνω μέλος της ομάδας του και για την εμπιστοσύνη που μου έδειχνε σε δύσκολες στιγμές.

Ευχαριστώ τον Αλέξανδρο Αργυροκαστρίτη και τη Μανουέλα Καφετάκη, παλαιότερα μέλη του Εργαστηρίου Μοριακής Γενετικής Ανθρώπου. Και οι δύο ήταν δάσκαλοί μου στο εργαστήριο στην περίοδο της διπλωματικής και μεταπτυχιακής μου εργασίας.

Ευχαριστώ την Αγγέλα Πασπαράκη, τεχνικό του εργαστηρίου μας κατά το μεγαλύτερο μέρος της διατριβής μου. Συμμαθήτριά μου από το Λύκειο και η πρώτη μου φίλη στο εργαστήριο. Εκτός όμως από την προσωπική μας φιλία, η φροντίδα της για το εργαστήριο, αλλά και η δουλειά της στη συντήρηση των κυτταροκαλλιεργειών ήταν, πιστεύω, σημαντική για τα αποτελέσματα αυτής της διατριβής.

Με τον Ηλία Παυλόπουλο αρχικά και με την Λία Κουτέλου στη συνέχεια, συνεργάστηκα σε μια πολύ δημιουργική προσπάθεια της ομάδας μας να κατανοήσουμε τη λειτουργία του γονιδίου *Neuralized* στα θηλαστικά. Η προσπάθεια αυτή συνεχίζεται και πέρα από αυτό το διδακτορικό, όμως ευχαριστώ και τους δύο για την άποψη συνεργασία μας και για το πολύ όμορφο συναίσθημα της συνεισφοράς στην αποκάλυψη του ρόλου της πρωτεΐνης *Neur1* σε βιολογικές διεργασίες, κρίσιμες για τη ζωή του κυττάρου. Ιδιαίτερα ευχαριστώ τη Λία για το έντονο ενδιαφέρον της σχετικά με την πρόοδο στη συγγραφή αυτής της διατριβής, αλλά και για την πολύ εποικοδομητική μας 'αλληλεπίδραση', που συνέπεσε με το τέλος του δικού μου διδακτορικού και το ξεκίνημα του δικού της.

Με τη Λιλίκα Σαραφίδου τυπικά συνεργάστηκα στο πρώτο μέρος της διατριβής, στα πλαίσια της κοινής προσπάθειας του εργαστηρίου μας για τη φυσική χαρτογράφηση της χρωμοσωμικής περιοχής 10q23-q26 του ανθρώπου, δουλειά που έθεσε τις βάσεις για την αλληλούχηση του χρωμοσώματος 10 και συνέβαλλε στην αποκάλυψη νέων γονιδίων. Ουσιαστικά όμως, η καθημερινή μας 'επικοινωνία' ήταν αναπόφευκτη και ευτυχώς ευχάριστη, εφόσον ήταν η 'διπλανή' μου στον πάγκο, με την οποία μοιράστηκα πολλά!

Στη διάρκεια του διδακτορικού είχα την τύχη να έχω δύο καλούς μαθητές, τη Δάφνη Μπαζοπούλου και τον Νίκο Τζιτζικα, στα πλαίσια της εκπόνησης των διπλωματικών τους εργασιών. Τους ευχαριστώ πολύ για τη συνεργασία μας και για όσα κι εγώ έμαθα στη διαδικασία της εκπαίδευσής τους. Επίσης, ευχαριστώ τον Σωτήρη Μπανάκο για τη συνεργασία μας στη διάρκεια του 'rotation' του στο εργαστήριό μας, αλλά και τα νεότερα μέλη του εργαστηρίου, τη Δέσποινα Αποστολοπούλου και τη Μαριάννα Φερετζάκη, οι οποίες αν και δεν συνεργαστήκαμε πειραματικά, εκδήλωναν συνεχώς το ενδιαφέρον τους για την πορεία του διδακτορικού μου.

Ένα πολύ καθοριστικό πείραμα του διδακτορικού, αυτό του διϋβριδικού συστήματος, πραγματοποιήθηκε στο Εργαστήριο Μοριακής Γενετικής Σακχαρομύκητα (IMBB-ITE), με την καθοδήγηση της καθηγήτριας Δέσποινας Αλεξανδράκη και του τότε μέλους του εργαστηρίου της Τάσσου Γεωργακόπουλου. Τους ευχαριστώ θερμά και τους δύο για τις συμβουλές και τη βοήθεια στη διάρκεια αυτού του πειράματος. Επίσης, ευχαριστώ την Αλεξάνδρα Βουσινά για τα όσα έμαθα κοντά της κατά τη διάρκεια του 'rotation' μου στο ίδιο εργαστήριο.

Ευχαριστώ τα μέλη της Τριμελούς μου Επιτροπής (Δ. Αλεξανδράκη και Χ. Λούη) για τη συμβολή τους στον καθορισμό της πορείας του διδακτορικού μου και τα μέλη της Επταμελούς μου Επιτροπής (Δ. Αλεξανδράκη, Χ. Λούη, Ι. Παπαμαθθαϊάκη, Χ. Δελιδάκη, Γ. Χαλεπάκη και Δ. Καρδάση) για τις χρήσιμες επιστημονικές τους κατά την εξέταση αυτής της διατριβής και το ενδιαφέρον τους για την όσο δυνατό καλύτερη παρουσίασή της.

Ευχαριστώ τη Σοφία Βρόντου και την αδελφή της Ελευθερία για τη φιλία τους και τις ώρες που μοιραστήκαμε συζητώντας, στο εργαστήριο ή εκτός, ποικιλία θεμάτων, προβληματισμούς αλλά και όνειρα για τη δουλειά και τη ζωή μας γενικότερα. Η επικοινωνία μου μαζί τους ήταν σημαντικό στήριγμα στην πορεία του διδακτορικού.

Ένα ευχαριστώ στους γονείς μου είναι πολύ λίγο... Στην ατέλειωτη αγάπη και στήριξη τους αφιερώνω ως ελάχιστο δώρο αυτή τη διατριβή.

Περιεχόμενα

ΠΕΡΙΛΗΨΗ

ABSTRACT

ΕΙΣΑΓΩΓΗ

1. Γενικά	1
2. Χαρτογράφηση και αλληλούχηση γονιδιωμάτων	3
α. Γενετικοί χάρτες	4
β. Φυσικοί χάρτες	6
γ. Αλληλούχηση του ανθρώπινου γονιδιώματος	7
3. Ταυτοποίηση γονιδίων	11
α. Αλληλούχηση κλώνων cDNA	
β. Πρόβλεψη γονιδίων με βιοϋπολογιστικές μεθόδους	
4. Οργάνωση γονιδιωματικής πληροφορίας	13
5. Συγκριτική γονιδιωματική ανάλυση	15
α. Συγκριτική ανάλυση γονιδιωμάτων ανθρώπου και ποντικού	17
6. Συγκριτική πρωτεωμική ανάλυση με στόχο τη διευκρίνιση της λειτουργίας	20
α. Λειτουργική σύγκριση πρωτεϊνών ανθρώπου και ποντικού	
β. Εξελικτική συντήρηση πρωτεϊνών ανθρώπου και ποντικού	
7. Το γονίδιο <i>Neuralized</i> ως μοντέλο μελέτης του βιοχημικού ρόλου γονιδίων άγνωστης λειτουργίας	26
α. Η αρχιτεκτονική της πρωτεΐνης <i>Neur1</i>	
8. Βιολογικός ρόλος του <i>neuralized</i> στη <i>Drosophila melanogaster</i> και στον <i>Xenopus laevis</i>	29
> <i>Neur1</i> και σηματοδότηση Notch	
α. Η πρωτεΐνη <i>neuralized</i> ρυθμίζει διαδικασίες πλευρικής αναστολής	31
β. Η πρωτεΐνη <i>neuralized</i> ρυθμίζει αποφάσεις κυτταρικής γενεαλογίας	33
γ. Το <i>neuralized</i> επάγει 'ασυμμετρία' πομπού-δέκτη για την ενίσχυση της σηματοδότησης Notch-Delta	34
> <i>Neur1</i> και ουβικουΐτινιλίωση	
α. Ουβικουΐτινιλίωση	35
Τροποποιήσεις ομόλογες της ουβικουΐτινιλίωσης	36
β. Η ουβικουΐτινιλίωση στη ρύθμιση της σηματοδότησης από τον υποδοχέα Notch	37
γ. Δομικά και λειτουργικά στοιχεία που υποστηρίζουν τη συμμετοχή του <i>neuralized</i> σε διαδικασίες ουβικουΐτινιλίωσης	38
9. Υποκυτταρικός εντοπισμός της πρωτεΐνης <i>Neuralized</i>	38
10. Λειτουργία του <i>Neuralized</i> στα θηλαστικά	40
α. Πρότυπο έκφρασης	40
β. Στοχευμένη απενεργοποίηση (<i>knock-out</i>) του <i>Neuralized</i> στον ποντικό	42
11. Αντικείμενο της παρούσας διατριβής	46

ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΙ

1. Βακτηριακά στελέχη	47
2. Στέλεχος <i>Sacharomyces cerevisiae</i> PJ69-4a	47
3. Κυτταρικές σειρές	48
4. Θρεπτικά μέσα	48
α. Θρεπτικά μέσα ανάπτυξης βακτηρίων	
β. Θρεπτικά μέσα ανάπτυξης βακτηριοφάγων	
γ. Θρεπτικά μέσα ανάπτυξης κυττάρων σακχαρομύκητα	
δ. Θρεπτικά μέσα ανάπτυξης κυτταρικών σειρών θηλαστικών	
5. Πλασμιδιακοί φορείς	49
6. Βιβλιοθήκες	53
α. Γονιδιωματική βιβλιοθήκη κλώνων PAC ανθρώπου	
β. Γονιδιωματική βιβλιοθήκη κλώνων BAC ανθρώπου	
γ. Γονιδιωματική βιβλιοθήκη ποντικού	
δ. cDNA βιβλιοθήκη εμβρύου ποντικού	
7. Ο κλώνος cDNA του γονιδίου <i>Neur1</i> του ποντικού	56
8. Ολιγονουκλεοτίδια	57
9. Ανιχνευτές γονιδίου <i>Neur1</i>	60
10. Αντισώματα	62
11. Διερεύνηση διατεταγμένων βιβλιοθηκών ανθρώπινου γονιδιωματικού DNA	63
➤ Διερεύνηση βιβλιοθήκης κλώνων PAC (RPCI-1)	
α. Παραγωγή και σήμανση ανιχνευτών	
β. Υβριδοποίηση	
γ. Ανάπτυξη κλώνων PAC	
δ. Ανάλυση κλώνων PAC	
➤ Διερεύνηση βιβλιοθήκης κλώνων BAC (RPCI-11)	
α. Παραγωγή και σήμανση ανιχνευτών	
β. Υβριδοποίηση	
γ. Ανάλυση κλώνων BAC	
12. Διερεύνηση βιβλιοθήκης βακτηριοφάγων	66
➤ Ανάπτυξη της βιβλιοθήκης σε στερεό θρεπτικό υπόστρωμα	
α. Παρασκευή κυττάρων δεκτικών για μόλυνση από βακτηριοφάγο λ	
β. Μόλυνση κυττάρων <i>E.coli</i> από βακτηριοφάγο λ	
➤ Μεταφορά των βακτηριοφάγων σε μεμβράνες υβριδοποίησης	
➤ Επιλογή θετικών κλώνων	
13. Απομόνωση και ανάλυση μορίων DNA	67
α. Απομόνωση DNA βακτηριοφάγων	
β. Απομόνωση πλασμιδιακού DNA από βακτήρια	
γ. Απομόνωση μίγματος γονιδιωματικού και πλασμιδιακού DNA από σακχαρομύκητα	
δ. Απομόνωση γονιδιωματικού DNA από εγκέφαλο ποντικού	
ε. Τεμαχισμός μορίων DNA με ενδονουκλεάσες περιορισμού	
στ. Ανάλυση γραμμικών μορίων DNA με ηλεκτροφόρηση σε πήκτωμα αγαρόζης	
- Ηλεκτροφόρηση παλλόμενου ηλεκτρικού πεδίου (PFGE)	
ζ. Ανάλυση DNA κατά Southern	
η. Αλυσιδωτή αντίδραση πολυμεράσης (Polymerase Chain Reaction, PCR)	
θ. Προσδιορισμός νουκλεοτιδικής αλληλουχίας μορίων DNA	
- Προσδιορισμός νουκλεοτιδικής αλληλουχίας με PCR (<i>cycle-sequencing</i>)	
- Αυτοματοποιημένη διαδικασία προσδιορισμού νουκλεοτιδικής αλληλουχίας με PCR	

14. Μέθοδοι ραδιοσήμανσης μορίων DNA	72
α. Επέκταση τυχαίων εκκινητών	
β. Σήμανση 5' άκρου	
γ. Σήμανση με PCR	
δ. Μέτρηση της ειδικής ενεργότητας ραδιοσημασμένου μορίου DNA	
15. Ανίχνευση τεμαχίων DNA με υβριδοποίηση	73
α. Υβριδοποίηση σε γονιδιωματικό DNA ποντικού	
β. Υβριδοποίηση σε κλωνοποιημένα τεμάχια DNA	
16. Κλωνοποίηση τεμαχίων DNA σε πλασμιδία	74
α. Απομόνωση θραύσματος DNA από πήκτωμα αγαρόζης	
β. Κατασκευή τυφλών άκρων σε γραμμικά μόρια DNA με προεξέχοντα 5' άκρα	
γ. Σύνδεση γραμμικών μορίων DNA (<i>ligation</i>)	
δ. Μετασχηματισμός βακτηρίων με υπερελικωμένο πλασμιδιακό DNA	
ε. Ταυτοποίηση ανασυνδυασμένων βακτηριακών κλώνων	
στ. Μετασχηματισμός κυττάρων σακχαρομύκητα ρJ69-4a με πλασμιδιακό DNA	
ζ. Διαχωρισμός πλασμιδίου pVP16 από το πλασμιδίο pGBT9	
17. Σύστημα δύο υβριδίων σακχαρομύκητα (<i>yeast two hybrid system</i>)	79
α. Ανάπτυξη του στελέχους σακχαρομύκητα ρJ69-4a	
β. Επιλογή αλληλεπιδρώντων	
γ. Πειράματα αρνητικού ελέγχου	
18. Έκφραση, ανάλυση και ανίχνευση πρωτεϊνών	82
α. Παροδική διαμόλυνση κυτταρικών σειρών με πλασμιδιακές κατασκευές	
β. Ανίχνευση πρωτεΐνης σε κύτταρα θηλαστικών με <i>in situ</i> ανοσοφθορισμό	
γ. Απομόνωση ολικού πρωτεϊνικού εκχυλίσματος από καλλιέργεια κυττάρων θηλαστικών	
δ. Παραγωγή πρωτεΐνης σε βακτήρια με επαγωγή	
ε. Ηλεκτροφορητικός διαχωρισμός πρωτεϊνικού εκχυλίσματος	
στ. Ανάλυση 'Western blot'	
ζ. Ανοσοανίχνευση πρωτεϊνών μονιμοποιημένων σε μεμβράνες	
η. Καθαρισμός της υπερεκφρασμένης πρωτεΐνης με χρωματογραφία συγγένειας	
θ. <i>In vitro</i> δοκιμή αλληλεπίδρασης πρωτεΐνης με πρωτεΐνη συντηγμένη με GST, σε χρωματογραφία συγγένειας (<i>GST-pulldown</i>)	
ι. Συν-ανοσοκατακρήμνιση (<i>co-immunoprecipitation</i>)	
19. Ανάλυση αλληλουχιών DNA και πρωτεϊνών <i>in silico</i>	89
α. Ανάκτηση αλληλουχιών DNA ή πρωτεϊνών από τη βάση δεδομένων GenBank	
β. Πηγές δεδομένων αλληλούχησης γονιδιωμάτων (με ελεύθερη πρόσβαση)	
γ. Σύγκριση αλληλουχιών DNA ή πρωτεϊνών με τις αλληλουχίες της GenBank	
δ. Στοιχισμός αλληλουχιών DNA ή πρωτεϊνών και κατασκευή εξελικτικών δέντρων	
ε. Μετάφραση αλληλουχιών DNA	
στ. Ταυτοποίηση θέσεων αναγνώρισης περιοριστικών ενζύμων	
ζ. Ταυτοποίηση θέσεων SUMΟλίωσης	

ΠΛΑΣΜΙΔΙΑΚΕΣ ΚΑΤΑΣΚΕΥΕΣ

1. Υποκλωνοποιήσεις γονιδιωματικών κλώνων	91
2. Κλωνοποίηση του <i>Neur1</i> cDNA σε φορείς έκφρασης	92
3. Κλωνοποίηση των cDNA-΄θηραμάτων΄ που προέκυψαν από το πείραμα δύο υβριδίων σε φορείς έκφρασης	95
4. Κλωνοποίηση του cDNA του γονιδίου <i>Sumo-1</i>	96
5. Κλωνοποίηση του cDNA του γονιδίου <i>Dlc-1</i>	96

ΑΠΟΤΕΛΕΣΜΑΤΑ

1. Συμβολή στη δημιουργία του φυσικού χάρτη του χρωμοσώματος 10	98
α. Επιλογή κλώνων PAC	98
β. Φυσικός χάρτης των κλώνων YAC και PAC στην περιοχή του D10S2311	102
γ. Επιλογή κλώνων BAC	103
δ. Φυσικός χάρτης κλώνων BAC και αλληλούχηση της περιοχής D10S541-D10S583	105
2. Ανάλυση γονιδιωματικής οργάνωσης του γονιδίου <i>Neuralized (Neurl)</i> του ποντικού	107
α. Χαρακτηρισμός των επιλεγμένων γονιδιωματικών κλώνων	107
β. Ανάλυση γονιδιωματικού DNA	110
γ. Προσδιορισμός ορίων εξονίων – ιντρονίων	112
δ. Αλληλούχηση των περιοχών 5'UTR και 3'UTR	115
ε. Ανάλυση της γονιδιωματικής οργάνωσης του <i>Neurl in silico</i>	117
στ. Σύγκριση γονιδιωματικής οργάνωσης των γονιδίων <i>Neurl</i> και <i>NEURL</i>	123
η. Σύγκριση γονιδιωματικής αλληλουχίας ανθρώπου και ποντικού στις περιοχές των γονιδίων <i>Neurl</i> και <i>NEURL</i>	124
3. Διερεύνηση πρωτεϊνικών αλληλεπιδράσεων του <i>Neurl</i> με το σύστημα 'δύο υβριδίων' στον σακχαρομύκητα	127
α. Διερεύνηση της βιβλιοθήκης cDNA με 'δόλωμα' το πλήρες <i>Neurl</i>	128
β. Ανάλυση νουκλεοτιδικής σύστασης των κλώνων <i>cDNA-prey/pVP16</i>	129
γ. Ομάδες επικαλυπτόμενων κλώνων	131
δ. Μοναδιαίοι κλώνοι	131
ε. 3 ^ο στάδιο επιλογής	133
στ. Έλεγχος αλληλεπίδρασης των επιλεγμένων πεπτιδίων με τα ελλειμματικά μόρια <i>Neurl(1-521)</i> και <i>Neurl(515-574)</i>	134
ζ. Επιλογή κλώνων για περαιτέρω ανάλυση	134
4. Έλεγχος αλληλεπιδράσεων της πρωτεΐνης <i>Neurl in vitro</i>	136
α. Έλεγχος αλληλεπιδράσεων του <i>Neurl</i> πλήρους μήκους	137
β. Έλεγχος αλληλεπίδρασης των ελλειμματικών μορίων <i>Neurl</i>	137
5. Διερεύνηση των πρωτεϊνικών αλληλεπιδράσεων του <i>Neurl</i> με οπτική μικροσκοπία	140
α. Εντοπισμός του <i>Neurl</i> σε διάφορες φάσεις της κυτταρικής διαίρεσης	141
β. Πυρηνικός εντοπισμός της πρωτεΐνης <i>Neurl</i>	142
γ. Συνεντοπισμός των πρωτεϊνών <i>Neurl</i> και <i>Ubc9</i>	143
δ. Συνεντοπισμός των πρωτεϊνών <i>Neurl</i> και <i>Dlc-1</i>	143
ε. Συνεντοπισμός των πρωτεϊνών <i>Neurl</i> και <i>Numa1</i>	144
στ. Συνεντοπισμός των πρωτεϊνών <i>Neurl</i> και <i>Dnmt3a</i>	145
ζ. Συνεντοπισμός <i>Neurl</i> και <i>SUMO-1</i>	145
6. Έλεγχος και χαρτογράφηση αλληλεπιδράσεων του <i>Neurl</i> σε κύτταρα θηλαστικών	146
α. Συνέκφραση της πρωτεΐνης <i>Neurl</i> και των πρωτεϊνών- 'θηραμάτων' σε κύτταρα COS7	147
β. Οι πρωτεΐνες- 'θηράματα' αλληλεπιδρούν με το πλήρες μήκους <i>Neurl</i>	148
γ. Οι ενδογενείς πρωτεΐνες <i>Neurl</i> και <i>Ubc9</i> αλληλεπιδρούν στα κύτταρα θηλαστικών	150
δ. Χαρτογράφηση των αλληλεπιδράσεων σε περιοχές της πρωτεΐνης <i>Neurl</i>	151
ε. Περαιτέρω ανάλυση των περιοχών του <i>Neurl</i> που αλληλεπιδρούν με <i>Ubc9</i>	153

ΣΥΖΗΤΗΣΗ

1. Το χρωμόσωμα 10 του ανθρώπου	156
α. Τα γονίδια της περιοχής D10S541-D10S583 (10q23.31-q23.33)	158
β. Εξελικτική συντήρηση των γονιδίων της περιοχής	161
2. Συντήρηση της γονιδιωματικής οργάνωσης του <i>Neurl</i>	164
3. Εξελικτική ιστορία της πρωτεΐνης Neuralized	166
4. Πρωτεΐνες που αλληλεπιδρούν με το <i>Neurl</i>	172
> Πρωτεΐνες που αλληλεπιδρούν με τον κυτταροσκελετό ή αποτελούν συστατικά ινιδίων	
α. Numa1	173
β. Dlc-1	177
γ. Growth Arrest Specific protein 8	181
δ. Envorlakín και Ankycorbin	182
> Πρωτεΐνες που σχετίζονται με διαδικασίες κυτταρικού πολλαπλασιασμού ή/και προγραμματισμένου κυτταρικού θανάτου	
ε. Lek1	183
στ. Peg3	184
> Πρωτεΐνες που συμμετέχουν ή υπόκεινται σε μεταμεταφραστικές τροποποιήσεις από τον ενζυμικό μηχανισμό ομοιοπολικής σύνδεσης μορίων SUMO	
ζ. Dnmt3a	187
η. Ubc9	188
5. Ο ρόλος του RING στις αλληλεπιδράσεις με E2	189
6. Ο βιολογικός ρόλος του <i>Neuralized</i> στα θηλαστικά – προτεινόμενα μοντέλα	191
7. Συμπεράσματα και προοπτικές	196

ΒΙΒΛΙΟΓΡΑΦΙΑ

197-210

ΔΗΜΟΣΙΕΥΣΕΙΣ

ΠΕΡΙΛΗΨΗ

Η παρούσα διατριβή αποτελείται από δύο μέρη. Στο πρώτο, περιγράφουμε την κατασκευή ενός υψηλής ευκρίνειας φυσικού χάρτη με στόχο την αλληλούχηση της πλούσιας σε γονίδια περιοχής q23.31-q23.33 του χρωμοσώματος 10 του ανθρώπου και αξιοποιούμε την πληροφορία που προκύπτει από την *in silico* ανάλυση της αντίστοιχης νουκλεοτιδικής αλληλουχίας. Στο δεύτερο, παρουσιάζουμε τη μελέτη δομικών, εξελικτικών και λειτουργικών χαρακτηριστικών του γονιδίου *Neuralized (Neurl)* του ποντικού, ορθόλογου του ανθρώπινου *NEURL*, που χαρτογραφείται πλησίον της παραπάνω χρωμοσωμικής περιοχής.

Χρησιμοποιώντας ως ανιχνευτές 36 STSs (Sequence Tag Sites), που χαρτογραφούνται στην περιοχή 10q23.31-q23.33, απομονώσαμε 22 και 316 χρωμοσωμικούς κλώνους PAC και BAC, αντίστοιχα. Ο δομικός χαρακτηρισμός των κλώνων έγινε με σύγκριση του προτύπου τεμαχισμού τους με διάφορα περιοριστικά ένζυμα και έτσι δημιουργήθηκε ένα ενιαίο, ιδιαίτερα αξιόπιστο contig κλώνων μήκους 4.38 Mb, που ορίζεται από τους μικροδορυφορικούς γενετικούς δείκτες D10S451 και D10S583, που τοποθετούνται εγγύς και άπω του κεντρομερούς, αντίστοιχα. Με βάση αυτό το contig, επιλέχθηκαν από το συνεργαζόμενο εργαστήριο του *Sanger Center* 35 κλώνοι BAC, οι οποίοι διατρέχουν τη συγκεκριμένη χρωμοσωμική περιοχή και έχουν τη μικρότερη δυνατή επικάλυψη. Η νουκλεοτιδική αλληλουχία που προέκυψε από την ανάλυση αυτών των κλώνων είναι ενιαία καθ' όλο το μήκος του contig. Όπως προσδιορίστηκε μετά από συστηματική ανάλυση *in silico*, η περιοχή που περικλείεται από τους ανωτέρω δείκτες περιέχει 51 γονίδια, 35 από τα οποία είναι άγνωστης λειτουργίας. Από τα τελευταία, 11 κωδικοποιούν πολυπεπίδια που φέρουν χαρακτηριστικές περιοχές ομολογίας με πρωτεΐνες γνωστής λειτουργίας.

Πρόσφατες αναφορές μιας σειράς ερευνητικών ομάδων, συμπεριλαμβανομένης και της δικής μας, έχουν προτείνει ότι η πρωτεΐνη *neuralized* της δροσόφιλας συμμετέχει στη σηματοδότηση Notch και ότι η ορθόλογη πρωτεΐνη *Neurl* του ποντικού πιθανόν να έχει ρυθμιστικό ρόλο σε διάφορες κυτταρικές λειτουργίες. Με σκοπό τον περαιτέρω χαρακτηρισμό του *Neurl*, απομονώσαμε 17 κλώνους βακτηριοφάγου από μια γονιδιωματική βιβλιοθήκη ποντικού (στέλεχος 129SV) και προσδιορίσαμε τη χρωμοσωμική οργάνωση του γονιδίου. Με βάση τα πειραματικά μας αποτελέσματα σε συνδυασμό με πληροφορίες από διάφορες βάσεις δεδομένων, το γονίδιο εκτείνεται σε 79 Kb χρωμοσωμικής αλληλουχίας, αποτελείται από 7 εξόνια και φαίνεται ότι ελέγχεται από δύο διαφορετικούς, εναλλακτικά χρησιμοποιούμενους υποκινητές, ο ένας από τους οποίους βρίσκεται εντός του πρώτου ιντρονίου, μήκους 47.3Kb. Το γονίδιο κωδικοποιεί δύο πολυπεπίδια *Neurl*, μήκους 574 και 557 αμινοξέων, που διαφέρουν στο αμινοτελικό τους άκρο, λόγω της χρήσης των εξονίων 1 και 1a από τους δύο εναλλακτικούς υποκινητές. Λεπτομερής συγκριτική ανάλυση στο νουκλεοτιδικό επίπεδο έδειξε ότι η χρωμοσωμική οργάνωση του γονιδίου είναι παρόμοια στα έντομα και στα θηλαστικά και επισήμανε πιθανά ρυθμιστικά στοιχεία του γονιδίου των θηλαστικών. Αντίστοιχα, από τη συστηματική σύγκριση της αμινοξικής αλληλουχίας των ορθόλογων πρωτεϊνών μιας σειράς οργανισμών (*C. elegans*, *D. melanogaster*, *A. gambiae*, *F. rubripes*, *D. rerio*, *X. Laevis*, *R. norvegicus*) αποκαλύφθηκε ότι η οικογένεια των πρωτεϊνών *neuralized* είναι αρκετά συντηρημένη στους ζωικούς οργανισμούς από τον νηματώδη και τα έντομα έως τα θηλαστικά, εμφανίζοντας υψηλότερη συντήρηση στις χαρακτηριστικές περιοχές NEUZ1, 2 και RING. Συγκεκριμένα, το NEUZ1 είναι η πιο συντηρημένη περιοχή του *Neurl* στα σπονδυλωτά, ενώ το RING είναι η πιο συντηρημένη περιοχή μεταξύ σπονδυλωτών και ασπόνδυλων. Το τελευταίο, πιθανόν, υποδηλώνει

την καθοριστική σημασία του RING στη λειτουργία της πρωτεΐνης και τον παρόμοιο βιοχημικό ρόλο του neutralized μεταξύ των ειδών.

Με στόχο την περαιτέρω διαλεύκανση του βιολογικού ρόλου του Neu1, χρησιμοποιήσαμε το σύστημα δύο υβριδίων του σακχαρομύκητα για την ταυτοποίηση πρωτεϊνών που αλληλεπιδρούν άμεσα με αυτό. Η διερεύνηση απέδωσε εννέα λιγότερο ή περισσότερο καλά χαρακτηρισμένες πρωτεΐνες, που συμμετέχουν σε σημαντικές λειτουργίες του κυττάρου, ήτοι στην ενδοκυττάρια κίνηση μέσω πρωτεϊνών-κινητήρων, στην κίνηση ή σταθερότητα του κυττάρου (Dlc-1, Gas8, Enp1 και Ankycorbin), στην κυτταρική διαίρεση ή απόπτωση (Numa1, Lek1, Dlc-1, Peg3), σε μετα-μεταφραστικές τροποποιήσεις πρωτεϊνών (Ubc9) και στη μεταγραφική ρύθμιση (Dnmt3a). Οι αλληλεπιδράσεις επιβεβαιώθηκαν περαιτέρω με δοκιμές πρωτεϊνικής αλληλεπίδρασης *in vitro* και μετά από παροδική συνέκφραση σε κύτταρα θηλαστικών, εφαρμόζοντας τις μεθόδους *GST-pull-down* και συν-ανοσοκατακρήμνισης (co-IP), αντίστοιχα. Πρόσθετες ενδείξεις που υποστηρίζουν αυτές τις αλληλεπιδράσεις προέκυψαν από: i. τον προσδιορισμό των επικαλύψεων στο πρότυπο έκφρασης των αντίστοιχων γονιδίων, με πειράματα υβριδοποίησης RNA *in situ* σε τομές εμβρύων ποντικού, και ii. την ανίχνευση του υποκυτταρικού συνεντοπισμού των ενδογενών ή/και παροδικά εκφραζόμενων ανασυνδυασμένων πρωτεϊνών, με παρατήρηση σε συνεστιακό μικροσκόπιο. Όπως έδειξαν πειράματα co-IP, στα οποία χρησιμοποιήθηκαν αλληλοσυμπληρούμενα ελλεμματικά πολυπεπίδια Neu1, οι αλληλεπιδράσεις του με τις άλλες πρωτεΐνες επιτυγχάνονται γενικά, μέσω των περιοχών NEUZ1 και NEUZ2, με ορισμένες όμως, διαφοροποιήσεις. Συγκεκριμένα, το NEUZ1 είναι απόλυτα αναγκαίο και ικανό για την αλληλεπίδραση του Neu1 με τις πρωτεΐνες Dlc-1 και Gas8 αλλά όχι με την Ubc9, ενώ το RING είναι απαραίτητο για τη βέλτιστη πρόσδεση του Neu1 στο Ubc9. Με βάση τα αποτελέσματά μας και σε συμφωνία με τον γενικό χαρακτηρισμό που αποδίδεται σε πρωτεΐνες που διαθέτουν περιοχή RING, προτείνουμε για το Neu1 τον ρόλο του 'μεσολαβητή', δηλαδή ενός ενδιάμεσου μορίου που φέρνει επιλεκτικά σε επαφή διάφορες πρωτεΐνες μεταξύ τους, συμμετέχοντας με αυτόν τον τρόπο σε λειτουργικά πολυ-πρωτεϊνικά σύμπλοκα. Έτσι, η αλληλεπίδρασή του με την Dlc-1 πιθανόν να εξηγεί τη αποδιοργανωμένη δομή του αξονήματος στα σπερματοζώαρια των αρσενικών ποντικών Neu1^{-/-}. Επιπλέον, η αλληλεπίδρασή του με το Ubc9 είναι πιθανόν να υποδηλώνει είτε τη μετα-μεταφραστική τροποποίησή του Neu1 με το πεπίδιο SUMO (Small Ubiquitin-like Modifier), ή τον ρόλο του ως λιγάση E3 σε διαδικασίες SUMΟλίωσης άλλων πρωτεϊνών. Αξίζει να σημειωθεί ότι αρκετές από τις πρωτεΐνες που αλληλεπιδρούν με το Neu1 εμφανίζουν κατάλληλες θέσεις για SUMΟλίωσης και θα μπορούσαν να θεωρηθούν ως υποψήφια υποστρώματα. Πράγματι, για μία από αυτές, την Dnmt3a, έχει ήδη επιβεβαιωθεί η τροποποίησή της με SUMO. Συμπερασματικά, θεωρούμε ότι η ταυτοποίηση των πρωτεϊνών που αλληλεπιδρούν με το Neu1, έχει θεμελιώδη σημασία για την αποσαφήνιση του βιολογικού του ρόλου και αποτελεί τη βάση για τον σχεδιασμό μιας μακράς σειράς σημαντικών *in vitro* και *in vivo* πειραμάτων για την περαιτέρω λειτουργική ανάλυση της πρωτεΐνης.

ABSTRACT

This thesis consists of two parts. In the first, we describe the construction of a high resolution, sequence-ready physical map of the gene-rich q23.31-q23.33 region of human chromosome 10 and evaluate the information derived from the *in silico* analysis of the corresponding DNA sequence. In the second part, we present our work on structural, evolutionary and functional characteristics of the murine *Neuralized* (*Neurl*) gene, the human *NEURL* ortholog located adjacently to the above region.

By using as probes 36 STSs (Sequence Tag Sites) mapped to 10q23.31-q23.33, we isolated 22 PAC and 316 BAC chromosomal clones and proceeded to their structural analysis by means of comparative DNA fingerprinting. Our data led to the construction of a well-supported, sequence-ready unique contig of about 4.38 Mb delimited by the microsatellite markers D10S541 and D10S583, located proximal and distal to the centromere, respectively. The contig was used by the collaborating laboratory of Sanger Center, for the selection of a tiling path of 35 BACs, which were then employed in DNA sequence determination. The corresponding finished DNA sequence is without gaps and, as revealed by extensive *in silico* analysis, contains 51 genes, 35 of which are of unknown function. However, 11 of the latter encode polypeptides with distinctive domains present in proteins of known function.

Recent studies from a number of groups including ours, have suggested that the drosophila neuralized (*neur*) participates in Notch signaling and that the mouse ortholog, *Neurl*, may play a regulatory role in various cellular functions. In order to further characterize *Neurl*, we isolated 17 phage clones from a mouse strain 129SV genomic library and determined the chromosomal organization of the gene. In combination with information from various databases, our experimental data suggested that *Neurl* spans 79 Kb of genomic DNA, consists of 7 exons and appears to be under the control of two distinct, alternatively used promoters, one of which is located in the 47.3 Kb long, intron 1. The gene encodes two *Neurl* polypeptides, 574 and 557 aa long, differing at their amino-terminus, as a result of the alternative use of exon 1 and the downstream exon 1a. Extensive comparative analysis at the nucleotide level suggested that the chromosomal organization of the gene is quite similar in insects and mammals and demonstrated possible *Neurl* regulatory elements in mammals. Moreover, systematic protein sequence comparisons among neuralized orthologs from several organisms, i.e.: *C. elegans*, *D. melanogaster*, *A. gambiae*, *F. rubripes*, *D. rerio*, *X. laevis*, *R. norvegicus*, revealed that the "neuralized" protein family is well-conserved in multicellular animal species from the worm and insects to mammals, prominently at the characteristic NEUZ 1, 2 and RING domains. In particular, NEUZ1 represents the most conserved domain among vertebrates whereas the RING is the most conserved domain between vertebrate and invertebrate species. The latter, possibly suggests similar biochemical function of the protein in all species and demonstrates the central role of the RING domain.

In order to further elucidate the biological role of *Neurl*, we employed the yeast two-hybrid system to identify proteins that directly interact with it. The screening resulted in the identification of 9 proteins known to participate in significant cell functions, i.e: motor-based intracellular transport, cell movement or stability (Dlc-1, Gas8, Evpl and ankyrin), cell division or apoptosis (Numa1, Lek1, Dlc-1, Peg3), post-translational protein modification (Ubc9) and transcriptional regulation (Dnmt3a). Protein interactions were

further verified by *in vitro* and cell culture experiments, using GST pull-downs and co-immunoprecipitation (co-IP) assays, respectively. Supporting evidence for these interactions, also, derived from: i. demonstrating overlapping expression patterns of the corresponding genes in mouse embryo sections, using comparative RNA *in situ* hybridization, and ii. detecting subcellular co-localization of the respective endogenous and/or transiently expressed recombinant proteins, using confocal microscopy. As shown by co-IP assays where a number of complementarily truncated polypeptides were employed, Neurl-protein interactions are in general mediated by NEUZ1 and NEUZ2. However, a number of differences exist. NEUZ1 is absolutely necessary and sufficient for binding with Dlc-1 and Gas8 but not with Ubc9, whereas the RING domain is required for adequate Neurl binding with Ubc9. In accordance to the general notion about RING containing proteins, our results suggest that Neurl may function as an "adaptor" molecule bringing, selectively, together various proteins and thus, participating as an essential component, in multi-protein cellular complexes. Accordingly, its interaction with Dlc-1 may explain the axonemal defects manifested in Neurl^{-/-} male mice. In addition, its interaction with Ubc9 may imply either its post-translational modification by SUMO (Small Ubiquitin-like Modifier) or its role as a SUMO E3 ligase. Interestingly, several of the Neurl interactors display appropriate sites for sumoylation and thus could be considered as candidate substrates. In fact one of them, Dnmt3a, is already known to be subjected in sumoylation. Conclusively, we believe that the identification of Neurl interacting proteins is fundamental towards the elucidation of its biological role and motivates the design of a long series of interesting *in vitro* and *in vivo* experiments for its further functional analysis.

ΕΙΣΑΓΩΓΗ

1. Γενικά

Η ανακάλυψη του *Oswald Avery* ότι το DNA είναι η χημική ουσία που φέρει και μεταβιβάζει τη γενετική πληροφορία έστρεψε το ενδιαφέρον των ερευνητών σε αυτό το μόριο σηματοδοτώντας το 1944 την εποχή της γονιδιωματικής ανάλυσης. Η συστηματική διερεύνηση της δομής και της λειτουργίας του γενετικού υλικού οδήγησε και σε άλλες θεμελιώδεις ανακαλύψεις, όπως αυτή της δομής της διπλής έλικας DNA [Watson & Crick, 1953] και των βασικών αρχών μεταγραφής του DNA σε RNA και σύνθεσης των πρωτεϊνών [Nirenberg *et al*, 1966 / Sharp & Matassi, 1994]. Η διαπίστωση ότι η γενετική πληροφορία βρίσκεται κωδικοποιημένη στην αλληλουχία των νουκλεοτιδικών βάσεων του DNA οδήγησε πολλές ερευνητικές ομάδες προς το φιλόδοξο εγχείρημα του προσδιορισμού αυτής της νουκλεοτιδικής αλληλουχίας για το γονιδίωμα του ανθρώπου και άλλων οργανισμών. Ιδιαίτερα για τον άνθρωπο, η ανάλυση του γονιδιώματός του ήταν απαραίτητη για την αποκάλυψη της μοριακής βάσης πληθώρας γενετικών νόσων (αναπτυξιακών ανωμαλιών, διαφόρων μορφών καρκίνου, νευρολογικών διαταραχών κλπ), των οποίων η συσχέτιση του φαινοτύπου με μοριακά χαρακτηριστικά δεν ήταν εμφανής.

Το ανθρώπινο γενετικό υλικό (3×10^9 bp ανά απλοειδές γονιδίωμα) είναι οργανωμένο σε 23 ζεύγη χρωμοσωμάτων, τα μεγέθη των οποίων κυμαίνονται από 50Mb για το χρωμόσωμα 21 έως 250Mb για το χρωμόσωμα 1. Τα περίπου 30.000 γονίδια του ανθρώπου είναι άνισα καταμεμημένα στα χρωμοσώματα. Η ανομοιόμορφη αυτή κατανομή φαίνεται από τα πρόσφατα δεδομένα αλληλούχησης [Lander *et al*, 2001] και την επεξεργασία της νουκλεοτιδικής αλληλουχίας του γονιδιώματος από τα προγράμματα γονιδιακών προβλέψεων, αλλά είχε διαπιστωθεί και από μια σειρά προηγούμενων ενδείξεων. Για παράδειγμα, μεγαλύτερη συγκέντρωση περιοχών που σχετίζονται με γενετικές ασθένειες παρατηρείται στα χρωμοσώματα X, 17 και 19, ενώ τα χρωμοσώματα Y, 4, 13 και 18 είναι τα πιο φτωχά σε γενετικούς τόπους που σχετίζονται με κλινικούς φαινοτύπους, σύμφωνα με τα στοιχεία της βάσης δεδομένων OMIM (*Online Mendelian Inheritance in Man*, www.ncbi.nlm.nih.gov/Omim/). Επιπλέον, πειράματα διαφορικής χρώσης των πλούσιων σε νουκλεοτίδια AT και GC χρωμοσωμικών περιοχών, έχουν δείξει ότι τα χρωμοσώματα 1, 9, 15, 16, 17, 19, 20 και 22 είναι τα πιο πλούσια σε AT και σε νησίδα CpG, που βρίσκονται κοντά στα περισσότερα ανθρώπινα γονίδια [Craig & Bickmore, 1994]. Αντίθετα, τα χρωμοσώματα 13, 18 και Y παρουσιάζονται εξαιρετικά φτωχά σε γονίδια. Όμως, και η κατανομή των γονιδίων κατά μήκος του ίδιου χρωμοσώματος είναι ανομοιόμορφη. Ορισμένες χρωμοσωμικές ζώνες συγκεντρώνουν το μεγαλύτερο ποσοστό των γονιδίων του χρωμοσώματος. Χαρακτηριστικό είναι το παράδειγμα της τελομερικής ζώνης 21q22.3, η οποία φέρει περισσότερα από το 45% των γονιδίων του χρωμοσώματος 21 ενώ αντιστοιχεί μόνο στο 20% του μεγέθους του μεγάλου του βραχίονα [Antonarakis, 1994].

Το διεθνές πρόγραμμα ανάλυσης του ανθρώπινου γονιδιώματος έχει σήμερα ολοκληρώσει την αλληλούχηση της ευχρωματινής των χρωμοσωμάτων του ανθρώπου. Παράλληλα με το ανθρώπινο γονιδίωμα, έχουν αναλυθεί γονιδιώματα και άλλων θηλαστικών (*Mus musculus*, *Rattus norvegicus*, *Pan troglodytes*, *Bos taurus*, *Sus scrofa*), του πτηνού *Gallus gallus*, του αμφιβίου *Xenopus laevis*, των ψαριών *Danio rerio* και *Fugu rubripes*, των εντόμων *Drosophila melanogaster* και *Anopheles gambiae*, του νηματώδη *Caenorhabditis elegans*, ορισμένων φυτών (*Arabidopsis thaliana*, *Medicago truncatula*, *Oryza sativa*), μυκήτων (*Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*) και περισσότερων των διακοσίων ειδών βακτηρίων και ιών. Τα αποτελέσματα της αλληλούχησης των γονιδιωμάτων είναι διαθέσιμα σε ηλεκτρονική μορφή από τους κόμβους βιοπληροφορικής *Ensembl* (www.ensembl.org) και *NCBI* (*National Center for Biotechnology Information*, www.ncbi.nlm.gov, βάση δεδομένων *RefSeq*), όπου παρέχεται η δυνατότητα ελεύθερης πρόσβασης στις νουκλεοτιδικές αλληλουχίες, στα υπολογιστικά προγράμματα ανάλυσης και παρουσίασης των αλληλουχιών, πληροφορίες για τη δομή και τη λειτουργία κάθε προβλεπόμενου ή γνωστού γονιδίου και οι σχετικές βιβλιογραφικές αναφορές.

Η πρόοδος στην ανάλυση των γονιδιωμάτων μεγάλης ποικιλίας οργανισμών συνέβαλλε στην ανάπτυξη του πεδίου της 'συγκριτικής γονιδιωματικής ανάλυσης' (*comparative genomics*). Η σύγκριση των γονιδιωμάτων διαφορετικών οργανισμών, εκτός από τα συμπεράσματα για την εξέλιξη των ειδών, μπορεί να συμβάλλει και στον προσδιορισμό των συντηρημένων περιοχών του DNA, με ιδιαίτερη λειτουργική ή ρυθμιστική σημασία, διαδραματίζοντας με αυτόν τον τρόπο καθοριστικό ρόλο στο αναπτυσσόμενο πεδίο της 'λειτουργικής γονιδιωματικής ανάλυσης' (*functional genomics*). Οργανισμοί όπως ο ποντικός (*Mus musculus*) και η δροσόφιλα (*D. melanogaster*), κλασικά πειραματόζωα με μεγάλο αριθμό φυσικών ή τεχνητών μεταλλάγων, αναγνωρίσιμων φαινοτύπων και τεχνικών γενετικής τροποποίησης, αποτελούν πολύτιμα πειραματικά μοντέλα για τη μελέτη της λειτουργίας ανθρώπινων γονιδίων. Επιπλέον, μονοκύτταροι ή ολιγοκύτταροι οργανισμοί όπως η ζύμη (*Saccharomyces cerevisiae*) και ο νηματώδης (*Caenorhabditis elegans*) προσφέρονται ως απλούστερα συστήματα για τη μελέτη βασικών κυτταρικών λειτουργιών ή μονοπατιών κυτταρικής σηματοδότησης.

Η ολοκλήρωση της αλληλούχησης των γονιδιωμάτων των *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster*, *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, σε συνδυασμό με τα μέχρι στιγμής αποτελέσματα από την ανάλυση των γονιδιωμάτων άλλων οργανισμών, δεν έχει φέρει επανάσταση μόνο στον τρόπο απομόνωσης γονιδίων αλλά, πολύ περισσότερο, έχει εισάγει την ιδέα ότι οι διαφορές στην πολυπλοκότητα μεταξύ των οργανισμών δεν εξηγούνται από τον αριθμό των γονιδίων τους. Για παράδειγμα, είναι χαρακτηριστική η διαπίστωση ότι πολύπλοκοι πολυκύτταροι οργανισμοί όπως ο *Homo sapiens* και ο *Mus musculus* έχουν στο γονιδίωμά τους μόνο 1,5 φορές περισσότερα γονίδια σε σχέση με τον *Caenorhabditis elegans*, τον οποίο αποτελούν λιγότερα από 1500 κύτταρα (Πίνακας 1). Φαίνεται, επομένως, ότι η αυξημένη πολυπλοκότητα του συστήματος των θηλαστικών οφείλεται στη ρύθμιση της γονιδιακής έκφρασης στο επίπεδο των διαφορετικών εναλλακτικών μεταγράφων που μπορούν να προκύψουν από κάθε γονίδιο αλλά και σε μετα-μεταφραστικές τροποποιήσεις των πρωτεϊνών, που καθορίζουν τις διαφορές στα επίπεδα, την τοπολογία και τις αλληλεπιδράσεις των πρωτεϊνικών μορίων, προκαλώντας πολυάριθμες διαφοροποιήσεις μεταξύ

των διαφορετικών κυτταρικών τύπων. Ως εκ τούτου, στο πεδίο της λειτουργικής ανάλυσης γονιδίων των θηλαστικών έχει αναγνωριστεί η ανάγκη λεπτομερούς μελέτης των μηχανισμών ρύθμισης της έκφρασης των γονιδίων, όσο και των αλληλεπιδράσεων μεταξύ των γονιδιακών προϊόντων.

Πίνακας 1. Μεγέθη γονιδιωμάτων και αριθμοί γονιδίων

Οργανισμός	Μέγεθος γονιδιώματος (Mb)	Αριθμός γονιδίων	Αριθμός γονιδίων/1Mb
<i>Homo sapiens</i>	3000	~30.000	20
<i>Mus musculus</i>	3000	~30.000	20
<i>Fugu rubripes</i>	400	~30.000	75
<i>Drosophila melanogaster</i>	180	13.600	110
<i>Arabidopsis thaliana</i>	125	25.500	250
<i>Caenorhabditis elegans</i>	97	19.100	200
<i>Saccharomyces cerevisiae</i>	12	6300	500
<i>Escherichia coli</i>	4.7	3200	700
<i>Haemophilus influenzae</i>	1.8	1700	1000

Σημ: Για κάθε οργανισμό του οποίου έχει ολοκληρωθεί η γονιδιωματική αλληλούχηση, αναφέρεται το μέγεθος του απλοειδούς γονιδιώματος (σε Mb), ο αριθμός των γονιδίων του και η μέση συχνότητα εμφάνισης των γονιδίων στο χρωμοσωμικό του DNA (αριθμός γονιδίων/ 1Mb).

2. Χαρτογράφηση και αλληλούχηση γονιδιωμάτων

Στο πρώτο μέρος αυτής της διατριβής ασχοληθήκαμε με την κατασκευή του φυσικού χάρτη υψηλής ευκρίνειας της περιοχής q23.3 του χρωμοσώματος 10. Μερικά εισαγωγικά στοιχεία για τις προσεγγίσεις στην κατασκευή χρωμοσωμικών χαρτών στον άνθρωπο, παρατίθενται στη συνέχεια.

Πρωταρχικό βήμα για την ανάλυση των γονιδιωμάτων είναι ο προσδιορισμός της αλληλουχίας των νουκλεοτιδίων που συνιστούν το γονιδιωματικό DNA κάθε οργανισμού. Οι στρατηγικές που ακολουθούνται για την αλληλούχηση διαφορετικών γονιδιωμάτων καθορίζονται κυρίως από το μέγεθος του γονιδιώματος

που αναλύεται. Για παράδειγμα, η αλληλούχηση γονιδιωμάτων μικρού μεγέθους, όπως αυτό του βακτηριοφάγου λ [Sanger *et al*, 1982] ή πολλών μικροβίων [Kasai *et al*, 1992 / Fleischmann *et al*, 1995] στηρίχθηκε στην κλωνοποίησή τους στους κλασσικούς φορείς (πλασμίδια, κοσμίδια και βακτηριοφάγοι λ), σε αντίθεση με τα γονιδιώματα μεγάλου μεγέθους, όπως αυτό των θηλαστικών (~3000 Mb) και της δροσόφιλας (180 Mb), που κλωνοποιήθηκαν σε τεχνητά χρωμοσώματα σακχαρομύκητα (YACs), βακτηριοφάγου P1 (PACs) ή βακτηρίων (BACs), φορείς με δυνατότητα ένθεσης μορίων DNA μεγέθους από 100Kb έως 1Mb. Επιπλέον, τα προγράμματα αλληλούχησης μεγάλων γονιδιωμάτων αναγκαστικά περιλαμβάνουν και την κατασκευή χρωμοσωμικών χαρτών, που χρησιμοποιούνται ως οδηγοί για την 'τοποθέτηση' των γονιδιωματικών κλώνων πάνω στα χρωμοσώματα, με σκοπό την πλήρη κάλυψη και τελικά την αλληλούχηση του γονιδιωματος. Εναλλακτικά, εφαρμόζεται και η προσέγγιση της ολικής γονιδιωματικής αλληλούχησης (*whole genome shotgun sequencing*), κατά την οποία το γονιδίωμα τεμαχίζεται τυχαία σε θραύσματα με μεγέθη 2-50Kb [Istrail *et al*, 2004], των οποίων στη συνέχεια προσδιορίζεται η νουκλεοτιδική σύσταση. Βέβαια, η τελευταία μέθοδος στηρίζεται αποκλειστικά και μόνο στη στοιχίση των επιμέρους αλληλουχιών για τη συγκρότηση της ολοκληρωμένης αλληλουχίας νουκλεοτιδίων του γονιδιωματος και είναι ευαίσθητη σε λανθασμένες στοιχίσεις που οφείλονται κυρίως στις επαναλήψεις νουκλεοτιδίων, οι οποίες είναι εξαιρετικά συχνές σε μεγάλα γονιδιώματα όπως αυτά των θηλαστικών [Lander *et al*, 2001]. Επομένως, φαίνεται ότι ανεξάρτητα από τη στρατηγική που επιλέγεται αρχικά για την αλληλούχηση, η συμβολή των χρωμοσωμικών χαρτών είναι ιδιαίτερα σημαντική, ώστε να εξασφαλιστεί η αρτιότητα της νουκλεοτιδικής αλληλουχίας που προκύπτει από κάθε πρόγραμμα γονιδιωματικής ανάλυσης. Οι χάρτες που χρησιμοποιούνται στα προγράμματα αλληλούχησης γονιδιωμάτων διακρίνονται σε γενετικούς και φυσικούς.

α. Γενετικοί χάρτες

Για την κατασκευή ενός γενετικού χάρτη, οι γενετικοί τόποι τοποθετούνται με βάση τα ποσοστά συγκληρονόμησής τους κατά τον ανασυνδυασμό των ομόλογων χρωμοσωμάτων στη μείωση. Οι πρώτοι γενετικοί χάρτες πειραματόζων με καλή γενετική, όπως η δροσόφιλα και ο ποντικός, στηρίζονταν σε γονίδια και κατασκευάζονταν διασταυρώνοντας διαφορετικά μεταλλάγματα, ώστε να διαπιστωθεί αν δύο γενετικοί τόποι είναι συνδεδεμένοι ή όχι. Αντίθετα, η κατασκευή γενετικών χαρτών για τα χρωμοσώματα του ανθρώπου στηρίχθηκε αναγκαστικά σε γενετικούς τόπους υψηλού πολυμορφισμού (κυρίως μικροδορυφόρους) [Dib *et al*, 1996], εφόσον η πιθανότητα διασταύρωσης ατόμων που φέρουν διαφορετικές γενετικές διαταραχές είναι εξαιρετικά μικρή.

Η κατασκευή γενετικών χαρτών ήταν απαραίτητη για το ανθρώπινο γονιδίωμα, εφόσον αποτελεί βασική προσέγγιση ανακάλυψης γονιδίων που ευθύνονται για γενετικές ασθένειες, και ιδιαίτερα αυτών που δεν συσχετίζονται με ευδιάκριτες χρωμοσωμικές ανωμαλίες. Η ισχυρή γενετική σύνδεση (μηδενικός ανασυνδυασμός) του φαινοτύπου μιας γενετικής νόσου με κάποιον γενετικό δείκτη κατά τη μείωση σε οικογένειες ασθενών αποτελεί σημαντική ένδειξη για την παρουσία του υπεύθυνου γονιδίου στην περιοχή αυτού του δείκτη και απαραίτητη προϋπόθεση για τον σχεδιασμό πειραμάτων κλωνοποίησης του γονιδίου με βάση τη θέση του στον γενετικό χάρτη (*positional cloning*). Είναι φανερό ότι το μέγεθος του διαστήματος, στο οποίο με βάση τις μελέτες γενετικής ανάλυσης σύνδεσης περιέχεται το υπεύθυνο γονίδιο, εξαρτάται απόλυτα από την πυ-

κνότητα του χάρτη σε γενετικούς δείκτες. Επομένως, η βελτίωση των γενετικών χαρτών των ανθρώπινων χρωμοσωμάτων ώστε να περιέχουν γενετικούς δείκτες σε αποστάσεις μικρότερες του 1Mb (~1cM) υπήρξε σημαντική προϋπόθεση για την ταυτοποίηση μιας πληθώρας γονιδίων που ευθύνονται για γενετικές νόσους.

Χαρακτηριστικό είναι το παράδειγμα της πορείας που ακολουθήθηκε για την ανακάλυψη της μοριακής βάσης της κυστικής ίνωσης (*Cystic Fibrosis, CF*), αποφρακτικής νόσου των πνευμόνων που συνοδεύεται από ανεπάρκεια παγκρεατικών ενζύμων και στειρότητα, εμφανίζοντας εξαιρετικά υψηλή συχνότητα στους Καυκάσιους (1/2000 ασθενείς και 1/22 ασυμπτωματικοί φορείς). Η ερευνητική προσπάθεια για τη γενετική χαρτογράφηση του φαινοτύπου της νόσου ξεκίνησε στις αρχές της δεκαετίας του 1980, όταν ακόμα οι γενετικοί χάρτες των ανθρώπινων χρωμοσωμάτων ήταν εξαιρετικά φτωχοί σε δείκτες. Μέχρι το 1985, εκατοντάδες γενετικών δεικτών ελέγχθηκαν για σύνδεση χωρίς επιτυχία, ώσπου διαπιστώθηκε η σύνδεση της νόσου με έναν ενζυμικό πολυμορφισμό της πρωτεΐνης του ορού παραξονάσης (PON). Στη συνέχεια, μελετώντας δεκάδες οικογένειες ασθενών ανιχνεύθηκε σύνδεση των γενετικών τόπων PON και CF με γενετικούς δείκτες της περιοχής q31-q32 του ανθρώπινου χρωμοσώματος 7. Τελικά, με τη βοήθεια των βελτιωμένων γενετικών χαρτών που κατασκευάστηκαν στα πλαίσια του διεθνούς προγράμματος ανάλυσης του ανθρώπινου γονιδιώματος, το διάστημα που περιείχε το υπεύθυνο για τη νόσο CF γονίδιο περιορίστηκε από τα 15cM στα 500Kb, που κλωνοποιήθηκαν σε λ βακτηριοφάγους και κοσμίδια. Στη νουκλεοτιδική αλληλουχία των κλώνων ανιχνεύθηκαν 4 καινούρια γονίδια από τα οποία το ένα ήταν μεταγραφικά ενεργό στους ιστούς που κυρίως προσβάλλονται από τη νόσο CF (πνεύμονες, πάγκρεας και ιδρωτοποιοί αδένες). Ήταν το γονίδιο *CFTR*, το οποίο ταυτοποιήθηκε ως υπεύθυνο για τη νόσο όταν σημειακές μεταλλάξεις του βρέθηκαν στους ασθενείς [Riordan *et al*, 1989]. Με τις μοριακές και βιοχημικές μελέτες που ακολούθησαν, διαπιστώθηκε ότι η πρωτεΐνη CFTR (μήκους 1480 αμινοξέων) είναι ένα διαμεμβρανικό κανάλι ιόντων χλωρίου, το οποίο λειτουργεί υδρολύοντας ATP και είναι απαραίτητο για την ομοίωση και τη φυσιολογική λειτουργία πολλών τύπων κυττάρων. Ελλείψεις συγκεκριμένων αμινοξικών καταλοίπων, όπως της φαινυλαλανίνης 508 στο 70% των ασθενών με CF, αναστέλλουν τον μεμβρανικό εντοπισμό της πρωτεΐνης ή τη δράση της ως ATPάσης, καθιστώντας την μη λειτουργική. Η κατανόηση της μοριακής βάσης της νόσου οδήγησε στην ανάπτυξη μεθόδων γονιδιακής θεραπείας, με στόχο την έκφραση της φυσιολογικής πρωτεΐνης CFTR στα επιθηλιακά κύτταρα των πνευμόνων των ασθενών [Rosenfeld *et al*, 1992].

Το παραπάνω είναι ένα από τα πολυάριθμα παραδείγματα που καταδεικνύουν τη σημασία της γενετικής χαρτογράφησης στην ταυτοποίηση γονιδίων ασθενειών, και ιδιαίτερα εκείνων για τα οποία εκ των προτέρων αγνοείται το γονιδιακό προϊόν και η χρωμοσωμική θέση. Επιπλέον, πρόσφατες προσπάθειες εφαρμόζουν την αρχή της γενετικής ανάλυσης σύνδεσης για τον εντοπισμό των γονιδίων που ευθύνονται για συχνές πολυπαραγοντικές νόσους, όπως σακχαρώδη διαβήτη, καρδιοπάθειες, άσθμα κλπ. Τα γονίδια αυτά δεν είναι εύκολο να εντοπιστούν με μελέτες σύνδεσης στις οποίες χρησιμοποιούνται κλασσικοί γενετικοί δείκτες (μικροδορυφόροι). Η δυσκολία σε αυτές τις μελέτες έγκειται στο ότι υπάρχουν περισσότερες από μία χρωμοσωμικές θέσεις που συνεισφέρουν στην ασθένεια, χωρίς η κάθε μία από αυτές να εμφανίζει απαραίτητα ισχυρή σύνδεση με τον φαινότυπο. Δηλαδή, η ασθένεια είναι αθροιστικό αποτέλεσμα ενός συνδυασμού αλληλομόρφων γονιδίων. Για το σκοπό αυτό έχουν αναπτυχθεί νέοι γενετικοί δείκτες με βάση μία από τις διαπι-

στώσεις της πλήρους αλληλούχησης του ανθρώπινου γονιδιώματος, ότι μεταξύ διαφορετικών ανθρώπινων γονιδιωμάτων υπάρχουν διαφορές νουκλεοτιδίων με συχνότητα περίπου 1/1000bp. Αυτές αντιπροσωπεύουν μονονουκλεοτιδικούς πολυμορφισμούς (*Single Nucleotide Polymorphisms, SNPs*). Ως γενετικοί δείκτες για τη διάκριση των χρωμοσωμάτων χρησιμοποιούνται οι απλότυποι που περιγράφουν τη νουκλεοτιδική σύσταση κάθε χρωμοσώματος σε ομάδες γειτονικών SNPs. Το πλεονέκτημα με αυτούς τους γενετικούς δείκτες είναι η υψηλή τους συχνότητα στο γονιδίωμα, γεγονός που τους επιτρέπει ισχυρή ανισορροπία σύνδεσης (*linkage disequilibrium*) με τα παθογόνα αλληλόμορφα γονιδίων, από τα οποία απέχουν απόσταση μερικών κιλοβάσεων, ώστε να μη διαχωρίζονται κατά τον μειωτικό ανασυνδυασμό. Με στόχο τη 'σήμανση' όλων των ανθρώπινων γονιδίων από γενετικούς δείκτες αυτού του είδους, αναπτύσσεται το πρόγραμμα χαρτογράφησης απλοτύπων (*HapMap Project*) των ανθρώπινων χρωμοσωμάτων [The International HapMap Consortium, 2003], το οποίο εκτός από την κατανόηση της μοριακής βάσης πολυπαραγοντικών ασθενειών είναι δυνατόν να συμβάλλει και στην αποκάλυψη των γενετικών παραγόντων που συνεισφέρουν σε φυσιολογικές καταστάσεις, όπως η ανθεκτικότητα σε παθογόνα, η προδιάθεση σε νοσήματα και η απόκριση σε φάρμακα, αποτελώντας τη νέα ευρείας κλίμακας ερευνητική τάση στο πεδίο της γονιδιωματικής ανάλυσης, μετά την ολοκλήρωση της φυσικής χαρτογράφησης του ανθρώπινου γονιδιώματος.

β. Φυσικοί χάρτες

Οι φυσικοί χάρτες χαρακτηρίζονται από μεγάλη ετερογένεια ως προς την αρχή κατασκευής τους. Τα διάφορα είδη φυσικών χαρτών διαφέρουν κυρίως ως προς τη διακριτική τους ικανότητα, δηλαδή το ελάχιστο διάστημα ανάμεσα σε διαδοχικούς γενετικούς τόπους που διακρίνεται σε κάθε χάρτη. Για το ανθρώπινο γονιδίωμα, ο χάρτης ελάχιστης διακριτικής ικανότητας είναι το πρότυπο ζώνωσης μετά από χρώση των 24 διαφορετικών χρωμοσωμάτων του, ενώ η μέγιστη διακριτική ικανότητα επιτυγχάνεται με την αλληλούχηση του γονιδιωματικού DNA (Πίνακας 2). Συνολικά, οι φυσικοί χάρτες που έχουν κατασκευαστεί για το ανθρώπινο γονιδίωμα περιλαμβάνουν:

i) *Κυτταρογενετικούς χάρτες*, στους οποίους με διάφορες τεχνικές χρώσης των χρωμοσωμάτων (πχ χρώση *Giemsa*) διακρίνονται μεταξύ τους τα διαφορετικά χρωμοσώματα ενός ανθρώπινου μεταφασικού κυττάρου, αλλά και γειτονικές χρωμοσωμικές περιοχές μεγέθους 5-10 Mb. Αυτοί οι χάρτες έχουν χρησιμοποιηθεί για τη χαρτογράφηση στα χρωμοσώματα γονιδιωματικών κλώνων YAC, PAC ή BAC, με χρήση της φθορίζουσας υβριδοποίησης *in situ* (*Fluorescence In Situ Hybridization, FISH*) [Bray-Ward *et al*, 1996].

ii) *Χάρτες χρωμοσωμικών θραύσεων*, οι οποίοι στηρίζονται στη χαρτογράφηση των τμημάτων του χρωμοσωμικού DNA που προκύπτουν από φυσική (έλλειμμα ή μετατόπιση χρωμοσωμικής περιοχής) ή επαγόμενη χρωμοσωμική θραύση. Οι τελευταίοι περιλαμβάνουν τους χάρτες που προκύπτουν με την τεχνολογία των *'radiation hybrids'*. Οι χάρτες αυτού του είδους είναι περιορισμένης διακριτικής ικανότητας, όμως υπήρξαν χρήσιμοι στη χαρτογράφηση μη πολυμορφικών γενετικών τόπων (*Sequence Tagged Sites, STSs*) και γονιδίων (*Expressed Sequence Tags, ESTs*) [Hudson *et al*, 1995].

iii) *Χάρτες επικαλυπτόμενων κλώνων γονιδιωματικού DNA (contig maps)*, οι οποίοι κατασκευάζονται συλλέγοντας τους γονιδιωματικούς κλώνους (YACs, BACs, PACs) που καλύπτουν μια χρωμοσωμική περιοχή και διατάσσοντάς τους με βάση τη σειρά ήδη χαρτογραφημένων γενετικών τόπων (STS, EST ή πολυμορφισμών). Η σημαντικότητα των χαρτών αυτού του είδους έγκειται στο ότι αυτοί προκύπτουν από την ενοποι-

ηση των δεδομένων όλων των παραπάνω χαρτών και τελικά προσφέρουν το DNA-μήτρα για την αλληλούχηση κάθε χρωμοσώματος. Στην Εικ.1 παρουσιάζονται ενδεικτικά ενοποιημένοι χάρτες του χρωμοσώματος 10 του ανθρώπου [Marzella, Kokkinaki et al, 1997/ Μ. Κοκκινάκη, Μεταπτυχιακή Διατριβή, Παν/μιο Κρήτης, 1996).

Πίνακας 2. Μέθοδοι φυσικής γονιδιωματικής χαρτογράφησης

Μέθοδος	Στόχος	Διακριτική ικανότητα (bp)
Καρυότυπος	Αντιστοίχιση με χρωμόσωμα	50-250 x10 ⁶
Υβρίδια σωματικών κυττάρων ανθρώπου και τρωκτικού (somatic cell hybrids)	Αντιστοίχιση με χρωμόσωμα	50-250 x10 ⁶
Radiation hybrids	Αντιστοίχιση με χρωμοσωμική περιοχή	5-10 x10 ⁶
Υβριδοποίηση <i>in situ</i> σε μεταφασικά χρωμοσώματα	Αντιστοίχιση με χρωμοσωμική περιοχή	1-2 x10 ⁶
Κλωνοποίηση σε YACs	Δημιουργία contig	0.1-1 x10 ⁶
Κλωνοποίηση σε BACs ή PACs	Δημιουργία contig	0.01 x10 ⁶
Ανάλυση περιοριστικών ενζύμων με σπάνιες θέσεις αναγνώρισης	Αδρή χαρτογράφηση	0.01 x10 ⁶
Υβριδοποίηση <i>in situ</i> σε μεσοφασικά χρωμοσώματα	Λεπτομερής χαρτογράφηση	0.003 x10 ⁶
Υβριδοποίηση <i>in situ</i> σε ινίδια αποσυμπυκνωμένης χρωματίνης (fiber FISH)	Λεπτομερής χαρτογράφηση	0.001 x10 ⁶
Αλληλούχηση	Νουκλεοτιδική αλληλουχία	1

Σημ: Για κάθε μέθοδο φυσικής χαρτογράφησης φαίνεται ο σκοπός τον οποίο εξυπηρετεί και η διακριτική της ικανότητα.

γ. Αλληλούχηση του ανθρώπινου γονιδιώματος

Η αναγκαιότητα της αλληλούχησης του ανθρώπινου γονιδιώματος υπήρξε εμφανής από τις πρώτες προσπάθειες μελέτης της μοριακής βάσης γενετικών νόσων που οφείλονταν σε αλλαγές στην αλληλουχία του DNA των ασθενών. Αυτή ήταν και η αφορμή για τη σύσταση το 1988 του προγράμματος ανάλυσης του ανθρώπινου γονιδιώματος (*Human Genome Project*, HGP), μιας παγκόσμιας ερευνητικής προσπάθειας για τη χαρτογράφηση του γονιδιώματος του ανθρώπου και τον προσδιορισμό της νουκλεοτιδικής του αλληλουχίας [Watson & Cook-Deegan, 1991].

Σύμφωνα με τους στόχους αυτού του προγράμματος, δόθηκε αρχικά προτεραιότητα στην αλληλούχηση των χρωμοσωμικών περιοχών που περιέχουν γονίδια, με βάση την εκτίμηση ότι οι κωδικές περιοχές του γονιδιώματος συγκεντρώνουν το μεγαλύτερο βιολογικό ενδιαφέρον και κυρίως λόγω του μικρού ποσοστού

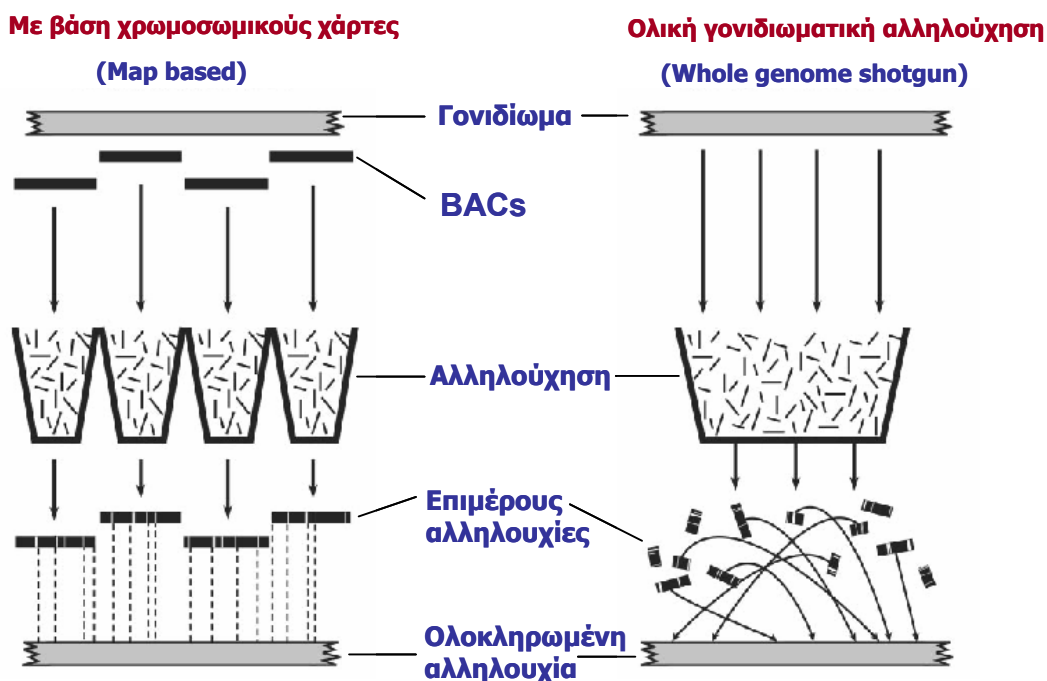
πεύονταν στις βιβλιοθήκες cDNA λόγω χαμηλής ή εξειδικευμένης (χρονικά ή ιστοειδικά) έκφρασης αλλά και τον προσδιορισμό μη κωδικών αλληλουχιών DNA με ιδιαίτερη βιολογική σημασία για τη ρύθμιση της έκφρασης των γονιδίων ή για τη λειτουργία των χρωμοσωμάτων.

Σημαντικό βήμα στην πορεία αυτού του προγράμματος ήταν η ίδρυση το 1988 του οργανισμού HUGO (*Human Genome Organization*), με σκοπό τον συντονισμό της πειραματικής διαδικασίας και την επικοινωνία μεταξύ των ομάδων που είχαν αναλάβει τη χαρτογράφηση διαφορετικών χρωμοσωμάτων (www.gdb.org/hugo/). Καθοριστική ήταν και η επιλογή των ερευνητών του HGP να ακολουθήσουν τη μέθοδο της αλληλούχησης βάσει χαρτών (*map based*) και όχι αυτήν της ολικής γονιδιωματικής αλληλούχησης θραυσμάτων DNA (*whole genome shotgun sequencing*), που είχε εφαρμοστεί μέχρι τότε στην ανάλυση μικρότερων γονιδιωμάτων όπως αυτά της *E.coli K-12* (4.6 Mb) [Blattner *et al*, 1997] και αργότερα του *Anopheles gambiae* (278 Mb) [Holt *et al*, 2002]. Πράγματι, αν και πρόκειται για πολύ πιο αργή διαδικασία στα αρχικά στάδια κατασκευής γενετικών και φυσικών χαρτών, η αλληλούχηση χαρτογραφημένων κλώνων γονιδιωματικού DNA παρουσιάζει βασικά πλεονεκτήματα έναντι της ολικής γονιδιωματικής αλληλούχησης κατά την τελική φάση του προγράμματος, στην οποία απαιτείται η στοιχίση των επιμέρους αλληλουχιών για την παραγωγή της ολοκληρωμένης νουκλεοτιδικής αλληλουχίας κάθε χρωμοσώματος (Εικ.2). Στη φάση αυτή, η αλληλούχηση βάσει χαρτών επιτρέπει την ασφαλέστερη τοποθέτηση των επιμέρους αλληλουχιών κατά μήκος του χρωμοσώματος, με βάση τις επικαλύψεις των γονιδιωματικών κλώνων και το περιεχόμενό τους σε ήδη χαρτογραφημένες αλληλουχίες STS και EST. Επομένως, η προσέγγιση αυτή είναι λιγότερο ευαίσθητη στις περιπτώσεις λανθασμένης στοιχίσης των επιμέρους τμημάτων, που οφείλονται κυρίως στην παρουσία επαναλαμβανόμενων αλληλουχιών οι οποίες καταλαμβάνουν περισσότερο από 50% του ανθρώπινου DNA [Lander *et al*, 2001].

Αποτέλεσμα της αρχικής προτεραιότητας στη χαρτογράφηση των κωδικών αλληλουχιών ήταν ο πρώτος χάρτης γονιδίων του ανθρώπου που κατασκευάστηκε χρησιμοποιώντας τη μέθοδο των ακτινοβολημένων υβριδίων σωματικών κυττάρων (*radiation hybrids*) για τη χαρτογράφηση περίπου 30.000 αλληλουχιών EST (*Expressed Sequence Tags*) στα χρωμοσώματα του ανθρώπου και δημοσιεύθηκε τον Οκτώβριο του 1998 [Deloukas *et al*, 1998, www.ncbi.nlm.nih.gov/genemap99/]. Εκτός από τις κωδικές αλληλουχίες, στον ίδιο χάρτη είχε ενσωματωθεί και μεγάλος αριθμός πολυμορφικών γενετικών δεικτών, ιδιαίτερα χρήσιμων για το συσχετισμό του με προηγούμενους γενετικούς χάρτες.

Μετά την κατασκευή γενετικών και φυσικών χαρτών, η αλληλούχηση του ανθρώπινου γονιδιώματος σε μεγάλη κλίμακα συνεχίστηκε με ρυθμό που αυξήθηκε θεαματικά τα τελευταία χρόνια, κυρίως λόγω του γεγονότος ότι η κλασική μέθοδος αλληλούχησης DNA με χρήση διδεοξυνουκλεοτιδίων [Sanger *et al*, 1977] που χρησιμοποιούνταν, υποστηρίχθηκε με τροποποιήσεις που ευνοούσαν την αυτοματοποίησή της, όπως η χρήση χρωστικών με φθορισμό και η ηλεκτροφόρηση των προϊόντων σε τριχοειδή [Starke *et al*, 1994]. Αποτέλεσμα αυτής της προσπάθειας ήταν η δημοσίευση από το *International Human Genome Sequencing Consortium* της προκαταρκτικής (*draft*) νουκλεοτιδικής αλληλουχίας, που αντιστοιχεί σε περισσότερο από το

Στρατηγική αλληλούχησης γονιδιώματος



Εικ.2 Οι δύο εναλλακτικές προσεγγίσεις για την αλληλούχηση γονιδιωμάτων, συστηματική αλληλούχηση με βάση χρωμοσωμικούς χάρτες ή ολική γονιδιωματική αλληλούχηση.

90% του γονιδιώματος του ανθρώπου, περιέχοντας κενά μέσου μήκους 500bp στο 4% των κλώνων [Lander *et al*, 2001] και η σταδιακή δημοσίευση της συνολικής αλληλουχίας των ανθρώπινων χρωμοσωμάτων 22 [Dunham *et al*, 1999], 21 [Hattori *et al*, 2000], 20 [Deloukas *et al*, 2001], 6 [Mungall *et al*, 2003], 7 [Hillier *et al*, 2003], 14 [Heilig *et al*, 2003], της ειδικής για το Y περιοχής [Skaletsky *et al*, 2003] και του χρωμοσώματος 10 [Deloukas *et al*, 2004].

Η εναλλακτική μέθοδος της τυχαίας αλληλούχησης γονιδιωματικού DNA (*whole genome shotgun*) δοκιμάστηκε και για το ανθρώπινο γονιδίωμα από την εταιρία *Celera Genomics*. Όπως είχε αρχικά υπολογιστεί, θα χρειαζόταν να 'διαβαστεί' το γονιδίωμα τουλάχιστον δέκα φορές για να προκύψει ολοκληρωμένη η αλληλουχία του με τη συγκεκριμένη προσέγγιση [Weber & Myers, 1997]. Όμως, η διαθέσιμη πληροφορία γονιδιωματικής χαρτογράφησης από το διεθνές πρόγραμμα HGP φαίνεται ότι συνέβαλλε σημαντικά στην τελική συναρμολόγηση της νουκλεοτιδικής αλληλουχίας των Venter *et al* (2001), για την οποία η ομάδα της *Celera Genomics* πραγματοποίησε τρεις επαναλήψεις τυχαίας αλληλούχησης του ανθρώπινου γονιδιώματος [Waterston *et al*, 2002b]. Αξίζει επιπλέον να σημειωθεί ότι για το αντίστοιχο διεθνές πρόγραμμα ανάλυσης του γονιδιώματος του ποντικού (*Mouse Genome Project*) χρησιμοποιήθηκε συνδυασμός των δύο μεθόδων, με αποτέλεσμα η αρχική αλληλουχία που προέκυψε να καλύπτει επτά φορές το γονιδίωμα του στελέχους C57BL/6J και να αντιστοιχεί στο 96% της ευχρωματίνης, περιέχοντας κενά μέσου μήκους 600bp στο 4% των κλώνων [Waterston *et al*, 2002a].

Αν και έχουν διατυπωθεί αντικρουόμενες απόψεις σχετικά με τον σχεδιασμό της στρατηγικής που ακολουθήθηκε για την ανάλυση του ανθρώπινου γονιδιώματος [Myers *et al*, 2002 / Waterston *et al*, 2003],

το ωφέλιμο στοιχείο από αυτή την προσπάθεια είναι η ίδια η νουκλεοτιδική αλληλουχία αλλά και οι γενετικοί και φυσικοί χάρτες των ανθρώπινων χρωμοσωμάτων που κατασκευάστηκαν στα πλαίσια του προγράμματος HGP και είναι διαθέσιμοι σε ηλεκτρονικές διευθύνσεις του διαδικτύου (Πίνακας 3), οι οποίες ενημερώνονται συνεχώς για την εξέλιξη της αλληλούχησης στις διάφορες χρωμοσωμικές περιοχές. Με αυτά τα δεδομένα είναι πλέον δυνατή η ταυτοποίηση γονιδίων *in silico*, γεγονός που διευκολύνει τη λειτουργική τους ανάλυση αλλά και επιτρέπει συγκριτικές μελέτες με γονιδιώματα άλλων οργανισμών.

Πίνακας 3. Πηγές πληροφόρησης για το ανθρώπινο γονιδίωμα

<p>University of California at Santa Cruz http://genome.ucsc.edu/ Η προκαταρκτική (draft) μορφή της αλληλουχίας που δημοσιεύτηκε το 2001 από τους Lander ES <i>et al/</i> και οι επόμενες βελτιωμένες μορφές της.</p>
<p>Washington University http://genome.wustl.edu/gsc/human/Mapping/ Συλλογή χαρτογραφημένων γενωμικών κλώνων του ανθρώπινου γονιδιώματος</p>
<p>EBI/Sanger Centre http://www.ensembl.org Γραφικές απεικονίσεις των χρωμοσωμάτων σε συνδυασμό με τις αλληλουχίες DNA και πρωτεϊνών και βασικά χαρακτηριστικά στοιχεία για αυτές (annotation)</p>
<p>NCBI (National Center for Biotechnology Information) http://www.ncbi.nlm.nih.gov/genome/guide/ Γραφικές απεικονίσεις χρωμοσωμάτων, χαρτών και γενετικών τόπων, οι οποίες συνδέονται και με άλλες πηγές πληροφόρησης του NCBI (πχ με βιβλιογραφικές αναφορές)</p>
<p>Gene map 99 http://www.ncbi.nlm.nih.gov/genemap99/ Ο χάρτης γονιδίων 'Gene map 99', στον οποίο περιέχονται δεδομένα και γραφικές απεικονίσεις για τη θέση αλληλουχιών STS και EST βάσει χαρτών ακτινοβολημένων υβριδίων σωματικών κυττάρων</p>
<p>Oak Ridge National Laboratory http://compbio.ornl.gov/channel/index.html Γραφικές απεικονίσεις δεδομένων χαρτογράφησης του ανθρώπινου γονιδιώματος</p>
<p>RIKEN , University of Tokyo http://hgrep.ims.u-tokyo.ac.jp/ Μια γενική επισκόπηση της δομής του ανθρώπινου γονιδιώματος</p>
<p>Online Mendelian Inheritance in Man (OMIM) http://www.ncbi.nlm.nih.gov/Omim/ Πληροφορία για τη σχέση γονιδίων με κληρονομικές ασθένειες στον άνθρωπο.</p>

3. Ταυτοποίηση γονιδίων

Οι δύο πιο διαδεδομένες προσεγγίσεις για τον προσδιορισμό των γονιδίων σε μια αλληλουχία νουκλεοτιδίων γονιδιωματικού DNA είναι η αλληλούχηση κλώνων cDNA και η χρήση βιοϋπολογιστικών προγραμμάτων για την πρόβλεψη κωδικών αλληλουχιών. Πρόσφατα έχουν αναπτυχθεί και άλλες μέθοδοι ή συνδυασμοί μεθόδων που έχουν στόχο να αυξήσουν την αξιοπιστία και την απόδοση των προβλέψεων.

α. Αλληλούχηση κλώνων cDNA

Αποτελεί τον πιο κλασικό τρόπο ταυτοποίησης γονιδίων και περιλαμβάνει την μερική ή πλήρη αλληλούχηση των μεταγράφων που έχουν κλωνοποιηθεί σε βιβλιοθήκες cDNA. Η προσέγγιση αυτή έχει συμ-

βάλλει στην ταυτοποίηση χιλιάδων γονιδίων του ανθρώπου και άλλων οργανισμών [Hillier *et al*, 1996 / Strausberg *et al*, 2002], όμως οι δυνατότητές της περιορίζονται από το γεγονός ότι η ανίχνευση ενός γονιδίου εξαρτάται σε μεγάλο βαθμό από τα επίπεδα έκφρασής του. Γονίδια με υψηλά επίπεδα έκφρασης συνήθως υπεραντιπροσωπεύονται στις βιβλιοθήκες cDNA και στις βάσεις καταχώρησης των αλληλουχιών EST, ενώ γονίδια που εκφράζονται σε χαμηλά επίπεδα, σε περιορισμένο αριθμό κυττάρων ή υπό ειδικές συνθήκες επαγωγής, πιθανόν να απουσιάζουν. Για το λόγο αυτό, είτε χρησιμοποιούνται κανονικοποιημένες βιβλιοθήκες [Patanjali *et al*, 1991 / Soares *et al*, 1994], είτε κατασκευάζονται βιβλιοθήκες από συγκεκριμένους κυτταρικούς τύπους, όπως έγινε από το *FANTOM Consortium* για το πρόγραμμα χαρακτηρισμού περισσότερων από 60.000 πλήρους μήκους μεταγράφων του ποντικού [Okazaki *et al*, 2002]. Ένα επιπλέον πρόβλημα αυτής της προσέγγισης είναι ότι τα προγράμματα αλληλούχησης κλώνων cDNA συνήθως μεροληπτούν εις βάρος της κωδικής αλληλουχίας, εφόσον αρχικά προσδιορίζεται η νουκλεοτιδική αλληλουχία των άκρων των κλώνων cDNA, που συχνά αντιστοιχούν στις μη μεταφραζόμενες περιοχές 3'UTR και 5'UTR. Σε μια προσπάθεια να προσδιοριστεί η ενδιάμεση αλληλουχία ανθρώπινων μεταγράφων των οποίων τα άκρα έχουν αλληλουχηθεί, οι Camargo *et al* (2001) σχεδίασαν εκκινητές PCR από τη διαθέσιμη αλληλουχία των άκρων, ώστε τα ενδιάμεσα τμήματα των μεταγράφων να ανακτηθούν με RT-PCR σε δείγματα RNA από 24 ανθρώπινους ιστούς. Το συγκεκριμένο πρόγραμμα οδήγησε στην παραγωγή περίπου 7×10^5 κωδικών αλληλουχιών (ORESTES, *Open Reading frame Expressed Sequence Tags*). Εκτιμώντας τα επίπεδα έκφρασης με βάση τον αριθμό των διαθέσιμων ESTs για κάθε γονίδιο, οι Camargo *et al* υπολογίζουν ότι στις αλληλουχίες ORESTES αντιπροσωπεύονταν το 80% των γονιδίων με υψηλή ή μέτρια έκφραση και το 40-50% των γονιδίων με περιορισμένη έκφραση.

β. Πρόβλεψη γονιδίων με βιοϋπολογιστικές μεθόδους

Η ανάπτυξη και η αξιολόγηση υπολογιστικών προγραμμάτων ικανών να αναγνωρίζουν τις κωδικές περιοχές μιας αλληλουχίας νουκλεοτιδίων γονιδιωματικού DNA ξεκίνησε πριν ακόμα ολοκληρωθεί η αλληλούχηση του ανθρώπινου γονιδιώματος [Burset & Guigo, 1996]. Βασικό χαρακτηριστικό αυτών των μεθόδων είναι ότι στηρίζονται μόνο στη νουκλεοτιδική αλληλουχία του DNA και επομένως δεν επηρεάζονται άμεσα από το επίπεδο έκφρασης κάθε γονιδίου. Τα προγράμματα που χρησιμοποιούνται διακρίνονται σε δύο κατηγορίες, ανάλογα με την αρχή λειτουργίας τους:

i) Τα προγράμματα που εκτελούν προβλέψεις γονιδίων με βάση την ομολογία των θεωρητικά μεταφραζόμενων περιοχών γονιδιωματικού DNA με χαρακτηρισμένα γονίδια ή πρωτεΐνες του ίδιου ή άλλων οργανισμών. Σε αυτήν την κατηγορία ανήκουν τα προγράμματα *GLASS/ROSETTA*, *GeneWise*, *Genie* και *ExoFish*, που έχουν εφαρμοστεί σε μελέτες γονιδιωματικής ανάλυσης στον άνθρωπο [Lander *et al*, 2001], στον ποντικό [Batzoglou *et al*, 2000], στη δροσόφιλα [Gopal *et al*, 2001] και στο ψάρι *Fugu* [Roest Crollius *et al*, 2000]. Η μέθοδος αυτή βασίζεται σε πειραματικά δεδομένα, όμως τα αποτελέσματά της επηρεάζονται από την παρουσία αλληλουχιών μη επεξεργασμένων mRNAs ή και γονιδιωματικού DNA στις βάσεις δεδομένων. Για τον λόγο αυτό, η δομή των προβλεπόμενων γονιδίων επιβεβαιώνεται συγκρίνοντας την γονιδιωματική με τις διαθέσιμες κωδικές αλληλουχίες και αναζητώντας σαφείς ενδείξεις γεγονότων συναρμογής εξονίων (*splicing*). Η σπουδαιότητα του τελευταίου κριτηρίου για την επιβεβαίωση των

προβλέψεων αυτής της μεθόδου, καθιστά δύσκολη την πρόβλεψη γονιδίων με ένα εξόνιο ή με μεγάλα εξόνια, εφόσον αυτά δεν διακρίνονται από το γονιδιωματικό DNA.

ii) Τα προγράμματα που εκτελούν προβλέψεις *'de novo'*, προβλέπουν τις θέσεις των εξονίων σε μια γονιδιωματική αλληλουχία DNA, με βάση τη διαφορά της νουκλεοτιδικής σύστασης εξονίων και ιντρονίων, καθώς και τα συντηρημένα νουκλεοτίδια που συμμετέχουν στη συναρμογή των εξονίων κατά την επεξεργασία του πρόδρομου μεταγράφου RNA. Επιπλέον, ορισμένα από τα προγράμματα αυτής της κατηγορίας (πχ *GenScan*, *Genie*, *GenFinder*) ορίζουν τα άκρα των γονιδίων, αναγνωρίζοντας συντηρημένα μοτίβα των υποκινητών (*TATA box*, *CAAT box*), τη θέση έναρξης της μετάφρασης (συντηρημένη αλληλουχία *Kojak*) και τα σήματα τερματισμού της μεταγραφής και πολυαδενυλίωσης. Οι *de novo* προβλέψεις είναι ανεξάρτητες τόσο από την έκφραση των γονιδίων όσο και από τη διαθεσιμότητα ομόλογων γονιδίων ή πρωτεϊνών. Όμως, επειδή τα αποτελέσματά τους βασίζονται σε μεγάλο βαθμό σε στατιστικά δεδομένα κατανομής των νουκλεοτιδίων στο μόριο του DNA, αναγνωρίζεται η ανάγκη αυτές οι προβλέψεις να επιβεβαιώνονται πειραματικά. Για την πειραματική επιβεβαίωση της έκφρασης των γονιδίων που προβλέπονται από βιοϋπολογιστικά προγράμματα έχουν αναπτυχθεί μέθοδοι ανίχνευσης των θεωρούμενων κωδικών αλληλουχιών σε βιβλιοθήκες cDNA ή δείγματα mRNA [Reboul *et al*, 2001] και ελέγχου της έκφρασής τους με χρήση μικροσυστοιχιών (*microarrays*) [Penn *et al*, 2000, Shoemaker *et al*, 2001]. Οι μικροσυστοιχίες γονιδιωματικού DNA έχουν αποδειχθεί ιδιαίτερα χρήσιμες για τον προσδιορισμό των άκρων των προβλεπόμενων γονιδίων (πρώτο και τελευταίο εξόνιο), καθώς και των περιοχών 5'UTR και 3'UTR, που συνήθως δεν καθορίζονται με ακρίβεια από τα βιοϋπολογιστικά προγράμματα γονιδιακής πρόβλεψης [Burge & Karlin, 1998]. Επιπλέον, προσφέρουν τη δυνατότητα ανίχνευσης γεγονότων εναλλακτικής συναρμογής (*alternative splicing*) εξονίων, φαινόμενο που αντιπροσωπεύει σημαντικό τρόπο μεταγραφικής ρύθμισης για τουλάχιστον το 40% των γονιδίων του ανθρώπου [Modrek *et al*, 2001] και του ποντικού [Okazaki *et al*, 2002].

4. Οργάνωση γονιδιωματικής πληροφορίας

Η νουκλεοτιδική αλληλουχία του γονιδιώματος αποτελεί το υπόβαθρο στο οποίο είναι δυνατόν να οργανωθεί η συσσωρευμένη πληροφορία σχετικά με τη βιολογία ενός οργανισμού. Οι γονιδιωματικές βάσεις δεδομένων κατασκευάζονται για να εξυπηρετήσουν τον παραπάνω στόχο και είναι απαραίτητες ιδιαίτερα για την οργάνωση των βιολογικών δεδομένων που αφορούν οργανισμούς με μεγάλα γονιδιώματα, όπως αυτά των θηλαστικών. Μια τέτοια βάση δεδομένων, η οποία συγκεντρώνει και οργανώνει στοιχεία για τη δομή και τη λειτουργία των γονιδίων του ανθρώπου αλλά και άλλων οργανισμών, έχει αναπτυχθεί από το Ευρωπαϊκό Ινστιτούτο Βιοπληροφορικής (EMBL) και το *Sanger Center* (*Hinxton, UK*) και ονομάζεται 'Πρόγραμμα *Ensembl*'. Επομένως, η βάση δεδομένων *Ensembl* του ανθρώπινου γονιδιώματος περιέχει ένα ευρύ φάσμα πληροφοριών για κάθε γονίδιο, από δομικά ή λειτουργικά δεδομένα μοριακής βιολογίας και συγκριτικής γονιδιωματικής ανάλυσης, έως στοιχεία κλινικής γενετικής της βάσης δεδομένων *OMIM* (*Online Mendelian Inheritance in Man*) και λειτουργικού χαρακτηρισμού πρωτεϊνικών μορίων της βάσης *InterPro* [Hubbard *et al*, 2002]. Η παραπάνω πληροφορία έχει οργανωθεί με βάση την αρχική πρόβλεψη του γονιδίου στην νουκλεο-

τιδική αλληλουχία του γονιδιώματος από το ίδιο το πρόγραμμα *Ensembl*. Το πρόγραμμα *Ensembl* χρησιμοποιεί συνδυασμό μεθόδων γονιδιακών προβλέψεων (*de novo* και βάσει ομολογίας), περιλαμβάνοντας τρία βασικά στάδια. Αρχικά, πραγματοποιούνται *de novo* προβλέψεις γονιδίων στη διαθέσιμη νουκλεοτιδική αλληλουχία από το πρόγραμμα *Genscan*. Ακολουθεί η επιβεβαίωση ορισμένων από αυτές με βάση την ομολογία με γνωστές πρωτεΐνες, πρωτεϊνικά *domains* και mRNAs από τις βάσεις δεδομένων *SwissProt*, *Pfam* και *EMBL*. Συγκεκριμένα, επιβεβαιώνονται τα ιντρόνια που γεφυρώνονται από γνωστές αλληλουχίες και τα εξόνια που περιβάλλονται από επιβεβαιωμένα ιντρόνια. Η αυξημένη αυστηρότητα αυτού του σταδίου οδηγεί συνήθως σε προβλέψεις τμημάτων των ανθρώπινων γονιδίων, που χρησιμοποιούνται από το πρόγραμμα *Genewise*, το οποίο εκτελεί *de novo* προβλέψεις γύρω από αυτά τα τμήματα, επεκτείνοντάς τα περισσότερο. Τέλος, τα παραπάνω αποτελέσματα συνδυάζονται με αυτά του προγράμματος *Genie*, το οποίο βασίζεται στην ομολογία της γονιδιωματικής αλληλουχίας με γνωστά mRNAs και ESTs, όπως και τα προηγούμενα, όμως συνδέει τα ESTs των 5' και 3' άκρων κλώνων cDNA, ώστε να παράγει προβλέψεις γονιδιωματικής οργάνωσης της πλήρους κωδικής αλληλουχίας, περιλαμβάνοντας συχνά και μη μεταφραζόμενες περιοχές (5' και 3' UTR) του κάθε γονιδίου, με βάση τις διαθέσιμες αλληλουχίες κλώνων cDNA. Με αυτόν τον τρόπο, εξασφαλίζεται η ελάχιστη πιθανότητα λάθους στις προβλέψεις των γονιδίων. Με παρόμοιες προσεγγίσεις προσδιορίζεται η θέση και η οργάνωση στο γονιδίωμα παράλογων ή ορθόλογων πρωτεϊνών, οι οποίες αντιπροσωπεύουν 'νέα' (*novel*) ανθρώπινα γονίδια. Με βάση τις παραπάνω μεθόδους, εκτιμάται ότι το ανθρώπινο γονιδίωμα περιέχει περίπου 11 γονίδια ανά Mb, ενώ υπολογίζεται ότι η ευαισθησία αυτών των προσεγγίσεων για την πρόβλεψη καινούριων γονιδίων είναι περίπου 68% [Lander *et al*, 2001].

Η παραπάνω πληροφορία για το γονιδίωμα του ανθρώπου και των άλλων οργανισμών είναι ελεύθερα διαθέσιμη στον δικτυακό τόπο του προγράμματος *Ensembl* (www.ensembl.org/). Ένα παράδειγμα απεικόνισης της πληροφορίας από το πρόγραμμα *Ensembl* για ένα γονίδιο του ανθρώπινου χρωμοσώματος 10, παρουσιάζεται στην Εικ.3. Η βάση δεδομένων δίνει τη δυνατότητα περιήγησης κατά μήκος των χρωμοσωμάτων και επιλογής συγκεκριμένων χρωμοσωμικών περιοχών για λεπτομερέστερη παρουσίαση. Για την περιγραφή των προβλεπόμενων γονιδίων χρησιμοποιούνται όροι από εξωτερικές βάσεις δεδομένων (πχ η ονοματολογία 'HUGO' για τα γνωστά γονίδια, τα ονόματα γενετικών δεικτών και οι κωδικοί πρόσβασης 'EMBL' για κλώνους cDNA), αλλά και κωδικά ονόματα της βάσης *Ensembl* για τα καινούρια γονίδια. Στον δικτυακό τόπο του προγράμματος επιτρέπονται εναλλακτικές απεικονίσεις της διαθέσιμης πληροφορίας για κάθε γονίδιο (Εικ.3), οι οποίες περιλαμβάνουν: α) Τη συνολική απεικόνιση κάθε χρωμοσώματος ή χρωμοσωμικής περιοχής (*mapview*), στην οποία το κυτταρογενετικό πρότυπο ζώνωσης συσχετίζεται με τη νουκλεοτιδική αλληλουχία μέσω γενετικών δεικτών, β) Την απεικόνιση κάθε προβλεπόμενου γονιδίου (*geneview*), στην οποία συσχετίζεται η προτεινόμενη οργάνωση εξονίων και ιντρονίων με τα μετάγραφα που την υποστηρίζουν, και γ) Την απεικόνιση κάθε προβλεπόμενης πρωτεΐνης (*proteinview*), στην οποία εμφανίζονται οι περιοχές ομολογίας με συντηρημένες δομές της βάσης δεδομένων *InterPro*. Επιπλέον, για κάθε γονίδιο αναφέρονται τα ορθόλογα που ανιχνεύονται στα γονιδιώματα άλλων οργανισμών. Για τα ήδη γνωστά ανθρώπινα γονίδια, τα παραπάνω συνοδεύονται και από βιβλιογραφικές αναφορές, καθώς και στοιχεία της βάσης δεδομένων OMIM, στις περιπτώσεις γονιδίων με κλινικό φαινότυπο (Εικ.3,ε-στ). Τέλος, στον ίδιο δικτυακό τόπο παρέχεται η δυνατότητα συγκρίσεων νουκλεοτιδικών ή αμινοξικών αλληλουχιών, με χρήση των αλγορίθμων BLAST (*Basic*

Local Alignment Search Tool) [Altschul *et al*, 1997] και SSAHA (*Sequence Search and Alignment by Hashing Algorithm*, www.sanger.ac.uk/Software/analysis/SSAHA/). Είναι, επομένως, φανερό η συμβολή των γονιδιω-ματικών βάσεων δεδομένων στην αξιοποίηση της νουκλεοτιδικής αλληλουχίας των γονιδιωμάτων και τη σύνδεσή της με πειραματικά δεδομένα και συγκρίσεις με άλλους οργανισμούς.

5. Συγκριτική γονιδιωματική ανάλυση

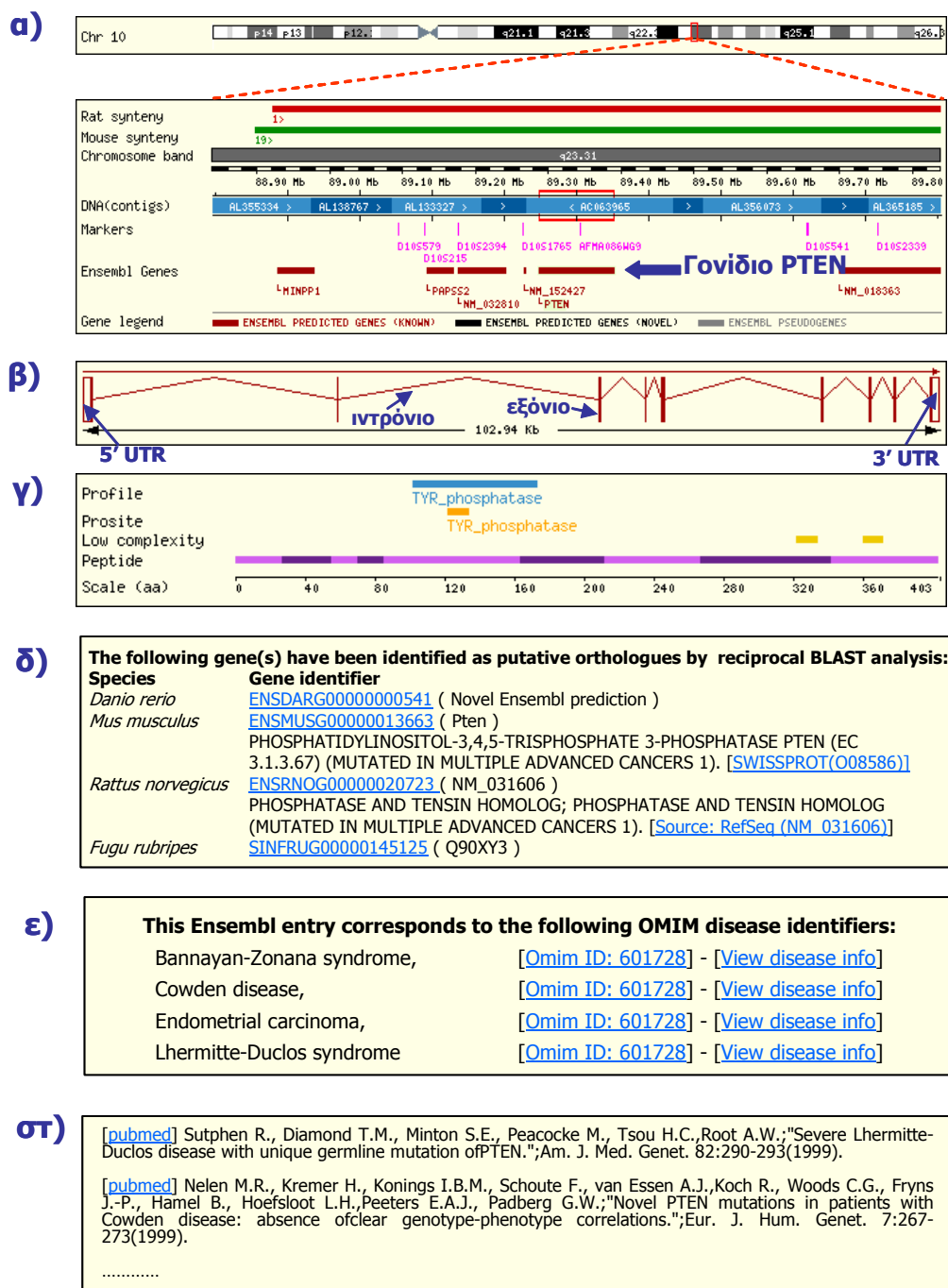
Η διαθεσιμότητα των αποτελεσμάτων της αλληλούχησης των γονιδιωμάτων πολλών οργανισμών παρέχει τη δυνατότητα συγκρίσεων, με τη χρήση κατάλληλων βιοϋπολογιστικών προγραμμάτων. Η εξελικτική συντήρηση των κωδικών αλληλουχιών DNA επιτρέπει την ταυτοποίηση ορθόλογων γονιδίων στα διαφορετικά γονιδιώματα.

Από τη σύγκριση του γονιδιώματος του ανθρώπου με γονιδιώματα μονοκύτταρων οργανισμών (πχ *S. cerevisiae*) εντοπίζονται γονίδια υπεύθυνα για στοχειώδεις κυτταρικές λειτουργίες (βασικός μεταβολισμός, μεταγραφή του DNA σε RNA, μετάφραση του RNA σε πρωτεΐνη, αντιγραφή του DNA κλπ), οι οποίες παρουσιάζονται συντηρημένες κατά την εξέλιξη των ειδών. Ο υψηλός βαθμός συντήρησης τέτοιων λειτουργιών εκφράζεται και από το υψηλό ποσοστό (46%) των πρωτεϊνών του *S. cerevisiae* που εμφανίζουν ομοιότητα με πρωτεΐνες του ανθρώπου.

Από τη σύγκριση με τα γονιδιώματα του *C. elegans* και της *D. melanogaster* διαπιστώνεται ότι το 60% των προβλεπόμενων με βιοϋπολογιστικά προγράμματα πρωτεϊνών του ανθρώπου έχουν αντίστοιχες ορθόλογες πρωτεΐνες στα δύο είδη οργανισμών. Αντίστοιχα, το 61% των πρωτεϊνών της *D. melanogaster* και το 43% των πρωτεϊνών του *C. elegans* εμφανίζουν ομολογία με ανθρώπινες πρωτεΐνες. Στις ορθόλογες πρωτεΐνες μεταξύ των τριών ειδών συγκαταλέγονται μόρια που ελέγχουν λειτουργίες ενδοκυτταρικής ή διακυτταρικής σηματοδότησης, βασικούς αναπτυξιακούς μηχανισμούς, προγραμματισμένο κυτταρικό θάνατο και μεταγραφική ρύθμιση γονιδίων. Επιπλέον, η σύγκριση των λειτουργικών περιοχών (*domains*) των πρωτεϊνών του ανθρώπου με αυτές του νηματώδη και της δροσόφιλας δείχνει ότι το 90% των *domains* των ανθρώπινων πρωτεϊνών αναγνωρίζονται και στα άλλα δύο είδη. Όμως, στον άνθρωπο παρατηρείται σχεδόν διπλάσιος αριθμός διαφορετικών συνδυασμών αυτών των περιοχών, δημιουργώντας έτσι μεγαλύτερη ποικιλία στην αρχιτεκτονική και στα λειτουργικά χαρακτηριστικά των πρωτεϊνών.

Εξαιρετικό ενδιαφέρον για τη συγκριτική γονιδιωματική ανάλυση έχουν και οι οργανισμοί με ομοιότητες στη φυσιολογία και συμπεριφορά, των οποίων τα γονιδιώματα έχουν εξελιχθεί τόσο ώστε τα μη λειτουργικά στοιχεία τους να έχουν αποκλίνει αρκετά. Αυτός είναι και ένας από τους λόγους που ο ποντικός (*Mus musculus*) ήταν το πρώτο θηλαστικό του οποίου το γονιδίωμα ολοκληρώθηκε μετά από αυτό του ανθρώπου. Ενδεικτικό της ομοιότητας των δύο οργανισμών είναι το υψηλό ποσοστό (80%) των γονιδίων του ποντικού για τα οποία υπάρχουν αντίστοιχα ορθόλογα στον άνθρωπο [Waterston *et al*, 2002a], όπως και οι παραπλήσιες τιμές στο συνολικό μέγεθος του γονιδιώματος και στον αριθμό των γονιδίων των δύο οργανισμών (Πί-

νακας 1). Επιπλέον, ομοιότητες παρατηρούνται και στο επίπεδο της οργάνωσης των χρωμοσωμάτων (Εικ.4), εφόσον το 96% των ορθόλογων γονιδίων βρίσκονται σε συνταϊνικές χρωμοσωμικές περιοχές [Waterston *et al*, 2002a].



Εικ.3 Η βάση δεδομένων Ensembl. Πρόβλεψη και χαρακτηρισμός του γονιδίου PTEN του χρωμοσώματος 10 του ανθρώπου. Παρουσιάζονται ενδεικτικά οι γραφικές απεικονίσεις της χρωμοσωμικής περιοχής (*mapview*) (α), του μεταγράφου (*geneview*) (β) και της πρωτεΐνης (*proteinview*) (γ). Επιπλέον, το πρόγραμμα *Ensembl* δίνει πληροφορίες για τις ορθόλογες πρωτεΐνες σε άλλους οργανισμούς (δ), για τη σχέση του γονιδίου PTEN με γενετικές νόσους (δεδομένα από OMIM) (ε), καθώς και τις σχετικές βιβλιογραφικές αναφορές (στ).

Αξιοσημείωτη υπήρξε η συμβολή της συγκριτικής γονιδιωματικής ανάλυσης στον προσδιορισμό της οργάνωσης εξονίων και ιντρονίων των γονιδίων του ανθρώπου. Συγκρίνοντας την νουκλεοτιδική αλληλουχία ανθρώπινων γονιδίων με το γονιδίωμα του ποντικού, διακρίνονται περιοχές με υψηλή συντήρηση, που συχνά

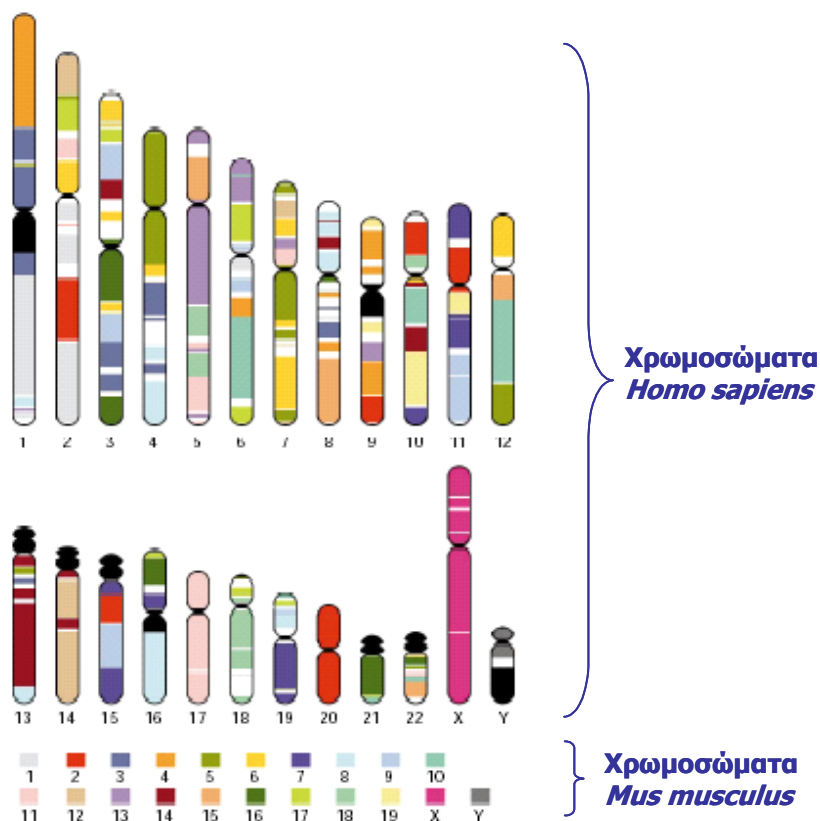
αντιστοιχούν σε εξόνια. Αυτή η προσέγγιση συμβάλλει στην περιγραφή της οργάνωσης γονιδίων που σπάνια αντιπροσωπεύονται σε βιβλιοθήκες cDNA ή εκείνων για τα οποία δεν υπάρχουν πειραματικά δεδομένα και διευκολύνει τη λειτουργία προγραμμάτων πρόβλεψης κωδικών αλληλουχιών από τα αντίστοιχα βιοϋπολογιστικά προγράμματα. Επιπλέον, η σύγκριση των γονιδιωμάτων ανθρώπου και ποντικού με αυτό του ψαριού *Fugu rubripes*, διευκολύνει τις μελέτες γονιδιακής οργάνωσης των θηλαστικών [Gilligan *et al*, 2002]. Το ψάρι *Fugu* έχει ένα από τα μικρότερα γονιδιώματα μεταξύ των σπονδυλωτών, με μέγεθος που αντιστοιχεί στο 1/8 του ανθρώπινου γονιδιώματος (400Mb), αλλά περίπου τον ίδιο αριθμό γονιδίων με τον άνθρωπο [Brenner *et al*, 1993 / Venkatesh *et al*, 2001]. Η διαφορά στο μέγεθος των δύο γονιδιωμάτων οφείλεται στην έλλειψη επαναλαμβανόμενου DNA στο γονιδίωμα του *Fugu* και στην πιο συμπαγή οργάνωση των γονιδίων του. Για τους λόγους αυτούς, το γονιδίωμα του *Fugu rubripes* είναι χρήσιμο μοντέλο συγκριτικής ανάλυσης, επιτρέποντας την ανίχνευση των συντηρημένων κωδικών ή ρυθμιστικών αλληλουχιών στο ανθρώπινο γονιδίωμα. Για παράδειγμα, από τη σύγκριση των δύο γονιδιωμάτων έχουν βρεθεί ανθρώπινα γονίδια που είτε δεν είχαν προηγουμένως εντοπιστεί με τις μεθόδους γονιδιακής πρόβλεψης, είτε αποτελούσαν λανθασμένες προβλέψεις [Gilligan *et al*, 2002]. Επιπλέον, η σύγκριση αυτή διευκολύνει την ταυτοποίηση των ρυθμιστικών αλληλουχιών (υποκινητών, ενισχυτών), λόγω της συντήρησής τους στα δύο σπονδυλωτά, αλλά και της μικρότερης απόστασής τους από την κωδική περιοχή του κάθε γονιδίου στο ψάρι *Fugu*. Χαρακτηριστικό είναι το παράδειγμα της ρυθμιστικής περιοχής του γονιδίου *PAX6*, το οποίο στο γονιδίωμα του *Fugu* απέχει 14Kb από την κωδική αλληλουχία, ενώ στον άνθρωπο η αντίστοιχη περιοχή (με ταυτότητα νουκλεοτιδίων 70%) βρίσκεται σε απόσταση 220Kb από το γονίδιο [Clark *et al*, 2001].

α. Συγκριτική ανάλυση γονιδιωμάτων ανθρώπου και ποντικού

Η ολοκλήρωση της αλληλούχησης του γονιδιώματος των δύο θηλαστικών [Lander *et al*, 2001/ Waterston *et al*, 2002a] επιτρέπει την πραγματοποίηση συγκριτικών αναλύσεων μεγάλης κλίμακας για τη διερεύνηση του ποσοστού συντήρησης στη δομή των χρωμοσωμάτων, στον αριθμό των γονιδίων, στην κωδική και σε μη κωδικές περιοχές των γονιδίων, στην οργάνωση εξονίων-ιντρονίων, σε ρυθμιστικές περιοχές του γονιδιώματος καθώς και στη δομή και τη λειτουργία των πρωτεϊνών.

Συνταινικές περιοχές των χρωμοσωμάτων ανθρώπου και ποντικού

Η κοινή καταγωγή των γονιδιωμάτων του ανθρώπου και του ποντικού διαπιστώνεται σε πρώτο επίπεδο στην οργάνωση των χρωμοσωμάτων των δύο οργανισμών. Αν και κατά τη διάρκεια των 75 εκατομμυρίων ετών που πέρασαν από τον διαχωρισμό των δύο ειδών έχουν συμβεί και στα δύο γονιδιώματα χρωμοσωμικές ανακατατάξεις, λόγω του χαμηλού ρυθμού αυτών των αλλαγών είναι δυνατόν σήμερα να αναγνωρίζονται περιοχές των χρωμοσωμάτων με την ίδια σειρά γονιδίων (συνταινικές) (Εικ.4). Στην πιο πρόσφατη μεγάλης κλίμακας συγκριτική μελέτη των δύο γονιδιωμάτων [Waterston *et al*, 2002a] ανιχνεύθηκαν 558.000 περιοχές μέσου μήκους 500bp, οι οποίες απέχουν μεταξύ τους κατά μέσο όρο 4.4Kb και εμφανίζουν σημαντική ομοιότητα στα δύο είδη. Αξιοσημείωτη είναι η συντήρηση στη διάταξη αυτών των περιοχών που παρατηρείται συγκρίνοντας τα χρωμοσώματα των δύο ειδών. Συνολικά, υπολογίζεται ότι το 90% του ανθρώπινου γονιδιώματος και το 93% του γονιδιώματος του ποντικού αντιστοιχούν σε συνταινικές περιοχές.



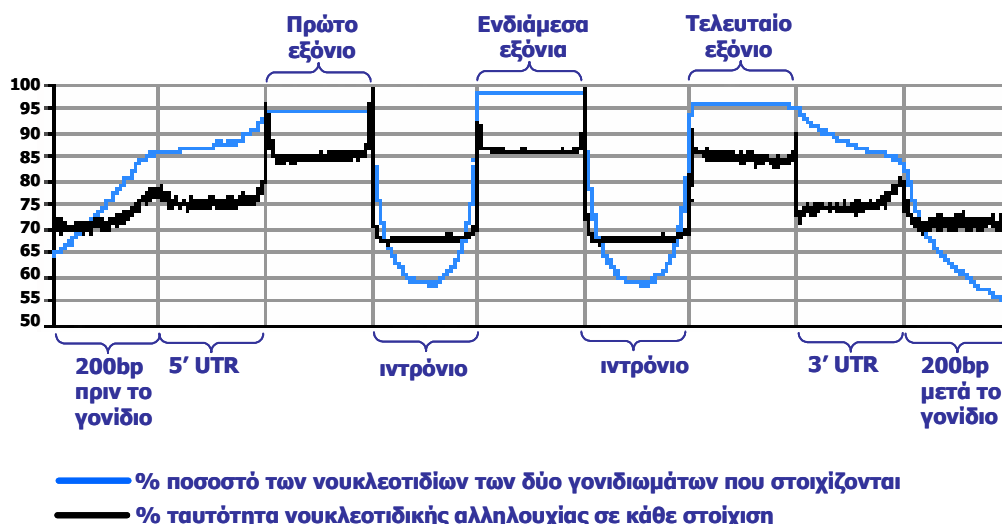
Εικ.4 Συνταινικές περιοχές χρωμοσωμάτων στα γονιδιώματα του ανθρώπου και του ποντικού. Τα 21 διαφορετικά χρώματα αντιστοιχούν σε καθένα από τα χρωμοσώματα του *Mus musculus*. Στα χρωμοσώματα του *Homo sapiens*, οι συνταινικές περιοχές υποδεικνύονται με το χρώμα του αντίστοιχου χρωμοσώματος του ποντικού [Lander ES *et al*, 2001].

Μέγεθος γονιδιώματος – Αριθμός γονιδίων

Το μέγεθος των γονιδιωμάτων του ανθρώπου και του ποντικού και ο αριθμός των γονιδίων που κωδικοποιούν πρωτεΐνες ή RNA στους δύο οργανισμούς ήταν τα πρώτα ερωτήματα μετά τη σχεδόν πλήρη αλληλούχηση των δύο γονιδιωμάτων. Το μέγεθος της ευχρωματινής (~90% της χρωματινής) στο ανθρώπινο γονιδίωμα προσδιορίστηκε στα 2950 Mb, ενώ το αντίστοιχο μέγεθος στον ποντικό είναι 2500 Mb. Για τον αριθμό των γονιδίων είχαν διατυπωθεί πολλές προβλέψεις τα προηγούμενα χρόνια. Οι τελευταίες εκτιμήσεις λαμβάνουν υπόψη τα δεδομένα από την αλληλούχηση των γονιδιωμάτων του ανθρώπου, του ποντικού αλλά και άλλων οργανισμών (δροσόφιλα, *C.elegans* κλπ), σε συνδυασμό με τα αποτελέσματα αλληλούχησης κλώνων cDNA και τις *de novo* προβλέψεις υπολογιστικών προγραμμάτων. Σύμφωνα με αυτές τις εκτιμήσεις, για το γονιδίωμα του ανθρώπου υπολογίζονται 30.000 – 40.000 γονίδια, ενώ για τον ποντικό 27.500 – 30.000 [Lander *et al*, 2001 / Waterston *et al*, 2002a].

Συντήρηση σε περιοχές γονιδίων

Από τις συγκρίσεις μεταξύ των γονιδιωμάτων ανθρώπου και ποντικού διαπιστώνεται, όπως αναμενόταν, μεγαλύτερη συντήρηση της νουκλεοτιδικής αλληλουχίας στις περιοχές των γονιδίων. Συγκεκριμένα, οι ερευνητές του MGSC (*Mouse Genome Sequencing Consortium*) επέλεξαν 3165 γονίδια με γνωστή θέση στο ανθρώπινο γονιδίωμα και για αυτά πραγματοποίησαν συγκρίσεις κωδικών και μη κωδικών περιοχών με το γονιδίωμα του ποντικού [Waterston *et al*, 2002a]. Αποτελέσματα αυτής της μελέτης φαίνονται στην Εικ.5. Στις κωδικές περιοχές των γονιδίων, το μέσο ποσοστό ταυτότητας νουκλεοτιδίων είναι 85%. Αξιοσημείωτο είναι



Εικ.5 Συντήρηση της νουκλεοτιδικής αλληλουχίας στην περιοχή των ορθόλογων γονιδίων μεταξύ των γονιδιωμάτων ανθρώπου και ποντικού. Παρουσιάζεται ο μέσος όρος των ποσοστών ταυτότητας νουκλεοτιδίων (μαύρη καμπύλη) κατά μήκος των ορθόλογων γονιδίων που συγκρίθηκαν στη μελέτη των Waterston *et al* (2002a). Μεγαλύτερα ποσοστά συντήρησης παρατηρούνται κατά μήκος των εξονίων και στα άκρα των ιντρονίων (ιδιαίτερα στο 5' άκρο).

και το υψηλό ποσοστό (52%) των κωδικών αλληλουχιών που εμφανίζονται δύο φορές στο ανθρώπινο γονιδίωμα, το οποίο πιθανόν εκφράζει την παρουσία παράλογων γονιδίων ή ψευδογονιδίων. Αντίθετα, στα ιντρόνια δεν παρατηρήθηκε ιδιαίτερη συντήρηση. Συγκεκριμένα, το ποσοστό νουκλεοτιδικής ταυτότητας στα δύο γονιδιώματα για τις περιοχές των ιντρονίων δεν διαφέρει σημαντικά από αυτό που προκύπτει από τις γενικευμένες συγκρίσεις των δύο γονιδιωμάτων (~70%). Υψηλά επίπεδα συντήρησης παρατηρούνται και στα σημεία συναρμογής των εξονίων (*splice junctions*) κατά την επεξεργασία του ανώριμου mRNA, είτε όταν αυτά είναι κανονικά (GT-AG) ή μη κανονικά (GC-AG και 5'-RTATCCTY). Τέλος, οι περιοχές 5'UTR και 3'UTR εμφανίζουν ενδιάμεση συντήρηση, με ποσοστό νουκλεοτιδικής ταυτότητας 75%. Σε μια ανάλογη σύγκριση 95 πειραματικά χαρακτηρισμένων ρυθμιστικών στοιχείων ιστοειδικής ή γενικευμένης γονιδιακής έκφρασης, παρατηρήθηκε συντήρηση σε επίπεδα συγκρίσιμα με αυτήν των περιοχών 5' και 3' UTR. Οι περιοχές αυτές βρέθηκαν στο μεγαλύτερο ποσοστό τους (85%) σε απόσταση μικρότερη των 2Kb από τον υποκινητή του αντίστοιχου γονιδίου, σε ιντρόνια (10%) ή σε απόσταση μεγαλύτερη των 2Kb από τον υποκινητή (5%). Επιπλέον, το 19% των ρυθμιστικών στοιχείων που μελετήθηκαν επικαλύπτονταν και στα δύο γονιδιώματα με νησίδια CpG.

Συντήρηση στον αριθμό και το μέγεθος των εξονίων

Από ανάλογες συγκριτικές μελέτες, διαπιστώνεται σημαντική συντήρηση στον αριθμό των κωδικών εξονίων μεταξύ ορθόλογων γονιδίων των δύο ειδών. Συγκεκριμένα, συγκρίνοντας 1506 ζεύγη ορθόλογων γονιδίων της βάσης δεδομένων *RefSeq* (NCBI), διαπιστώνεται ότι το 86% των ζευγών έχουν ίδιο αριθμό κωδικών εξονίων. Αντίθετα, μελετώντας το σύνολο των εξονίων, το αντίστοιχο ποσοστό είναι 62%, λόγω της μειωμένης συντήρησης στον αριθμό των μη κωδικών εξονίων. Το συνολικό μήκος της κωδικής αλληλουχίας είναι ίσο για το 46% των ζευγών ορθόλογων γονιδίων, ενώ από τη σύγκριση των 10.061 ζευγών εξονίων που περιέχονται στα γονίδια με ίσο αριθμό κωδικών εξονίων, βρέθηκε ότι το 91% των εξονίων έχουν ταυτόσημο μήκος σε άνθρωπο και ποντικό. Από τα 930 ζεύγη εξονίων με διαφορετικό μέγεθος, τα 858 έχουν διαφορετικό αριθμό νουκλεοτιδίων πολλαπλάσια του τρία, όπως αναμενόταν, λόγω των περιορισμών της μετάφρασης.

Αντίθετα, μόνο το 1% των ιντρονίων ορθόλογων γονιδίων έχουν ταυτόσημο μήκος, αν και διατηρείται η αναλογία των μεγεθών των διαφορετικών ιντρονίων ενός γονιδίου στα δύο είδη. Διαπιστώθηκε ότι το μέσο μέγεθος των ιντρονίων του ποντικού (3.9Kb) είναι μικρότερο από αυτό του ανθρώπου (4.6Kb), γεγονός που εν μέρει δικαιολογεί και τη διαφορά μεγέθους των δύο γονιδιωμάτων.

Συντήρηση σε μη κωδικές αλληλουχίες

Στα γονιδιώματα του ανθρώπου και του ποντικού έχουν πρόσφατα εντοπιστεί αλληλουχίες που παρόλο που δεν κωδικοποιούν πρωτεΐνη ή RNA εμφανίζουν εξαιρετική συντήρηση μεταξύ των θηλαστικών. Συγκεκριμένα, στο χρωμόσωμα 21 του ανθρώπου ανιχνεύθηκαν 191 αλληλουχίες CNG (*Conserved Non-Genic*), οι οποίες είναι συντηρημένες στα 14 είδη θηλαστικών που μελετήθηκαν και τα οποία περιελάμβαναν απομακρυσμένες φυλογενετικές ομάδες, από πρωτεύοντα μέχρι μαρσιποφόρα και μονοτρήματα [Dermitzakis *et al*, 2003]. Σύμφωνα με την ίδια μελέτη, η συντήρηση που παρατηρείται στις αλληλουχίες CNG είναι υψηλότερη από αυτήν των κωδικών αλληλουχιών και συγκρίσιμη με αυτή γνωστών θέσεων πρόσδεσης πρωτεϊνών. Ο ρόλος των CNGs στο γονιδίωμα, που θα δικαιολογούσε και την υψηλή τους συντήρηση, είναι άγνωστος. Σύμφωνα όμως με τα παραπάνω, είναι πιθανόν η λειτουργία αυτού του νέου είδους αλληλουχιών να σχετίζεται με την πρόσδεση πρωτεϊνών σε αυτές.

6. Συγκριτική πρωτεωμική ανάλυση με στόχο τη διευκρίνιση της λειτουργίας

Η ολοκλήρωση της αλληλούχησης του γονιδιώματος πολλών οργανισμών, σε συνδυασμό με την οργάνωση αυτής της πληροφορίας σε βάσεις δεδομένων και την επεξεργασία της με κατάλληλα υπολογιστικά προγράμματα, επιτρέπει την πραγματοποίηση συγκρίσεων και στο επίπεδο των γονιδιακών προϊόντων και την ταξινόμηση των πρωτεϊνών ενός οργανισμού ανάλογα με το είδος των λειτουργικών περιοχών (*domains*) που περιέχουν ή την εξελικτική τους συντήρηση κατά τη φυλογένεση, που υποδεικνύεται από την ομοιότητά τους με πρωτεΐνες άλλων οργανισμών. Για τον σκοπό αυτό δημιουργήθηκε η βάση δεδομένων *InterPro* (www.ebi.ac.uk/interpro/), η οποία αποτελεί έναν κατάλογο των γνωστών οικογενειών πρωτεϊνών και των λειτουργικών πρωτεϊνικών περιοχών (*domains*).

α. Λειτουργική σύγκριση πρωτεϊνών ανθρώπου και ποντικού

Με βάση τα αποτελέσματα της διερεύνησης της *InterPro* και τα τυχόν διαθέσιμα πειραματικά δεδομένα, οι πρωτεΐνες που κωδικοποιούνται από τα γονιδιώματα του ανθρώπου και του ποντικού ταξινομούνται σε λειτουργικές ομάδες σύμφωνα με τους όρους του συστήματος ταξινόμησης πρωτεϊνών '*Gene Ontology*' (GO, www.geneontology.org/), που χρησιμοποιούνται για να περιγράψουν το κυτταρικό διαμέρισμα εντοπισμού της κάθε πρωτεΐνης, τον μοριακό της ρόλο και τη βιολογική λειτουργία στην οποία συμμετέχει [Waterston *et al*, 2002a]. Συγκρίνοντας την κατανομή των πρωτεϊνών του ανθρώπου και του ποντικού σε αυτές τις λειτουργικές ομάδες διαπιστώνουμε σημαντικές ομοιότητες στα δύο είδη (Εικ.6). Αναλυτικότερα, παρατηρούνται και στα δύο είδη υψηλά ποσοστά (40-50%) μεμβρανικών πρωτεϊνών, ενώ οι περισσότερες

πρωτεΐνες και των δύο οργανισμών φαίνεται να συμμετέχουν σε μεταβολικές διαδικασίες. Τέλος, τόσο οι πρωτεΐνες που κωδικοποιούνται από το γονιδίωμα του ανθρώπου, όσο και αυτές που κωδικοποιούνται από το γονιδίωμα του ποντικού, προσδένουν ή μεταφέρουν μόρια-συνδέτες (*ligands*), σε ποσοστό 65% και 50% αντίστοιχα, ή εμφανίζουν ενζυμική δράση σε ποσοστό 40-45%.

β. Εξελικτική συντήρηση πρωτεϊνών ανθρώπου και ποντικού

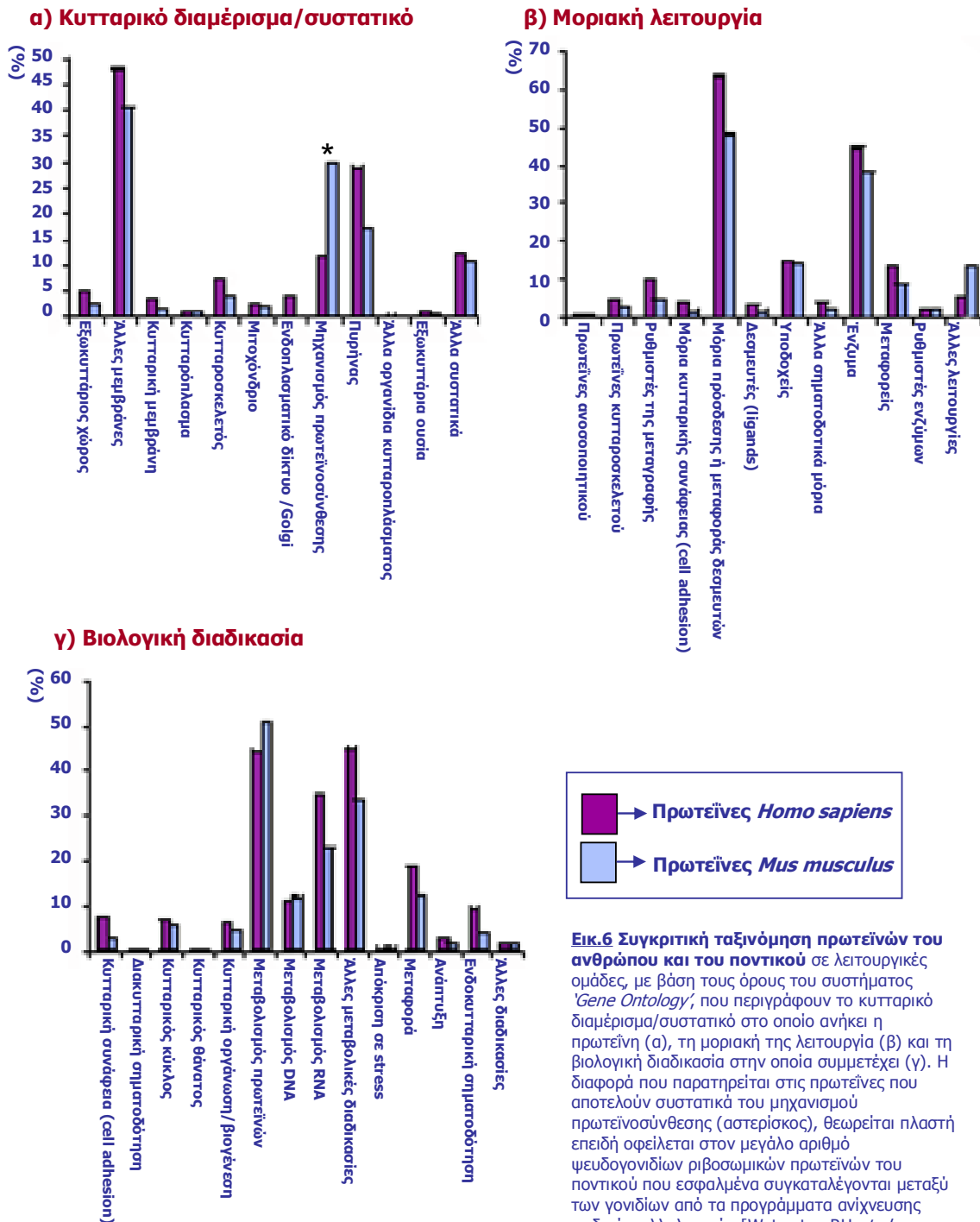
Στους ευκαρυωτικούς οργανισμούς η εξέλιξη των πρωτεϊνών φαίνεται να είναι αποτέλεσμα κυρίως δύο μηχανισμών: i) του συνδυασμού συντηρημένων λειτουργικών περιοχών (*domains*) με διαφορετικούς εναλλακτικούς τρόπους για τη δημιουργία νέων δομών, και ii) της επέκτασης των γονιδιακών οικογενειών με τη δημιουργία νέων μελών, που προκύπτουν από διπλασιασμό των προγονικών γονιδίων. Η μελέτη και των δύο αυτών μηχανισμών διευκολύνεται συγκρίνοντας 'συγγενικά' γονιδιώματα, όπως αυτά του ποντικού και του ανθρώπου. Η μεγάλη ομοιότητα των δύο γονιδιωμάτων επιτρέπει τη σύγκριση μεταξύ των ορθόλογων πρωτεϊνών, οι οποίες προέκυψαν από το ίδιο γονίδιο του κοινού προγόνου των δύο θηλαστικών, επιτρέποντας τη διερεύνηση των εξελικτικών πιέσεων που ασκήθηκαν στις διάφορες ομάδες αυτών των πρωτεϊνών κατά την ειδογένεση, και οι οποίες οδήγησαν στον διαχωρισμό των πρωτευόντων και των τρωκτικών. Ενδιαφέρον παρουσιάζει και η σύγκριση των πρωτεϊνών του ανθρώπου ή του ποντικού με τις πρωτεΐνες άλλων οργανισμών (πχ των βακτηρίων, του σακχαρομύκητα, του νηματώδη ή της δροσόφιλας), από την οποία μπορούν να εντοπιστούν οι πρωτεΐνες με συντηρημένη λειτουργία κατά την εξέλιξη, αλλά και αυτές που διαφοροποιούν τα σπονδυλωτά από τα υπόλοιπα ζώα. Οι διαφοροποιήσεις μεταξύ των πρωτεϊνών του ανθρώπου και του ποντικού, κυρίως στον τομέα της επέκτασης των γονιδιακών οικογενειών, είναι δυνατόν να προβάλουν τις διαφορές στη φυσιολογία, ανατομία και συμπεριφορά των δύο ειδών, στο επίπεδο του γονιδιώματος.

Με βάση τα παραπάνω, έχουν πραγματοποιηθεί συγκρίσεις όλων των προβλεπόμενων ή γνωστών πρωτεϊνών του ποντικού με τις πρωτεΐνες άλλων οργανισμών της βάσης δεδομένων 'nr' (*non-redundant*) του NCBI [Waterston *et al*, 2002a], και αντίστοιχη ανάλυση για τις πρωτεΐνες που κωδικοποιούνται από το ανθρώπινο γονιδίωμα [Lander *et al*, 2001]. Οι ομόλογες πρωτεΐνες που εντοπίστηκαν στις δύο μελέτες, ταξινομούνται ανάλογα με τη φυλογενετική ομάδα του οργανισμού στον οποίο ανήκουν. Και οι δύο συγκριτικές μελέτες καταλήγουν σε συγκρίσιμα πρότυπα κατανομής των ομόλογων πρωτεϊνών του ανθρώπου και του ποντικού στις διάφορες φυλογενετικές ομάδες. Κοινή διαπίστωση των δύο μελετών είναι ο χαμηλός ρυθμός ανακάλυψης νέων πρωτεϊνών ή νέων *domains* για κάθε φυλογενετική ομάδα. Ενδεικτικό είναι το εξαιρετικά χαμηλό ποσοστό (<1%) των πρωτεϊνών του ποντικού που εμφανίζονται αποκλειστικά και μόνο στα τρωκτικά (Εικ.7), αλλά και το μικρό ποσοστό (7%) των πρωτεϊνών του ανθρώπου που εμφανίζονται αποκλειστικά στα σπονδυλωτά, τα οποία στις συγκεκριμένες μελέτες αντιπροσωπεύονται από τον άνθρωπο και τον ποντικό. Η τελευταία κατηγορία των πρωτεϊνών που φαίνεται να 'εφευρέθηκαν' από τη φύση ειδικά για τα σπονδυλωτά, περιλαμβάνει μόρια με δράση στο ανοσοποιητικό και στο νευρικό σύστημα των ζώων [Lander *et al*, 2001].

Σε αντίθεση με τη χαμηλή συχνότητα 'νεοεμφανιζόμενων' πρωτεϊνών στα σπονδυλωτά, διαπιστώνεται υψηλή συχνότητα πρωτεϊνών με καινούρια αρχιτεκτονική, οι οποίες εμφανίζουν νέους συνδυασμούς των προϋπάρχοντων συντηρημένων *domains*. Συγκεκριμένα, στις πρωτεΐνες του ανθρώπου παρατηρούνται σχεδόν διπλάσιοι συνδυασμοί των *domains* σε σχέση με αυτές της δροσόφιλας ή του νηματώδη, και εξαπλάσιοι σε σχέση με τις πρωτεΐνες του σακχαρομύκητα. Αυτή η διαφορά είναι περισσότερο εμφανής στην πρόσφατη εξέλιξη νέων διαμεμβρανικών ή εξωκυττάρων δομών στις πρωτεΐνες του ανθρώπου και του ποντικού.

Η δεύτερη διαπίστωση για την εξέλιξη των πρωτεϊνών των σπονδυλωτών που προκύπτει από τις πρόσφατες συγκριτικές μελέτες είναι η μεγαλύτερη τάση επέκτασης γονιδιακών οικογενειών που παρατηρείται στα σπονδυλωτά σε σχέση με τους υπόλοιπους ευκαρυωτικούς οργανισμούς. Συγκεκριμένα, διαπιστώνεται ότι υψηλό ποσοστό (60%) των γονιδιακών οικογενειών που εμφανίζονται στα ευκαρυωτικά, απαρτίζονται από μεγαλύτερο αριθμό μελών σε σχέση με τον αριθμό των μελών τους σε άλλα είδη. Επομένως, φαίνεται ότι οι διπλασιασμοί γονιδίων αποτέλεσαν σημαντικό εξελικτικό μηχανισμό για τα σπονδυλωτά. Πολλές από τις οικογένειες πρωτεϊνών που με αυτόν τον μηχανισμό έχουν αποκτήσει περισσότερα μέλη στα σπονδυλωτά σε σχέση με τη δροσόφιλα ή τον νηματώδη, συμμετέχουν σε διαδικασίες που χαρακτηρίζουν τη φυσιολογία των σπονδυλωτών. Ενδεικτικό παράδειγμα είναι τα πολυάριθμα μέλη της οικογένειας των πρωτεϊνών με *domains* ανοσοσφαιρίνης (Ig) στον άνθρωπο και στον ποντικό, σε σχέση με την απουσία της συγκεκριμένης οικογένειας από τον σακχαρομύκητα και τα φυτά και την παρουσία του *domain* Ig μόνο σε ορισμένα μόρια στην επιφάνεια των κυττάρων των ασπονδύλων. Η επέκταση αυτής της οικογένειας στα σπονδυλωτά συνοδεύτηκε και με λειτουργική απόκλιση των μελών της, ώστε να δημιουργηθεί μεγαλύτερη ποικιλία στη λειτουργία των πρωτεϊνών με *domain* Ig, με αποτέλεσμα αυτή η οικογένεια μορίων να ανταποκριθεί στην αυξημένη πολυπλοκότητα που απαιτείται στο ανοσοποιητικό σύστημα των σπονδυλωτών για την αντιμετώπιση πολυάριθμων αντιγόνων. Αντίστοιχα, το γονιδίωμα του ανθρώπου και του ποντικού κωδικοποιεί περίπου 30 μέλη της οικογένειας του αυξητικού παράγοντα ινοβλαστών (*Fibroblast Growth Factor*, FGF), σε αντίθεση με τα δύο μέλη της οικογένειας στη δροσόφιλα και τον νηματώδη. Τα μόρια αυτά στα σπονδυλωτά ρυθμίζουν την οργανογένεση πολλών διαφορετικών οργάνων, μεταξύ των οποίων και του πνεύμονα, ενώ στη μύγα η μία από τις δύο πρωτεΐνες-μέλη της οικογένειας FGF (η *branchless*) ρυθμίζει την ανάπτυξη του αναπνευστικού συστήματος των εμβρύων. Επομένως, οι αναπτυξιακές διεργασίες που οδήγησαν στη μορφογένεση των σπονδυλωτών φαίνεται να έχουν εξελιχθεί από συγγενικά αλλά απλούστερα συστήματα των ασπονδύλων. Τέλος, ενδιαφέρον παρουσιάζει και μια ομάδα πρωτεϊνών που αντιστοιχεί στο 1% περίπου του ανθρώπινου πρωτεώματος, οι οποίες εμφανίζονται κοινές μόνο μεταξύ σπονδυλωτών και προκαρυωτικών οργανισμών, χωρίς να υπάρχουν σε κανένα άλλο ευκαρυωτικό σύστημα. Η θεωρία που επικρατεί για τα γονίδια που τις κωδικοποιούν είναι ότι αυτά εισήλθαν στο γονιδίωμα των σπονδυλωτών ως αποτέλεσμα οριζόντιας μεταφοράς από τα βακτήρια [Lander *et al*, 2001]. Οι περισσότερες από αυτές τις πρωτεΐνες είναι ένζυμα που καταλύουν τον μεταβολισμό χημικών ουσιών και πιθανώς να πρόσφεραν εξελικτικό πλεονέκτημα στα κύτταρα που αρχικά δέχτηκαν τα αντίστοιχα γονίδια. Ένα παράδειγμα που δείχνει το πόσο έχει παγιωθεί ο ρόλος αυτών των πρωτεϊνών στο σύστημα των σπονδυλωτών ρυθμίζοντας κρίσιμες λειτουργίες είναι αυτό της οξειδάσης μονοαμινών (MAO), ενζύμου της εξωτερικής μεμβράνης των μιτοχονδρίων, που ευθύνεται για τον μεταβολισμό

νευροδιαβιβαστών και είναι γνωστό για τις μεταλλαγές που προκαλούν νευρολογικές διαταραχές στον άνθρωπο.



Για τη διερεύνηση πιο πρόσφατων εξελικτικών γεγονότων, που οδήγησαν στη διαφοροποίηση των συστημάτων του ανθρώπου και του ποντικού, είναι χρήσιμα τα συμπεράσματα από τις συγκρίσεις ανάμεσα στις πρωτεΐνες των δύο ειδών. Στις μελέτες αυτές παρατηρείται περίπου ίδιος αριθμός μελών στις οικογένειες πρωτεϊνών ή λειτουργικών περιοχών (*domains*) μεταξύ των δύο ειδών, ενώ φαίνεται ότι ο μοριακός ρόλος

του κάθε είδους *domain* επηρεάζει την εξελικτική του πορεία. Αναλυτικότερα, από τα συγκριτικά δεδομένα των Waterston *et al* (2002a) προκύπτει ότι μεγαλύτερη συντήρηση στα δύο είδη εμφανίζουν οι καταλυτικές



Εικ.7. Ταξινόμηση των πρωτεϊνών άλλων οργανισμών που εμφανίζουν ομολογία με αυτές του ποντικού. Μόνο ένα μικρό ποσοστό πρωτεϊνών (<1%) είναι ειδικές για τα τρωκτικά, σε σύγκριση με αυτές που είναι κοινές μεταξύ των θηλαστικών (και μη τρωκτικών, 14%), των χορδωτών (και μη θηλαστικών, 6%), των μεταζώων (και μη χορδωτών, 27%), των ευκαρυωτικών οργανισμών (και μη μεταζώων, 29%) και των προκαρυωτικών ή άλλων οργανισμών (και μη ευκαρυωτικών, 23%) [Waterston RH *et al*, 2002a]. Παρόμοιο πρότυπο κατανομής ακολουθούν και οι ομόλογες πρωτεΐνες με αυτές του ανθρώπου [Lander ES *et al*, 2001].

περιοχές των ενζύμων σε σχέση με τις μη καταλυτικές, ενώ λιγότερο συντηρημένες είναι οι λειτουργικές περιοχές εκκρινόμενων πρωτεϊνών, που αντιπροσωπεύουν κυρίως λειτουργίες του ανοσοποιητικού συστήματος, στις οποίες παρατηρείται διαφοροποίηση των δύο συστημάτων λόγω διαφορετικής έκθεσης σε αντιγόνα κατά την εξέλιξη των δύο ειδών (Πίνακας 4). Μεγαλύτερες διαφορές μεταξύ των δύο ειδών παρατηρούνται σε οικογένειες πρωτεϊνών με λειτουργίες στο ανοσοποιητικό και στο αναπαραγωγικό σύστημα των δύο οργανισμών. Για παράδειγμα, στο γονιδίωμα του ποντικού έχουν εντοπιστεί 25 οικογένειες πρωτεϊνών, οι οποίες στον άνθρωπο αντιπροσωπεύονται από ένα μόνο μέλος. Σε αυτές περιλαμβάνονται 14 οικογένειες πρωτεϊνών που συμμετέχουν στην αναπαραγωγή των τρωκτικών. Σε αυτήν την κατηγορία ανήκουν μόρια που ρυθμίζονται μεταγραφικά από ανδρογόνα και συμμετέχουν στη βιοσύνθεση ή στον μεταβολισμό ορμονών, στην ανάπτυξη του πλακούντα, στην ωρίμανση των γονάδων και των σπερματοζωαρίων ή στην σηματοδότηση από φερομόνες για την επιλογή συντρόφου. Το γεγονός ότι οι περισσότερες οικογένειες πρωτεϊνών που εμφανίζουν επιλεκτική επέκταση στον ποντικό σχετίζονται με την αναπαραγωγή πιθανόν αντικατοπτρίζει τις φυσιολογικές διαφορές των δύο ειδών, που είναι πολύ πιο εμφανείς στο συγκεκριμένο σύστημα και αφορούν κυρίως τη δομή του πλακούντα, τα μεγέθη των νεογέννητων και την περίοδο κύησης. Πιθανόν τα ιδιαίτερα χαρακτηριστικά του αναπαραγωγικού συστήματος του ποντικού να ήταν υπεύθυνα για την εξελικτική πίεση που ασκήθηκε στο γονιδίωμά του και η ανάγκη για δημιουργία νέων γονιδίων να καλύφθηκε από τον διπλασιασμό ήδη υπάρχοντων.

Συμπερασματικά, οι συγκριτικές μελέτες υποδηλώνουν τη σημαντική ομοιότητα τόσο στη δομή όσο και στη λειτουργία των πρωτεϊνών του ανθρώπου και του ποντικού, η οποία εκφράζει και τις ομοιότητες στη φυσιολογία των δύο οργανισμών. Οι αναλύσεις αυτές στηρίζουν περισσότερο την επιλογή του *Mus musculus* ως το ιδανικό πειραματικό μοντέλο για τη μελέτη της λειτουργίας μεγάλου αριθμού γονιδίων του ανθρώπου, ιδιαίτερα εκείνων που σχετίζονται με λειτουργίες χαρακτηριστικές των σπονδυλωτών. Επιπλέον, από τις συγκρίσεις των πρωτεϊνών σπονδυλωτών και ασπονδύλων διαπιστώνεται αυξημένη πολυπλοκότητα του πρωτεώματος των σπονδυλωτών κατά 2-3 φορές, όχι μόνο λόγω του μεγέθους του γονιδιώματός τους, αλλά και ως

συνέπεια επέκτασης συγκεκριμένων οικογενειών γονιδίων [Lander *et al*, 2001]. Το κατά πόσο αυτή η αύξηση στην πρωτεωμική πολυπλοκότητα μπορεί να εξηγήσει και τη διαφορά στη φαινοτυπική πολυπλοκότητα των δύο συστημάτων είναι ένα ενδιαφέρον ερώτημα. Πιθανώς η πραγματική εξήγηση να περιλαμβάνει μηχανισμούς που πολλαπλασιάζουν αυτές τις αρχικές διαφορές, όπως είναι η εναλλακτική συναρμογή των εξονίων κατά την ωρίμανση του πρόδρομου mRNA και ορισμένες μετα-μεταφραστικές τροποποιήσεις των πρωτεϊνών, που συναντώνται με μεγαλύτερη συχνότητα στα σπονδυλωτά, αυξάνοντας περισσότερο την πολυπλοκότητα σε αυτό το σύστημα.

Πίνακας 4. Ομοιότητα πρωτεϊνών ανθρώπου και ποντικού στις περιοχές των *domains*

Περιοχές πρωτεϊνών που συγκρίνονται	Μέση % ταυτότητα αμινοξέων ανθρώπου - ποντικού
Πλήρους μήκους πρωτεΐνες	78.5
Περιοχές με <i>domains</i>	93.5
Περιοχές χωρίς <i>domains</i>	71.1
Καταλυτικά <i>domains</i>	96.6
Μη καταλυτικά <i>domains</i>	94.9
Εκκρινόμενα <i>domains</i>	88.9
Κυτταροπλασματικά <i>domains</i>	96.7
Πυρηνικά <i>domains</i>	98.6

Σημ: Ο μοριακός ρόλος κάθε *domain* φαίνεται να επηρεάζει τη συντήρησή του μεταξύ των δύο ειδών. Περισσότερο συντηρημένα είναι τα καταλυτικά σε σχέση με τα μη καταλυτικά *domains*, όπως και τα *domains* των ενδοκυττάρων πρωτεϊνών σε σχέση με αυτά των εκκρινόμενων.

Επομένως, είναι φανερό ότι η λειτουργική μελέτη πολλών γονιδίων των σπονδυλωτών για τα οποία υπάρχουν γνωστά ορθόλογα σε ασπόνδυλα δεν είναι αποτέλεσμα απλής αναγωγής. Είναι αναγκαίο να λαμβάνονται υπόψη πιθανά γεγονότα επέκτασης του συγκεκριμένου γονιδίου και να διερευνάται η ρύθμιση της έκφρασης και οι αλληλεπιδράσεις των αντίστοιχων πρωτεϊνών της οικογένειας, ώστε να κατανοηθεί πλήρως ο ρόλος τους στη φυσιολογία των σπονδυλωτών. Το σύστημα του ποντικού προσφέρεται ως μοντέλο για τη λειτουργική μελέτη των ανθρώπινων γονιδίων, καθώς επιτρέπει την πραγματοποίηση πειραμάτων για την περιγραφή του αναπτυξιακού και ιστοειδικού προτύπου γονιδιακής έκφρασης, συγκριτικές μελέτες με πρότυπα έκφρασης άλλων γονιδίων αλλά και πειράματα στοχευμένης απενεργοποίησης (*knock-out*) ή εισαγωγής τροποποιημένων αλληλομόρφων (*knock-in*) του γονιδίου [Müller, 1999/ Sauer & Henderson, 1988].

Στα πλαίσια αυτά, ασχοληθήκαμε με την μελέτη της δομής και τη λειτουργική ανάλυση του γονιδίου *Neuralized* (*Neurl*) του ποντικού. Στη συνέχεια παρουσιάζονται εισαγωγικά στοιχεία για το γονίδιο αυτό, η μελέτη του οποίου απετέλεσε ένα μεγάλο μέρος αυτής της διατριβής.

7. Το γονίδιο *Neutralized*

Ένα μεγάλο μέρος της παρούσας διατριβής αφορά τη δομή και τη λειτουργία του γονιδίου *Neutralized* (*Neurl*) του ποντικού, ως μοντέλο ανάλυσης του ανθρώπινου γονιδίου *NEURL*. Το *NEURL* χαρτογραφείται στη ζώνη 10q24.33 και στην αντίστοιχη συνταϊνική ζώνη D2 του χρωμοσώματος 19 του ποντικού (Εικ.8). Το *Neurl* αποτελεί μέλος της οικογένειας των γονιδίων *Neutralized* που εμφανίζονται συντηρημένα στους πολυκύτταρους ζωικούς οργανισμούς. Τα μόρια *Neutralized* που έχουν χαρακτηριστεί μοριακά ή/και λειτουργικά μέχρι σήμερα είναι της δροσόφιλας [Price *et al*, 1993 / Pavlouros *et al*, 2001 / Lai *et al*, 2001], του ανθρώπου [Nakamura *et al*, 1998], του ποντικού [Pavlouros, Kokkinaki *et al*, 2002], του αρουραίου [Timmusk *et al*, 2002] και του βατράχου [Deblandre *et al*, 2001]. Επιπλέον, με βάση τα αποτελέσματα της αλληλούχησης γονιδιωμάτων άλλων οργανισμών, στις βάσεις δεδομένων εντοπίζονται και τα δομικά ομόλογα μόρια του ψαριού *Fugu rubripes*, του εντόμου *Anopheles gambiae* και του νηματώδη *Caenorhabditis elegans*. Εισαγωγικά, παρουσιάζονται παρακάτω τα γνωστά μόρια της οικογένειας *Neutralized* και τα στοιχεία που είναι γνωστά για τη λειτουργία τους σε κάθε σύστημα-μοντέλο.

α. Η αρχιτεκτονική της πρωτεΐνης *Neurl*

Η πρωτεΐνη *Neurl* εμφανίζεται αρκετά συντηρημένη στους ζωικούς οργανισμούς από τον νηματώδη και τη δροσόφιλα μέχρι τα θηλαστικά. Συγκρίνοντας τις ορθόλογες πρωτεΐνες *Neurl* των θηλαστικών *Mus musculus* ή *Homo sapiens* με αυτές του εντόμου *Drosophila melanogaster* και του νηματώδη *Caenorhabditis elegans*, διαπιστώνεται ταυτότητα αμινοξέων σε ποσοστό 27% και 30% του συνολικού μήκους της πρωτεΐνης, αντίστοιχα (Εικ.9). Στις πρωτεΐνες της οικογένειας *Neutralized* διακρίνονται τρεις συντηρημένες περιοχές. Οι δύο από αυτές είναι επαναλήψεις 122-124 αμινοξέων, γνωστές ως περιοχές NEUZ1 και NEUZ2 ενώ η τρίτη περιοχή, μήκους 39-41 αμινοξέων στο C-τελικό άκρο του μορίου, αποτελεί χαρακτηριστικό μοτίβο πρόσδεσης ψευδαργύρου σε αμινοξικά κατάλοιπα κυστεΐνης (C) και ιστιδίνης (H), γνωστό ως 'δακτύλιος RING' με δομή C₃HC₄.

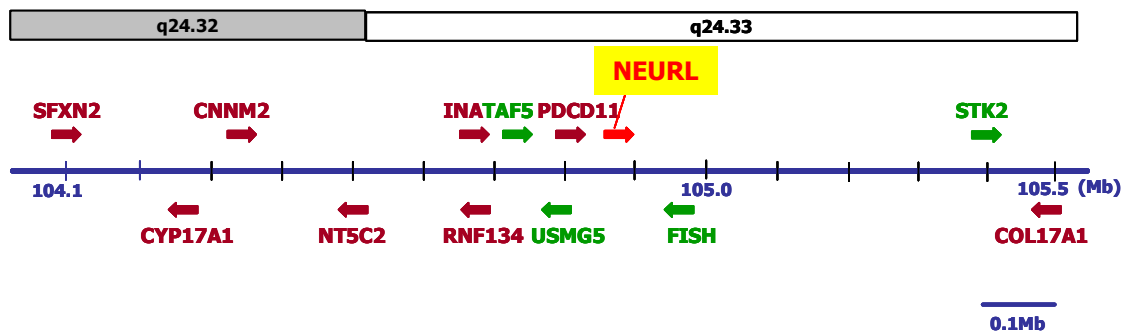
Περιοχές NEUZ1 και NEUZ2

Ο ακριβής ρόλος των περιοχών NEUZ1 και NEUZ2 στη λειτουργία των μορίων *Neutralized* δεν είναι πλήρως γνωστός. Σύμφωνα με πειράματα παροδικής έκφρασης χιμαιρικών μορίων *Neurl* του ποντικού σε κύτταρα θηλαστικών COS7, η παρουσία της περιοχής NEUZ1 είναι απαραίτητη για το ειδικό πρότυπο υποκυτταρικού εντοπισμού του μορίου κυρίως στο κυτταρόπλασμα και στην περιπυρηνική περιοχή [Pavlouros, Kokkinaki *et al*, 2002]. Επιπλέον, αποτελέσματα πειραμάτων μελέτης πρωτεϊνικών συμπλόκων του *neur* της δροσόφιλας [Lai *et al*, 2001], σε συμφωνία και με δικές μας παρατηρήσεις, υποστηρίζουν ότι οι περιοχές NEUZ είναι απαραίτητες για τις αλληλεπιδράσεις του μορίου *neutralized* (βλ. Αποτελέσματα).

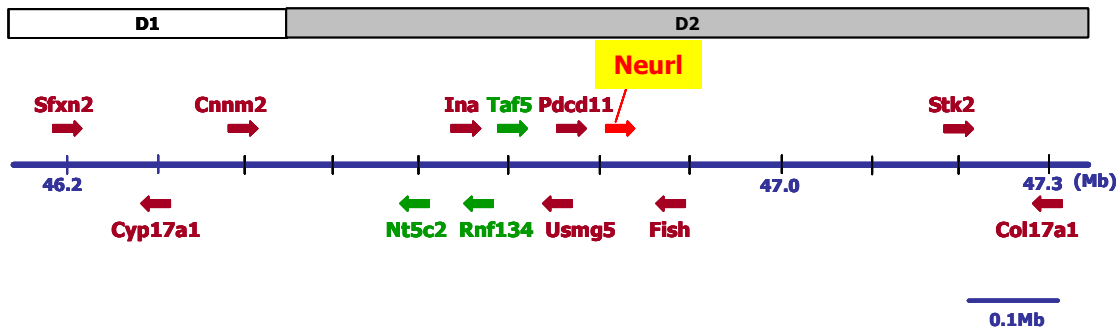
Περιοχές επαναλήψεων NEUZ συναντώνται και σε άλλες πρωτεΐνες εκτός της οικογένειας των μορίων *Neutralized*, όπως φαίνεται από την έρευνα στη βάση δεδομένων SMART (Εικ.10). Συγκεκριμένα, 3-6 αντιγραφα των επαναλήψεων NEUZ συναντώνται σε πρωτεΐνες που έχουν προκύψει από τη μετάφραση ανοικτών πλαισίων ανάγνωσης των γονιδιωμάτων θηλαστικών, εντόμων και ψαριών. Αξιοσημείωτη είναι και η παρουσία μίας τουλάχιστον πρωτεΐνης του *Mus musculus* (της Q9D0S4), η οποία φέρει την επανάληψη NEUZ σε συν-

δυσασμό με την περιοχή SOCS (*Suppressors Of Cytokine Signalling*) [Η. Παυλόπουλος, Διδακτορική Διατριβή, Παν/μιο Κρήτης 2002]. Η συντηρημένη περιοχή SOCS ταυτοποιήθηκε πρωταρχικά στο καρβοξυτελικό άκρο οικογένειας πρωτεϊνών (SOCS 1-8), που έφεραν στο αμινοτελικό τους άκρο την περιοχή SH2, υπεύθυνη για τις αλληλεπιδράσεις με τους ενεργοποιημένους υποδοχείς κυτοκινών ή τις κινάσες που αλληλεπιδρούν με αυτούς, καταστέλλοντας τη σηματοδότηση από τις κυτοκίνες. Η περιοχή SOCS στα μόρια αυτά αλληλεπιδρά με το πρωτεϊνικό σύμπλοκο VCB (*VHL/Elongin B/Elongin C*), το οποίο λειτουργεί ως E3 λιγάση ουβικουΐτινης, επάγοντας την πρωτεόλυση σηματοδοτικών μορίων του μονοπατιού των κυτοκινών [Rui *et al*, 2002].

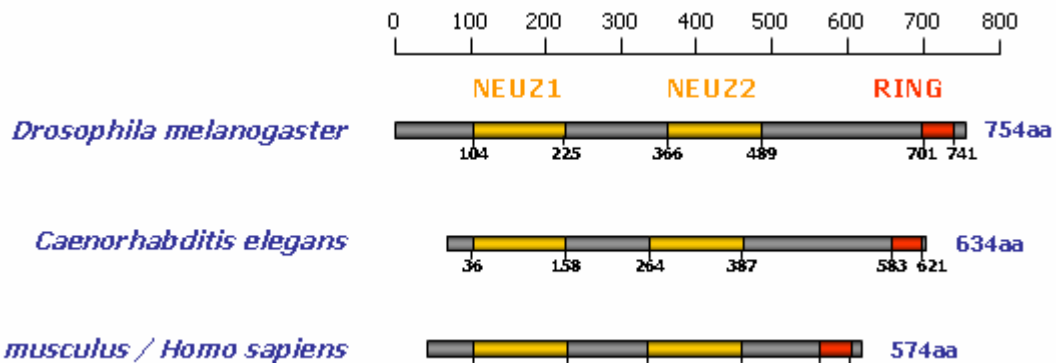
A) *Homo sapiens*, χρωμόσωμα 10



B) *Mus musculus*, χρωμόσωμα 19

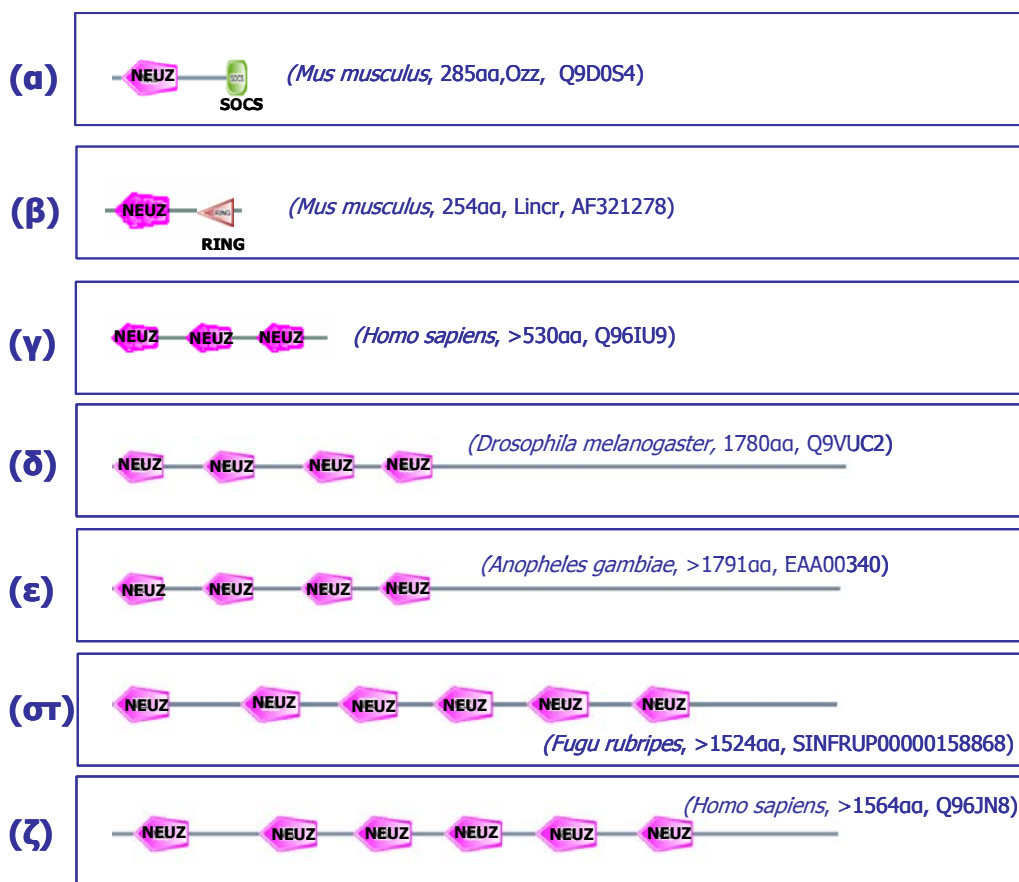


Εικ.8. Οι συντηνικές περιοχές των χρωμοσωμάτων του ανθρώπου (A) και του ποντικού (B), που περιέχουν το γονίδιο *Neuralized*. Φαίνεται η ομοιότητα στη διάταξη των γονιδίων της περιοχής. Με κόκκινο παρουσιάζονται τα γνωστά γονίδια σε κάθε οργανισμό, ενώ με πράσινο τα προβλεπόμενα βάσει ομολογίας με τα αντίστοιχα ορθόλογα.



Εικ.9. Η πρωτεΐνη *Neuralized*. Στα ορθόλογα μόρια των οργανισμών της εικόνας φαίνονται οι χαρακτηριστικές επαναλήψεις NEUZ1 και NEUZ2 (κίτρινο) και το μοτίβο δακτυλίου RING (κόκκινο). Τα ποσοστά αμινοξικής ταυτότητας μεταξύ των μορίων της δροσόφιλας και του νηματώδη με αυτά των θηλαστικών είναι 27% και 30%, αντίστοιχα [Ρανίλοπουλος *et al*, 2002], σε όλο το μήκος των πρωτεϊνών. Μεγαλύτερη ομοιότητα των τριών μορίων παρατηρείται στις περιοχές NEUZ1 και RING (βλ. και Αποτελέσματα).

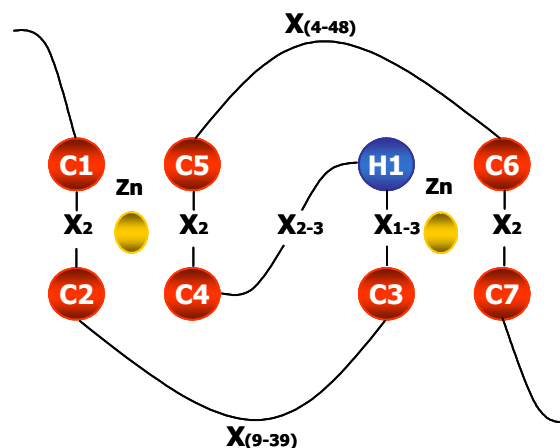
Πρόσφατη λειτουργική μελέτη της πρωτεΐνης Q9D0S4 του ποντικού, την οποία οι Nastasi *et al* (2004) ονομάζουν Ozz, δείχνει ότι επάγει την ουβικουϊτινίωση και πρωτεόλυση της μεμβρανικής β-κατενίνης των μυικών ινών κατά τη μυογένεση. Συγκεκριμένα, η Ozz αλληλεπιδρά με την β-κατενίνη μέσω της αμινοτελικής περιοχής NEUZ και με το υπερμοριακό σύμπλοκο E3 λιγάσης ουβικουϊτίνης των Elongin B, Elongin C, Rbx1 και Cullin 5, μέσω του καρβοξυτελικού *domain* SOCs, επάγοντας με αυτόν τον τρόπο την ουβικουϊτινίωση της β-κατενίνης [Nastasi *et al*, 2004]. Αξίζει να σημειωθεί, τέλος, και η παρουσία μιας πρωτεΐνης μήκους 254 αμινοξέων στον άνθρωπο και στον ποντικό, η οποία φέρει μία περιοχή NEUZ στο αμινοτελικό της άκρο και μία περιοχή RING του τύπου C3HC4 στο καρβοξυτελικό. Η πρωτεΐνη αυτή ονομάστηκε LincR (Lung Inducible Neuralized Related C3HC4 RING), λόγω της επαγωγής της έκφρασης του αντίστοιχου γονιδίου στους πνεύμονες ποντικών με φλεγμονή, μετά τη χορήγηση γλυκοκορτικοειδών [Smith JB *et al*, 2002]. Αν και η βιοχημική λειτουργία του μορίου LincR είναι άγνωστη, οι Smith *et al* (2002) προτείνουν ότι κατατάσσεται μεταξύ των παραγόντων που επάγονται από τα γλυκοκορτικοειδή, διευκολύνοντας την αντιφλεγμονώδη δράση τους. Η λειτουργία των υπόλοιπων πρωτεϊνών (Εικ.10) είναι άγνωστη, όμως η παρουσία των επαναλήψεων NEUZ ανεξάρτητα από την ύπαρξη 'δακτυλίου RING' σε αυτά τα μόρια υποδηλώνει μια γενικότερη λειτουργία αυτών των περιοχών που πιθανόν να σχετίζεται με τη διευκόλυνση πρωτεϊνικών αλληλεπιδράσεων στα κύτταρα των σπονδυλωτών.



Εικ.10 Περιοχές 'NEUZ' σε πρωτεΐνες εκτός της οικογένειας των μορίων Neuralized [Pavlopoulos *et al*, αδημοσίευτες παρατηρήσεις]. (α-ζ): Γραφικές απεικονίσεις των μορίων που φέρουν NEUZ, σύμφωνα με τη βάση δεδομένων SMART (dylan.embl-heidelberg.de). Στις παρενθέσεις φαίνονται στοιχεία για κάθε πρωτεΐνη (Οργανισμός προέλευσης, αριθμός αμινοξέων (aa), όνομα, κωδικός πρόσβασης SMART).

Περιοχή 'RING'

Η τρίτη συντηρημένη περιοχή των πρωτεϊνών Neuralized είναι ένα μοτίβο πρόσδεσης ψευδαργύρου του τύπου C_3HC_4 , γνωστό ως 'δακτύλιος RING'. Η περιοχή RING αρχικά ταυτοποιήθηκε στο γονίδιο *RING1* (*Really Interesting New Gene 1*) και αποτελείται από 40-60 αμινοξέα, μεταξύ των οποίων και συντηρημένα κατάλοιπα κυστεΐνης και ιστιδίνης που προσδένουν δύο άτομα ψευδαργύρου, σχηματίζοντας την τριτοταγή διαμόρφωση που φαίνεται στην Εικ.11 [Borden, 2000]. Το μοτίβο αυτό έχει πρόσφατα συσχετιστεί με τη διαδικασία μεταμεταφραστικής τροποποίησης πρωτεϊνών από ουβικουΐτινη, λόγω της δράσης του ως E3 λιγάσης σε αυτή τη διαδικασία [Joazeiro *et al*, 1999]. Συγκεκριμένα, πρωτεΐνες με μοτίβα RING όπως η Rbx1 και η Arc11 αποτελούν υπομονάδες των υπερμοριακών συμπλόκων SCF (*Skp1/Culin/F-box*), APC (*Anaphase Promoting Complex*) και VCB (*VHL/Elongin B/Elongin C*), τα οποία εμφανίζουν δράση E3 λιγάσης, μεταφέροντας την ουβικουΐτινη από το ένζυμο E2 στο υπόστρωμα [Tyers & Willems, 1999]. Επιπλέον, μία μεγάλη οικογένεια πρωτεϊνών με μοτίβα RING, όπως οι Cbl, Siah-1 και Mdm2, οι οποίες δεν έχουν άλλη δομική ομοιότητα με τις Rbx1 και Arc11, εμφανίζουν δράση E3 λιγάσης ουβικουΐτινης, για την οποία ευθύνεται η παρουσία του μοτίβου RING σε αυτές [Joazeiro *et al*, 1999/ Hu *et al*, 1997/ Lorick *et al*, 1999].



Εικ.11 Το μοτίβο δακτυλίου RING, τύπου C_3HC_4 [Borden, 2000], που υπάρχει στο μόριο Neuralized. Στο σχήμα φαίνεται η διάταξη των συντηρημένων αμινοξικών καταλοίπων (7 κυστεϊνών και μίας ιστιδίνης), που προσδένουν δύο ιόντα ψευδαργύρου (Zn), (X=οποιοδήποτε αμινοξύ).

8. Βιολογικός ρόλος του *neuralized* στη *Drosophila melanogaster* και στον *Xenopus laevis*

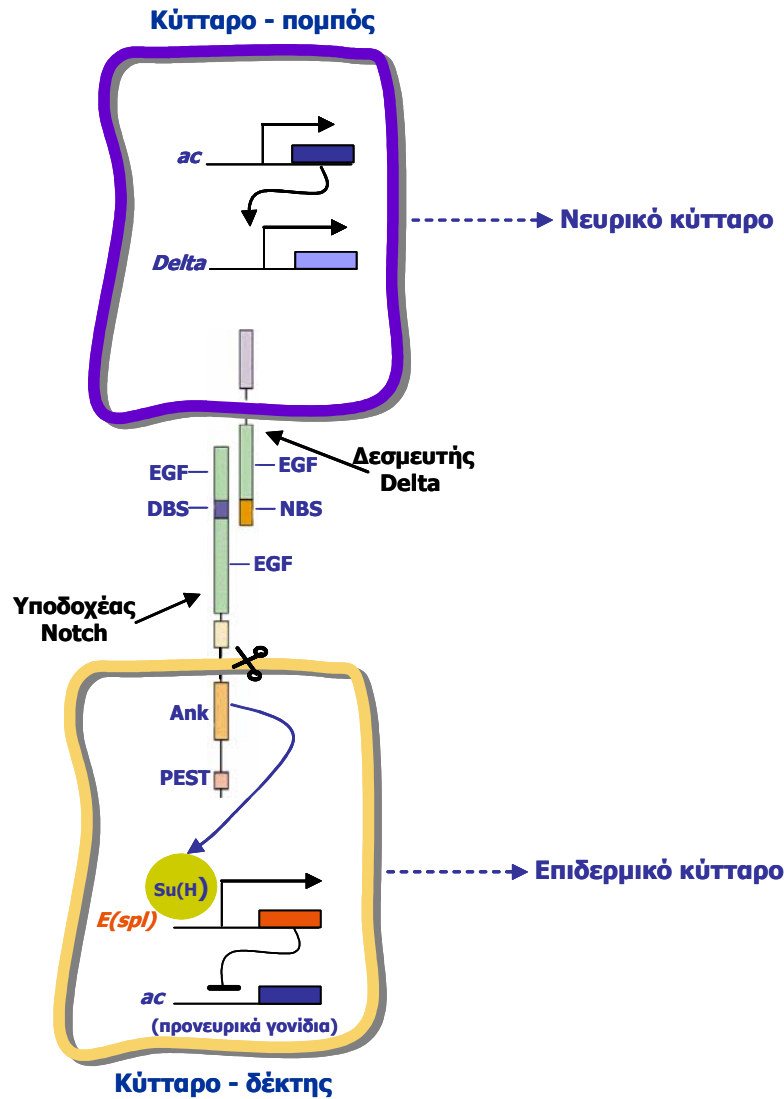
Για το γονίδιο *neuralized* της *Drosophila melanogaster* (*neur*) είναι γνωστό από γενετικά πειράματα ότι συμμετέχει σε αναπτυξιακές διαδικασίες αποφάσεων κυτταρικής τύχης, όπως αυτή της 'πλευρικής αναστολής', που ρυθμίζονται από το σηματοδοτικό μονοπάτι του υποδοχέα Notch, μέσω του δεσμευτή Delta. Πρόσφατα, αποδείχθηκε και για το ορθόλογο γονίδιο του *Xenopus laevis* (*Xneur*) ότι ρυθμίζει το σηματοδοτικό μονοπάτι των μορίων Notch και Delta. Μελέτες και στα δύο συστήματα διαπιστώνουν ότι το μόριο *neuralized* ενεργοποιεί τη σηματοδότηση του Notch επάγοντας την ενδοκύτωση του Delta, ενώ σε μοριακό επίπεδο διερευνάται η λειτουργία του ως E3 λιγάσης ουβικουΐτινης.

➤ **Neur1 και σηματοδότηση Notch**

Το γονίδιο *neur* συγκαταλέγεται μεταξύ των νευρογόνων γονιδίων *Notch*, *Delta*, *Enhancer of split*, *big brain*, *mastermind* και *almondex* [Lehmann *et al*, 1983]. Τα γονίδια αυτά λειτουργούν κατά την εμβρυϊκή και προνυμφική ζωή της μύγας συμμετέχοντας στην επιλογή των νευροβλαστών ή των πρόδρομων κυττάρων αισθητηρίων οργάνων (*Sensory Organ Precursors*, SOPs) από κύτταρα με ισοδύναμο αναπτυξιακό δυναμικό της νευρογόνου περιοχής των εμβρύων ή των 'προνευρικών' αθροισμάτων των αναπτυσσόμενων εμβρυϊκών δίσκων [Greenwald, 1998].

Τα νευρογόνα γονίδια ρυθμίζουν την αναπτυξιακή διαδικασία πλευρικής αναστολής (*lateral inhibition*), κατά την οποία τα πρόδρομα νευρικά κύτταρα (νευροβλάστες ή SOPs) αναστέλλουν το νευρικό δυναμικό των γειτονικών κυττάρων μέσω σηματοδότησης του υποδοχέα Notch. Συγκεκριμένα, από τα αναπτυξιακά ισοδύναμα κύτταρα της νευρογόνου κοιλιακής περιοχής του εμβρύου της *Drosophila*, μόνο το 25% διατηρούν την εξ ορισμού νευρική τους ταυτότητα και διαφοροποιούνται σε νευροβλάστες, ενώ τα υπόλοιπα διαφοροποιούνται σε επιδερμοβλάστες, ως αποτέλεσμα αναστολής της νευρικής αναπτυξιακής τους πορείας [Campos-Ortega, 1988]. Ο ίδιος μηχανισμός λειτουργεί και κατά την ανάπτυξη του περιφερικού νευρικού συστήματος στην προνύμφη και συγκεκριμένα στην αναπτυξιακή διαδικασία σχηματισμού των αισθητηρίων οργάνων του ενήλικου [Hartenstein & Posakony, 1990 / Ghysen *et al*, 1993]. Η σηματοδότηση Notch στη διαδικασία της πλευρικής αναστολής πραγματοποιείται ανάμεσα στο μελλοντικό νευροβλάστη ή SOP (κύτταρο-πομπός) και τους μελλοντικούς επιδερμοβλάστες (κύτταρα-δέκτες). Συγκεκριμένα, ο μεμβρανικός υποδοχέας Notch του κυττάρου-δέκτη, ενεργοποιείται με την αλληλεπίδραση της εξωκυττάριας περιοχής του με τους δεσμευτές Delta ή Serrate του κυττάρου-πομπού και την επαγόμενη πρωτεολυτική απελευθέρωση της κυτταροπλασματικής περιοχής του Notch [Artavanis-Tsakonas *et al*, 1991, 1995, 1999], η οποία μετατοπίζεται στον πυρήνα [Struhl & Adachi, 1998] και ενεργοποιεί τη μεταγραφή γονιδίων-στόχων, που κωδικοποιούν καταστολείς του νευρικού φαινοτύπου στο κύτταρο-δέκτη (Εικ.12, πχ γονίδια του συμπλέγματος *E(spl)*, Bailey & Posakony, 1995).

Μεταλλαγές απώλειας λειτουργίας οποιουδήποτε από τα νευρογόνα γονίδια, συμπεριλαμβανομένου του *neur*, προκαλούν υπερπλασία του νευρικού συστήματος [Artavanis-Tsakonas *et al*, 1995]. Η απώλεια λειτουργίας του γονιδίου *neur* οδηγεί σε εμβρυϊκή θνησιμότητα λόγω του ισχυρού νευρογόνου φαινοτύπου της. Από πειράματα γενετικών αλληλεπιδράσεων έχει διαπιστωθεί ότι η δράση του *neur* προηγείται της σηματοδότησης του υποδοχέα, εφόσον ο φαινότυπος απώλειας της λειτουργίας του διασώζεται μερικώς από την αύξηση της έκφρασης του *Notch* ή των γονιδίων στόχων (*E(spl)*) [de la Concha *et al*, 1988]. Πειράματα απενεργοποίησης του *neur* σε μιτωτικούς κλώνους δείχνουν ότι η ενεργότητά του είναι απαραίτητη σε ένα υποσύνολο των αναπτυξιακών διαδικασιών που ρυθμίζονται από το Notch [Lai & Rubin, 2001a / Pavlouroulos *et al*, 2001]. Επιπλέον, έχει παρατηρηθεί ότι η υπερέκφραση του *neur* σε εμβρυϊκούς δίσκους μπορεί να οδηγήσει σε ενίσχυση [Pavlouroulos *et al*, 2001] αλλά και σε καταστολή [Lai & Rubin, 2001a] της σηματοδότησης από τον υποδοχέα Notch, διαφορά που πιθανόν οφείλεται στα επίπεδα της υπερέκφρασης του *neur*. Συγκεκριμένα, οι Lai και Rubin (2001a) έχουν προτείνει ότι σχετικά χαμηλά επίπεδα υπερέκφρασης του *neur* ενισχύουν τη σηματοδότηση, ενώ υψηλά επίπεδα *neur* την καταστέλλουν.



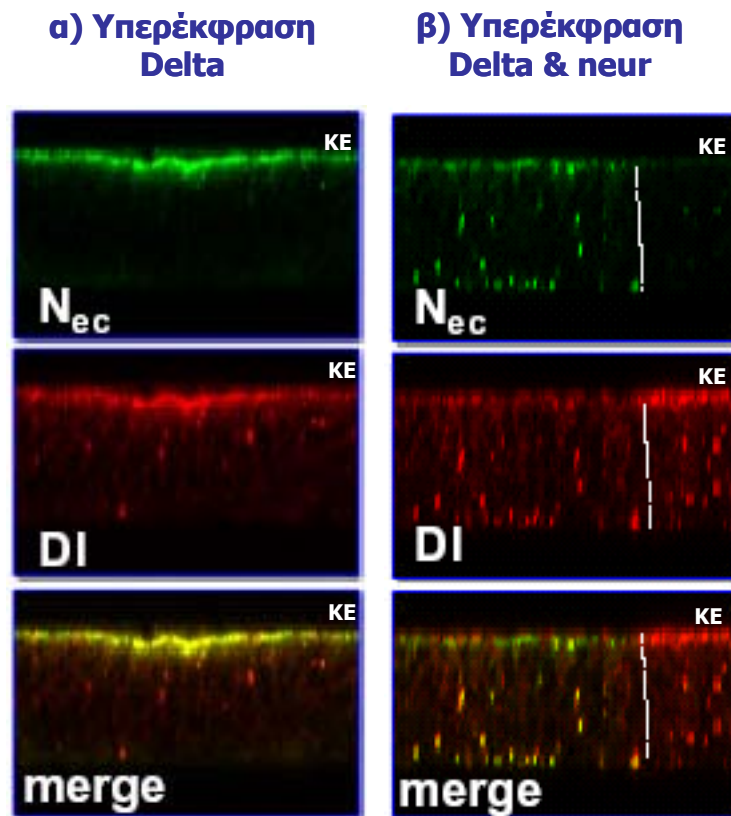
Εικ.12 Συνοπτική περιγραφή της σηματοδότησης Notch-Delta κατά την πλευρική αναστολή. Η αλληλεπίδραση του δεσμευτή Delta με τον υποδοχέα Notch διευκολύνει την πρωτεολυτική επεξεργασία του Notch και την απελευθέρωση της ενδοκυττάριας περιοχής του, η οποία εισέρχεται στον πυρήνα όπου ενεργοποιεί τη μεταγραφική γονιδίων-στόχων, όπως αυτά του συμπλέγματος *E(spl)*, σε συνεργασία με τον παράγοντα πρόσδεσης στο DNA Suppressor of Hairless [Su(H)]. Οι πρωτεΐνες *E(spl)* καταστέλλουν τη μεταγραφή προνευρικών γονιδίων (πχ του *achaete*, *ac*). Αντίθετα, στο κύτταρο-πομπό το γονίδιο *ac* είναι ενεργό, προκαλώντας την ενεργοποίηση του *Delta* και άλλων νευρογόνων γονιδίων, με αποτέλεσμα την εκδήλωση του νευρικού φαινοτύπου από το συγκεκριμένο κύτταρο. (EGF= επαναλήψεις EGF(Epidermal Growth Factor)-like, DBS=Delta Binding Site, NBS=Notch Binding Site, Ank= επαναλήψεις αγκυρίνης, PEST=μοτίβο PEST.)

α. Η πρωτεΐνη *neuralized* ρυθμίζει διαδικασίες πλευρικής αναστολής

Το *neur*, όπως και τα νευρογόνα γονίδια *Notch* και *Delta*, συμμετέχει σε αναπτυξιακές διαδικασίες πλευρικής αναστολής τόσο στη δροσόφιλα όσο και στον βάτραχο. Επιπλέον, σε αντίθεση με τα μόρια Notch και Delta που εκφράζονται ομοιόμορφα μεταξύ των κυττάρων των προνευρικών αθροισμάτων στη μύγα, η έκφραση του *neur* κατά την εμβρυογένεση περιορίζεται σταδιακά στο νευροεκτόδερμα και ειδικότερα στους πρόδρομους νευροβλάστες, ενώ κατά την ανάπτυξη του νευρικού συστήματος στο ενήλικο άτομο μετάγραφα του *neur* ανιχνεύονται στα πρόδρομα κύτταρα αισθητηρίων οργάνων που εντοπίζονται στους προνυμφικούς δίσκους του φτερού, του ποδιού, της κεραίας και του ματιού, από τα οποία προκύπτουν τα αισθητήρια μηχανικών ερεθισμάτων (σμήριγγες), τα χορδοτονικά αισθητήρια, το όργανο *Johnston* και οι φωτούποδοχείς, αντίστοιχα [Boulianne *et al*, 1991]. Επιπλέον, το γονίδιο είναι μεταγραφικά ενεργό στους νευροβλάστες του ενήλικου κεντρικού νευρικού συστήματος. Κατά ανάλογο τρόπο, η έκφραση του ορθόλογου γονιδίου *Xneur* στον

Xenopus laevis ανιχνεύεται ομοιόμορφη και σε χαμηλά επίπεδα στα έμβρυα πρώιμων σταδίων και βαθμιαία περιορίζεται στο αναπτυσσόμενο νευρικό σύστημα και στο δέρμα, εντοπισμένη σε υψηλά επίπεδα στους πρωτογενείς νευρώνες και τα πρόδρομα μαστιγοφόρα κύτταρα του δέρματος [Deblandre *et al*, 2001], των οποίων η διαφοροποίηση ρυθμίζεται από τη σηματοδότηση μέσω του υποδοχέα Notch [Chitnis *et al*, 1995 / Deblandre *et al*, 1999 / Wettstein *et al*, 1997].

Όπως έχει αποδειχθεί με βάση πρόσφατες μελέτες στη δροσόφιλα, η παρουσία της πρωτεΐνης *neuralized* στο μελλοντικό κύτταρο του νευροβλάστη ενισχύει την ικανότητά του να στέλνει σήματα αναστολής του νευρικού δυναμικού μέσω του δεσμευτή Delta, προς τον υποδοχέα Notch των γειτονικών κυττάρων [Yeh *et al*, 2000/ Lai & Rubin, 2001b/ Pavlopoulos *et al*, 2001]. Συγκεκριμένα, από πειράματα εκτοπικής



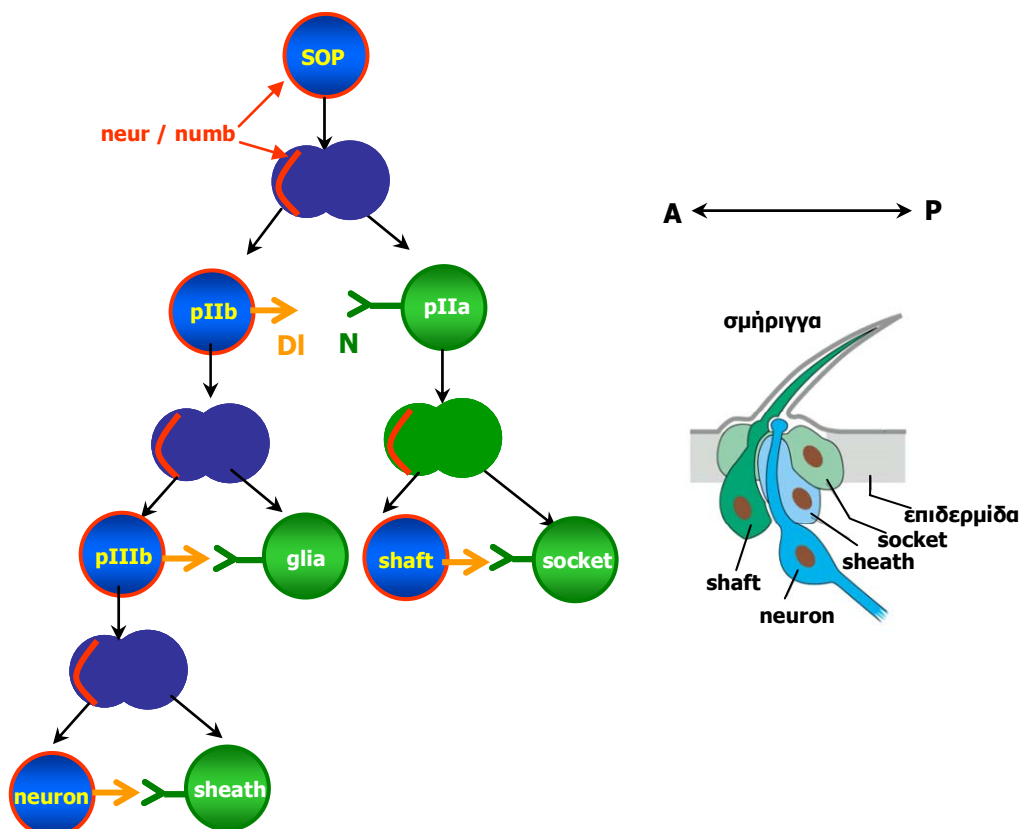
Εικ.13. Η πρωτεΐνη *neur* της δροσόφιλας επάγει ενδοκύττωση του Delta και της εξωκυττάριας περιοχής του Notch. Εγκάρσια οπτική τομή σε επιθηλιακά κύτταρα της περιοχής 'omb' προνυμφικού δίσκου φτερού δροσόφιλας, όπου υπερεκφράζεται μέσω του συστήματος GAL4-UAS ο δεσμευτής Delta (DI), είτε μόνος του (α), είτε παρουσία υπερεκφρασμένου *neur* (β). Η υπερέκφραση του *neuralized* αυξάνει σημαντικά τον αριθμό των κυστιδίων που φέρουν DI και την αλληλεπιδρούσα εξωκυττάρια περιοχή του Notch (Nec) και προέρχονται από την κορυφαία επιφάνεια (KE) των κυττάρων. Με κόκκινο φαίνεται η υπερεκφρασμένη πρωτεΐνη Delta (DI) και με πράσινο η χρώση ειδικού αντισώματος για την εξωκυττάρια περιοχή του υποδοχέα Notch (Nec) [Pavlopoulos *et al*, 2001].

υπερέκφρασης των μορίων *neur* και Delta, διαπιστώνεται ότι το γονίδιο *neur* επάγει την ενδοκύττωση του Delta στο κύτταρο του νευροβλάστη (Εικ.13), γεγονός που φαίνεται να διευκολύνει την πρωτεολυτική ενεργοποίηση του Notch στο απέναντι κύτταρο και επομένως καθορίζει το κύτταρο που εκφράζει *neur* ως κύτταρο-πομπό [Pavlopoulos *et al*, 2001/ Parks *et al*, 2000/ Le Borgne & Schweisguth, 2003b]. Η εκτοπική έκφραση του *neur* σε μιτωτικούς κλώνους κυττάρων προνυμφικών δίσκων φτερού ενισχύει την ενδοκύττωση του Delta αλλά και του εξωκυττάριας τμήματος του Notch (N-EC) στα ίδια κυστίδια, στις περιοχές όπου τα μόρια Notch και Delta συνεκφράζονται. Όπου εκφράζεται μόνο ο υποδοχέας Notch και όχι ο δεσμευτής Delta, η έκφραση

του *neur* δεν προκαλεί ενδοκύττωση του τμήματος N-EC, γεγονός που αποδεικνύει ότι το *neur* επηρεάζει τη σηματοδότηση μόνο μέσω του Delta και αποσαφηνίζει τον κυτταρικά μη αυτόνομο μηχανισμό δράσης του στο κύτταρο-πομπό του σήματος (Εικ.13). Τέλος, η ενδοκύττωση του Delta φαίνεται ότι ακολουθείται από πρωτεόλυσή του, εφόσον η εκτοπική έκφραση της πρωτεΐνης *neur* σε εμβρυϊκούς δίσκους φτερού, ελαττώνει τα επίπεδα του ενδογενούς Delta [Pavliopoulos *et al*, 2001/ Lai *et al*, 2001]. Η ίδια παρατήρηση υποστηρίζεται και από πειράματα συνέκφρασης των μορίων Delta και *Xneur* σε έμβρυα του *Xenopus laevis* [Deblandre *et al*, 2001].

β. Η πρωτεΐνη *neuralized* ρυθμίζει αποφάσεις κυτταρικής γενεαλογίας

Η διαδικασία της πλευρικής αναστολής καταλήγει στον καθορισμό του πρόδρομου κυττάρου (νευροβλάστη ή SOP) από τις διαιρέσεις του οποίου θα προκύψουν οι διαφοροποιημένοι νευρώνες και τα κύτταρα των αισθητηρίων οργάνων. Σε κάθε μία από αυτές τις διαιρέσεις απαιτούνται αποφάσεις κυτταρικής γενεαλογίας που καθορίζουν ποιο από τα θυγατρικά κύτταρα θα ακολουθήσει τη νευρική πορεία. Το σύστημα-μοντέλο στο οποίο έχει μελετηθεί γενετικά και κυτταρολογικά ο ρόλος της σηματοδότησης του Notch σε αυτά τα αναπτυξιακά γεγονότα είναι η διαφοροποίηση των εξωτερικών αισθητηρίων οργάνων των σμηρίγγων του ενήλικου από τα πρόδρομα κύτταρα SOP που επιλέγονται μέσω πλευρικής αναστολής στους εμβρυϊκούς δι-



Εικ.14 Οι αποφάσεις κυτταρικής γενεαλογίας κατά την ανάπτυξη της σμηρίγγας ρυθμίζονται από την ασύμμετρη κατανομή των πρωτεϊνών *neur* και *numb* κατά τις διαιρέσεις του πρόδρομου (SOP) κυττάρου [Le Borgne & Schweisguth, 2003a]. Οι μίτωτικές διαιρέσεις του κυττάρου SOP που οδηγούν στον σχηματισμό της σμηρίγγας. Από τον οπίσθιο απόγονο της πρώτης ασύμμετρης διαιρέσης (κύτταρο pIIa) προκύπτουν με μία επιπλέον ασύμμετρη διαιρέση τα κύτταρα 'shaft' (πρόσθιο) και 'socket' (οπίσθιο). Αντίστοιχα, από το πρόσθιο κύτταρο pIIb προκύπτει το πρόσθιο pIIIb κύτταρο και ένα οπίσθιο κύτταρο που διαφοροποιείται σε μεταναστευτικό κύτταρο γλοίας. Τέλος, η ασύμμετρη διαιρέση του pIIIb δίνει τον νευρώνα (πρόσθιο κύτταρο) και το κύτταρο 'sheath' (οπίσθιο). Ο επιλεκτικός εντοπισμός των πρωτεϊνών *neur* και *numb* στον πρόσθιο απόγονο κάθε μίτωσης φαίνεται με κόκκινο. Οι τέσσερις τύποι κυττάρων συνιστούν το αισθητήριο της σμηρίγγας στην ενήλικη μύγα. Το διπλό βέλος δείχνει τον πρόσθιο(A)-οπίσθιο(P) άξονα του ζώου. (DI=δεσμευτής Delta, N=υποδοχέας Notch).

σκους του φτερού. Φαίνεται ότι η ασύμμετρη κυτταρική διαίρεση του SOP ευθύνεται για τον επιλεκτικό εντοπισμό στο κύτταρο που πρόκειται να γίνει νευρικό (κύτταρο-πομπό) των πρωτεϊνών που επάγουν τη σηματοδότηση αναστολής του νευρικού δυναμικού στον άλλο απόγονο της ίδιας μίτωσης. Συγκεκριμένα, πρόσφατες μελέτες δείχνουν ότι η επιλεκτική παρουσία της πρωτεΐνης *neur* στο κύτταρο-πομπό είναι αποτέλεσμα της ασύμμετρης κατανομής της στον διαιρούμενο πρόγονο αυτού του κυττάρου [Le Borgne & Schweisguth, 2003a]. Αναλυτικότερα, κατά τη διαίρεση του προγονικού κυττάρου SOP κατά μήκος του πρόσθιο-οπίσθιου άξονα του ζώου, προκύπτει το πρόσθιο κύτταρο *rIIb* και το οπίσθιο *rIIa*, από τα οποία μόνο το πρώτο φέρει την πρωτεΐνη *neur*. Κατά συνέπεια, μόνο στο κύτταρο *rIIb* ενισχύεται η ενδοκύττωση του δεσμευτή Delta. Στο ίδιο κύτταρο εντοπίζεται επιλεκτικά και η πρωτεΐνη *numb*, η οποία λειτουργεί ως αναστολέας του μεμβρανικού υποδοχέα Notch επάγοντας την ενδοκύττωσή του. Η συγχρονισμένη δράση των πρωτεϊνών *neur* και *numb* για την ενεργοποίηση του δεσμευτή και την καταστολή του υποδοχέα αντίστοιχα, καθιστά το κύτταρο *rIIb* πομπό του σήματος καταστολής του νευρικού δυναμικού προς το οπίσθιο κύτταρο *rIIa* [Le Borgne & Schweisguth, 2003a / Yeh *et al*, 2000]. Σύμφωνα με το επικρατέστερο μοντέλο που περιγράφει τη γενεαλογία των κυττάρων *rI*, η ασύμμετρη κατανομή των πρωτεϊνών *neur* και *numb* κατά τις επόμενες διαιρέσεις των κυττάρων *rIIa* και *rIIb*, οδηγεί με ανάλογο τρόπο στον καθορισμό της ταυτότητας των τεσσάρων κυττάρων που συνιστούν τη σμήριγγα, μεταξύ των οποίων και του νευρώνα, ο οποίος είναι μιτωτικός απόγονος του κυττάρου *rIIb* (Εικ.14).

γ. Το neuralized επάγει 'ασυμμετρία' πομπού-δέκτη για την ενίσχυση της σηματοδότησης Notch-Delta

Αν και το *neur* είναι απαραίτητο συστατικό του μονοπατιού σηματοδότησης από τον υποδοχέα Notch σε διαδικασίες πλευρικής αναστολής μεταξύ ισοδύναμων κυττάρων και σε αποφάσεις κυτταρικής γενεαλογίας που λαμβάνονται κατά την ανάπτυξη του περιφερικού νευρικού συστήματος στη δροσόφιλα, ωστόσο η δράση του δεν είναι απαραίτητη σε άλλες αναπτυξιακές διαδικασίες που ρυθμίζονται με σηματοδότηση Notch-Delta και πλευρική αναστολή, όπως η ανάπτυξη φλεβών και ο καθορισμός του ραχαιοκοιλιακού συνόρου του φτερού [Lai & Rubin, 2001a / 2001b] ή η διαφοροποίηση των θυλακοκυττάρων των ωοθηκών [Lopez-Schier & St Johnston, 2001]. Επιπλέον, η έκφραση των μορίων Notch και Delta προηγείται και είναι ευρύτερη από την έκφραση του *neur* στα προνευρικά αθροίσματα κυττάρων του εμβρυϊκού δίσκου του φτερού. Επομένως, αλληλεπίδραση Notch-Delta και πιθανόν χαμηλά επίπεδα σηματοδότησης είναι δυνατόν να συμβαίνουν και απουσία του *neur*. Το *neur*, επάγοντας ασυμμετρία στις επιφάνειες των κυττάρων πομπού και δέκτη με την ενδοκύττωση του Delta στο κύτταρο-πομπό, ταυτόχρονα ενισχύει και τη σηματοδότηση προς τον δέκτη. Η ενίσχυση της σηματοδότησης από το *neur* είναι απαραίτητη για τη μεταγραφική ενεργοποίηση των γονιδίων-στόχων *E(spl)*, που κωδικοποιούν τους ισχυρότερους ανταγωνιστές των προνευρικών πρωτεϊνών και τα οποία δεν ενεργοποιούνται απουσία του *neur* [Jennings *et al*, 1994 / Pavlopoulos *et al*, 2001]. Είναι, επομένως, πιθανόν η συμμετοχή του *neur* στη σηματοδότηση Notch-Delta να μην είναι απαραίτητη στις περιπτώσεις που το *neur* αντικαθίσταται λειτουργικά από άλλο μόριο ή όταν τα μόρια Notch και Delta εκφράζονται εξ αρχής σε διαφορετικούς πληθυσμούς κυττάρων. Για παράδειγμα, για τον καθορισμό του ραχαιοκοιλιακού ορίου του φτερού, η ασυμμετρία υποδοχέα-δεσμευτή επάγεται από την πρωτεΐνη Fringe, η οποία εκφράζεται ειδικά στα ραχιαία κύτταρα και γλυκοζυλιώνει τον υποδοχέα Notch, καθιστώντας τον ικανό να αλληλεπιδρά μόνο με τον δεσμευτή Delta των κοιλιακών κυττάρων και όχι με το Serrate των ραχιαίων [Blair, 2000]. Επιπλέον, επαγωγή

ασυμμετρίας δεν είναι απαραίτητη μεταμεταφραστικά κατά τη διαφοροποίηση των κυττάρων των φλεβών του φτερού γιατί η έκφραση των γονιδίων *Notch* και *Delta* σε αυτό το σύστημα δεν είναι επικαλυπτόμενη [de Celis *et al*, 1997 / Huppert *et al*, 1997]. Τέλος, η διαπίστωση ότι η υπερέκφραση του πλήρους μήκους *neur* ή του ελλειμματικού μορίου *neurΔRING* επηρεάζει την ενεργοποίηση του *Notch* ακόμα και στις περιπτώσεις που αυτή δεν εξαρτάται από την ενεργότητα του *neur*, δείχνει ότι το μόριο αυτό σχετίζεται με βασικά συστατικά του μονοπατιού σηματοδότησης *Notch-Delta*.

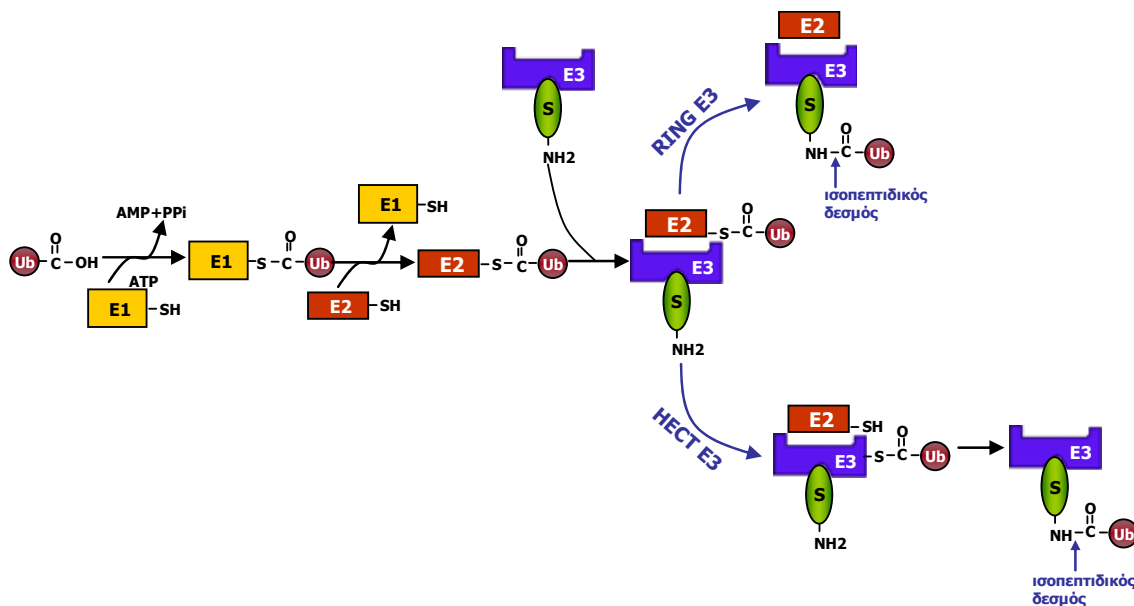
➤ **Neur1 και ουβικουϊτινιλίωση**

Σύμφωνα με πρόσφατα δημοσιευμένες μελέτες, η περιοχή *RING* των μορίων *neuralized* της *Drosophila* και του *Xenopus* εμφανίζει, όπως ήταν αναμενόμενο, ενεργότητα E3 λιγάσης ουβικουϊτίνης σε πειράματα *in vitro* ουβικουϊτινιλίωσης. Η αυτο-ουβικουϊτινιλίωση της πρωτεΐνης *neuralized* και των δύο οργανισμών και η συμμετοχή της στην αποικοδόμηση του *Delta* μέσω του πρωτεασώματος υποστηρίζεται και από *in vivo* πειράματα υπερέκφρασης σε έμβρυα ή σε κυτταρικές σειρές θηλαστικών [Yeh *et al*, 2001/ Lai *et al*, 2001/ Deblandre *et al*, 2001/ Pavlopoulos *et al*, 2001].

α. Ουβικουϊτινιλίωση

Η ουβικουϊτινιλίωση είναι μετα-μεταφραστική τροποποίηση που οδηγεί σε μερική πρωτεόλυση ή αποικοδόμηση πρωτεϊνών των ευκαρυωτικών κυττάρων. Με αυτόν τον τρόπο ρυθμίζει μια ποικιλία βιολογικών διεργασιών, όπως ο κυτταρικός κύκλος και η σηματοδότηση μέσω μεμβρανικών υποδοχέων [Hershko & Ciechanover, 1998/ Levkowitz *et al*, 1999/ Tyers & Jorgensen, 2000]. Η τροποποίηση πρωτεϊνικών υποστρωμάτων με ένα ή περισσότερα μόρια ουβικουϊτίνης πραγματοποιείται σε τρία βήματα, τα οποία καταλύονται αντίστοιχα από το ενεργοποιητικό ένζυμο E1, τα συζευκτικά ένζυμα E2 και συνδετικά ένζυμα E3 [Hershko & Ciechanover, 1998]. Συγκεκριμένα, μόρια ελεύθερης ουβικουϊτίνης συνδέονται ομοιοπολικά και μεταφέρονται διαδοχικά από τα ένζυμα E1, E2 και E3 σε κατάλοιπα λυσίνης του υποστρώματος (Εικ.15). Τα ένζυμα που έχουν ρόλο E3 λιγάσης προσφέρουν εξειδίκευση στη διαδικασία αναγνωρίζοντας το υπόστρωμα.

Ενδοκυττάρειες πρωτεΐνες που τροποποιούνται με αλυσίδες περισσότερων των τεσσάρων μορίων ουβικουϊτίνης οδηγούνται κατά κανόνα στο πρωτεάσωμα για αποικοδόμηση. Αντίθετα, η μονο- ή η ολιγο-ουβικουϊτινιλίωση συνήθως ρυθμίζει τα επίπεδα μεμβρανικών πρωτεϊνών (πχ των υποδοχέων EGFR και MetR), προκαλώντας την ενδοκυττάρωση και την καταστροφή τους στο λυσόσωμα [Dupre *et al*, 2004/ Grovdal *et al*, 2004/ Carter *et al*, 2004/ Hicke & Dunn, 2003]. Επιπλέον, αναφέρονται και μη πρωτεολυτικές λειτουργίες της μονο- ή ολιγο-ουβικουϊτινιλίωσης, που περιλαμβάνουν τη ρύθμιση του υποκυτταρικού εντοπισμού των υποστρωμάτων [Shcherbik & Haines, 2004] και της μεταγραφής [Muratani & Tansey, 2003]. Φαίνεται ότι η λυσίνη της ουβικουϊτίνης που επιλέγεται για τον πολυμερισμό της, είναι καθοριστική για το αποτέλεσμα της τροποποίησης. Έτσι, αλυσίδες ουβικουϊτίνης που συνδέονται στην K48 οδηγούν τα υποστρώματα στο πρωτεάσωμα, ενώ οι 2-3 ουβικουϊτίνες συνδεδεμένες στην K63 έχουν, όπως και η μονο-ουβικουϊτινιλίωση, μη πρωτεολυτικές συνέπειες [Aguilar RC & Wendland B, 2003].

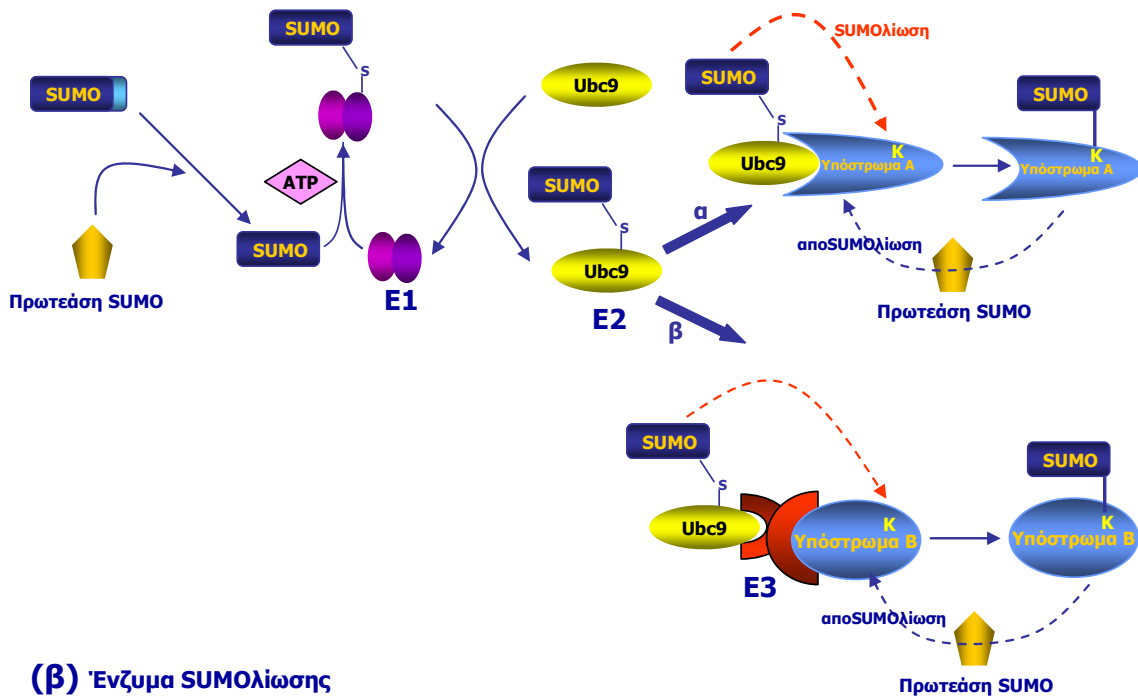


Εικ.15α Η διαδικασία της ουβικουΐτινίωσης. Η ελεύθερη ουβικουΐτινη (Ub) ενεργοποιείται παρουσία ATP με τον σχηματισμό θειοεστερικού δεσμού ανάμεσα στο ένζυμο E1 και το καρβοξυτελικό της άκρο. Στη συνέχεια, η Ub μεταφέρεται σε ένα από τα πολλά διαφορετικά ένζυμα σύζευξης (E2). Τα ένζυμα E2 αλληλεπιδρούν με διάφορα ένζυμα σύνδεσης (E3), τα οποία προσδένουν το υπόστρωμα (S). Για τα ένζυμα E3 που φέρουν περιοχή HECT (*Homologous to E6-AP Carboxy-Terminus*), η Ub μεταφέρεται από το E2 αρχικά στην κυστεΐνη του ενεργού κέντρου της περιοχής HECT και στη συνέχεια σε μία λυσίνη του υποστρώματος, με την οποία σχηματίζει ισοπεπτιδικό δεσμό. Για τα ένζυμα E3 που φέρουν περιοχή RING, η μεταφορά της Ub από το E2 είναι δυνατόν να γίνει απευθείας στο υπόστρωμα. Η επανάληψη της παραπάνω διαδικασίας οδηγεί στον σχηματισμό αλυσίδων πολύ-ουβικουΐτινης, συνδεδεμένων με το υπόστρωμα.

Τροποποιήσεις ομόλογες της ουβικουΐτινίωσης

Επιπρόσθετα, αναφέρονται και τροποποιήσεις πρωτεϊνών από μόρια ομόλογα της ουβικουΐτινης (*Ubiquitin Like molecules*, UBLs). Οι πιο καλά μελετημένες από αυτές αφορούν τα μόρια SUMO-1 (*Small Ubiquitin-like Modifier 1*) και Nedd8. Και αυτές οι τροποποιήσεις πραγματοποιούνται σε τρία ενζυμικά στάδια, όπως και η ουβικουΐτινίωση, τα οποία καταλύονται από ειδικά ένζυμα E1, E2 και E3 (Εικ.15β). Η πρωτεΐνη SUMO-1 (101 αμινοξέα) είναι κατά 18% ταυτόσημη της ουβικουΐτινης και συντηρημένη από τον σακχαρομύκητα μέχρι τα θηλαστικά. Οι συνέπειες της SUMOλίωσης διαφέρουν για κάθε υπόστρωμα, όμως γενικά επηρεάζουν τις πρωτεϊνικές αλληλεπιδράσεις και τον υποκυτταρικό του εντοπισμό. Για παράδειγμα, η SUMOλίωση της RanGAP1 (*Ran GTPase-Activating Protein 1*) ευθύνεται για τον εντοπισμό της στους πυρηνικούς πόρους [Matunis *et al*, 1998], ενώ η ίδια τροποποίηση προκαλεί τη συγκέντρωση της PML (*Promyelocytic Leukemia protein*) σε πυρηνικές δομές, γνωστές ως 'PML bodies' [Muller *et al*, 1998]. Επιπλέον, η SUMOλίωση μεταγραφικών παραγόντων επηρεάζει την ενεργότητά τους, όπως έχει αποδειχθεί για τους παράγοντες c-Jun, p53 και c-Myc [Bies *et al*, 2002/ Muller *et al*, 2000/ Rodriguez *et al*, 1999]. Αν και το αποτέλεσμα της SUMOλίωσης διαφέρει σημαντικά από αυτό της πολυ-ουβικουΐτινίωσης, σημαντικός αριθμός πρωτεϊνών ρυθμίζεται και από τους δύο μηχανισμούς [Gill, 2004/ Johnson, 2004]. Οι κανόνες που καθορίζουν την συγχρονισμένη λειτουργία των μηχανισμών ουβικουΐτινίωσης και SUMOλίωσης δεν έχουν πλήρως περιγραφεί για τα περισσότερα υποστρώματα. Για παράδειγμα, είναι γνωστό ότι η SUMOλίωση δρα ανταγωνιστικά με την ουβικουΐτινίωση σε ορισμένα υποστρώματα (IκBa, PCNA, Smad4), ενώ σε άλλα (PML, PML-RARα) είναι αναγκαία προϋπόθεση για να ακολουθήσει η ουβικουΐτινίωσή τους [Kracklauer & Schmidt, 2003/ Lallemand-Breitenbach *et al*, 2001]. Σε αυτή την εργασία ταυτοποιήσαμε την Ubc9, το μοναδικό γνωστό ένζυμο E2 της SUMOλίωσης, μεταξύ των πρωτεϊνών που αλληλεπιδρούν με το Neu1 του ποντικού. Περισσότερα στοιχεία για τον ρόλο της αναφέρονται στη 'Συζήτηση'. Τέλος, η τροποποίηση υποστρωμάτων με

(α) Μηχανισμός SUMΟλίωσης



(β) Ένζυμα SUMΟλίωσης

	E1	E2	E3	Πρωτεάση SUMO
<i>S.c.</i>	Aos1-Uba2	Ubc9	Siz1 Siz2 (Nfi1)	Ulp1 Ulp2 (Smt4)
<i>H.s.</i> <i>M.m.</i> <i>R.n.</i>	Aos1-Uba2 (SAE1-SAE2)	Ubc9	PIAS1(GuBP) PIAS3 (KChaP) PIASαβ (ARIP3, Miz1) PIASγ RanBP2 (Nup358) Pc2	SENP1 SENP2 (Axam, SuPr-1, SMT3IP2) SENP3 (SMT3IP1) SENP6 (SUSP1)

Εικ.15β Η διαδικασία τροποποίησης πρωτεϊνών από SUMO (SUMΟλίωση). α) Τα τρία ενζυμικά στάδια (E1, E2, E3) του μηχανισμού SUMΟλίωσης. Το SUMO ενεργοποιείται με πρωτεόλυση του C-τελικού του άκρου και μεταφέρεται παρουσία ATP διαδοχικά στα ένζυμα E1 (διμερές SAE1-SAE2) και E2 (Ubc9), με θειοεστερικό δεσμό. Από το E2 μεταφέρεται στη λυσίνη-στόχο του υποστρώματος. Το υπόστρωμα A αλληλεπιδρά με το Ubc9 και τροποποιείται απουσία της λιγάσης E3, η οποία έχει ρόλο συνδετικό, διευκολύνοντας την αλληλεπίδραση του Ubc9 με το υπόστρωμα B. Η διαδικασία της SUMΟλίωσης είναι αντιστρεπτή, λόγω της παρουσίας ισοπεπτιδασών του SUMO που αποSUMΟλιώνουν τα τροποποιημένα υποστρώματα. β) Τα ένζυμα E1, E2, οι γνωστές λιγάσες E3 και οι ισοπεπτιδάσες (πρωτεάσες SUMO) που συμμετέχουν στον μηχανισμό της SUMΟλίωσης, στον σακχαρομύκητα (*S.c.*) και στα θηλαστικά (*H.s.*, *M.m.*, *R.n.*).

την πρωτεΐνη Nedd8, γνωστή και ως RUB (*Related to Ubiquitin*), προκαλεί την αποικοδόμησή τους. Τα μόνα γνωστά υποστρώματα της Nedd8 είναι οι κουλίνες (cullins), βασικά συστατικά των πολυπρωτεϊνικών συμπλόκων λιγασών ουβικουϊτίνης SCF (*Skp1/Culin/F-box*). Για παράδειγμα, η τροποποίηση της Cul-1 από την Nedd8, ενεργοποιεί την πρωτεόλυση του IκB μέσω του συμπλόκου SCF [Read *et al*, 2000].

β. Η ουβικουϊτινιλίωση στη ρύθμιση της σηματοδότησης από τον υποδοχέα Notch

Η ουβικουϊτινιλίωση αποτελεί ρυθμιστικό μηχανισμό και για τη σηματοδότηση από τον υποδοχέα Notch. Τόσο στη δροσόφιλα όσο και σε άλλους οργανισμούς έχουν βρεθεί μέχρι τώρα 4 πρωτεΐνες που δρουν ως θετικοί ή αρνητικοί ρυθμιστές της σηματοδότησης Notch, επάγοντας την ουβικουϊτινιλίωση και πρωτεόλυση σηματοδοτικών μορίων αυτού του μονοπατιού. Επιπλέον, η παρουσία στο καρβοξυτελικό άκρο του υποδοχέα Notch ενός μοτίβου PEST (Εικ.12), το οποίο χαρακτηρίζει πρωτεΐνες μικρού χρόνου ημιζωής [Rechsteiner & Rogers, 1996], υποδηλώνει τη ρύθμιση του ίδιου του υποδοχέα από το σύστημα της ουβικουϊτίνης.

Αναλυτικότερα, στα θηλαστικά έχει βρεθεί η πρωτεΐνη SEL-10, η οποία επάγει την ουβικουϊτινιλίωση του Notch4, αλληλεπιδρώντας με το φωσφορυλιωμένο καρβοξυτελικό του άκρο μέσω των επαναλήψεων WD40. Επιπλέον, αλληλεπιδρά με τον παράγοντα SKP1 του συμπλόκου E3 λιγασής ουβικουϊτίνης SCF (Skp1-Cullin-F box), μέσω ενός F-box στο μόριο SEL-10 [Wu *et al*, 2001]. Επίσης, η πρωτεΐνη *Suppressor of deltex* (*Su(dx)*) της δροσόφιλας, για την οποία είναι γνωστή η γενετική αλληλεπίδραση με το Notch, φέρει την περιοχή HECT, χαρακτηριστική για πρωτεΐνες με δράση E3 λιγασής (Εικ.15) [Cornell *et al*, 1999]. Αν και δεν έχει αποδειχθεί άμεση σχέση του μορίου *Su(dx)* στην ουβικουϊτινιλίωση του Notch στη δροσόφιλα, ωστόσο η δομικά ομόλογη πρωτεΐνη των θηλαστικών *Itch* αλληλεπιδρά με το Notch και το ουβικουϊτινιλιώνει, τόσο *in vitro* όσο και σε ανθρώπινες κυτταρικές σειρές [Qiu *et al*, 2000]. Επιπλέον, μεταλλαγές σε υπομονάδες του 20S πρωτεασώματος προκαλούν φαινοτύπους υπερέκφρασης του Notch κατά την ανάπτυξη των αισθητηρίων οργάνων και αυξάνουν τη σταθερότητα της ενεργοποιημένης μορφής του υποδοχέα σε έμβρυα δροσόφιλας [Schweisguth, 1999]. Αντίθετα, η πρωτεΐνη LNX των θηλαστικών λειτουργεί ως θετικός ρυθμιστής της σηματοδότησης από τον υποδοχέα Notch, επάγοντας την ουβικουϊτινιλίωση και πρωτεόλυση της πρωτεΐνης Numb, η οποία έχει κατασταλτική δράση για τον υποδοχέα Notch. Η δράση της LNX ως E3 λιγασής ουβικουϊτίνης οφείλεται στις περιοχές RING και PDZ στο αμινοτελικό της άκρο [Nie *et al*, 2002]. Τέλος, η πρωτεΐνη *mind-bomb* (*mib*), η οποία είναι γνωστή για τον νευρογόνο φαινότυπο της μεταλλαγής της στο ψάρι *Danio rerio*, ρυθμίζει τη σηματοδότηση από τον υποδοχέα Notch, κατά την επιλογή των πρόδρομων νευρικών κυττάρων στα έμβρυα με πλευρική αναστολή [Bingham *et al*, 2003]. Πρόσφατες μελέτες δείχνουν ότι η *mib* φέρει στο καρβοξυτελικό άκρο δύο μη τυπικούς και έναν τυπικό δακτύλιο RING, ο οποίος ευθύνεται για τη δράση της ως E3 λιγασής ουβικουϊτίνης. Συγκεκριμένα, πειράματα διαμολύνσεων κυτταρικών σειρών θηλαστικών και *in vitro* δοκιμές δείχνουν ότι η πρωτεΐνη *mib* επάγει την ουβικουϊτινιλίωση και πρωτεόλυση του δεσμευτών *deltaD* (του *Danio rerio*) και *Delta1* (του *Xenopus laevis*). Επιπλέον, φαίνεται να συμμετέχει σε μηχανισμούς ενδοκύττωσης, προκαλώντας την ενδοκύττωση των μορίων *Delta1* και *deltaD* [Itoh *et al*, 2003 / Chen & Casey Corliss, 2004], ενισχύοντας με αυτόν τον τρόπο τη σηματοδότηση από τον υποδοχέα Notch [Le Borgne & Schweisguth, 2003b]. Για το ορθόλογο μόριο των θηλαστικών *DIP-1* (*DAPK kinase Interacting Protein 1*), δεν έχει ακόμα διερευνηθεί ο πιθανός του ρόλος στη σηματοδότηση Notch. Όμως, είναι γνωστό ότι αλληλεπιδρά

με την πρωτεϊνική κινάση σερίνης/θρεονίνης DAPK (*Death Associated Protein Kinase*), την οποία και ρυθμίζει λειτουργώντας ως E3 λιγάση για την ουβικουΐτινιλίωσή της τόσο *in vitro*, όσο και σε κυτταρικές σειρές [Jin *et al*, 2002].

γ. Δομικά και λειτουργικά στοιχεία που υποστηρίζουν τη συμμετοχή του neuralized σε διαδικασίες ουβικουΐτινιλίωσης

Όπως αναφέρθηκε και παραπάνω, τα μόρια της οικογένειας Neuralized φέρουν τη συντηρημένη περιοχή RING, η οποία πρόσφατα έχει συσχετιστεί με τη λειτουργία των E3 λιγασών στη διαδικασία της ουβικουΐτινιλίωσης [Lorick *et al*, 1999/ Jackson *et al*, 2000]. Επιπλέον, από γενετικά και βιοχημικά πειράματα με τα ορθόλογα μόρια neuralized της *Drosophila* και του *Xenopus* προκύπτουν τα εξής συμπεράσματα: α) η περιοχή του RING είναι απαραίτητη για τη λειτουργία του neur, τουλάχιστον όσον αφορά στην ενίσχυση της ενδοκύτωσης του Delta και την ακόλουθη πρωτεόλυσή του [Pavlopoulos *et al*, 2001/ Lai *et al*, 2001/ Deblandre *et al*, 2001], β) ελλειμματικά μόρια neurΔRING εμφανίζουν αυξημένη σταθερότητα όταν εκφράζονται σε κυτταρικές σειρές ή σε εμβρυϊκούς δίσκους σε σχέση με το πλήρους μήκους neur [Lai *et al*, 2001 / Deblandre *et al*, 2001], γ) μεταλλαγή του E2 ενζύμου UbcD1 καταστέλλει μερικώς το φαινότυπο της εκτοπικής υπερέκφρασης των μορίων neur και neurΔRING [Lai *et al*, 2001], δ) η μείωση της ενεργότητας του προτεασώματος με κατευθυνόμενη έκφραση μεταλλαγμένης υπομονάδας του σε εμβρυϊκούς δίσκους φτερού στη δροσόφιλα, αυξάνει τα επίπεδα του ενδογενούς Delta και μειώνει σημαντικά το αποτέλεσμα της εκτοπικής έκφρασης του neur [Lai *et al*, 2001], ε) σε *in vitro* πειράματα ουβικουΐτινιλίωσης, οι πρωτεΐνες Neuralized της δροσόφιλας και του βατράχου που έχουν εκφραστεί σε κύτταρα *E.coli*, εμφανίζουν δράση E3 λιγάσης ουβικουΐτίνης, η οποία εξαρτάται από την παρουσία της περιοχής του RING και στα δύο μόρια [Yeh *et al*, 2001 / Lai *et al*, 2001 / Deblandre *et al*, 2001]. Τέλος, έχει ανιχνευθεί το πρωτεϊνικό σύμπλοκο στο οποίο συμμετέχουν οι πρωτεΐνες neur και Delta της δροσόφιλας *in vivo*. Η παρουσία του neur σε αυτό το σύμπλοκο οφείλεται σε αλληλεπιδράσεις που πραγματοποιούνται μέσω της περιοχής NEUZ1 [Lai *et al*, 2001].

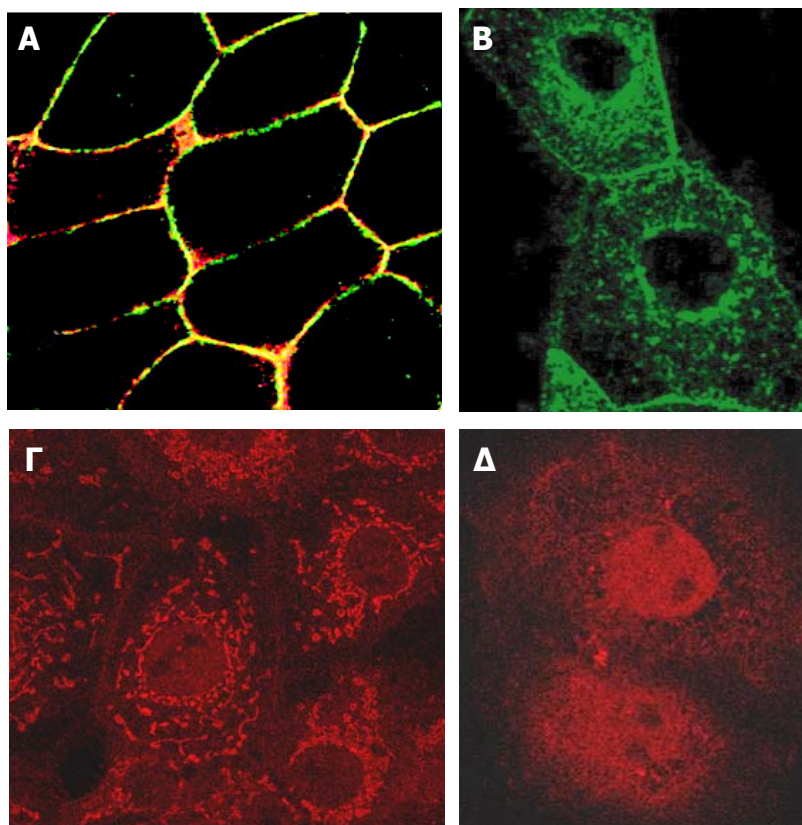
Τα παραπάνω υποστηρίζουν ότι η δράση της πρωτεΐνης neur στο μονοπάτι σηματοδότησης του Notch περιλαμβάνει και διαδικασίες ουβικουΐτινιλίωσης για τη ρύθμιση των επιπέδων του ίδιου του μορίου neur, ή και για την τροποποίηση άλλων συστατικών του μονοπατιού, όπως του δεσμευτή Delta. Βέβαια, τα παραπάνω αποτελέσματα δεν αποδεικνύουν άμεση αλληλεπίδραση των μορίων Neur1 και Delta *in vivo*, ούτε συνδέουν άμεσα την παρατηρούμενη μείωση των επιπέδων του Delta με ουβικουΐτινιλίωση του ίδιου μορίου στα κύτταρα της δροσόφιλας ή του βατράχου. Το στοιχείο ωστόσο που αποδεικνύεται είναι η αναγκαιότητα της παρουσίας του RING για τη ρύθμιση των επιπέδων της πρωτεΐνης neuralized.

9. Υποκυτταρικός εντοπισμός της πρωτεΐνης Neuralized

Οι ορθόλογες πρωτεΐνες της οικογένειας Neuralized δεν φέρουν χαρακτηριστικά μοτίβα ή σινιάλα υποκυτταρικού εντοπισμού. Η θέση της πρωτεΐνης Neur1 υποκυτταρικά έχει διερευνηθεί τόσο στη δροσόφιλα όσο και στα θηλαστικά με χρήση ανοσοϊστοχημικών και βιοχημικών μεθόδων. Συγκεκριμένα, χιμαιρικά πολυπεπίδια δροσοφιλικού neur που υπερεκφράζονται εκτοπικά σε κύτταρα εμβρυϊκών δίσκων, ανιχνεύονται στην

περιφέρεια της κυτταρικής μεμβράνης [Lai & Rubin, 2001b/ Yeh *et al*, 2000/ Pavliopoulos *et al*, 2001] (Εικ.16). Επιπλέον, βιοχημικά πειράματα δείχνουν ότι η πρωτεΐνη εντοπίζεται στην κυτταροπλασματική πλευρά της κυτταρικής μεμβράνης και ότι αυτός ο εντοπισμός οφείλεται σε ασθενείς αλληλεπιδράσεις με άλλα μόρια [Pavliopoulos *et al*, 2001].

Αντίστοιχες μελέτες υποκυτταρικού εντοπισμού της πρωτεΐνης στα θηλαστικά δείχνουν ότι πρόκειται για ένα μόριο που μετακινείται από την κυτταροπλασματική πλευρά της κυτταρικής μεμβράνης μέχρι και το εσωτερικό του πυρήνα. Συγκεκριμένα, σε πειράματα υπερέκφρασης χιμαιρικών πολυπεπτιδίων *Neur1* [Pavliopoulos, Kokkinaki *et al*, 2002], καθώς και ανοσοανίχνευσης του ενδογενούς *Neur1* σε κυτταρικές σειρές θηλαστικών [Ε. Κουτέλου, αδημοσίευτα αποτελέσματα], η πρωτεΐνη κυρίως ανιχνεύεται σε στικτές δομές, διάσπαρτες στο κυτταρόπλασμα, σχηματίζοντας πολλές φορές ένα δακτύλιο γύρω από τον πυρήνα, περισσότερο ή λιγότερο συμπαγή, ενώ σε σημαντικό ποσοστό των κυττάρων παρατηρείται πυρηνικός εντοπισμός της πρωτεΐνης. Σε μικρότερο ποσοστό κυττάρων (5-10%), η ενδογενής πρωτεΐνη *Neur1* εντοπίζεται στην πλασματική μεμβράνη (Εικ.16). Βιοχημικά πειράματα δείχνουν ότι η πρωτεΐνη εντοπίζεται στο μεμβρανικό κλάσμα εκχυλισματος διαμολυσμένων κυττάρων και συγκεκριμένα στην κυτταροπλασματική επιφάνεια μεμβρανικών κυστιδίων, όπου συγκρατείται με ασθενείς αλληλεπιδράσεις. Ο εντοπισμός αυτός υποστηρίζει το μοντέλο της συμμετοχής του μορίου *Neur1* σε μηχανισμούς ενδοκύττωσης, όπως προτείνεται με βάση προκαταρκτικά πειράματα μικροσκοπίας σε συνθήκες καταστολής της ενδοκύττωσης [Η. Παυλόπουλος, Διδακτορική Διατριβή, Παν/μιο Κρήτης, 2002]. Επιπλέον, η ενδογενής πρωτεΐνη *Neur1* εντοπίζεται τόσο στο κυτταροπλασματικό όσο και στο πυρηνικό κλάσμα κυττάρων θηλαστικών, με βάση πρόσφατα βιοχημικά πειράματα κυτταρικής κλασμάτωσης [Ε. Κουτέλου, αδημοσίευτα αποτελέσματα]. Ο πυρηνικός εντοπισμός της πρωτεΐνης έχει διερευνηθεί και από την ομάδα των Timmusk *et al* (2002), οι οποίοι παρατήρησαν πυρηνικό εντοπισμό στο 40% των κυττάρων της κυτταρικής σειράς *Neuro2a* που είχαν διαμολύνει με χιμαιρικά πολυπεπτιδία *Neur1*. Με βάση παρατηρήσεις της ίδιας ομάδας, τα αμινοξέα 60-120 που εντοπίζονται στην περιοχή NEUZ1 και είναι πλούσια σε αργινίνη και λυσίνη, ευθύνονται για τον πυρηνικό εντοπισμό. Επιπλέον, παρατηρώντας τα διαμολυσμένα κύτταρα παρουσία λεπτομυκίνης Β (*Leptomycin B*), η οποία δρα ως αναστολέας εξόδου από τον πυρήνα μέσω του συστήματος CRM1/exportin1 [Hood & Silver, 1999], διαπίστωσαν αποκλειστικά πυρηνικό πρότυπο υποκυτταρικού εντοπισμού της πρωτεΐνης *Neur1*, γεγονός που σημαίνει ότι φυσιολογικά η πρωτεΐνη κυκλοφορεί μέσα και έξω από τον πυρήνα [Timmusk *et al*, 2002], χωρίς ακόμα να έχει διερευνηθεί η βάση αυτού του φαινομένου. Γενική διαπίστωση σε όλες τις παραπάνω μελέτες ήταν ότι η περιοχή RING της πρωτεΐνης *Neur1* δεν επηρεάζει τον υποκυτταρικό της εντοπισμό. Αντίθετα, καθοριστική για τη θέση του μορίου στα κύτταρα είναι η περιοχή NEUZ1 [Pavliopoulos, Kokkinaki *et al*, 2002/ Timmusk *et al*, 2002].



Εικ.16 Πειράματα υποκυτταρικού εντοπισμού της πρωτεΐνης *Neuralized*. **(Α)**. Ανοσοεντοπισμός χιμαιρικής πρωτεΐνης *neur* (κόκκινο) υπερεκφρασμένης σε κύτταρα σιελογόνων αδένων της *D.melanogaster*. Με κίτρινο φαίνεται ο συνεντοπισμός της πρωτεΐνης *neur* με την α -σπεκτρίνη (πράσινο), μια περιφερειακή μεμβρανική πρωτεΐνη [Yeh E *et al*, 2000]. **(Β)**. Χιμαιρική πρωτεΐνη EGFP-*Neur1* ποντικού, εκφρασμένη στην κυτταρική σειρά θηλαστικών COS7 [Pavliouros *et al*, 2002]. **(Γ,Δ)**. Πειράματα ανοσοεντοπισμού της ενδογενούς πρωτεΐνης *Neur1* με χρήση πολυκλωνικού αντισώματος σε κύτταρα COS7, όπου παρατηρείται κυτταροπλασματικό (Γ) ή πυρηνικό (Δ) πρότυπο [Κουτέλου Ε, αδημοσίευτα αποτελέσματα].

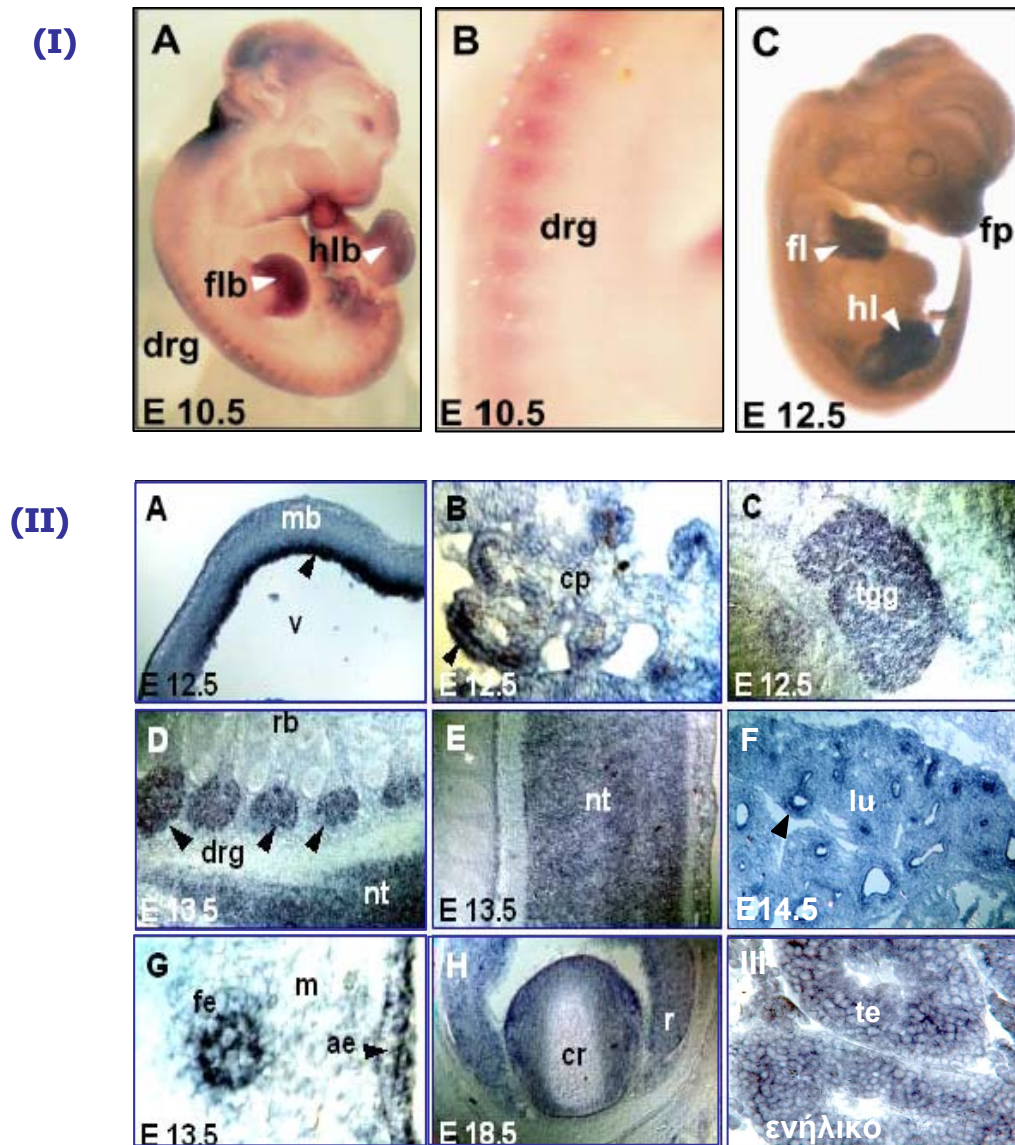
10. Λειτουργία του *Neuralized* στα θηλαστικά

Σύμφωνα με τα δεδομένα από τη λειτουργική μελέτη του *neuralized* στη δροσόφιλα και τις ενδείξεις για τη λειτουργία του ορθόλογου γονιδίου στον βάτραχο, το *neur* φαίνεται να ρυθμίζει θεμελιώδεις αναπτυξιακούς μηχανισμούς κυτταρικής επικοινωνίας, όπως αυτόν της σηματοδότησης από τον υποδοχέα Notch, που συναντώνται σε εξελικτικά συντηρημένες διαδικασίες διαφοροποίησης κυττάρων κατά την ανάπτυξη του εμβρύου και τη ζωή του ενήλικου. Η ίδια ένδειξη προκύπτει και από τη μελέτη του προτύπου έκφρασης του κατά την ανάπτυξη των θηλαστικών. Φαίνεται, επομένως, ότι η λειτουργία του *Neur1* στα θηλαστικά σχετίζεται με γεγονότα που συμβαίνουν νωρίς κατά την εμβρυογένεση. Η πιθανή συμμετοχή του *Neur1* στη σηματοδότηση Notch υποστηρίζεται από τη συγκριτική ανάλυση του προτύπου έκφρασης του *Neur1* και των συνδεδών του Notch στα θηλαστικά *Dll1*, *Dll4*, *Jag1* και *Jag2* [E. Κουτέλου, Μεταπτυχιακή Διατριβή, Παν/μιο Κρήτης, 2001].

α. Πρότυπο έκφρασης

Η έκφραση του *Neuralized* των θηλαστικών έχει μελετηθεί σε καλλιεργούμενες κυτταρικές σειρές, σε έμβρυα και σε ιστούς ενηλίκων στον άνθρωπο, στον αρουραίο και στον ποντικό. Το γονίδιο *Neur1* είναι μεταγραφικά ενεργό από τα αρχικά στάδια της εμβρυϊκής ανάπτυξης του ποντικού σε συγκεκριμένες κυτταρικές ομάδες, κυρίως επιθήλια, σε ποικιλία εσωτερικών και αισθητηρίων οργάνων καθώς και στο νευρικό και στηρικτικό σύστημα. Στο ενήλικο, το γονίδιο εμφανίζει πιο περιορισμένη έκφραση και ανιχνεύεται σε υψηλά επίπεδα

σε εγκέφαλο, ήπαρ, σκελετικό μυ και στους όρχεις. Ειδικότερα, με ανάλυση *Northern* και RT-PCR παρατηρείται έντονη έκφραση του ανθρώπινου ορθόλογου (*NEURL*) σε φυσιολογικό εγκέφαλο και απουσία ή πολύ χαμηλά επίπεδα των μεταγράφων του σε κυτταρικές σειρές γλοιοβλαστώματος, υποδεικνύοντας ενδεχόμενη συσχέτιση της χαμηλής έκφρασης του *NEURL* με την εξέλιξη των όγκων [Nakamura *et al*, 1998]. Υψηλά επίπεδα έκφρασης του γονιδίου στον αρουραίο (*Rneu*) έχουν ανιχνευθεί με πειράματα υβριδοποίησης *in situ* σε τομές ενήλικου εγκεφάλου. Επιπλέον, διαπιστώνεται ότι τα μεταγγραφα του *Rneu* σε συγκεκριμένες περιοχές του εγκεφάλου όπως είναι ο ιππόκαμπος, εντοπίζονται εκτός των νευρικών σωματίων και στους δενδρίτες των νευρώνων, υποδηλώνοντας πιθανή συμμετοχή του *Rneu* στη λειτουργία των συνάψεων [Timmusk *et al*, 2002]. Στον ποντικό, πειράματα υβριδοποίησης σε ολόκληρα έμβρυα ηλικίας E9.5–E14.5 δείχνουν ότι το γονίδιο είναι μεταγραφικά ενεργό στα ραχιαία ριζικά γάγγλια, στα αναπτυσσόμενα άκρα και στην περιοχή του προσώπου [Pavliopoulos, Kokkinaki *et al*, 2002] (Εικ.17). Σε τομές εμβρύων ποντικού ηλικίας E12.5–E18.5 διαπιστώνεται έκφραση του *Neurl* σε ιστούς νευρικής προέλευσης, όπως στην κοιλιακή ζώνη του μεσεγκεφάλου, στην επιθηλιακή στοιβάδα του χοριοειδούς πλέγματος, στο τρίδυμο νεύρο, στον νευρικό σωλήνα και στα ραχιαία ριζικά γάγγλια. Υψηλά επίπεδα μεταγράφων ανιχνεύονται και σε αναπτυσσόμενα αισθητήρια όργανα και συγκεκριμένα στα επιθηλιακά κύτταρα των εκβλαστημάτων των μουστακίων, στον αμφιβληστροειδή χιτώνα και στα κύτταρα του φακού του ματιού καθώς και στις επιθηλιακές στιβάδες της ρινικής κοιλότητας και των ρινικών αδένων. Επιπλέον, παρατηρείται έκφραση του *Neurl* σε διακριτές κυτταρικές ομάδες των εσωτερικών οργάνων και συγκεκριμένα στο βρογχικό επιθήλιο των πνευμόνων, στη στιβάδα των σωληνοειδών επιθηλιακών κυττάρων των νεφρικών σπειραμάτων, στο βλεννογόνο επιθήλιο του στομάχου, στο ήπαρ, στο επιθήλιο του μεσονεφρικού βολφιανού ουρογεννητικού αγωγού και στον καρδιακό ιστό. Μεταγραφική ενεργότητα του γονιδίου ανιχνεύεται και στις επιθηλιακές στιβάδες του δέρματος, σε επιθηλιακά και μεσεγγυματικά κύτταρα της περιοχής των αναπτυσσόμενων δοντιών καθώς και στο αναπτυσσόμενο στηρικτικό σύστημα, όπως σε μύες, στο περιχόνδριο των πλευρών, στα οστά των γνάθων και των πρόσθιων και οπίσθιων άκρων [Pavliopoulos, Kokkinaki *et al*, 2002] (Εικ.17). Τέλος, με χρήση της αντίδρασης RT-PCR σε δείγματα RNA από ιστούς ενήλικου ποντικού και από κυτταρικές σειρές, έχει διαπιστωθεί η έκφραση του γονιδίου στον εγκέφαλο, στο ήπαρ και στους όρχεις ενήλικου ποντικού, καθώς και στις εμβρυονικές κυτταρικές σειρές ινοβλαστών και τερατοκαρκινωμάτων NIH/3T3 και P19, αντίστοιχα [Pavliopoulos, Kokkinaki *et al*, 2002]. Επιπλέον, έχει ανιχνευθεί ένα εναλλακτικό μεταγγραφο *Neurl* να εκφράζεται ειδικά σε σκελετικό μυ σε ανάλυση προστασίας από RNase και σε ανάλυση *Northern* [Timmusk *et al*, 2002 /Vollrath *et al*, 2001].



Εικ.17 Πειράματα μελέτης του αναπτυξιακού προτύπου έκφρασης του γονιδίου *Neurl* στον ποντικό, με τη μέθοδο της υβριδοποίησης RNA *in situ*. **(I)** Υβριδοποίηση σε ολόκληρα έμβρυα ηλικίας E10.5 (IA) και E12.5 (IB, IC). Μετάγραφα του *Neurl* ανιχνεύονται στα αναπτυσσόμενα πρόσθια και οπίσθια άκρα (IA: flb, hlb και IC: fl, hl), στα ραχιαία ριζικά γάγγλια (IB: drg) και σε δομές του προσώπου (IC: fp). **(II)** Υβριδοποίηση σε τομές εμβρύων ηλικίας E12.5 – E18.5 (IIA-H) ή οργάνων ενηλίκων (III). Έκφραση του *Neurl* ανιχνεύεται στην κοιλιακή ζώνη του μεσεγκεφάλου (IIA: mb), στην επιθηλιακή στοιβάδα του χοριοειδούς πλέγματος (IIB: cp), στο τρίδυμο νεύρο (IIC: tagg), στα ραχιαία ριζικά γάγγλια (IID: drg), στον νευρικό σωλήνα (IIE: nt), στο βρογχικό επιθήλιο των πνευμόνων (IIF: lu), στο θυλακοειδές επιθήλιο των εκβλαστημάτων των μωστακίων (IIG: fe), στον αμφιβληστροειδή χιτώνα και στα κύτταρα του φακού του ματιού (IIH: r και cr) [Pavloroulos E *et al*, 2002]. Τέλος, εικονίζεται η έκφραση σε τομή όρχεων ενήλικου ποντικού (III:te) [Μπαζοπούλου Μ-Δ, διπλωματική εργασία, Παν/μιο Κρήτης, 2002].

β. Στοχευμένη απενεργοποίηση (*knock-out*) του *Neuralized* στον ποντικό

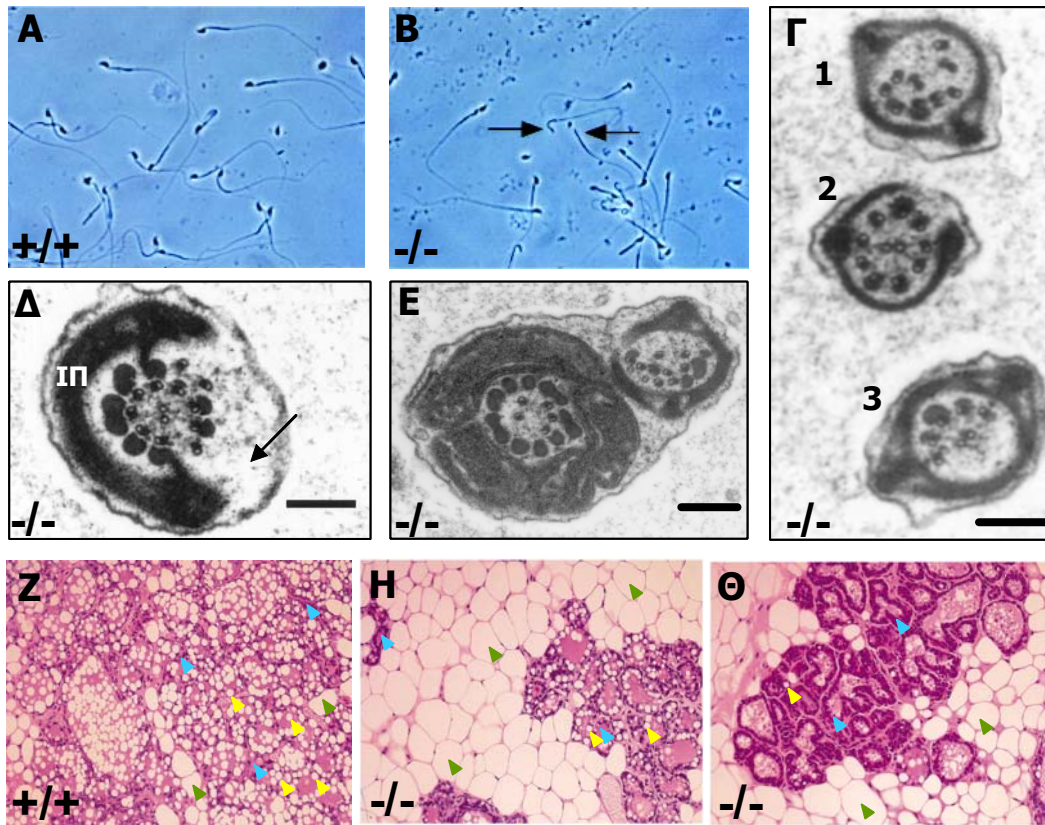
Η μέθοδος της στοχευμένης απενεργοποίησης (*knock-out*) στηρίζεται στη δυνατότητα ομόλογου ανασυνδυασμού ανάμεσα στη χρωμοσωμική περιοχή που φέρει το γονίδιο-στόχο και σε μια πλασμιδιακή κατασκευή που φέρει κλωνοποιημένες τις περιοχές εκατέρωθεν του γονιδίου-στόχου, με αποτέλεσμα την αντικατάσταση ολόκληρης ή μέρους της κωδικής περιοχής του από ένα γονίδιο-μάρτυρα (πχ *neo^R*). Ο ομόλογος ανασυνδυασμός συμβαίνει σε πολυδύναμα κύτταρα βλαστοκύστης ποντικού πριν την εμφύτευση (κύτταρα ES), τα οποία διατηρούν την ικανότητα ακόμα και μετά από καλλιέργεια να ενσωματώνονται στη γαμετική σειρά χιμαιρικών ποντικών [Bradley *et al*, 1984], όταν μεταφέρονται σε βλαστοκύστες. Από τη διασταύρωση των χιμαιρικών

κών ζώων που φέρουν το μεταλλαγμένο αλληλόμορφο στη γαμετική τους σειρά, προκύπτουν άτομα ομόζυγα για το απενεργοποιημένο γονίδιο σε ποσοστό 25% των απογόνων της F1 γενεάς. Την τελευταία δεκαετία, η προσέγγιση αυτή γνωρίζει σημαντική πρόοδο, λόγω των πειραμάτων στοχευμένης απενεργοποίησης (*knock-out*) του 10-15% των γονιδίων του ποντικού [Bradley, 2002].

Με στόχο τη διερεύνηση της λειτουργίας του *Neurl*, έχουν πραγματοποιηθεί δύο ανεξάρτητα πειράματα στοχευμένης απενεργοποίησής του. Συγκεκριμένα, οι Ruan *et al* (2001) αντικατέστησαν το τμήμα του γονιδιώματος που περιέχει τα τρία τελευταία εξόνια (E4-E6) του *Neurl* με την κασέτα της νεομυκίνης, ενώ οι Vollrath *et al* (2001) πραγματοποίησαν ένθεση στο τέταρτο εξόνιο του *Neurl* της κασέτας που περιελάμβανε την αλληλουχία IRES (*I*nternal *R*ibosome *E*ntry *S*ite) μπροστά από το γονίδιο δείκτη GFP (*G*reen *F*luorescent *P*rotein) και το γονίδιο της νεομυκίνης με τα 5' ρυθμιστικά του στοιχεία. Επομένως, και στα δύο πειράματα αφαιρέθηκε η περιοχή του μεταγράφου που προκύπτει από τα εξόνια E4-E6 και κωδικοποιεί την δεύτερη περιοχή NEUZ (NEUZ2) και το RING της πρωτεΐνης *Neurl*.

Φαίνεται ότι οι τροποποιήσεις στις πλασμιδιακές κατασκευές στόχευσης του γονιδίου, αλλά και το διαφορετικό γενετικό υπόβαθρο των ζώων που χρησιμοποιήθηκαν από τις δύο ομάδες, οδήγησαν σε διαφορετικούς φαινοτύπους έλλειψης του *Neurl* σε κάθε πείραμα. Συγκεκριμένα, στο πείραμα των Ruan *et al*, τα πολυδύναμα κύτταρα ES του τύπου JM1 που έφεραν τον επιθυμητό ομόλογο ανασυνδυασμό μεταφέρθηκαν σε βλαστοκύστες στελέχους C57BL/6, από τις οποίες επιλέχθηκαν τα αρσενικά χιμαιρικά ποντίκια που εμφάνιζαν τον ομόλογο ανασυνδυασμό στη γαμετική σειρά. Στη συνέχεια, αυτά διασταυρώθηκαν με αγρίου τύπου θηλυκά του ίδιου στελέχους (C57BL/6) για να δώσουν απογόνους ετερόζυγους για τη μεταλλαγή του *Neurl*. Από τη διασταύρωση των τελευταίων προέκυψαν τα ομόζυγα για τη μεταλλαγή ποντίκια (*Neurl* *-/-*). Αντίστοιχα, στο πείραμα των Vollrath *et al*, τα πολυδύναμα κύτταρα ES του τύπου TC1 που έφεραν τον επιθυμητό ομόλογο ανασυνδυασμό μεταφέρθηκαν σε βλαστοκύστες στελέχους C57BL/6, από τις οποίες επιλέχθηκαν τα αρσενικά χιμαιρικά ποντίκια που εμφάνιζαν τον ομόλογο ανασυνδυασμό στη γαμετική σειρά. Στη συνέχεια, αυτά διασταυρώθηκαν με αγρίου τύπου θηλυκά του στελέχους 129/SvEv για να δώσουν απογόνους ετερόζυγους για τη μεταλλαγή του *Neurl*. Από τη διασταύρωση των τελευταίων προέκυψαν τα ομόζυγα για τη μεταλλαγή ποντίκια (*Neurl* *-/-*). Επιπλέον, αξιοσημείωτο είναι και το γεγονός ότι στο πείραμα των Ruan *et al*, με ανάλυση *Northern* δειγμάτων RNA από τα ποντίκια *Neurl* *-/-* δεν ανιχνεύεται το μετάγραφο που παράγεται από τα εξόνια E1-E3 που δεν επηρεάζονται από τον ομόλογο ανασυνδυασμό, πράγμα που πιθανόν οφείλεται σε αστάθεια του ελλειμματικού αυτού μορίου RNA. Αντίθετα, στο πείραμα των Vollrath *et al*, η παρουσία μιας εσωτερικής θέσης πρόσδεσης σε ριβόσωμα (αλληλουχία IRES) συντηγμένης με την πρωτεΐνη GFP καθώς και σημάτων τερματισμού της μεταγραφής που ακολουθούν το γονίδιο *GFP* φαίνεται ότι σταθεροποιούν το μεταλλαγμένο μετάγραφο, με αποτέλεσμα αυτό να ανιχνεύεται σε ανάλυση *Northern*. Κατά συνέπεια, από το πείραμα των Vollrath *et al* δεν αποκλείεται να παράγεται υπομορφικό και όχι πλήρως απενεργοποιημένο αλληλόμορφο του *Neurl*, πιθανότητα που η ίδια ερευνητική ομάδα θεωρεί εξαιρετικά μικρή με βάση τη στρατηγική στόχευσης συντηρημένων περιοχών του γονιδίου.

Ενδεχομένως, οι διαφορετικές στρατηγικές στόχευσης του γονιδίου που εφαρμόστηκαν από τις δύο ομάδες αλλά και η διαφορά του γενετικού υπόβαθρου στα στελέχη των ποντικών που χρησιμοποιήθηκαν σε κάθε πείραμα απενεργοποίησης του *Neurl*, οδήγησαν σε διαφορετικούς φαινοτύπους, αποκαλύπτοντας όμως έτσι διαφορετικές λειτουργίες στις οποίες απαιτείται η ενεργότητα της πρωτεΐνης *Neurl*, με τρόπο που πιθανόν επηρεάζεται από τη διαφορική έκφραση άλλων τροποποιητικών γονιδίων (*modifier genes*) στα διαφορετικά στελέχη ποντικών *Neurl*^{-/-}, ή από τη διαφορική παρουσία πρωτεϊνών που δρουν συμπληρωματικά ή υποκαθιστούν τη λειτουργία του *Neurl*. Συγκεκριμένα, οι Ruan *et al* προσδιόρισαν συγκεκριμένες λειτουργίες του εγκεφάλου για τις οποίες είναι απαραίτητη η ενεργότητα του *Neurl*. Αναλυτικότερα, τα ποντίκια *Neurl*^{-/-} είναι βιώσιμα και χωρίς μεγάλες μορφολογικές ανωμαλίες, αλλά εμφανίζουν υστέρηση στη διάκριση οσμών και αδυναμία συντονισμού των κινήσεών τους παρουσία αιθανόλης στο αίμα τους. Επιπλέον, σύμφωνα με τους Ruan *et al*, αντιδρούν φυσιολογικά σε διάφορες δοκιμές μάθησης και μνήμης. Και στο πείραμα των Vollrath *et al* τα ποντίκια *Neurl*^{-/-} επιβιώνουν και αναπτύσσονται φυσιολογικά, χωρίς να παρουσιάζουν ανωμαλίες στην ανάπτυξη του εγκεφάλου, του σκελετού και των μυών, διαδικασίες που ρυθμίζονται από τη σηματοδότηση *Notch* στα θηλαστικά. Όμως, τα αρσενικά ποντίκια *Neurl*^{-/-} των Vollrath *et al* εμφανίζουν στειρότητα, η οποία οφείλεται σε έλλειψη κινητικότητα των σπερματοζωαρίων τους και όχι σε ανωμαλίες κατά τη διαδικασία της σπερματογένεσης. Συγκεκριμένα, παρατηρείται διαταραγμένη δομή του αξονήματος στο κεντρικό τμήμα (*principal piece*) των μαστιγίων (Εικ.18). Τα θηλυκά ποντίκια *Neurl*^{-/-} των Vollrath *et al* είναι γόνιμα, όμως εμφανίζουν ανώριμους γαλακτοφόρους αδένες κατά το θηλασμό, με αποτέλεσμα την αυξημένη θνησιμότητα των απογόνων τους. Συγκεκριμένα, παρατηρήθηκε σημαντική μείωση του αριθμού των κυψελίδων του γαλακτοφόρου αδένου που φυσιολογικά διαπερνούν τον λιπώδη ιστό του μαστού, αλλά και ελαττωμένος λιπώδης ιστός στους μαστούς των θηλυκών ζώων *Neurl*^{-/-}, σε σχέση με τα αγρίου τύπου (Εικ.18). Τέλος, στη μελέτη των Vollrath *et al* τονίζεται ότι τόσο στα ομόζυγα μεταλλαγμένα ζώα *Neurl*^{-/-}, όσο και στα ετερόζυγα *Neurl*^{+/-}, δεν παρατηρείται αυξημένη συχνότητα εμφάνισης νεοπλασιών του κεντρικού νευρικού συστήματος σε σχέση με τα αγρίου τύπου ζώα, γεγονός που δεν υποστηρίζει το προτεινόμενο μοντέλο των Nakamura *et al* (1998), για συσχέτιση της μειωμένης έκφρασης του γονιδίου με την εμφάνιση γλοιοβλαστωμάτων στον άνθρωπο.



Εικ.18 Φαινότυπος των *Neur1*^{-/-} ζώων. (A): Σπερματοζωάρια απομονωμένα από ποντίκια αγρίου τύπου, (B): Σπερματοζωάρια των ομόζυγων μεταλλαγμένων ζώων *Neur1*^{-/-}, τα οποία εμφανίζουν μειωμένη κινητικότητα και μορφολογικές ανωμαλίες (βέλη), (Γ-Ε): Παρατήρηση σε ηλεκτρονικό μικροσκόπιο εγκάρσιων τομών του κεντρικού τμήματος (*principal piece*) του μαστιγίου σπερματοζωαρίων των *Neur1*^{-/-} ζώων. Αν και σε ορισμένες τομές (Γ2) εμφανίζεται η φυσιολογική δομή (9+2) των μικροσωληνίσκων του αξονήματος, στις περισσότερες παρατηρούνται αναδιατάξεις (Γ1, Δ, Ε) ή έλλειμμα του μισού αξονήματος (Γ3). Συχνά λείπει μέρος του ινώδους περιβλήματος (ΙΠ) του μαστιγίου (Δ) ή τα μαστίγια είναι διπλασιασμένα με αποδιοργανωμένο αξονήμα (Ε). (Ζ-Θ): Χρώση αιματοξυλίνης-εοζίνης γαλακτοφόρων αδένων ποντικών αγρίου τύπου (Ζ) ή ομόζυγων μεταλλαγμένων (Η,Θ) κατά την πρώτη ημέρα θηλασμού. Στα *Neur1*^{-/-} ζώα παρατηρούνται είτε λιγότερες κυψελίδες των γαλακτοφόρων αδένων μέσα στον λιπώδη ιστό (Η), είτε μειωμένη παραγωγή λιπιδίων στους αγωγούς των αδένων (Θ), σε σχέση με τα αγρίου τύπου ζώα. (Με βέλη υποδεικνύονται: ο λιπώδης ιστός (πράσινο), οι κυψελοειδείς αγωγοί του γαλακτοφόρου αδέν (μπλε) και τα συσσωματώματα των παραγόμενων λιπιδίων στο εσωτερικό των αγωγών (κίτρινο)).

Τα δεδομένα από τη μελέτη του προτύπου έκφρασης και την απενεργοποίηση του γονιδίου στα θηλαστικά παρέχουν χρήσιμες πληροφορίες για το χρονικό και χωρικό πλαίσιο της λειτουργίας του *Neur1* σε επίπεδο οργανισμού, όμως δεν προσδιορίζουν τη λειτουργία του σε μοριακό επίπεδο. Για το σκοπό αυτό χρειάζεται να ταυτοποιηθούν τα μόρια που συμμετέχουν στις ίδιες λειτουργίες αλληλεπιδρώντας με την πρωτεΐνη *Neur1*.

Εδώ πρέπει να ληφθεί υπόψη η αυξημένη πολυπλοκότητα σε σχέση με το σύστημα της δροσόφιλας. Μπορεί οι μηχανισμοί ανάπτυξης να περιλαμβάνουν συντηρημένα σηματοδοτικά μονοπάτια, όπως αυτό του υποδοχέα *Notch*, όμως τα σηματοδοτικά μόρια είναι περισσότερα και κατά συνέπεια είναι μεγαλύτερος και ο αριθμός των διαφορετικών συνδυασμών με τους οποίους αυτά αλληλεπιδρούν. Χαρακτηριστικό είναι το παράδειγμα των δεσμευτών του *Notch*, οι οποίοι στη δροσόφιλα αντιστοιχούν σε δύο μόρια (*Delta* και *Serrate*), ενώ στα θηλαστικά κωδικοποιούνται από πέντε διαφορετικά γονίδια (*Jagged 1* [Lindsell *et al*, 1995], *Jagged 2* [Shawber *et al*, 1996], *Delta-like 1* [Bettenhausen *et al*, 1995], *Delta-like 3* [Dunwoodie *et al*, 1997] και *Delta-like 4* [Shutter *et al*, 2000]). Επομένως, η πρωτεΐνη *Neur1* είναι δυνατόν να αλληλεπιδρά με πολύ περισσότερα μόρια στο σύστημα των θηλαστικών σε σχέση με αυτό της δροσόφιλας, συμμετέχοντας σε πολυπλοκότερα ή

πολυπληθέστερα λειτουργικά συστήματα. Επιπλέον, δεν αποκλείεται η πρωτεΐνη *Neurl* να επιτελεί μια συγκεκριμένη λειτουργία στα θηλαστικά αλλά με πολλούς διαφορετικούς στόχους, ανάλογα με τον ιστό και τον κυτταρικό τύπο στον οποίο εκφράζεται. Τέλος, είναι πιθανόν και ο ρόλος του *Neurl* στο σύστημα των θηλαστικών, όπως συμβαίνει και με άλλες πρωτεΐνες, να επικαλύπτεται από τη δράση λειτουργικά ομόλογων μορίων.

11. Αντικείμενο της παρούσας διατριβής

Η χρονική περίοδος κατά την οποία διεξήχθη αυτή η διατριβή αλλά και το αντικείμενό της την εντάσσουν στη μεταβατική εποχή μεταξύ γονιδιωματικής και μεταγονιδιωματικής ανάλυσης. Έχοντας συμβάλλει στην ολοκλήρωση του φυσικού και μεταγραφικού χάρτη υψηλής ευκρίνειας του χρωμοσώματος 10 του ανθρώπου, μετά τη χαρτογράφηση της περιοχής 10q23-qter, η ομάδα μας εστιάζεται στην ανάλυση μιας σειράς γονιδίων της περιοχής 10q23.31-qter με στόχο τη διερεύνηση της μοριακής βάσης μενδελικών νοσημάτων (γονίδια *LGI1* και *FRA10AC1*) [Θ. Σαραφίδου, Διδακτορική Διατριβή, υπό συγγραφή] και τη λειτουργική ανάλυση γονιδίων με ιδιαίτερο βιολογικό ενδιαφέρον. Η λειτουργία του γονιδίου *Neuralized (NEURL)* του ανθρώπου μελετήθηκε χρησιμοποιώντας τα ετερόλογα συστήματα της δροσόφιλας, του ποντικού και κυτταρικές σειρές θηλαστικών (Η. Παυλόπουλος, Διδακτορική Διατριβή, Παν/μιο Κρήτης, 2002/ Ε. Κουτέλου, Διδακτορική Διατριβή, Παν/μιο Κρήτης, υπό εξέλιξη και παρούσα Διδακτορική Διατριβή). Σε αυτά τα πλαίσια, οι στόχοι για τη συγκεκριμένη διατριβή ήταν:

- Η συμβολή στην κατασκευή του φυσικού και μεταγραφικού χάρτη μιας από τις πιο πλούσιες σε γονίδια περιοχές του χρωμοσώματος 10 του ανθρώπου (10q23.31-q23.33), μεταξύ των δεικτών D10S541 και D10S583.
- Η ανάλυση της γονιδιωματικής οργάνωσης και της δομής του γονιδίου *Neurl* του ποντικού
- Ο προσδιορισμός των πρωτεϊνών που αλληλεπιδρούν άμεσα με την πρωτεΐνη *Neurl*.

ΥΛΙΚΑ & ΜΕΘΟΔΟΙ

1. Βακτηριακά στελέχη

Χρησιμοποιήθηκαν τα εξής στελέχη του βακτηρίου *Escherichia coli* :

DH5a: Η χρήση τους πραγματοποιήθηκε σε πειράματα μετασχηματισμού για την κλωνοποίηση θραυσμάτων DNA σε πλασμιδιακούς φορείς και την απομόνωση υπερελικωμένου πλασμιδιακού DNA. Τα βακτήρια αυτού του στελέχους χαρακτηρίζονται από το ότι επιδέχονται μετασχηματισμό σε υψηλό ποσοστό, καθώς και από τη μεταλλαγή που φέρουν στο γονίδιο της β-γαλακτοσιδάσης (*LacZ*), ώστε να επιτρέπεται η επιλογή των αποικιών που έχουν μετασχηματιστεί με πλασμίδια που φέρουν το γονίδιο [Sambrook *et al*, 1989].

BL21-Codon Plus-DE3-RIL (Stratagene Cat.no. 230245) : Χρησιμοποιήθηκε για την επαγόμενη από IPTG (*Isopropyl-1-thio-β-D-galactopyranoside*) έκφραση πεπτιδίων. Τα βακτήρια αυτού του στελέχους επιτρέπουν την έκφραση ετερόλογων πρωτεϊνών σε υψηλά επίπεδα, ιδιαίτερα εκείνων που ελέγχονται από υποκινητή T7, γιατί φέρουν γονίδιο που επάγεται από IPTG και εκφράζει T7 RNA πολυμεράση. Επιπλέον, φέρουν το πλασμίδιο RIL, το οποίο εκτός του γονιδίου ανθεκτικότητας σε χλωραμφαινικόλη, φέρει και τα γονίδια για τα tRNAs της αργινίνης (*argU*), της ισολευκίνης (*ileY*) και της λευκίνης (*leuW*), τα οποία συναντώνται σε χαμηλά επίπεδα στα βακτήρια.

K802 : χρησιμοποιήθηκαν για την ανάπτυξη λ βακτηριοφάγων.

KC8 (Clontech, Cat.no.640434): Χρησιμοποιήθηκαν στα πειράματα μετασχηματισμού με πλασμιδιακό DNA που είχε απομονωθεί από τις διπλά μετασχηματισμένες αποικίες σακχαρομύκητα που επιλέγονταν στο πείραμα των δύο υβριδίων. Αυτό το στέλεχος βακτηρίων παρουσιάζει αυξοτροφία σε ουρακίλη, τρυπτοφάνη, ιστιδίνη, αδενίνη και λευκίνη, επιτρέποντας την επιλογή των πλασμιδίων που φέρουν τα αντίστοιχα γονίδια, όταν αναπτύσσεται σε ελάχιστο θρεπτικό μέσο M9. Τα βακτήρια στελέχους KC8 φέρουν τη μεταλλαγή *hsdR*, η οποία τους δίνει τη δυνατότητα να μεθυλιώνουν το DNA, γι'αυτό το πλασμιδιακό DNA που απομονώνεται από τα κύτταρα KC8 είναι ακατάλληλο για πειράματα προσδιορισμού νουκλεοτιδικής αλληλουχίας ή τεμαχισμού με περιοριστικά ένζυμα. Επιπλέον, τα βακτήρια KC8 φέρουν το μεταθετό στοιχείο *Tn5*, το οποίο προσδίδει ανθεκτικότητα στην καναμυκίνη.

2. Στέλεχος *Sacharomyces cerevisiae* PJ69-4a

Για το πείραμα των δύο υβριδίων χρησιμοποιήθηκαν κύτταρα σακχαρομύκητα του στελέχους PJ69-4a με γονότυπο: *MATa trp1-901 leu2-3,112 ura3-52 his3-200 gal4Δ gal80Δ GAL2-AD2 LYS2::GAL1-HIS3 met2::GAL7-lacZ* [James *et al*, 1996]. Το στέλεχος αυτό φέρει τρία διαφορετικά γονίδια αναφοράς, επαγόμενα από τον μεταγραφικό παράγοντα GAL4 μέσω διαφορετικών υποκινητών:

i) Η κατασκευή *GAL1-HIS3* έχει εντεθεί στην 3' πλευρική περιοχή του γονιδίου *LYS2*.

ii) Η κατασκευή *GAL2-ADE2* αντικατέστησε τον γενετικό τόπο *ADE2*, ώστε ο υποκινητής του γονιδίου *ADE2* να αντικατασταθεί από τον υποκινητή GAL2.

iii) Η κατασκευή *GAL7-lacZ* έχει εντεθεί στην κωδική περιοχή του γονιδίου *MET2*, προκαλώντας έλλειμμα του μεγαλύτερου μέρους του *MET2*.

Η ομόλογη περιοχή μεταξύ των τριών υποκινητών (*GAL1*, *GAL2*, *GAL7*) περιορίζεται στις θέσεις πρόσδεσης του παράγοντα GAL4, επιτρέποντας τον ασφαλέστερο αποκλεισμό ψευδώς θετικών αποτελεσμάτων.

3. Κυτταρικές σειρές

Τα χαρακτηριστικά και οι συνθήκες ανάπτυξης των κυτταρικών σειρών που χρησιμοποιήθηκαν σε πειράματα διαμόλυνσης ή ανοσοανίχνευσης πρωτεϊνών αναφέρονται στον κατάλογο κυτταρικών σειρών και υβριδωμάτων 'American Type Culture Collection' (ATCC, 1992):

CV-1: Η κυτταρική σειρά CV-1 προήλθε από κύτταρα νεφρού αρσενικού πράσινου Αφρικάνικου πιθήκου (*Cercopithecus aethiops*), που μετασχηματίστηκαν από τον ιό RSV και επιτρέπει τον πολλαπλασιασμό πλασμιδίων που φέρουν αρχή αντιγραφής SV40. Αναπτύσσεται σε μονοστιβάδες κυττάρων με μορφολογικά χαρακτηριστικά ινοβλαστών, σε θρεπτικό μέσο DMEM 90%, fetal bovine serum 10%, γενταμικίνη 50μg/ml, στους 37°C και σε ατμόσφαιρα 5% CO₂ (ATCC, CCL 70).

COS-7: Κυτταρική σειρά με μορφολογικά χαρακτηριστικά ινοβλαστών, προερχόμενη από τον μετασχηματισμό κυττάρων CV-1 με ιό SV40 μεταλλαγμένο στην περιοχή έναρξης της αντιγραφής του (ori), ο οποίος όμως παράγει T αντιγόνο. Η σειρά αυτή είναι κατάλληλη για πειράματα διαμόλυνσης με υψηλά ποσοστά επιτυχίας, ειδικά στην περίπτωση πλασμιδίων που φέρουν την αρχή αντιγραφής του SV40. Τα κύτταρα αναπτύσσονται σχηματίζοντας μονοστιβάδες σε θρεπτικό μέσο DMEM 90%, fetal bovine serum 10%, γενταμικίνη 50μg/ml, στους 37°C και σε ατμόσφαιρα 5% CO₂ (ATCC, CRL 1651).

Caco-2: Είναι ανθρώπινη υπερτετραπλοειδής κυτταρική σειρά που προέρχεται από αδενοκαρκίνωμα άρρενος ατόμου της Καυκάσιας φυλής. Αναπτύσσεται σε μονοστιβάδες, με μορφολογικά χαρακτηριστικά επιθηλιακών κυττάρων, σε θρεπτικό μέσο DMEM 80%, fetal bovine serum 20%, γενταμικίνη 50μg/ml στους 37°C και σε ατμόσφαιρα 5% CO₂ (ATCC, HTB 37).

4. Θρεπτικά μέσα

Για την ανάπτυξη των διαφόρων στελεχών βακτηρίων, των βακτηριοφάγων, των κυττάρων σακχαρομύκητα καθώς και των κυττάρων θηλαστικών, χρησιμοποιήθηκαν κατάλληλα θρεπτικά μέσα, των οποίων η σύσταση περιγράφεται παρακάτω.

α. Θρεπτικά μέσα ανάπτυξης βακτηρίων

Για την ανάπτυξη διαφορετικών στελεχών *E.coli* χρησιμοποιήθηκαν τα παρακάτω θρεπτικά μέσα:

Για την ανάπτυξη των στελεχών *E.coli* DH5a και BL21-Codon Plus-DE3-RIL

Θρεπτικό μέσο LB: 1% NaCl, 1% *Bacto-Tryptone*, 0.5% *Yeast extract*, pH 7.0. Για στερεό θρεπτικό μέσο προστίθεται άγαρ τελικής συγκέντρωσης 1.5% [Sambrook *et al*, 1989].

Θρεπτικό μέσο TYM: 0.6% NaCl, 0.25% MgSO₄, 2% *Bacto-Tryptone*, 0.5% *Yeast extract*, pH 7.0. Λόγω της μικρής συγκέντρωσης αλάτων, αυτό το θρεπτικό μέσο χρησιμοποιείται για την παρασκευή δεκτικών κυττάρων στα πειράματα μετασχηματισμού με τη μέθοδο της ηλεκτρικής εκκένωσης [Sambrook *et al*, 1989].

Ελάχιστο θρεπτικό μέσο M9 με Leu, His, Ura, Ade (για την ανάπτυξη στελεχών KC8)

Σε αποστειρωμένο διάλυμα 2% άγαρ προστίθεται διάλυμα 20% γλυκόζης (αποστειρωμένο με φίλτρο διαμέτρου πόρων 0.2μm) μέχρι τελικής συγκέντρωσης γλυκόζης 0.4%. Στη συνέχεια, προστίθεται διάλυμα φωσφορικών (0.6% κβ Na₂HPO₄, 0.05% κβ KH₂PO₄, 0.1% κβ NH₄Cl), 1mM MgSO₄, 0.1mM CaCl₂, Leu, His, Ura, Ade και αμπικιλίνη 100μg/ml [Sambrook *et al*, 1989].

Θρεπτικό μέσο ανάπτυξης βακτηρίων για μόλυνση από λ βακτηριοφάγο (*E.coli* K802)

Είναι το θρεπτικό μέσο LB, το οποίο περιέχει επιπλέον 0.2% μαλτόζη και 10mM MgSO₄.

β. Θρεπτικά μέσα ανάπτυξης βακτηριοφάγων

NZYM επιφάνειας: 0.5% NaCl, 0.25% MgSO₄·7H₂O, 0.5% *yeast extract*, 1% *NZamine (casein hydrolysate)*, 0.7% άγαρ, pH 7.0-7.5, χρησιμοποιείται προθερμασμένο στους 45°C, ώστε να βρίσκεται σε ρευστή κατάσταση [Sambrook *et al*, 1989].

NZYM βάσης: 0.5% NaCl, 0.2% MgSO₄·7H₂O, 0.5% *yeast extract*, 1% *NZamine*, 1.5% άγαρ, pH 7.0-7.5.

γ. Θρεπτικά μέσα ανάπτυξης κυττάρων σακχαρομύκητα*

Η σύσταση των θρεπτικών μέσων που χρησιμοποιήθηκαν για την ανάπτυξη κυττάρων σακχαρομύκητα περιγράφεται στα σχετικά εργαστηριακά εγχειρίδια [Ausubel *et al*, 2000 / Guthrie & Fink, 1991].

Συνθετικό θρεπτικό μέσο SC: *Yeast nitrogen base* (χωρίς αμινοξέα) 0.67%, γλυκόζη 2%, μίγμα αμινοξέων 0.2% (2gr από όλα τα αμινοξέα εκτός : Leu, Trp, His). Άγαρ (για στερεό θρεπτικό μέσο) 2%.

Στερεό θρεπτικό μέσο SC με Xgal: Στο στερεό θρεπτικό μέσο SC προστίθεται μετά την αποστείρωση σε κλίβανο, διάλυμα φωσφορικού καλίου τελικής συγκέντρωσης 70mM (43.05mM K₂HPO₄ και 26.95mM KH₂PO₄) και X-gal 20μg / ml.

Πλούσιο θρεπτικό μέσο YEP: *Yeast extract* 1%, γλυκόζη 2%, *bactopeptone* 2%, άγαρ (για στερεό θρεπτικό μέσο) 2%.

* Συγκεντρώσεις αμινοξέων και νουκλεοτιδίων στα θρεπτικά μέσα ανάπτυξης σακχαρομύκητα και βακτηρίων *E.coli* KC8: Αδενίνη (Ade) 0.25 mM, Ουρακίλη (Ura) 0.20 mM, Ιστιδίνη (His) 0.33 mM, Τρυπτοφάνη (Trp) 0.55 mM, Λευκίνη (Leu) 1.80 mM.

δ. Θρεπτικά μέσα ανάπτυξης κυτταρικών σειρών θηλαστικών

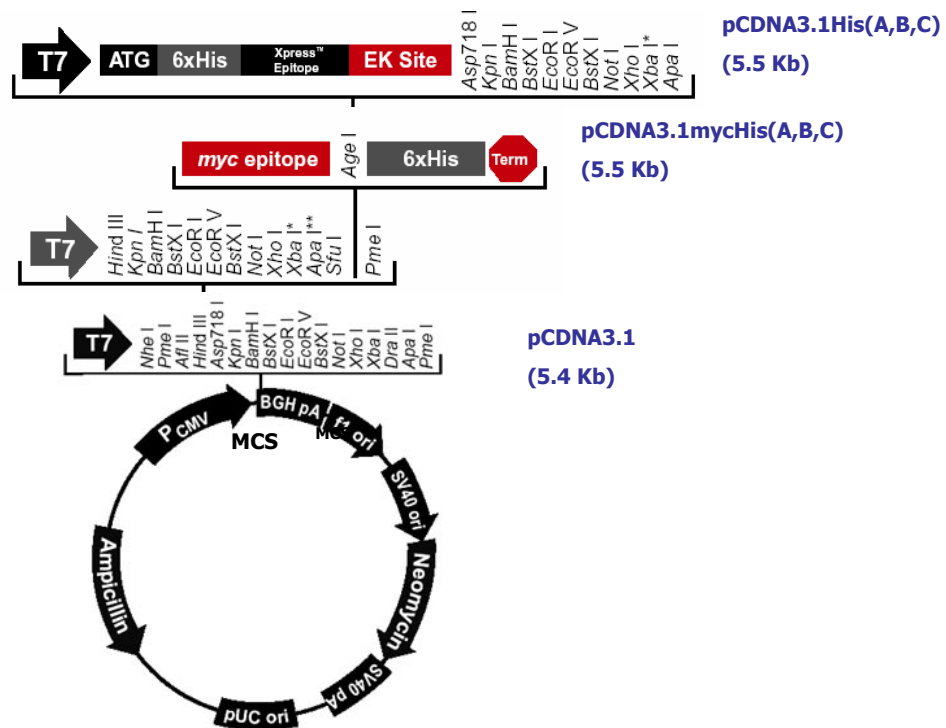
Για την ανάπτυξη των κυτταρικών σειρών θηλαστικών χρησιμοποιούνταν το θρεπτικό μέσο DMEM (*Dulbecco's Modified Eagle's Medium*: 4 mM *L-glutamine*, 0.45% *glucose*, 0.15% *sodium bicarbonate*) σε συγκέντρωση 80 ή 90% και ορός FCS (*Fetal Calf Serum*), CS (*Calf Serum*) ή FBS (*Fetal Bovine Serum*) της εταιρίας Gibco σε συγκέντρωση 20 ή 10%, ανάλογα με τον τύπο των κυττάρων που αναπτύσσονταν. Για την παρεμπόδιση βακτηριακών μόλυνσεων προσθέτονταν στο θρεπτικό μέσο γενταμικίνη συγκέντρωσης 50μg/ml. Τα κύτταρα αναπτύσσονταν στους 37°C, σε ατμόσφαιρα 5% CO₂.

5. Πλασμιδιακοί φορείς

Χρησιμοποιήθηκαν οι εξής πλασμιδιακοί φορείς:

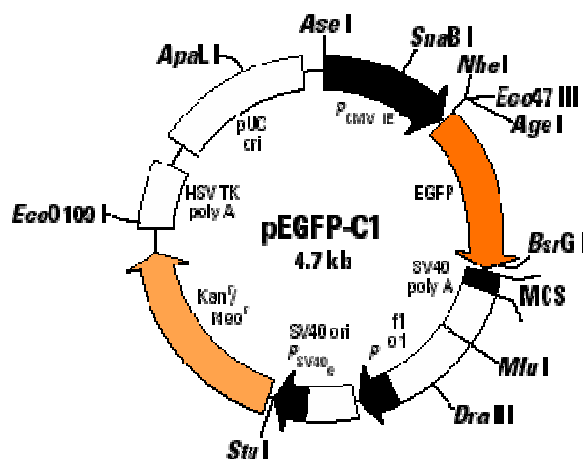
pcDNA3.1(+), (Invitrogen Cat. no. V79020). Πρόκειται για πλασμιδιακό φορέα που παρέχει τη δυνατότητα υψηλών επιπέδων έκφρασης σε κύτταρα θηλαστικών, μέσω του υποκινητή CMV (Cytomegalovirus), καθώς και *in vitro* μεταγραφής και μετάφρασης των ενθεμάτων από RNA πολυμεράση που αναγνωρίζει τον υποκινητή T7 (Εικ.19).

pcDNA3.1mycHis(A,B,C) και pcDNA3.1His(A,B,C), Invitrogen (Cat.no. V80020, V38520). Οι φορείς αυτοί φέρουν όλα τα χαρακτηριστικά του pcDNA3.1 και επιπλέον παρέχουν τη δυνατότητα παραγωγής ανασυνδυασμένων πεπτιδίων που φέρουν τους επίτοπους myc και His₆ ή 6xHis και Xpress, αντίστοιχα, ώστε να επιτρέπεται η ανοσοανίχνευση (για τους επίτοπους myc, His₆ και Xpress) ή ο βιοχημικός καθαρισμός (για τον επίτοπο 6xHis) της παραγόμενης πρωτεΐνης (Εικ.19).



Εικ.19 Γραφική αναπαράσταση των πλασμιδίων **pCDNA3.1**, **pCDNA3.1mycHis(A,B,C)** και **pCDNA3.1His(A,B,C)**. Οι φορείς φέρουν υποκινητή CMV για υψηλά επίπεδα έκφρασης σε κύτταρα θηλαστικών, υποκινητή T7 για *in vitro* μεταγραφή/μετάφραση, πολλαπλές θέσεις κλωνοποίησης των ενθεμάτων (MCS) και ανθεκτικότητα σε νεομικίνη και αμπικιλίνη για επιλογή των μόνιμα διαμολυσμένων ευκαρυωτικών κυττάρων και των μετασχηματισμένων βακτηριακών αποικιών, αντίστοιχα. Τα πλασμίδια pCDNA3.1mycHis και pCDNA3.1His φέρουν επιπλέον και τους επιτόπους myc και His₆ ή 6xHis και Xpress, αντίστοιχα. Και οι δύο φορείς παρέχονται σε τρία αναγνωστικά πλαίσια (A, B, C), για την κλωνοποίηση οποιουδήποτε ενθέματος.

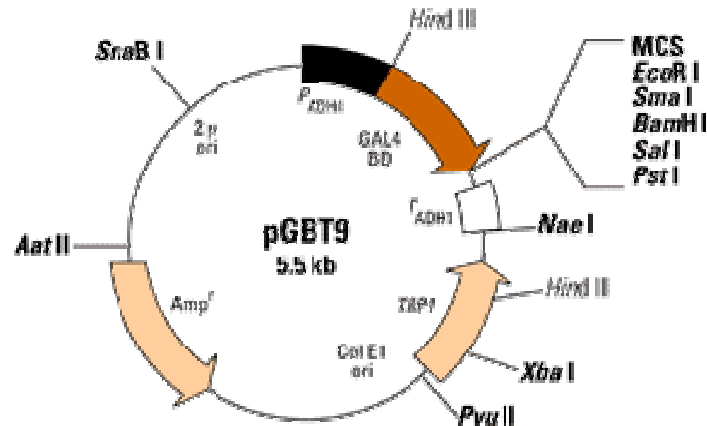
pEGFP-C1 (BD Biosciences Clontech, Cat.no. 6084-1). Πρόκειται για φορέα που παρέχει τη δυνατότητα υψηλών επιπέδων έκφρασης σε κύτταρα θηλαστικών μέσω του υποκινητή CMV καθώς και σήμανσης του αμινοτελικού άκρου των παραγόμενων πεπτιδίων με ενισχυμένη μορφή της πράσινης φθορίζουσας πρωτεΐνης (EGFP), η οποία εμφανίζει μέγιστο απορρόφησης στα 488nm και εκπομπής στα 507nm (Εικ.20).



Εικ.20 Γραφική αναπαράσταση του πλασμιδιακού φορέα **pEGFP-C1**. Ο φορέας περιλαμβάνει ισχυρό υποκινητή (P_{CMV}), πολλαπλές θέσεις κλωνοποίησης των ενθεμάτων (MCS), καθώς και γονίδια ανθεκτικότητας σε καναμικίνη και νεομικίνη (Kan^r / Neo^r) για την επιλογή των μόνιμα διαμολυσμένων κυττάρων θηλαστικών και των μετασχηματισμένων βακτηρίων, αντίστοιχα.

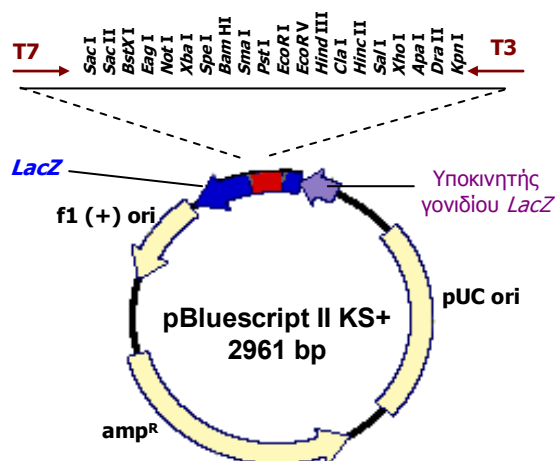
pGBT9 (BD Biosciences Clontech, Cat.no. K1605-A). Πρόκειται για φορέα χαμηλών επιπέδων έκφρασης σε κύτταρα σακχαρομύκητα. Από την κλωνοποίηση ενθεμάτων σ'αυτόν τον φορέα προκύπτουν υβριδικές πρωτεΐνες, οι οποίες περιέχουν στο αμινοτελικό τους άκρο την υπομονάδα πρόσδεσης στο DNA του μεταγραφικού παράγοντα GAL4 (GAL4-BD, 147 αμινοξέα). Η υβριδική πρωτεΐνη εκφράζεται σε κύτταρα σακχαρομύκητα από τον συστατικό υποκινητή *ADHI* και

κατευθύνεται στον πυρήνα των κυττάρων από σήματα πυρηνικού εντοπισμού που βρίσκονται στην αμινοξική αλληλουχία της GAL4-BD. Το πλασμίδιο pGBT9 έχει ικανότητα αυτόνομου διπλασιασμού τόσο σε κύτταρα *S. cerevisiae* (αρχή αντιγραφής 2μ), όσο και σε *E. coli* (αρχή αντιγραφής Col E1). Φέρει το γονίδιο Amp^r για ανθεκτικότητα σε αμπικιλίνη και επιλογή των μετασχηματισμένων βακτηρίων και το γονίδιο βιοσύνθεσης της τρυπτοφάνης *TRP1*, για την επιλογή του σε μεταλλαγμένα στελέχη σακχαρομύκητα που αναπτύσσονται σε επιλεκτικό συνθετικό μέσο, απουσία τρυπτοφάνης (Εικ. 21)



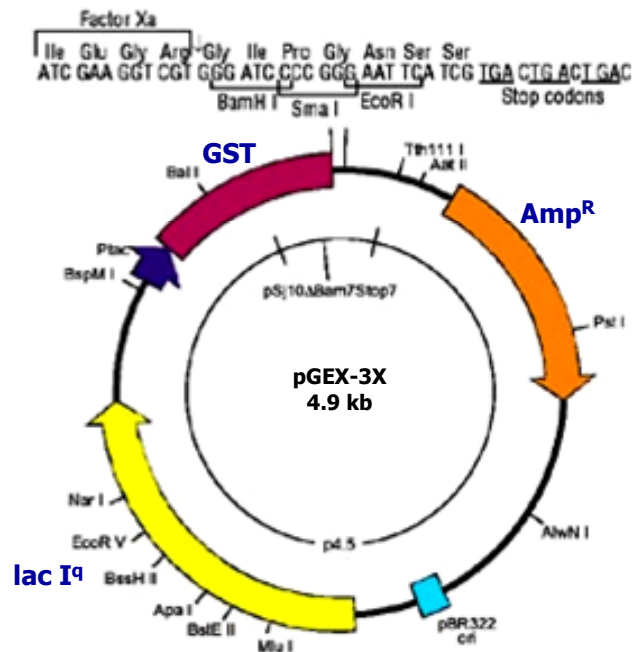
Εικ.21 Γραφική αναπαράσταση του πλασμιδιακού φορέα pGBT9, ο οποίος φέρει πολλαπλές θέσεις κλωνοποίησης των ενθεμάτων (MCS), στο αναγνωστικό πλαίσιο της υπομονάδας GAL4-BD. Η έκφραση του υβριδικού πεπτιδίου στα κύτταρα *S. cerevisiae* πραγματοποιείται μέσω του υποκινητή P_{ADHI}, ενώ η επιλογή του πλασμιδίου στα ανασυνδυασμένα κύτταρα σακχαρομύκητα ή *E.coli* γίνεται μέσω των γονιδίων *TRP1* και Amp^r, αντίστοιχα.

pBluescript II KS(+) (Stratagene). Ο φορέας αυτός χαρακτηρίζεται από τη μεγάλη ποικιλία θέσεων αναγνώρισης περιοριστικών ενζύμων για την κλωνοποίηση των ενθεμάτων και την παρουσία των υποκινητών/εκκινητών T3 και T7 εκατέρωθεν της θέσης κλωνοποίησης (Εικ. 22). Γι'αυτό, χρησιμοποιείται ευρέως για την κλωνοποίηση μορίων DNA με σκοπό είτε την ανάλυση του προτύπου τεμαχισμού τους από περιοριστικές ενδονουκλεάσες, είτε τον προσδιορισμό της νουκλεοτιδικής τους σύστασης με εκκινητές T3 και T7, είτε την *in vitro* μεταγραφή τους από RNA πολυμεράσες που αναγνωρίζουν αυτούς τους υποκινητές. Επιπλέον, η θέση κλωνοποίησης βρίσκεται στην κωδική περιοχή του γονιδίου *LacZ*, με αποτέλεσμα να επιτρέπεται η επιλογή των ανασυνδυασμένων βακτηριακών αποικιών ανιχνεύοντας τη β-γαλακτοσιδάση (παρουσία X-gal) στις μη ανασυνδυασμένες αποικίες (μπλε χρώση).



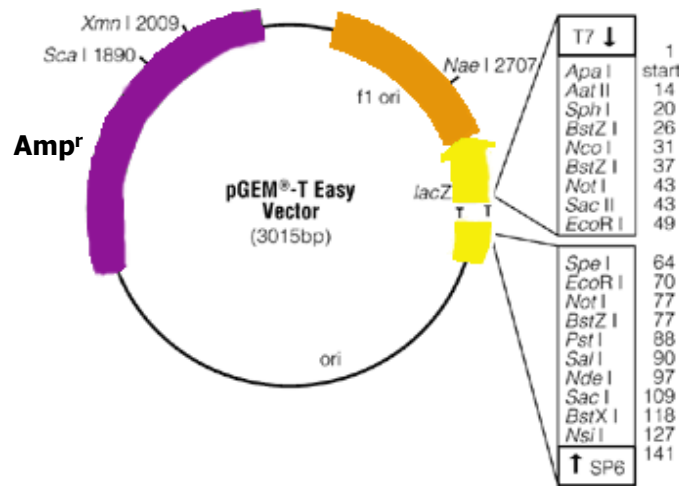
Εικ.22 Γραφική αναπαράσταση του πλασμιδιακού φορέα pBluescript II KS(+). Ο φορέας φέρει πολλαπλές θέσεις κλωνοποίησης των ενθεμάτων (MCS) στην κωδική περιοχή του γονιδίου της β-γαλακτοσιδάσης (*LacZ*), υποκινητές/εκκινητές T3 και T7, θέσεις έναρξης της αντιγραφής (f1 και PUC), υποκινητή του γονιδίου *LacZ* και γονίδιο ανθεκτικότητας στην αμπικιλίνη (amp^R), ως δείκτη επιλογής.

pGEX-3x (Amersham Biosciences Cat.no. 27-4803-01). Πρόκειται για φορέα που παρέχει τη δυνατότητα επαγόμενης έκφρασης πεπτιδίων σε βακτήρια και σύντηξης στο αμινοτελικό τους άκρο με τη θειοτρανσφεράση της γλουταθειόνης (GST, 26KD), για τον βιοχημικό καθαρισμό τους από χρωματογραφία συγγένειας με γλουταθειόνη. Το παραγόμενο πεπτιδίο μπορεί στη συνέχεια να αποκοπεί από την GST με επίδραση του παράγοντα Xa. Η έκφραση των υβριδικών πεπτιδίων ελέγχεται από τον υποκινητή *tac*, ο οποίος σε συνθήκες μη επαγωγής καταστέλλεται από τον καταστολέα $IacI^q$. Η επαγωγή πραγματοποιείται με προσθήκη 0.1-1mM IPTG (χημικό ανάλογο της γαλακτόζης), το οποίο προσδένεται στον καταστολέα και προκαλεί αλλοστερικά την αποσύνδεσή του από τον υποκινητή *tac*. Ο ίδιος ο φορέας φέρει και το γονίδιο που κωδικοποιεί τον καταστολέα, ώστε η επαγωγή να είναι δυνατή σε μεγάλη ποικιλία στελεχών *E.coli*. (Εικ.23).



Εικ.23 Γραφική αναπαράσταση του πλασμιδιακού φορέα **pGEX-3x**, ο οποίος φέρει πολλαπλές θέσεις κλωνοποίησης των ενθεμάτων (MCS) στο 3' άκρο του αναγνωστικού πλαισίου της πρωτεΐνης GST, για την έκφραση υβριδικών μορίων υπό τον έλεγχο του υποκινητή *tac*. Επιπλέον, φέρει θέση αναγνώρισης του παράγοντα Xa, το γονίδιο που κωδικοποιεί τον καταστολέα ($IacI^q$), θέση έναρξης της αντιγραφής (pBR322 ori) για τον αυτόνομο διπλασιασμό του σε βακτήρια, καθώς και γονίδιο ανθεκτικότητας στην αμπικιλίνη (Amp^R), ως δείκτη επιλογής.

pGEM-T Easy (Promega, Cat.no. A3600). Πρόκειται για φορέα ο οποίος χρησιμοποιείται ευρέως στην κλωνοποίηση και ανάλυση προϊόντων της αντίδρασης PCR. Κύριο χαρακτηριστικό του είναι ότι φέρει προεξέχοντα νουκλεοτίδια θυμιδίνης στα 3' άκρα του. Κατά την κλωνοποίηση, αυτά υβριδοποιούνται με τα προεξέχοντα νουκλεοτίδια αδενίνης που προστίθενται από πολλές θερμοανθεκτικές πολυμεράσες DNA στα 3' άκρα των προϊόντων PCR. Επιπλέον, χαρακτηρίζεται από τη μεγάλη ποικιλία θέσεων αναγνώρισης περιοριστικών ενζύμων και την παρουσία των υποκινητών/εκκινητών T7 και SP6 εκατέρωθεν της θέσης κλωνοποίησης, η οποία βρίσκεται στην κωδική περιοχή του ενζύμου της β-γαλακτοσιδάσης, ώστε να επιτρέπεται η χρωματική επιλογή των ανασυνδυασμένων βακτηριακών αποικιών (Εικ. 24). Για την κλωνοποίηση ενθεμάτων σ'αυτόν τον φορέα χρησιμοποιήθηκε το σύστημα της κατασκευάστριας εταιρίας (pGEM-T Vector System), σύμφωνα με τις οδηγίες που αναφέρονται σε ειδικό εγχειρίδιο (*Promega, Technical Manual No.042*).

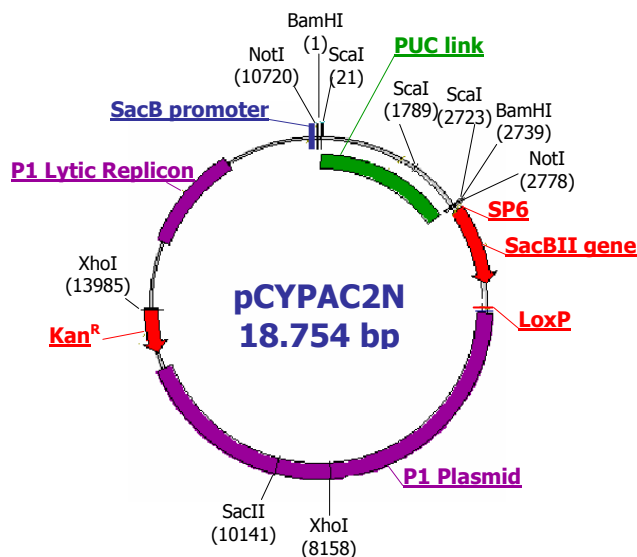


Εικ.24 Γραφική αναπαράσταση του πλασμιδιακού φορέα **pGEM-T Easy**. Στην κωδική περιοχή του γονιδίου *LacZ* έχει γίνει η ένθεση προεξέχοντων νουκλεοτιδίων θυμιδίνης (T) στα 3' άκρα. Εκατέρωθεν της θέσης κλωνοποίησης υπάρχουν πολλαπλές θέσεις αναγνώρισης περιοριστικών ενζύμων καθώς και οι υποκινητές/εκκινητές T7 και SP6. Το πλασμίδιο έχει δυνατότητα ανεξάρτητου διπλασιασμού και υψηλού αριθμού αντιγράφων (*f1 ori*) στα κύτταρα *E.coli*, στα οποία προσφέρει ανθεκτικότητα στην αμικικιλίνη (*Amp^r*).

6. Βιβλιοθήκες

α. Γονιδιωματική βιβλιοθήκη κλώνων PAC ανθρώπου

Η βιβλιοθήκη κλώνων PAC (*P1 bacteriophage Artificial Chromosomes*) που χρησιμοποιήθηκε για την κατασκευή του φυσικού χάρτη του χρωμοσώματος 10 ήταν η RPCI-1 (*Roswell Park Cancer Institute-1*), η προμήθεια της οποίας έγινε από το HGMPRC (*Human Genome Mapping Project Resource Center, UK*). Η βιβλιοθήκη κατασκευάστηκε από γονιδιωματικό DNA που απομονώθηκε από κυτταρική σειρά ανθρώπινων ινοβλαστών (HSF7), τεμαχίστηκε μερικώς με την περιοριστική ενδονουκλεάση *MboI* και κλωνοποιήθηκε στη συμβατή θέση *BamHI* του φορέα *pCYPAC2N* (Εικ.25) [Ιοαννου & de Jong, 1996]. Για τη διερεύνηση της βιβλιοθήκης με ραδιοσημασμένους ανιχνευτές, χρησιμοποιήθηκε η διατεταγμένη (*gridded*) μορφή της σε 7 νάυλον φίλτρα διαστάσεων 22x22cm, που περιέχουν συνολικά 12.000 κλώνους με μέσο μέγεθος ενθέματος 110Kb, υπερκαλύπτοντας κατά 4 περίπου φορές το ανθρώπινο γονιδίωμα. Σύμφωνα με πειράματα ελέγχου που πραγματοποιήθηκαν από τους κατασκευαστές, η βιβλιοθήκη δεν εμφανίζει φαινόμενα αστάθειας ή χιμαιρισμό των ενθεμάτων της.



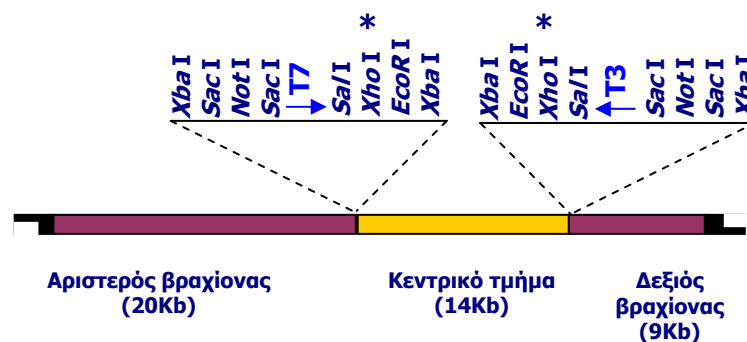
Εικ.25 Γραφική αναπαράσταση του πλασμιδιακού φορέα **pCYPAC2N**. Τα ενθέματα της βιβλιοθήκης RPCI-1 έχουν κλωνοποιηθεί στις θέσεις *NotI*, αντικαθιστώντας το τμήμα 'PUC link'. Ο φορέας ανθεκτικότητα σε καναμικίνη (*Kan^R*).

β. Γονιδιωματική βιβλιοθήκη κλώνων BAC ανθρώπου

Η βιβλιοθήκη κλώνων BAC (*Bacterial Artificial Chromosomes*) που χρησιμοποιήθηκε για την κατασκευή του φυσικού χάρτη του χρωμοσώματος 10 ήταν η RPCI-11 (*Roswell Park Cancer Institute-11*), η οποία ήταν διαθέσιμη από το *Sanger Center*. Η βιβλιοθήκη κατασκευάστηκε από γονιδιωματικό DNA που απομονώθηκε από τα λευκά αιμοσφαίρια ανδρός και κλωνοποιήθηκε στον φορέα pBACe3.6 [Osoegawa *et al*, 2001]. Για τη διερεύνηση της βιβλιοθήκης με ραδιοσημασμένους ανιχνευτές, χρησιμοποιήθηκε η διατεταγμένη (*gridded*) μορφή της σε 36 νάυλον φίλτρα διαστάσεων 8x12cm, που περιέχουν συνολικά 220.000 κλώνους με μέσο μέγεθος ενθέματος 180Kb, υπερκαλύπτοντας κατά 30 περίπου φορές το ανθρώπινο γονιδίωμα. Σύμφωνα με πειράματα ελέγχου, η συγκεκριμένη βιβλιοθήκη χαρακτηρίζεται από μεγάλη σταθερότητα των ενθεμάτων της, εμφανίζοντας μικρές αναδιατάξεις μόνο στο 11% των κλώνων της, ενώ δεν ανιχνεύονται σε αυτήν χιμαιρικοί κλώνοι.

γ. Γονιδιωματική βιβλιοθήκη ποντικού (Stratagene 946305, προσφορά από Γ.Μαυροθαλασσίτη)

Πρόκειται για βιβλιοθήκη του γονιδιώματος του ποντικού σε βακτηριοφάγο Lambda FIXII (Εικ.26), η οποία έχει κατασκευαστεί από γονιδιωματικό DNA ποντικού στελέχους 129SV, ηλικίας 4-8 εβδομάδων και φύλου θηλυκού. Το στέλεχος 129SV χρησιμοποιείται ευρέως για την ανάπτυξη ποντικών με απενεργοποιημένα ενδογενή γονίδια (*knock-out*) με τη μέθοδο του ομόλογου ανασυνδυασμού. Το DNA τεμαχίστηκε μερικώς με τα ένζυμα BamHI, MboI, BglII ή Sau3A και κλωνοποιήθηκε στη θέση XhoI του φάγου, μετά από κατάλληλες τροποποιήσεις των άκρων. Η βιβλιοθήκη περιέχει κλώνους με ενθέματα μεγέθους 9-23Kb. Οι βακτηριοφάγοι Lambda FIXII αναπτύσσονται μολύνοντας κύτταρα *E.coli* του στελέχους K802.



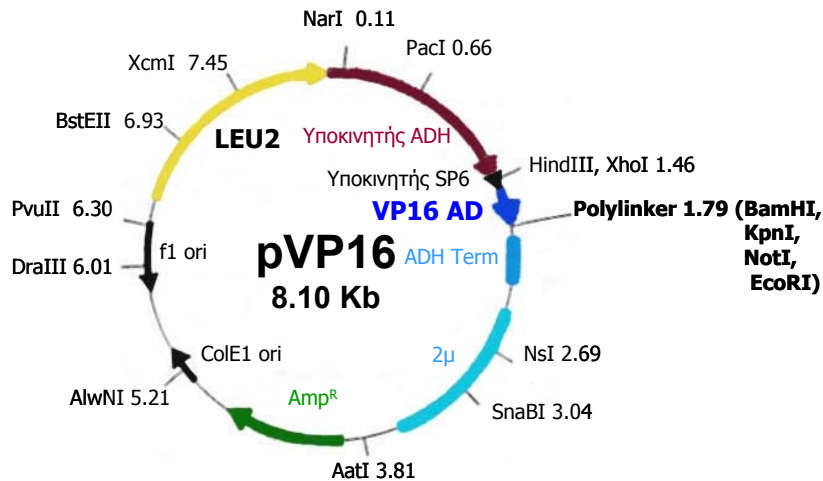
Εικ.26 Γραφική αναπαράσταση του φάγου λFIXII, στον οποίο κλωνοποιήθηκαν τα θραύσματα γονιδιωματικού DNA ποντικού για την κατασκευή της βιβλιοθήκης που χρησιμοποιήθηκε. Τα ενθέματα (μεγέθους 9-23Kb) είχαν κλωνοποιηθεί ανάμεσα στους δύο βραχίονες του φάγου, στη θέση αναγνώρισης του ενζύμου XhoI, αφού προηγουμένως είχε αφαιρεθεί το κεντρικό τμήμα.

δ. cDNA βιβλιοθήκη εμβρύου ποντικού (προσφορά από Γ. Χαλεπάκη και P. Gruss)

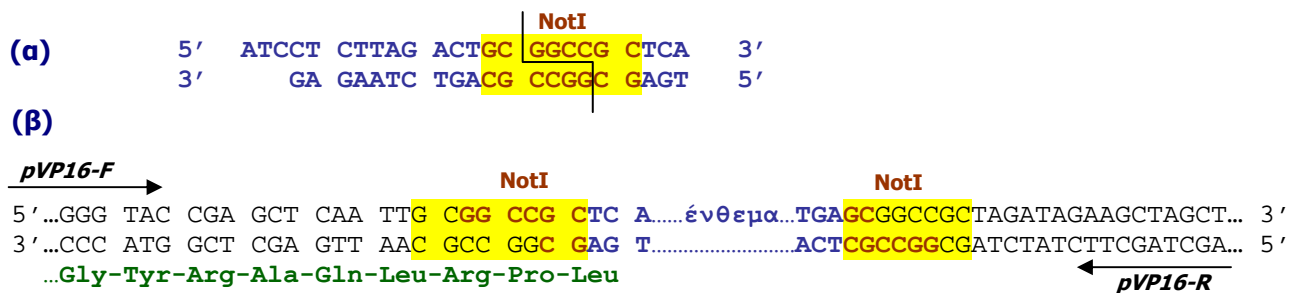
Πρόκειται για κλώνους μορίων cDNA εμβρύου ποντικού E9.5 στον πλασμιδιακό φορέα pVP16 (Εικ.27). Η βιβλιοθήκη είχε κατασκευαστεί με σύνθεση των μορίων cDNA από τυχαίους εκκινητές, σύνδεση των προϊόντων με συνδέτες-NotI (Εικ.28), επιλογή των μορίων μεγέθους 350-700 νουκλεοτιδίων, πολλαπλασιασμό με PCR, πέψη με NotI και κλωνοποίηση στη θέση NotI του pVP16. Η βιβλιοθήκη περιέχει περίπου 5×10^6 διαφορετικούς κλώνους cDNA [Hollenberg *et al*, 1995].

Ο pVP16 είναι φορέας έκφρασης σε κύτταρα σακχαρομύκητα και τα ενθέματα cDNA, που κλωνοποιούνται στο αναγνωστικό πλαίσιο της υπομονάδας ενεργοποίησης της μεταγραφής του παράγοντα VP16 (VP16-AD), εκφράζονται ως υβριδικά μόρια σε σύντηξη με αυτήν την υπομονάδα και κατευθύνονται στον πυρήνα των κυττάρων σακχαρομύκητα από σήματα πυρηνικού εντοπισμού που βρίσκονται σε αυτήν. Το πλασμίδιο pVP16 φέρει θέσεις έναρξης της αντιγραφής για τον αυτόνομο διπλασιασμό του τόσο σε κύτταρα *E.coli* (f1, ColE1), όσο και σε *S. cerevisiae* (2μ). Επιπλέον, φέρει το γονίδιο *LEU2*, για την επιλογή του λόγω βιοσύνθεσης της λευκίνης στα *LEU2*- στελέχη σακχαρομύκητα που αναπτύσσονται

απουσία λευκίνης, και το γονίδιο ανθεκτικότητας στην αμπικιλίνη (Amp^R) για την επιλογή των μετασχηματισμένων με $pVP16$ βακτηρίων (Εικ.27). Τέλος, για την αλληλούχηση των ενθεμάτων της βιβλιοθήκης φέρει τους εκκινητές $pVP16-F$ (5'-GAGTTTGAGCAGATGTTTACC-3') και $pVP16-R$ (5'-TGACCGGCAGCAAAATGTTG-3'), εκατέρωθεν της θέσης κλωνοποίησης ($NotI$) (Εικ.28).



Εικ.27 Γραφική αναπαράσταση του πλασμιδιακού φορέα $pVP16$, στον οποίο είχαν κλωνοποιηθεί τα cDNAs της βιβλιοθήκης εμβρύου ποντικού E9.5. Ο φορέας φέρει πολλαπλές θέσεις κλωνοποίησης (*polylinker*) στο 3' άκρο του αναγνωστικού πλαισίου της υπομονάδας ενεργοποίησης της μεταγραφής του VP16 (VP16AD), υποκινητές ADH και SP6, θέσεις έναρξης της αντιγραφής (f1, ColE1, 2μ), γονίδιο ανθεκτικότητας στην αμπικιλίνη (Amp^R) και το γονίδιο για τη βιοσύνθεση της λευκίνης (*LEU2*).



Εικ.28 Κλωνοποίηση ενθεμάτων βιβλιοθήκης στον $pVP16$. (α) Οι συνδέτες- $NotI$ (*NotI* linkers), με τους οποίους συνδέθηκαν τα cDNAs για την κλωνοποίησή τους στη θέση $NotI$ του $pVP16$. (β) Νουκλεοτιδική αλληλουχία των κλώνων της βιβλιοθήκης στη θέση κλωνοποίησης $NotI$. Φαίνονται τα κωδικόνια του γονιδίου *VP16-AD* και η κατεύθυνση των εκκινητών $pVP16-F$ και $pVP16-R$, που χρησιμοποιήθηκαν για την αλληλούχηση των ενθεμάτων. (Με έντονους χαρακτήρες παρουσιάζεται η νουκλεοτιδική αλληλουχία του ενθέματος.)

7. Ο κλώνος cDNA του γονιδίου *Neurl* του ποντικού

Το cDNA ολικού μήκους του γονιδίου *Neurl* του ποντικού [Pavlouros *et al*, 2002], χρησιμοποιήθηκε για την παρασκευή των ανιχνευτών στα πειράματα ανάλυσης της γονιδιωματικής οργάνωσης του *Neurl*, καθώς και για την κλωνοποίηση της κωδικής περιοχής του γονιδίου σε πλασμιδιακούς φορείς έκφρασης. Το μόριο cDNA του *Neurl* είχε απομονωθεί από φαγική βιβλιοθήκη (λExCell *EcoRI* CIP), της οποίας οι κλώνοι έφεραν ενθέματα με 'συνδέτες' (*linkers*) *EcoRI*-*NotI*. Στη συνέχεια, κλωνοποιήθηκε στους πλασμιδιακούς φορείς *pBluescript*, *pRSETC* και *pEGFP* (Η. Παυλόπουλος, διδακτορική διατριβή 2002, Παν/μιο Κρήτης) ως εξής:

α) *Neurl(2433)/pBluescript* : Το cDNA αποκόπηκε από τον φάγο με ενζυματικό τεμαχισμό στις θέσεις *EcoRI* των συνδετών και συγκολλήθηκε με τον γραμμοποιημένο φορέα *pBluescript/EcoRI*. Το ένθεμα του πλασμιδίου περιέχει 1713 από τα 1725 νουκλεοτίδια κωδικής αλληλουχίας του γονιδίου γιατί φέρει στο 5' άκρο του έλλειμα των 12 αρχικών νουκλεοτιδίων (4 κωδικόνια) του *Neurl*, τα οποία προσδιορίστηκαν σε ένα προϊόν RT-PCR της περιοχής [Pavlouros *et al*, 2002]. Η συνολική αλληλουχία (2445bp) του cDNA είναι διαθέσιμη στη βάση δεδομένων *Genbank* (Εικ. 29) με αριθμό καταχώρησης Y15160.

β) *Neurl(5-574)/pRSETC* : Το cDNA αποκόπηκε (με *EcoRI*) από τον *pBluescript*, τα προεξέχοντα άκρα του μετατράπηκαν σε τυφλά με *Klenow* και συγκολλήθηκε με το πλασμιδίο *pRSETC* (φορέας βακτηριακής έκφρασης που φέρει θέση έναρξης της μετάφρασης, *Invitrogen-V35120*), το οποίο είχε υποστεί γραμμοποίηση (με *HindIII*) και κατασκευή τυφλών άκρων με *Klenow*. Η κατασκευή αυτή κωδικοποιεί στα βακτήρια τα αμινοξέα 5-574 της πρωτεΐνης *Neurl*.

γ) *Neurl(5-574)/pEGFPC1* : Το cDNA αποκόπηκε (με *EcoRI* και *KpnI*) από τον *pBluescript* και συνδέθηκε στις ίδιες θέσεις του πλασμιδιακού φορέα έκφρασης *pEGFPC1*, στο καρβοξυτελικό άκρο της φθορίζουσας πρωτεΐνης EGFP. Η κατασκευή που προέκυψε κωδικοποιεί την ανασυνδυσμένη πρωτεΐνη GFP-*Neurl*(5-574).

δ) *Neurl(5-537)/pEGFPN3* : Το cDNA αποκόπηκε (με *BglII* και *MscI*) από τον *pRSETC* και συνδέθηκε στις θέσεις *BglII* και *SmaI* του πλασμιδιακού φορέα έκφρασης *pEGFPN3*, στο αμινοτελικό άκρο της φθορίζουσας πρωτεΐνης EGFP. Η κατασκευή που προέκυψε κωδικοποιεί την ανασυνδυσμένη πρωτεΐνη GFP-*Neurl*(5-537).

ε) *Neurl(5-269)/pEGFPN2* : Το cDNA αποκόπηκε (με *XhoI* και *NaeI*) από τον *pRSETC* και συνδέθηκε στις θέσεις *XhoI* και *SmaI* του πλασμιδιακού φορέα έκφρασης *pEGFPN2*, στο αμινοτελικό άκρο της φθορίζουσας πρωτεΐνης EGFP. Η κατασκευή που προέκυψε κωδικοποιεί την ανασυνδυσμένη πρωτεΐνη GFP-*Neurl*(5-269).

στ) *Neurl(227-574)/pEGFPC1* : Το cDNA αποκόπηκε (με *NotI* και *KpnI*) από τον *pBluescript* και συνδέθηκε στις ίδιες θέσεις του πλασμιδιακού φορέα έκφρασης *pEGFPC1*, στο καρβοξυτελικό άκρο της φθορίζουσας πρωτεΐνης EGFP. Η κατασκευή που προέκυψε κωδικοποιεί την ανασυνδυσμένη πρωτεΐνη GFP-*Neurl*(227-574).

Από τη σύγκριση της αλληλουχίας του *Neurl* cDNA (*Genbank*, Y15160) με τη διαθέσιμη αλληλουχία του γονιδιώματος του ποντικού, προκύπτει ότι ο κλώνος αυτός περιέχει την κωδική αλληλουχία του γονιδίου *Neurl* (1725bp), όπως φαίνεται από την ταυτότητα της αλληλουχίας 1-1749bp του cDNA με στελώς ή και πλήρως χαρακτηρισμένες περιοχές του χρωμοσώματος 19 του ποντικού. Όμως, για το τμήμα του cDNA από το νουκλεοτίδιο 1750 μέχρι το 2445, δεν υπάρχει ομολογία με το χρωμόσωμα 19, αλλά με το χρωμόσωμα 2 του ποντικού. Αυτό σημαίνει ότι το cDNA Y15160 είναι χιμαιρικό στο 3' άκρο του, περιέχοντας τα αρχικά 24bp της 3' μη μεταφραζόμενης περιοχής του *Neurl*.

LOCUS **MMNEURALI** 2445 bp mRNA linear ROD 07-JAN-2000
DEFINITION Mus musculus mRNA homologous to protein encoded by Drosophila neuralized gene.
ACCESSION Y15160
ORGANISM Mus musculus
AUTHORS Pavlopoulos,E., Prinos,P., Kilpatrick,M., Tsipouras,P. and Moschonas,N.K.
 Submitted (17-OCT-1997) N.K. Moschonas, Dept of Biology, University of Crete and IMBB-FORTH, P.O.Box 1527, Heraklion 711 10, Crete, GREECE
REMARK Revised by author 23-MAR-1998
FEATURES Location/Qualifiers
 source
 1..2445
 /organism="Mus musculus"
 /tissue_type="brain"
 /clone_lib="E14.5"
 /dev_stage="embryonic"
 CDS
 1..1725
 /codon_start=1
 1 atgggtaata acttctccag tgtctctct ctgcagcgag gaaaccocgag ccgcgcgctcg
 61 cggggccacc cccagaacct caaagactcc atcgggggct ccttcccggg gccctctcac
 121 cgatgccatc acaagcagaa gcattgcccg cctacgctgt caggtggggg gctcccggcc
 181 acgcccgtgc tcttccacc ccacactaag ggctcccaga tctcatgga cctcagccac
 241 aaggccgtca agagggcaggc cagctttctgc aatgccatca ccttcagtaa ccgcccggtg
 301 ctcatctacg agcaagtccag gctgaagatc accaagaagc aatgctgctg gagcggggcc
 361 ctgcgacttg gcttccaccg caaggaccct tcccgcattc accccgactc gctgcccagg
 421 tacgctgccc ctgacctggg gtctcagagt ggcttctggg ccaaagcatt gcctgaggag
 481 tttgccaacg agggcaacat cattgccttc tgggtggaca agaagggccg cgtcttctac
 541 cggatcaatg agtcagctgc tatgcttttc ttcagtgggg tccggacggc ggaccggctc
 601 tggggcctgg tggacgtcta cggcctcacg cggggtgtcc agctgctaga cagcagctg
 661 gtgctgcccg actgcctgcg gccgcgctcc ttcaccgcgc tgcggcggcc gtcgctgagg
 721 tgcgaggcgg atgaggcggc cctgtcgggt agcctgtgcg acctcaactg gccgggagcc
 781 cagcggcagc acggcgcacc gcctgcccgc tgcggatcc cgcagaactc gctcaattct
 841 cagcacagcc gcgcgctgcc ggcgcagctc gacggcgacc tgcgcttcca cgcgcttcgc
 901 cgcgcgcgcg acgtccgcat cctggacgag cagacgggtg cgcgcgtgga gcacggggcg
 961 gacgagcggc cgctcgtctt caccagccgg ccggtgcccg tggccgagac catcttcatc
 1021 aaggtcaacg gctcgggtgg cgcgcggccg ggcgcgctgt cgttcggggg caccacgtgt
 1081 gaccocggca cgctgcccgc cgcggacctg cccttcagcc ccgagggccct ggtggaccgc
 1141 aaggagtctt gggcgggtg tgcgctgccc gggcctctgc acagcggcga catcctggg
 1201 ctggtggtca acgcccagcg agagctgcac ctgagtcaca acggcgcggc ggcggcagtg
 1261 cagctgtgcg tggatgccc cgcagcccctc tggatgctc tcagcctgca tggcggccat
 1321 acgcaaggcc gcatcctcgg ctccaccatc atgactgaac ggggtggccc atctctccc
 1381 tgctcacctg cctccactcc aaactcaacc agtgccctgg gcatccgctc ctctgacccc
 1441 ctgctcaagc cctgcccgtc tggggccccta ggtggctctg ctggagggac agcccccaac
 1501 tcacctgtga gcctgcccga gtcaccgggtg accccaggte tgggcccagt gagtgatgaa
 1561 tgcaccattt gctatgaaca cgcagtggtt acagtcattc acacgtgtgg ccacatgtgc
 1621 ctgtgctact cctgtggcct gcgcctcaag aaggccctgc acgctgtgct ccccatctgc
 1681 cgtcgcocca tcaaggacat catcaagacc taccgcagct cctagcccac tgcagagccc
 1741 cacctgcaca acgtctccct catccacttt gctctgtggt ttgtagcatt gactttctct
 1801 ctttgtggtg caccctgtcc cccacaatta ccagggagtc tggatagggt cacaaagaac
 1861 tggcatgagc agactacatg atgccaactc ctcttctgtg atcctgcttg gctgctctgt
 1921 gtcttgggtt cacacacaca cacatacaca cacacactgc ccgaggctgg tcacagcaag
 1981 catgtatgcc acccttgacc caggcaggaa ctgggtgtgt tacactagtg agccactggt
 2041 actcagcctt gttggtgagg gggacacaaa acaacgtgtc caggtgacca ggtaccatg
 2101 gggggccacac gggctgtgccc tgtgaactgg tggatataagg tttgaaggca gttggggtgt
 2161 cttccaaatc cccttgacca gcagactggc tcatacagcc caccctgact cctctgtacc
 2221 ccctactcct ctactccacc cctcccactg ctgctcaaaag tctgaggggc ctggattagg
 2281 tagggtcagg gaaggactgc tgtttgacct aggcctataa ctggaccagc ctgggacttg
 2341 agaagcagcc actctccacc atgctcactc tgggtgagca ccagcagggg gcgtggcagg
 2401 tgcagggcagc taagttcaaa attcagaaca acctctagtt gtctgt

Εικ.29 Η καταχώρηση του cDNA του *Neur1* στη βάση δεδομένων Genbank (απόσπασμα)

8. Ολιγονουκλεοτίδια

Τα συνθετικά ολιγονουκλεοτίδια που χρησιμοποιήθηκαν σε αντιδράσεις PCR για την επιλογή και ανάλυση των κλώνων ανθρώπινου γονιδιωματικού DNA (PAC, BAC), στα πλαίσια της φυσικής χαρτογράφησης του ανθρώπινου χρωμοσώματος 10, παρουσιάζονται στους Πίνακες 5 και 6. Επιπλέον, συνθετικά ολιγονουκλεοτίδια χρησιμοποιήθηκαν σε αντιδράσεις PCR ή προσδιορισμού νουκλεοτιδικής σύστασης μορίων DNA για την ανάλυση της οργάνωσης εξονίων και ιντρονίων του γονιδίου *Neur1* του ποντικού (Πίνακας 7). Τέλος, σχεδιάστηκαν κατάλληλοι εκκινητές PCR, οι οποίοι έφεραν στο 5' άκρο τους θέσεις αναγνώρισης περιοριστικών ενζύμων για την κλωνοποίηση ολόκληρου ή μέρους του cDNA του *Neur1* και άλλων γονιδίων σε διάφορους πλασμιδιακούς φορείς (Πίνακες 8, 9). Η σύνθεση των ολιγονουκλεοτιδίων έγινε από την εταιρία MWG-Biotech, εκτός από τα ολιγονουκλεοτίδια του Πίνακα 5, που παραχωρήθηκαν από το *Sanger Center* (Hinxton, UK).

Πίνακας 5. Εκκινητές PCR για ESTs του ανθρώπινου χρωμοσώματος 10

Sanger ID	Genbank ID	Εκκινητής 1	Εκκινητής 2	Μέγεθος (bp)
Stsg 41319	D12036	TGTTGAAAGGGATTTTGAAGG	CAGGCTTTATTTCAACCAGACA	154
Stsg 26781	D20012	TTGTCTAAGTGAGAGATTGCGA	TGGCTTAGTTCCTGTCTGCA	149
Stsg 41741	H14604	TCTGACACCTCAACCCATGA	TCCCATGCCACAGAATATGA	131
Stsg 32258	W95888	CTTGACACAGAGCAACCCAA	CACCGTCAACATCAGACACC	168
Stsg 30644	AA169811	GGATAATCAAGTTTTTGGCACA	AATTCCTCACACCTTGACC	154
Stsg 16023	R97513	CCTTAGAGTATGCAAGCACGC	TGTGGCTGTCTCCCAATGT	207
Stsg 55729	A004R39	TCAAAGAAACTACATTAGAGAAG	TAGATACTCCAGGAAGAGTTAA	139
Stsg 55566	L16782	GCCAAGTGATTTAATGGACC	ACAGGATGCAAGTTCCAATG	148
Stsg 2721	Z40249	CTTTCACGGCATTGACAGA	TTGGACAATTAGCAACAAGTCA	190
Stsg 21644	N51865	TGGTACTGAGTGACTTTGGGC	CATTCATTGAGGGCCTGTTT	162
Stsg 21645	N51873	CACAACAAACAGAGCCCTT	GCGGGCATACCAAGTGTC	129
Stsg 29905	AA143609	GAGCAGTTGTTTCGCTACAGG	GACACGCTGTGGCTCATCT	170
Stsg 22113	N63988	GCACCCTGATGACCTCAAAT	GGAAAATGTTGCCTGTTGG	186
Stsg 3666	H01317	AACTTTGTAAAGCAGCTGAGGG	GTTTTGTAATCTGTGAGCACCG	166
Stsg 15752	R78938	CATCAGAAGCAACATCTCTTGC	ATTCTCCAGGGTCACACAG	214
Stsg 2551	Z39336	TCITTC AATTGACAGAGATGCC	GCTAGCTCCTCCCTGGTTG	166
Stsg 32004	N67037	ATGGGGACTCACTATGTTCC	ATTTGGTATTTTTCAGCAGC	164
Stsg 12880	H57760	GAATGAGTGGGCGAGTGACT	TGAGGGAAGAAGTGTGATCA	112
Stsg 8438	H64542	AATGAAATATGATGTCTGGGG	AGCCATGCTATCTAATATGGCA	125
Stsg 39120	L21195	CTGTGGGTGGAGAGGACATT	TTCCGGTTGATATTCCGGTA	140
Stsg 55704	H10009	CGTTGCTTATTTTCAGTACAAA	TGGAATCCTGACCATGGAT	290
Stsg 55850	SHGC-35401	GGCATTTTGAAGGCATGG	CCAGATGGATGATCATGAAGG	222
Stsg 31377	AA115498	GGAGGCATTCATTTCAAGGA	TCCCGTTGACCTCAGGATAC	136
Stsg 15491	R63338	ATTTAGGATTACCTTCCACC	GAAGCTCTTTACTGGCCTCA	191
Stsg 8384	H60064	GTTTATTGCAGGTTGGAGAAGG	GGAGACCCTCTGAGGAGCTT	175
Stsg 55807	A006A33	ATACTTGAAGGTAAGAAACTG	TCAGGTTATACTACCTGACTAAA	127
Stsg 35189	R49668	CTGCCAGGATGTTATATCC	AAATATTGGAAGGAGTGGCC	121
Stsg 30011	AA149620	GTGAGCAAAACATCTCTGTTCCG	CAACAAATCTTTGTTGCTGAGG	137
Stsg 31710	AA194926	TGTTCCATGAACATGGTACATC	AGCAAGCAACAAAATCCAGG	124
Stsg 2803	R36999	GGACCTGTGTTACCCAGGG	TAAATAAATGTGCCCCCAA	130
Stsg 12559	H04015	TGGAAAAGTTGGAATGTGTGA	GCTTCTGAGTGGGAAACAGG	115
Stsg 35957	N68018	GTACAGCTGAGGTCTCGATGGTTG	CATGCCTGTGTGCATACATATG	133

Σημ: Τα oligονουκλεοτίδια που χρησιμοποιήθηκαν για τη σύνθεση των ραδιοσημασμένων ανιχνευτών στο πείραμα διερεύνησης της βιβλιοθήκης ανθρώπινου γονιδωματικού DNA RPCI-11. Για κάθε EST φαίνεται το όνομα με βάση το *Sanger Center* (Sanger ID) και τη βάση δεδομένων *Genbank* (Genbank ID), τα ζεύγη των εκκινητών PCR (εκκινητές 1 και 2) και το μέγεθος του αναμενόμενου προϊόντος (bp).

Πίνακας 6. Εκκινητές επαναλήψεων Alu

Εκκινητής Alu	Αλληλουχία εκκινητή (5'-3')
CL1	AGGGTTTCACGACCCTAATGTCC
CL2	ACTGCACTCCAGCCTGGG
CL3	GTTGGCCAGGCTGGTCT
Alu-S	GAGGTTGCAGTGAGCCGAGAT
Alu-J	GAGGCTGCAGTGAGCCGTGAT
Alu-end	ACAGAGCGAGACTCCGTCTC
Alu-4723	GAGGGTTTCACGACCCTAATGTCC

Σημ: Οι εκκινητές που χρησιμοποιήθηκαν σε αντιδράσεις Alu-PCR κατά τη σύνθεση ανιχνευτών ειδικών για το ανθρώπινο γονιδίωμα ή για την ανάλυση των κλώνων YAC και PAC.

Πίνακας 7. Ολιγονουκλεοτίδια του γονιδίου *Neurl*

Εκκινητής	Αλληλουχία (5' – 3')	Κατεύθυνση (F ή R)	Θέση στο <i>Neurl</i> cDNA (Y15160)
<i>Neurl-1</i>	CTTTGAGGTTCTGGGGGTG	R	85-67
<i>Neurl-10</i>	TCTCACCGATGCCATCACAAGC	F	115-136
<i>Neurl-2</i>	CAATGCTTCTGCTTGTGATGG	R	146-126
<i>Neurl-3</i>	CAATGAGTCAGCTGCTATGC	F	546-565
<i>Neurl-5</i>	GCATAGCAGCTGACTCATTG	R	565-546
<i>Neurl-11</i>	TGCTGAGAATTGAGCGAGTTC	R	845-825
<i>Neurl-19</i>	GTCAACGCGGACGGAGAG	F	1207-1224
<i>Neurl-8</i>	GCTCACCTGCCTCCACTCAA	F	1382-1402
<i>Neurl-20</i>	GAGTGGAGGCAGGTGAGC	R	1399-1382
<i>Neurl-7</i>	GCACCATTGCTATGAACAC	F	1562-1581
<i>Neurl-12</i>	TCCACTGCGTGTTCATAGC	R	1589-1571
<i>Neurl-A</i>	GCAGTCCTGCCCGGCCT	F	(-57) – (-40)
<i>Neurl-17</i>	AAACGGAGATGGCCAAGC	F	EST AA667491 (11-26)

Σημ: Ολιγονουκλεοτίδια που χρησιμοποιήθηκαν στα πειράματα ανάλυσης της γονιδιοματικής οργάνωσης του *Neurl*. Όλα υβριδοποιούνται στο *Neurl* cDNA Y15160, εκτός τα *Neurl-A* και *Neurl-17*, που έχουν σχεδιαστεί από την περιοχή του 5'UTR και από το EST AA667491, αντίστοιχα.

Πίνακας 8. Εκκινητές PCR για την κλωνοποίηση του *Neurl*

Εκκινητής	Αλληλουχία (5' – 3')	Κατεύθυνση (F ή R)	Θέση υβριδοποίησης
<i>BamHI</i> -pRSETC-F	AAGGATCC ATGGATCCGACC	F	190-209
<i>EcoRI</i> - <i>Neurl</i> -R	CCGAATTC GCGGTAGGTCTTGATGATG	R	1716-1698
<i>Neurl</i> -2hyb-F	AACAAAGGTCAAAGACAGTTGACTGTA TCGATGGGTAATAACTTCTCCAGTGTC TCCTCTCTGC	F	(845-874) (1-34)
<i>Neurl</i> -2hyb-R	TAAGAAATTCGCCCGGAATTAGCTTGG CTGCTAGGAGCTGCCGTAGGTCTTGATG	R	(903-932) (1725-1701)
<i>EcoRI</i> - <i>Neurl</i> del-F	CCGAATTC ATGGGTAATAACTTCTCC	F	1-18
<i>BamHI</i> - <i>Neurl</i> del-R	CGGGATCC GCATTCATCACTCCACTG	R	1563-1546
<i>EcoRI</i> -RZD-F	CCGAATTC CAGTGGAGTGATGAATGC	F	1545-1563
<i>BamHI</i> -RZD-R	CGGGATCC CTAGGAGCTGCCGTAGGT	R	1725-1708

Σημ: Τα ολιγονουκλεοτίδια που χρησιμοποιήθηκαν στις αντιδράσεις PCR για την κλωνοποίηση ολόκληρου ή μέρους του *Neurl* cDNA στους πλασμιδιακούς φορείς έκφρασης. Παρουσιάζονται οι αλληλουχίες και οι θέσεις υβριδοποίησης των εκκινητών στο *Neurl* cDNA Y15160 (κόκκινο), και στα πλασμίδια pGBT9 (U07646, πράσινο) και pRSETC (Invitrogen V35120, μπλε). Με υπογραμμισμένους χαρακτήρες φαίνονται οι θέσεις αναγνώρισης περιοριστικών ενζύμων που έχουν προστεθεί στους εκκινητές.

Πίνακας 9. Εκκινητές PCR για την κλωνοποίηση των γονιδίων *Sumo-1* και *Dlc-1*

Όνομα εκκινητή	Αλληλουχία εκκινητή (5'-3')	Κατεύθυνση (F ή R)	Θέση υβριδοποίησης
<i>BamHI</i> - <i>Sumo</i> -F	CGGGATCC TGTCTGACCAGGAGGCAAAC	F	<i>Sumo-1</i> (120-140)
<i>ScaI</i> - <i>Sumo</i> -R	AAAAGTACT CTA ACCCCCCGTTTGTTCCTG	R	<i>Sumo-1</i> (409-392)
<i>BamHI</i> - <i>DLC1</i> -F	CGGGATCC TGTGCGACCGAAAGGCCGTG	F	<i>Dlc-1</i> (95-114)
<i>EcoRI</i> - <i>DLC1</i> -R	GGAATTC TTAACCAGATTTGAACAG	R	<i>Dlc-1</i> (363-346)

Σημ: Οι εκκινητές PCR που χρησιμοποιήθηκαν για την κλωνοποίηση των κωδικών περιοχών των γονιδίων *Sumo-1* και *Dlc-1* του ποντικού σε πλασμιδιακούς φορείς (βλ. Πλασμιδιακές κατασκευές). Με υπογραμμισμένους χαρακτήρες φαίνονται οι θέσεις αναγνώρισης των περιοριστικών ενζύμων. Οι θέσεις υβριδοποίησης των εκκινητών αναφέρονται στις αλληλουχίες της βάσης δεδομένων Genbank, με αριθμούς πρόσβασης NM_009460 (*Sumo-1*) και NM_003746 (*Dlc-1*). Σημειώνεται ότι οι εκκινητές για το γονίδιο *Dlc-1* σχεδιάστηκαν με βάση τη διαθέσιμη νουκλεοτιδική αλληλουχία του ανθρώπινου γονιδίου (NM_003746), όμως χρησιμοποιήθηκαν για την κλωνοποίηση του αντίστοιχου γονιδίου του ποντικού, λόγω της υψηλής τους ομολογίας.

9. Ανιχνευτές γονιδίου *Neurl*

Στα πειράματα αρχικής διερεύνησης της βιβλιοθήκης γονιδιωματικού DNA ποντικού καθώς και σε ακόλουθα επιβεβαιωτικά πειράματα χρησιμοποιήθηκαν ως ανιχνευτές το ένθεμα του cDNA κλώνου *Neurl(2433bp)/pBluescript*, καθώς και 10 επικαλυπτόμενα τμήματά του, τα οποία κατασκευάστηκαν είτε με ενζυματικό τεμαχισμό του αρχικού κλώνου, είτε ως προϊόντα PCR με κατάλληλους εκκινητές. Οι διαφορετικοί ανιχνευτές cDNA και γονιδιωματικού DNA που χρησιμοποιήθηκαν για τη διερεύνηση της φαγικής βιβλιοθήκης και την ανάλυση ολικού γονιδιωματικού DNA αναφέρονται στον Πίνακα 10. Συγκεκριμένα, οι ανιχνευτές 1, 1A, 2, 3, 4 και 5 προέκυψαν από τον τεμαχισμό του *Neurl(2433bp)/pBluescript*

με τα ένζυμα που φαίνονται στην παρένθεση (Πίνακας 10). Οι ανιχνευτές 7, 8 και 9 αντιστοιχούν σε προϊόντα PCR που προέκυψαν από τον κλώνο *Neur1(2433bp)/pBluescript* με τους αντίστοιχους εκκινητές PCR. Ο ανιχνευτής 6 αντιστοιχεί στο προϊόν PCR των εκκινητών A και 1, το οποίο παρασκευάστηκε από προϊόν αντίδρασης αντίστροφης μεταγραφής του εκκινητή 5 σε ολικό RNA εγκεφάλου ποντικού. Ο ανιχνευτής 10 (προϊόν PCR 17-2), προέκυψε από αντίδραση PCR με τους εκκινητές 17 και 2 σε προϊόν αντίδρασης αντίστροφης μεταγραφής του εκκινητή 5. Ο ανιχνευτής *SacII-PstI* (400bp) (Πίνακας 10) προήλθε από γονιδιωματικό κλώνο (1.4Kb *PstI*) που χαρτογραφείται στην περιοχή 5'UTR του γονιδίου (βλ. Αποτελέσματα).

Πίνακας 10. Ανιχνευτές cDNA και γονιδιωματικού DNA

(A)	Ανιχνευτής	Αντίστοιχο τμήμα <i>Neur1</i> cDNA	Μέγεθος (bp)
	2433	13-2445 bp (EcoRI)	2433
	1	13-617 bp (EcoRI-AccI)	605
	1A	13-218 bp (EcoRI-BstYI)	206
	2	606-1466 bp (ApaI)	861
	3	1466-2270 bp (ApaI)	805
	4	679-1103 bp (NotI – SacII)	425
	5	1103-1734 (SacII – PstI)	632
	6 (προϊόν PCR A-1)	-57 - 85	142
	7 (προϊόν PCR 5-10)	115 - 565	451
	8 (προϊόν PCR 3-11)	546 - 845	300
	9 (προϊόν PCR 8-12)	1382 - 1589	208
	10 (προϊόν PCR 17-2)	-	126

(B)	Ανιχνευτής γονιδιωματικού DNA	Προέλευση	Μέγεθος (bp)
	<i>SacII-PstI</i>	Το 3' ακραίο τμήμα του κλώνου 1.4Kb <i>PstI</i>	400

Σημ: Οι ανιχνευτές cDNA (A) και γονιδιωματικού DNA (B) που χρησιμοποιήθηκαν για την απομόνωση και τον χαρακτηρισμό των γονιδιωματικών κλώνων και για την ανάλυση ολικού γονιδιωματικού DNA. Η αρίθμηση των θέσεων αναγνώρισης των περιοριστικών ενζύμων και υβριδοποίησης των εκκινητών PCR είναι με βάση την αλληλουχία *Neur1* cDNA Y15160.

10. Αντισώματα

Τα αντισώματα που χρησιμοποιήθηκαν σε πειράματα ανοσοανίχνευσης πρωτεϊνών μονιμοποιημένων σε μεμβράνες (ανάλυση *Western*) ή σε μονιμοποιημένα κύτταρα (ανοσοφθορισμός *in situ*) και ανοσοκατακρήμνισης πρωτεϊνικών συμπλόκων από εκχυλίσματα κυττάρων αναφέρονται στον Πίνακα 11.

Πίνακας 11α. Πρωτογενή αντισώματα

Αντίσωμα	Πρωτεΐνη που ανιχνεύεται	Αραίωση	Χρόνος επώασης	Προέλευση
Rabbit anti- Neur1	Neur1	1: 1000 (W) 1: 100 (ΑΦ)	3 hrs (W) 1hr (ΑΦ)	E. Κουτέλου, Διπλωματική Εργασία, Παν/μιο Κρήτης, 1999
Mouse anti-Lamin A/C	Lamin A/C	1: 500 (ΑΦ)	45 min (ΑΦ)	Santa Cruz Biotech. (Cat.no. sc-7292)
Mouse anti-Xpress	Συνθετικός επίτοπος Xpress	1: 2500 (W) 1: 500 (ΑΦ)	1 hr (W, ΑΦ)	Invitrogen (Cat.no. R910-25)
Mouse anti-FLAG	Συνθετικός επίτοπος FLAG	1: 3000 (W) 1: 250 (ΑΦ)	1 hr (W, ΑΦ)	Sigma (Cat.no. F3165)
Goat anti-Ubc9	Ubc9	1:2000 (W) 1: 200 (ΑΦ)	2 hrs (W, ΑΦ)	Santa Cruz Biotech. (Cat.no. sc-5229)
Rat anti-Dlc1	Dlc-1	1:100 -1:1000 (W) 1: 100 (ΑΦ)	2 hrs (W, ΑΦ)	Alexis Biochemicals (Cat.no. 804-341-C100)
Mouse anti-NuMA	NuMA1	1:50 (ΑΦ)	90 min (ΑΦ)	NeoMarkers (Cat.no. MS-496-P0)
Mouse anti-β tubulin	β-tubulin	1:100 (ΑΦ)	90 min (ΑΦ)	Santa Cruz Biotech. (Cat.no. sc-5274)
Mouse anti-SUMO-1	SUMO-1	1: 1000 (W) 1 : 100 (ΑΦ)	2 hrs (W, ΑΦ)	Santa Cruz Biotech. (Cat.no. sc-5308)
Mouse anti-GST	GST	1:5000 (W)	1hr (W)	Santa Cruz Biotech. (Cat.no. sc-138)
Mouse anti-GFP	GFP	1:3000 (W)	1hr (W)	Santa Cruz Biotech. (Cat.no. sc-9996)

Σημ: Τα χαρακτηριστικά των πρωτογενών αντισωμάτων που χρησιμοποιήθηκαν. Για κάθε αντίσωμα αναφέρεται η εφαρμογή στην οποία χρησιμοποιήθηκε (W= ανάλυση Western, ΑΦ=ανοσοφθορισμός σε μονιμοποιημένα κύτταρα), η αραίωση και ο χρόνος επώασης.

Πίνακας 11β. Δευτερογενή αντισώματα

Αντίσωμα	Πρωτεΐνη που ανιχνεύεται	Αραίωση	Χρόνος επώασης	Προέλευση
ΔΕΥΤΕΡΟΓΕΝΗ ΑΝΤΙΣΩΜΑΤΑ				
Συζευγμένα με Horse Radish Peroxidase (HRP)				
Donkey anti-mouse	Mouse IgG	1:2500 (W)	2hrs (W)	Jackson Immuno. (Cat.no. 715-035-150)
Donkey anti-goat	Goat IgG	1:1000 (W)	2hrs (W)	Santa Cruz Biotech. (Cat.no. sc-2020)
Goat anti-rabbit	Rabbit IgG	1:12000 (W)	1hr (W)	Jackson Immuno. (Cat.no. 111-005-003)
Goat anti-rat	Rat IgG	1:3000 (W)	2hrs (W)	Jackson Immuno. (Cat.no. 112-035-003)
Συζευγμένα με φθορίζουσες ουσίες (Alexa Fluor)				
Alexa Fluor 488 goat anti-mouse	Mouse IgG	1:1000 (ΑΦ)	45min (ΑΦ)	Molecular Probes (Cat.no. A-11029)
Alexa Fluor 568 goat anti-mouse	Mouse IgG	1:1000 (ΑΦ)	45min (ΑΦ)	Molecular Probes (Cat.no. A-11031)
Alexa Fluor 568 goat anti-rabbit	Rabbit IgG	1:1000 (ΑΦ)	45min (ΑΦ)	Molecular Probes (Cat.no. A-11036)
Alexa Fluor 647 goat anti-rabbit	Rabbit IgG	1:1000 (ΑΦ)	45min (ΑΦ)	Molecular Probes (Cat.no. A-21244)
Alexa Fluor 488 donkey anti-goat	Goat IgG	1:1000 (ΑΦ)	45min (ΑΦ)	Molecular Probes (Cat.no. A-11055)
Alexa Fluor 568 donkey anti-goat	Goat IgG	1:1000 (ΑΦ)	45min (ΑΦ)	Molecular Probes (Cat.no. A-11057)
Alexa Fluor 488 donkey anti-rat	Rat IgG	1:1000 (ΑΦ)	45min (ΑΦ)	Molecular Probes (Cat.no. A-21208)

Σημ: Τα χαρακτηριστικά των δευτερογενών αντισωμάτων που χρησιμοποιήθηκαν. Για κάθε αντίσωμα αναφέρεται η εφαρμογή στην οποία χρησιμοποιήθηκε (W= ανάλυση Western, ΑΦ=ανοσοφθορισμός σε μονιμοποιημένα κύτταρα), η αραίωση και ο χρόνος επώασης.

11. Διερεύνηση διατεταγμένων βιβλιοθηκών ανθρώπινου γονιδιωματικού DNA

Για τη δημιουργία του φυσικού χάρτη του χρωμοσώματος 10, συλλέχθηκαν κλώνοι ανθρώπινου γονιδιωματικού DNA από βιβλιοθήκες PAC (RPCI-1) και BAC (RPCI-11). Η διερεύνηση των βιβλιοθηκών πραγματοποιήθηκε με υβριδοποίηση κατάλληλων ραδιοσημασμένων ανιχνευτών. Αναλυτικότερα, η βιβλιοθήκη RPCI-1 υβριδοποιήθηκε με ραδιοσημασμένα προϊόντα αντιδράσεων Alu-PCR, στις οποίες εκκινητές που αναγνωρίζουν τις χαρακτηριστικές για το ανθρώπινο DNA επαναλήψεις Alu χρησιμοποιήθηκαν με μήτρα το DNA κλώνων YAC που χαρτογραφούνται στην υπό μελέτη χρωμοσωμική περιοχή. Επιπλέον, η βιβλιοθήκη κλώνων BAC (RPCI-11) υβριδοποιήθηκε με ομάδες ραδιοσημασμένων αλληλουχιών EST που προέκυψαν από ανθρώπινο γονιδιωματικό DNA ως προϊόντα αντιδράσεων PCR με ειδικούς εκκινητές.

➤ Διερεύνηση βιβλιοθήκης κλώνων PAC (RPCI-1)

Η διαδικασία που ακολουθήθηκε για τη διερεύνηση της βιβλιοθήκης κλώνων PAC (RPCI-1) περιελάμβανε τα εξής στάδια:

α. Παραγωγή και σήμανση ανιχνευτών

Για τη δημιουργία ραδιοσημασμένων ανιχνευτών αντιπροσωπευτικών της μελετούμενης χρωμοσωμικής περιοχής, πραγματοποιήθηκαν αντιδράσεις PCR με εκκινητές των επαναλαμβανόμενων στοιχείων Alu (Πίνακας 6) και μήτρα κλώνους YAC που είχαν χαρτογραφηθεί στη συγκεκριμένη περιοχή από προηγούμενη μελέτη [Marzella, Kokkinaki *et al*, 1997]. Αναλυτικότερα, σε κάθε αντίδραση Alu-PCR χρησιμοποιήθηκαν 100ng DNA κλώνου YAC, 25μM από κάθε εκκινητή Alu, 1x ρυθμιστικού διαλύματος Taq DNA πολυμεράσης της εταιρίας Minotech, 100μg/ml BSA, 1mM dNTPs, 1.5mM MgCl₂ και 3units ενζύμου, σε τελικό όγκο 100μl. Αρχική αποδιάταξη του DNA στους 96°C για 3 min, ακολουθούνταν από 30 κύκλους που περιελάμβαναν αποδιάταξη για 1 min στους 96°C, υβριδοποίηση των εκκινητών για 30sec στους 58°C και ενζυμικό πολλαπλασιασμό των μεταξύ των επαναλήψεων Alu αλληλουχιών στους 72°C για 2 min. Ακολουθούσε στάδιο τελικής επέκτασης των προϊόντων PCR στους 72°C για 5 min. Στη συνέχεια, απομακρύνονταν οι εκκινητές PCR χρησιμοποιώντας το σύστημα καθαρισμού προϊόντων PCR της εταιρίας Millipore (Cat. No: UFC7PCR50). Τα μεγέθη των προϊόντων της αντίδρασης για κάθε συνδυασμό εκκινητή Alu και κλώνου YAC κυμαίνονταν από 100bp έως και 5Kb. Σε αντιδράσεις που προέκυπταν σχεδόν αποκλειστικά μικρού μεγέθους προϊόντα, η σήμανση πραγματοποιούνταν στο 5' άκρο τους με πολυνουκλεοτιδική κινάση, ενώ για τη σήμανση μεγαλύτερων προϊόντων επιλέγονταν η μέθοδος των τυχαίων εκκινητών. Σε κάθε περίπτωση, για την υβριδοποίηση της βιβλιοθήκης χρησιμοποιήθηκαν ραδιοσημασμένοι ανιχνευτές με ειδική ενεργότητα τουλάχιστον 2x10⁸cpm/μg, η οποία προσδιορίστηκε όπως περιγράφεται παρακάτω.

β. Υβριδοποίηση

Οι μεμβράνες της βιβλιοθήκης RPCI-1 επώαστηκαν αρχικά σε διάλυμα 2x SSC για 10 min και στη συνέχεια σε 10 ml διαλύματος προϋβριδοποίησης [50% φορμαμίδη, 10% διάλυμα Denhart's (1% BSA, 1% PVP 40 και 1% Ficoll 400), 6x SSC, 0.5% SDS] στους 42°C για 2 hrs σε γυάλινους κυλίνδρους υβριδοποίησης Hybaid (2 μεμβράνες / κύλινδρο) παρουσία μίγματος αποδιαταγμένων θραυσμάτων γονιδιωματικού DNA από *E.coli* και σπέρμα σολωμού (100μg/ml διαλύματος προϋβριδοποίησης) για την παρεμπόδιση μη ειδικής υβριδοποίησης. Ακολουθούσε προσθήκη 10ml διαλύματος υβριδοποίησης ίδιας σύστασης με το προηγούμενο, το οποίο περιείχε αποδιαταγμένο και τον ραδιοσημασμένο ανιχνευτή (τουλάχιστον 10⁶cpm/ml). Μετά από υβριδοποίηση για 16hrs στις ίδιες συνθήκες, για την απομάκρυνση της περίσσειας ή των μη ειδικά δεσμευμένων ανιχνευτών ακολουθούσαν πλυσίματα των μεμβρανών σε συνθήκες ελαττούμενης αλατότητας και αυξανόμενης θερμοκρασίας. Συγκεκριμένα, οι μεμβράνες επώαστηκαν δύο φορές για 10 min με ήπια ανακίνηση σε καθεμία από τις ακόλουθες συνθήκες: α) διάλυμα 2x SSC σε θερμοκρασία δωματίου, β) διάλυμα 2x SSC και 0.5% SDS, στους 65°C, γ) διάλυμα 0.5x SSC και 0.5% SDS, στους 65°C, δ) διάλυμα 0.1x SSC και 0.5% SDS, στους 65°C. Ακολουθούσε αυτοραδιογραφία των μεμβρανών και έκθεση στους -80°C για περίπου 12hrs. Στη συνέχεια, προσδιορίζονταν οι συντεταγμένες κάθε σήματος και καθορίζονταν ο αντίστοιχος κλώνος PAC, με βάση τον κώδικα που παρέχονταν για τη συγκεκριμένη βιβλιοθήκη από το HGMP *Resource Center*. Για τη χρήση των μεμβρανών σε επόμενα πειράματα υβριδοποίησης, η αποϋβριδοποίηση γινόταν σε διάλυμα 0.4N NaOH με ήπια ανακίνηση για 20min σε θερμοκρασία δωματίου, ακολουθούσε εξουδετέρωση σε διάλυμα 5x SSC για 10 min και φύλαξη των μεμβρανών στους 4°C.

γ. Ανάπτυξη κλώνων PAC

Η προμήθεια των θετικών κλώνων της προηγούμενης διαδικασίας έγινε από το HGMP *Resource Center* με τη μορφή αποικιών βακτηριακών κυττάρων *E.coli* DH10B αναπτυγμένων σε θρεπτικό μέσο LB, παρουσία 25μg/ml καναμικίνης. Ακολουθούσε απομόνωση πλασμιδιακού DNA (PAC) με αλκαλική λύση και ανάλυσή του με διάφορες μεθόδους.

δ. Ανάλυση κλώνων PAC

Οι κλώνοι PAC αναλύονταν με στόχο να επιβεβαιωθεί η προέλευσή τους από τη συγκεκριμένη χρωμοσωμική περιοχή αλλά και να προσδιοριστούν οι μεταξύ τους επικαλύψεις, ώστε να ομαδοποιηθούν σε *contigs*. Η διαδικασία περιελάμβανε: i) Ενζυμικό τεμαχισμό των κλώνων με περιοριστικές ενδονουκλεάσες σπάνιας (NotI) ή πιο συχνής (BamHI, EcoRI) θέσης αναγνώρισης και ανάλυσή τους σε ηλεκτροφόρηση παλλόμενου ηλεκτρικού πεδίου. ii) Αντιδράσεις PCR στις οποίες ανιχνεύονταν η παρουσία μικροδορυφορικών δεικτών του ανθρώπινου χρωμοσώματος 10 (Πίνακας 12), για τους οποίους ήταν γνωστό από προηγούμενη μελέτη [Marzella, Kokkinaki *et al*, 1997] ότι περιέχονταν στα YACs που είχαν χρησιμοποιηθεί για την παρασκευή των ανιχνευτών και την επιλογή των κλώνων PAC από τη βιβλιοθήκη.

Πίνακας 12. Εκκινητές PCR πολυμορφικών μικροδορυφορικών δεικτών

ΓΕΝΕΤΙΚΟΣ ΔΕΙΚΤΗΣ	Εκκινητής 1	Εκκινητής 2	Ελάχιστο μέγεθος (bp)
D10S2311	GAATGGGTTGAATAAACATGC	GAGATACAATTCAAGTCGAG	224
D10S564 *	TGGGAATGTGTCTTTATCCA	AGCTCTAACATAGAGGCCAGAT	262
D10S1753 *	CTGCTGCCACCAACCTAA	CAAGTGGGAGACTCGATGACA	263
D10S2309	GTAATGACGAGATAACAGGTG	CTCCTAACAAATGATTATCAC	355

Σημ: Για κάθε μικροδορυφορικό γενετικό δείκτη, φαίνεται το ζεύγος των αντίστοιχων εκκινητών PCR (Εκκινητές 1 και 2) και το ελάχιστο μέγεθος του προϊόντος PCR σε ανθρώπινο γονιδιωματικό DNA. Στοιχεία από Genome Data Base (*) και δικά μας πειράματα [Κοκκινάκη Μ, Μεταπτυχιακή διατριβή, Παν/μιο Κρήτης, 1996]

➤ Διερεύνηση βιβλιοθήκης κλώνων BAC (RPCI-11)

Η διαδικασία που ακολουθήθηκε για τη διερεύνηση της βιβλιοθήκης κλώνων BAC (RPCI-11) περιελάμβανε τα εξής στάδια:

α. Παραγωγή και σήμανση ανιχνευτών

Ως ανιχνευτές για τη διερεύνηση της βιβλιοθήκης κλώνων BAC χρησιμοποιήθηκαν αλληλουχίες EST (*Expressed Sequenced Tags*), οι οποίες χαρτογραφούνται στην υπό μελέτη περιοχή του χρωμοσώματος 10, σύμφωνα με τους διαθέσιμους χάρτες ακτινοβολημένων κυτταρικών υβριδίων (*radiation cell hybrids*). Τα ESTs επελέγησαν ώστε να καλύπτουν ομοιόμορφα τη χρωμοσωμική περιοχή και να ανήκουν σε διαφορετικά γονίδια, σύμφωνα με την ομαδοποίησή τους σε μεταγραφικές μονάδες της βάσης δεδομένων *'Unigene'*. Τα ζεύγη των εκκινητών PCR για τα ESTs που χρησιμοποιήθηκαν στη μελέτη (Πίνακας 5) παραχωρήθηκαν από το *Sanger Center (Hinxton, UK)*, στα πλαίσια της συνεργασίας μας για την αλληλούχηση του χρωμοσώματος 10. Η ραδιοσήμανση των προϊόντων PCR περιγράφεται παρακάτω. Ο καθαρισμός από την περίσσεια ραδιενεργών νουκλεοτιδίων έγινε με χρωματογραφία Sephadex G-50. Στη συνέχεια, τα ραδιοσημασμένα προϊόντα PCR χρησιμοποιήθηκαν ως ανιχνευτές για τη διερεύνηση της βιβλιοθήκης σε ομάδες των 10, ανά πείραμα υβριδοποίησης. Η ειδική ενεργότητα των ραδιοσημασμένων μορίων ήταν της τάξεως του 10^9 cpm/μg προϊόντος PCR, μετά από κατακρήμνιση με τριχλωροξικό οξύ.

β. Υβριδοποίηση

Τα φίλτρα της βιβλιοθήκης RPCI-11 επώαστηκαν αρχικά σε διάλυμα 2x SSC για 10min και στη συνέχεια σε 10 ml διαλύματος προϋβριδοποίησης [100mM NaH₂PO₄, 50mM Na₄P₂O₇, 10mM EDTA, 7% SDS, 1% BSA (pH 7.0-7.5)] στους 65°C για 2hrs σε γυάλινους κυλίνδρους υβριδοποίησης Hybaid (12 μεμβράνες / κύλινδρο) παρουσία μίγματος

αποδιαταγμένων θραυσμάτων γονιδιωματικού DNA από σπέρμα σολωμού (100μg/ml διαλύματος προϋβριδοποίησης) και με 10 ml διαλύματος υβριδοποίησης ίδιας σύστασης, το οποίο περιείχε αποδιαταγμένο και τον ραδιοσημασμένο ανιχνευτή (τουλάχιστον 10^6 cpm/ml) για 16hrs στις ίδιες συνθήκες. Ακολουθούσε η ίδια διαδικασία πλυσίματος που περιγράψαμε για τη βιβλιοθήκη RPCI-1.

γ. Ανάλυση κλώνων BAC

Οι θετικοί κλώνοι BAC, ενσωματώθηκαν στην ομάδα των ειδικών για το χρωμόσωμα 10 κλώνων, που προέκυψαν από παρόμοιες μελέτες για άλλες περιοχές του ίδιου χρωμοσώματος είτε στο εργαστήριό μας είτε στο συνεργαζόμενο *Sanger Center* και απετέλεσαν το υλικό για τη δημιουργία της σειράς φίλτρων με διατεταγμένους (*gridded*) κλώνους του χρωμοσώματος 10 (*Sanger Center*). Στη συνέχεια, οι κλώνοι αυτοί υβριδοποιήθηκαν με κάθε EST ξεχωριστά και, επιπλέον, με πολυμορφικούς γενετικούς δείκτες και STSs που σχεδιάστηκαν από τον προσδιορισμό των ακραίων αλληλουχιών των κλώνων BAC της συγκεκριμένης βιβλιοθήκης, που έγινε στο κέντρο αλληλούχησης του Πανεπιστημίου της *Washington (St. Louis)*. Στο *Sanger Center* πραγματοποιήθηκε και η σύγκριση των προτύπων τεμαχισμού των κλώνων BAC με την περιοριστική ενδονουκλεάση *HindIII (fingerprinting)*, από τα αποτελέσματα της οποίας δημιουργούνταν η βάση δεδομένων FPC για το χρωμόσωμα 10 (www.sanger.ac.uk/cgi-bin/humace/fpcwebmap.cgi). Με βάση τα παραπάνω, δημιουργούνταν *contigs* επικαλυπτόμενων κλώνων, σταδιακά αυξανόμενου μεγέθους και επιλέγονταν οι κλώνοι που κάλυπταν επαρκώς την χρωμοσωμική περιοχή σχηματίζοντας το ελάχιστο *'tiling path'*, ώστε να χρησιμοποιηθούν σε πειράματα προσδιορισμού της νουκλεοτιδικής τους αλληλουχίας με τη μέθοδο της τυχαίας (*shotgun*) αλληλούχησης για κάθε κλώνο [Bentley *et al*, 2001].

12. Διερεύνηση βιβλιοθήκης βακτηριοφάγων

Για τη συλλογή κλώνων γονιδιωματικού DNA ποντικού από την περιοχή του γονιδίου *Neuralized (Neurl)*, διερευνήθηκε βιβλιοθήκη γονιδιωματικών κλώνων ποντικού στελέχους 129SV, κλωνοποιημένη σε βακτηριοφάγους λFIXII, με μεγέθη ενθεμάτων 9-23Kb. Ο τίτλος της προσδιορίστηκε ότι ήταν της τάξεως του 10^8 πλάκες βακτηριοφάγων/ml. Η διαδικασία που ακολουθήθηκε περιγράφεται αναλυτικά στα σχετικά πρωτόκολλα [Sambrook *et al*, 1989]. Αναπτύχθηκαν περίπου 6×10^5 βακτηριοφάγοι σε στερεό υπόστρωμα, μεταφέρθηκαν σε μεμβράνες και υβριδοποιήθηκαν με ραδιοσημασμένο ανιχνευτή. Τα θετικά σήματα αναπτύχθηκαν εκ νέου σε στερεό μέσο και επαναυβριδοποιήθηκαν με τον ανιχνευτή, μέχρι να είναι δυνατή η απομόνωση μοναδιαίων θετικών πλακών βακτηριοφάγων.

➤ Ανάπτυξη της βιβλιοθήκης σε στερεό θρεπτικό υπόστρωμα

Ο βακτηριοφάγος λFIXII αναπτύχθηκε μολύνοντας κύτταρα *Escherichia coli* στελέχους K802. Αρχικά παρασκευάστηκαν κύτταρα δεκτικά για μόλυνση από βακτηριοφάγο και στη συνέχεια τα μολυσμένα βακτήρια αναπτύχθηκαν σε στερεό θρεπτικό μέσο:

α. Παρασκευή κυττάρων δεκτικών για μόλυνση από βακτηριοφάγο λ

Μοναδιαία αποικία κυττάρων *E. coli* στελέχους K802 καλλιεργήθηκε σε 50ml θρεπτικού μέσου LB με 0.2% μαλτόζη και 10mM $MgSO_4$, για 14-16 hrs ($OD_{600} \sim 1-1.5$). Ακολούθησε φυγοκέντρηση των κυττάρων στις 2000 rpm για 15 min στους 4°C και επαναιώρηση του ιζήματος σε 20ml (4/10 του όγκου της αρχικής καλλιέργειας) 10mM $MgSO_4$, ώστε το κυτταρικό εναιώρημα που προκύπτει να περιέχει περίπου 1.6×10^9 βακτήρια ανά ml, συγκέντρωση που αντιστοιχεί σε οπτική απορρόφηση OD_{600} περίπου ίση με 2.

β. Μόλυνση κυττάρων *E.coli* από βακτηριοφάγο λ

Πρόσληψη των λ φάγων από τα βακτήρια επιτυγχάνεται παρουσία ιόντων Mg^{+2} , μέσω υποδοχέα της εξωτερικής μεμβράνης των βακτηρίων που κωδικοποιείται από το γονίδιο *LamB*, το οποίο επάγεται από μαλτόζη. Η διείσδυση του φαγικού DNA στο βακτήριο γίνεται αποτελεσματικά στους 37°C. Αναλυτικότερα, σε 4 δοκιμαστικούς σωλήνες Falcon 50ml, περίπου 2×10^9 δεκτικά βακτήρια (1.2ml εναιωρήματος) αναμείχθηκαν με 1.5×10^5 φάγους. Μετά από επώαση στους 37°C για 20 min, ακολούθησε προσθήκη 40ml θρεπτικού μέσου NZYM επιφάνειας προθερμασμένου στους 45°C, ώστε να βρίσκεται σε ρευστή κατάσταση. Το μίγμα μεταφέρθηκε σε 4 τρυβλία διαστάσεων 23cm x 23cm με στερεό θρεπτικό υπόστρωμα NZYM βάσης. Τα 4 τρυβλία επώαστηκαν για 12-14 hrs στους 37°C, όπου οι φάγοι πολλαπλασιάστηκαν λύνοντας τα βακτήρια και παράγοντας συνολικά 6×10^5 πλάκες και περίπου 6×10^{10} φαγικές μονάδες (*pfus: plaque forming units*, 1 πλάκα= $\sim 10^5$ φαγικές μονάδες που φέρουν το ίδιο ένθεμα).

➤ Μεταφορά των βακτηριοφάγων σε μεμβράνες υβριδοποίησης

Οι βακτηριοφάγοι μεταφέρθηκαν από τα τρυβλία σε μεμβράνες υβριδοποίησης (Hybond-N⁺, Amersham) σύμφωνα με τη μέθοδο των Grunstein και Hogness (1975). Οι μεμβράνες τοποθετήθηκαν επάνω στα τρυβλία των φάγων για 2 min και στη συνέχεια μεταφέρθηκαν για 5 min σε διάλυμα 1.5M NaCl και 0.5M NaOH για την αποδιάταξη του DNA των δεσμευμένων φάγων, για 7 min σε διάλυμα εξουδετέρωσης (1M Tris-Cl pH 8.0 και 1.5M NaCl) και τέλος σε 2x SSC για 5-10 min. Το δεσμευμένο DNA μονιμοποιήθηκε με έκθεση σε υπεριώδη ακτινοβολία (*UV cross-linking*) και θέρμανση των μεμβρανών στους 80°C για 2 hrs. Ακολούθησαν πειράματα υβριδοποίησης.

➤ Επιλογή θετικών κλώνων

Οι μεμβράνες υβριδοποιήθηκαν με ανιχνευτές που είχαν σημανθεί ραδιενεργά με τη μέθοδο των τυχαίων εκκινήτων ή με PCR. Συγκεκριμένα, ανά μεμβράνη χρησιμοποιήθηκαν 10ml διαλύματος υβριδοποίησης [100mM NaH_2PO_4 , 50mM $Na_4P_2O_7$, 10mM EDTA, 7% SDS, 1% BSA (pH 7.0-7.5)] και 1mg αποδιαταγμένου carrier DNA. Η περίσσεια ραδιενέργειας μετά την υβριδοποίηση απομακρύνθηκε με πλύσιμο των μεμβρανών σε διάλυμα 20mM NaH_2PO_4 , 10mM $Na_4P_2O_7$, 2mM EDTA, 0.3M NaCl, 5% SDS. Για κάθε θετικό σήμα εντοπίστηκε η αντίστοιχη θέση στο τρυβλίο και από αυτήν συλλέχθηκε ένα κυλινδρικό τεμάχιο στερεού θρεπτικού υποστρώματος διαμέτρου περίπου 1cm, που περιείχε περίπου 200 πλάκες ($\sim 2 \times 10^7$ βακτηριοφάγους). Κάθε δείγμα βακτηριοφάγων μεταφέρθηκε σε 500μl διαλύματος SM (0.1M NaCl, 8mM $MgSO_4$, 50mM Tris-Cl pH 7.5), που περιείχε και 50μl χλωροφορμίου για την απομάκρυνση υπολειμμάτων των βακτηριακών κυττάρων. Ακολούθησαν 2 ή και 3 πειράματα υβριδοποίησης για κάθε μία από τις ομάδες κλώνων, στα οποία οι φάγοι αναπτύχθηκαν σε μικρότερα τρυβλία (διαμέτρου 10cm) και σε μικρότερη πυκνότητα (500-1000 φάγοι/τρυβλίο) μολύνοντας περίπου $1.5-2 \times 10^8$ βακτήρια (100μl εναιωρήματος), ώστε τελικά να είναι δυνατή η απομόνωση μοναδιαίας θετικής πλάκας. Στη συνέχεια, απομονώθηκε DNA από τους θετικούς κλώνους, υδρολύθηκε με περιοριστικές ενδονουκλεάσες, αναλύθηκε κατά Southern, ενώ συγκεκριμένα τεμάχια του φαγικού ενθέματος υποκλωνοποιήθηκαν σε πλασμιδιακούς φορείς και προσδιορίστηκε μέρος της νουκλεοτιδικής τους σύστασης.

13. Απομόνωση και ανάλυση μορίων DNA

Χρησιμοποιήθηκαν οι εξής μέθοδοι απομόνωσης: 1) DNA βακτηριοφάγων με λύση των καψιδίων, 2) πλασμιδιακού DNA με λύση βακτηριακών κυττάρων, 3) γονιδιωματικού DNA κυττάρων θηλαστικού με ομογενοποίηση και λύση του αντίστοιχου ιστού, καθώς και 4) μίγματος γονιδιωματικού και πλασμιδιακού DNA σακχαρομύκητα. Ακολούθησαν πειράματα ανάλυσης του DNA.

α. Απομόνωση DNA βακτηριοφάγων

Για την απομόνωση DNA από βακτηριοφάγους εφαρμόστηκε ο καθαρισμός με φυγοκέντρηση σε διαβάθμιση συγκεντρώσεων γλυκερόλης [Sambrook *et al*, 1989], με κατάλληλες τροποποιήσεις. Αναλυτικότερα, η κάθε μοναδιαία πλάκα βακτηριοφάγων εκλούστηκε σε διάλυμα SM (0.1M NaCl, 8mM MgSO₄, 50mM Tris-Cl pH 7.5), προσδιορίστηκε ο τίτλος του εναιωρήματος και στη συνέχεια 6x10⁵ βακτηριοφάγοι χρησιμοποιήθηκαν για τη μόλυνση και την πλήρη λύση περίπου 2x10⁹ βακτηριακών κυττάρων *E. coli*, τα οποία απλώθηκαν σε τρυβλία διαστάσεων 23x23cm με θρεπτικό μέσο NZYM. Μετά από 16 hrs επώασης στους 37°C και πλήρη λύση της βακτηριακής χλόης, προστέθηκαν στο τρυβλίο 50ml διαλύματος SM. Ακολούθησε έκλυση των φάγων με ανακίνηση στις 100 στροφές / min για 6 hrs στους 25°C και το εναιώρημα που προέκυψε επώαστηκε στους 37°C για 30 min με DNase I σε τελική συγκέντρωση 1μg/ml, για την υδρόλυση του βακτηριακού DNA. Στη συνέχεια, προστέθηκε στερεό NaCl σε τελική συγκέντρωση 1M και ακολούθησε επώαση στον πάγο για 1 hr και φυγοκέντρηση για 30 min στις 3000 rpm, για την απομάκρυνση των υπολειμμάτων των βακτηριακών κυττάρων. Οι βακτηριοφάγοι κατακρημνίστηκαν από το υπερκείμενο προσθέτοντας στερεά πολυαιθυλενική γλυκόλη μοριακού βάρους 8000 (PEG 8000) σε τελική συγκέντρωση 10% και επώαστηκαν για 1 hr στον πάγο. Ακολούθησε φυγοκέντρηση στις ίδιες συνθήκες και επαναιώρηση του ιζήματος σε 3ml διαλύματος TM (50mM Tris-Cl pH 7.8 και 10mM MgSO₄). Το εναιώρημα των φάγων αρχικά εκχυλίστηκε 2-3 φορές με χλωροφόρμιο για την απομάκρυνση μεγάλου μέρους της PEG και στη συνέχεια τοποθετήθηκε στην κορυφή μιας διαβάθμισης συγκεντρώσεων γλυκερόλης, η οποία παρασκευάζονταν σε πολυανθρακικό σωλήνα υπερφυγοκέντρου Beckman SW41 προσθέτοντας διαδοχικά 3ml διαλύματος 40% γλυκερόλης σε TM και 4ml διαλύματος 5% γλυκερόλης σε TM. Ακολούθησε υπερφυγοκέντρηση στις 35.000 rpm για 90 min στους 4°C. Το υπερκείμενο απομακρύνθηκε πλήρως και το ιζήμα επαναιωρήθηκε σε 500μl TM. Το βακτηριακό RNA απομακρύνθηκε επώαζοντας το εναιώρημα στους 37°C για 30 min με RNase A σε τελική συγκέντρωση 1μg/ml. Στη συνέχεια προστέθηκαν διαλύματα EDTA pH 8.0, SDS και πρωτεϊνάσης K σε τελικές συγκεντρώσεις 20mM, 0.5% και 50μg/ml αντίστοιχα, και ακολούθησε επώαση στους 55 °C για 1 hr. Το DNA των βακτηριοφάγων εκχυλίζονταν διαδοχικά με φαινόλη, φαινόλη-χλωροφόρμιο (1:1) και χλωροφόρμιο και κατακρημνίζονταν με 0.1 όγκους 3M CH₃COONa pH 7.0 και 2 όγκους απόλυτης αιθανόλης, μετά από επώαση για 30 min στον πάγο και φυγοκέντρηση για 20 min στις 13.000 rpm στους 4°C. Τέλος, διαλυτοποιούνταν σε 100μl διαλύματος TE (10mM Tris-Cl pH 8.0, 1mM EDTA pH 8.0).

β. Απομόνωση πλασμιδιακού DNA από βακτήρια

Για την απομόνωση πλασμιδιακού DNA σε μικρή (5-10μg) ή μεγαλύτερη (100-500μg) κλίμακα, χρησιμοποιήθηκε η μέθοδος της αλκαλικής λύσης [Sambrook *et al*, 1989], με ορισμένες τροποποιήσεις. Συγκεκριμένα, για την επαναιώρηση της βακτηριακής πάστας χρησιμοποιήθηκε το διάλυμα 1 (15mM Tris-Cl pH 8.0, 10mM EDTA pH 8.0), για τη λύση των κυττάρων και την κατακρήμνιση των πρωτεϊνών χρησιμοποιήθηκαν τα διαλύματα 2 (0.2N NaOH, 1% SDS) και 3 (3M CH₃COOK pH 5.5), σε συνθήκες και όγκους διαλυμάτων που περιγράφονται αναλυτικά στην ίδια αναφορά. Στις περιπτώσεις που ήταν αναγκαία η απομόνωση πλασμιδιακού DNA σε μεγαλύτερη κλίμακα (100-500μg), στο διάλυμα 1 προσθέτονταν επιπλέον λυσοζύμη σε τελική συγκέντρωση 1mg/ml και ακολουθούσε επώαση για 20 min στον πάγο, ενώ μετά την προσθήκη και του διαλύματος 3, το υπερκείμενο μετά τη φυγοκέντρηση διάλυμα νουκλεϊκών οξέων διαχωρίζονταν από τη στερεή φάση με διήθηση μέσα από υαλοβάμβακα. Ακολουθούσε κατακρήμνιση με 0.7 όγκους ισοπροπανόλης, επαναδιάλυση στον επιθυμητό όγκο διαλύματος TE (10mM Tris-Cl, 1mM EDTA, pH 8.0), διαδοχικές εκχυλίσεις με μίγμα φαινόλης-χλωροφορμίου (1:1) και χλωροφορμίου, κατακρήμνιση με διπλάσιο όγκο παγωμένης απόλυτης αιθανόλης και 0.1 όγκους 3M CH₃COONa pH 5.2, επαναδιάλυση σε TE και επώαση για 30 min στους 37°C παρουσία 10μg ριβονουκλεάσης A (RNase A), τελικής συγκέντρωσης 100ng/μl. Μετά από 1-2 επιπλέον εκχυλίσεις με μίγμα φαινόλης-χλωροφορμίου (1:1) και κατακρήμνιση, το DNA επαναδιαλύονταν τελικά σε TE.

Για την παρασκευή πλασμιδιακού DNA υψηλότερης καθαρότητας, π.χ. για χρήση σε αντιδράσεις προσδιορισμού νουκλεοτιδικής αλληλουχίας ή σε πειράματα διαμόλυνσης ευκαρυωτικών κυττάρων, ακολουθήθηκε η

παραπάνω διαδικασία ξεκινώντας από βακτηριακή καλλιέργεια όγκου 0.8-1 lt και εφαρμόστηκε η διαδικασία απομόνωσης με προσθήκη ίσου όγκου 2x ρυθμιστικού διαλύματος (0.04% SDS, 200mM NaCl, 25mM Tris-Cl pH 7.5, 4mM EDTA pH 7.5) και πρωτεΐνάσης K σε τελική συγκέντρωση 100ng/μl, ενώ ακολούθησε επώαση για 1 hr στους 55°C. Ακολουθούσαν 1-2 εκχυλίσσεις με φαινόλη-χλωροφόρμιο (1:1) και χλωροφόρμιο, κατακρήμνιση με 2 όγκους απόλυτης αιθανόλης (χωρίς προσθήκη άλατος), 'ψάρεμα' του νήματος DNA και επαναδιάλυσή του στον επιθυμητό όγκο διαλύματος TE.

γ. Απομόνωση μίγματος γονιδιωματικού και πλασμιδιακού DNA από σακχαρομύκητα

Ολονύκτια καλλιέργεια όγκου 2ml κυττάρων σακχαρομύκητα στελέχους rJ69-4a, μετασχηματισμένων με πλασμίδιο, φυγοκεντρήθηκε στις 10.000 rpm για 3 min, αφαιρέθηκε το υπερκείμενο και το ίζημα των κυττάρων ξεπλύθηκε μία φορά με 500ml αποιονισμένου H₂O, επαναιωρήθηκε σε 200μl διαλύματος λύσης (2% Triton X-100, 1% SDS, 100mM NaCl, 10mM Tris-Cl pH 8.0, 1mM EDTA pH 8.0) και στη συνέχεια προστέθηκαν 200μl σφαιρίδια γυαλιού, 200μl διαλύματος φαινόλης-χλωροφορμίου (1:1) και ακολούθησε ισχυρή ανάδευση (*vortex*) για 5 min. Τέλος, προστέθηκαν 200μl αποιονισμένου H₂O και το δείγμα φυγοκεντρήθηκε (12.000 rpm) για 5 min. Τα νουκλεϊκά οξέα κατακρημνίστηκαν από το υπερκείμενο με 0.1 όγκους 3M CH₃COONa pH 5.2, 2.5 όγκους παγωμένης απόλυτης αιθανόλης, παραμονή στους -80°C για 15 min και φυγοκέντρωση στις 12.000 στροφές/min για 30 min στους 4°C. Το ίζημα, μετά από ένα πλύσιμο με 70% αιθανόλη, επαναδιαλύθηκε σε 20μl αποιονισμένου H₂O. Το μίγμα νουκλεϊκών οξέων που προέκυψε με αυτή τη μέθοδο, περιείχε κυρίως γονιδιωματικό DNA σακχαρομύκητα και πολύ περιορισμένη ποσότητα πλασμιδιακού DNA, η οποία δεν ήταν δυνατόν να ανιχνευθεί με ηλεκτροφόρηση. Ακολούθησε διαδικασία επιλογής του πλασμιδιακού DNA (*plasmid rescue*), μετασχηματίζοντας βακτήρια *E. coli* DH5a με 1-5μl του παραπάνω μίγματος και αναπτύσσοντάς τα σε κατάλληλο θρεπτικό μέσο, παρουσία του κατάλληλου αντιβιοτικού, επιλεκτικού για το συγκεκριμένο πλασμίδιο.

δ. Απομόνωση γονιδιωματικού DNA από εγκέφαλο ποντικού

Ο ιστός αφαιρέθηκε από το ζώο, ξεπλύθηκε από το αίμα σε διάλυμα 1x PBS (40mM Na₂HPO₄, 5mM NaH₂PO₄, 150mM NaCl, pH 7.4) και μεταφέρθηκε αμέσως σε υγρό άζωτο. Ακολούθησε κονιορτοποίηση του ιστού σε περιβάλλον υγρού αζώτου. Η παγωμένη σκόνη μεταφέρθηκε σε 2 όγκους διαλύματος TSE (150mM NaCl, 100mM EDTA pH 8.0, 20mM Tris-Cl pH 8.0) και αρχικά ομογενοποιήθηκε μηχανικά σε ομογενοποιητή Dounce. Στη συνέχεια, χρησιμοποιήθηκαν διαδοχικά σύριγγες 21g (5ml) και 26g (1ml) για την πλήρη ομογενοποίηση του ιστού. Σε περίπτωση που το εναιώρημα ήταν πολύ ιξώδες ώστε να μη διέρχεται εύκολα από τις βελόνες των συριγγών, ο όγκος του αυξάνονταν με προσθήκη επιπλέον διαλύματος TSE. Στη συνέχεια, προσθέτονταν διαλύματα SDS και πρωτεΐνάσης K σε τελικές συγκεντρώσεις 1% και 100μg/ml αντίστοιχα, και το δείγμα επώαζονταν για 6 hrs στους 65°C. Ακολουθούσαν διαδοχικές εκχυλίσσεις με ίσους όγκους φαινόλης, φαινόλης – χλωροφορμίου (1:1) και χλωροφορμίου – ισοαμλικής αλκοόλης (24:1). Το γονιδιωματικό DNA κατακρημνίζονταν υπό μορφή νήματος με την προσθήκη 0.1 όγκων 3M CH₃COONa pH 5.2 και 2 όγκων παγωμένης απόλυτης αιθανόλης στην υδατική φάση, απόπου συλλέγονταν με περιτύλιξη σε πιπέττα Pasteur και ξεπλένονταν με 70% αιθανόλη. Στη συνέχεια, αφήνονταν να στεγνώσει στον αέρα και επαναδιαλυτοποιούνταν σε 5ml TE (10mM Tris-Cl, 1mM EDTA, pH 8.0).

ε. Τεμαχισμός μορίων DNA με ενδονουκλεάσες περιορισμού

Για τον τεμαχισμό μορίων DNA χρησιμοποιήθηκαν περιοριστικές ενδονουκλεάσες, σύμφωνα με τις οδηγίες των κατασκευαστριών εταιριών (Minotech και New England Biolabs). Συγκεκριμένα, χρησιμοποιήθηκε αναλογία 1μg DNA με 5 μονάδες ενζύμου σε ρυθμιστικό διάλυμα όπου το ένζυμο παρουσίαζε την υψηλότερη ενεργότητα, σε συγκέντρωση DNA 100ng/μl για περίπου 3-12 hrs (για τεμαχισμό μορίων DNA πλασμιδίων ή λ βακτηριοφάγων που προορίζονταν για κλωνοποίηση ή για τεμαχισμό μορίων γονιδιωματικού DNA ποντικού ή DNA λ βακτηριοφάγου ή πλασμιδιακού κλώνου PAC, που προορίζονταν για ηλεκτροφόρηση και ανάλυση κατά Southern). Στις τελευταίες περιπτώσεις, στην αντίδραση τεμαχισμού προσθέτονταν σπερμιδίνη σε τελική συγκέντρωση 1mM.

στ. Ανάλυση γραμμικών μορίων DNA με ηλεκτροφόρηση σε πήκτωμα αγαρόζης

Τα μόρια DNA, λόγω του αρνητικού φορτίου τους, μεταναστεύουν σε πήκτωμα αγαρόζης υπό την επίδραση ηλεκτρικού πεδίου. Για τα γραμμικά μόρια DNA μεγέθους μέχρι 20Kb, η ταχύτητα μετανάστευσης είναι ανάλογη του μεγέθους τους. Οι συνθήκες ηλεκτροφόρησης (πυκνότητα και μέγεθος του πηκτώματος, εφαρμοζόμενη τάση πεδίου, διάρκεια ηλεκτροφόρησης) επιλέγονταν σε κάθε περίπτωση με βάση τα μεγέθη των μορίων που επρόκειτο να διαχωριστούν [Sambrook *et al*, 1989]. Το διάλυμα ηλεκτροφόρησης που χρησιμοποιήθηκε ήταν το 1x TAE (40mM Tris-acetate, 1mM EDTA pH 8.0), με προσθήκη βρωμιούχου αιθιδίου (125μg/lit διαλύματος) για τη χρώση των μορίων DNA κατά την ηλεκτροφόρηση και την παρατήρησή τους σε υπεριώδη φωτισμό.

Ηλεκτροφόρηση παλλόμενου ηλεκτρικού πεδίου (PFGE)

Για την ανάλυση μορίων DNA μεγαλύτερων από 20Kb, όπως ήταν τα προϊόντα του τεμαχισμού κλώνων PAC ανθρώπινου γονιδιωματικού DNA με το περιοριστικό ένζυμο BamHI, εφαρμόστηκε η ηλεκτροφόρηση παλλόμενου ηλεκτρικού πεδίου (*Pulse Field Gel Electrophoresis*), η οποία στηρίζεται στην εφαρμογή ηλεκτρικού πεδίου διαδοχικά και επαναλαμβανόμενα σε δύο εναλλακτικές κατευθύνσεις κατά την διάρκεια της ηλεκτροφόρησης, πράγμα που ευνοεί τον διαχωρισμό μορίων μεγαλύτερων από 20Kb [Schwartz & Cantor, 1984]. Χρησιμοποιήθηκε η πειραματική διάταξη του συστήματος ηλεκτροφόρησης CHEF-DRII (BioRad), με βάση το σχετικό πρωτόκολλο της κατασκευάστριας εταιρίας και σε συνθήκες που ρυθμιζόνταν ώστε να επιτευχθεί καλύτερος διαχωρισμός στα επιθυμητά μεγέθη. Αναλυτικότερα, η ηλεκτροφόρηση πραγματοποιήθηκε σε πήκτωμα αγαρόζης 1% και διάλυμα 0.5x TBE (45mM Tris-Cl, 45mM βορικό οξύ, 1mM EDTA, pH 8.0), υπό την επίδραση ηλεκτρικού πεδίου τάσεως 200V, για 12 hrs στους 4°C, με συνεχή ανακύκλωση του διαλύματος ηλεκτροφόρησης. Ο χρόνος εναλλαγής του ηλεκτρικού πεδίου (*switch time*) ρυθμιζόνταν να κυμαίνεται μεταξύ 0.3sec (αρχική τιμή) και 1.3sec (τελική τιμή). Μετά την ηλεκτροφόρηση, ακολουθούσε χρώση του πηκτώματος σε υδατικό διάλυμα βρωμιούχου αιθιδίου (125μg/lit διαλύματος) για 30min με ήπια ανακίνηση, 2-3 πλυσίματα με απιονισμένο νερό για 20min και παρατήρηση σε υπεριώδη φωτισμό.

ζ. Ανάλυση DNA κατά Southern

Μετά τον ηλεκτροφορητικό διαχωρισμό σε πήκτωμα αγαρόζης, γραμμικά μόρια DNA μπορούν να μεταφερθούν και να μονιμοποιηθούν σε μεμβράνες νάυλον με τη βοήθεια τριχοειδών φαινομένων [Southern, 1975], ώστε να ακολουθήσουν πειράματα υβριδοποίησης με ραδιοσημασμένους ανιχνευτές και αυτοραδιογραφία. Αναλυτικότερα, ο ρυθμός μεταφοράς διαφορετικών τεμαχίων DNA εξαρτάται από το μέγεθός τους. Μικρά μόρια DNA (μικρότερα από 2Kb) μεταφέρονται σε 1-3 hrs, ενώ για τα μεγαλύτερα ο χρόνος μεταφοράς αυξάνεται σε περισσότερες από 10 hrs, και στην περίπτωση μορίων μεγαλύτερων από 10-15Kb, ένα πολύ μικρό ποσοστό μεταφέρεται στη μεμβράνη. Για την αντιμετώπιση αυτού του προβλήματος, τα μόρια DNA αποδιάσσονται πριν τη μεταφορά επώζοντας το πήκτωμα αγαρόζης σε αλκαλικό διάλυμα. Επιπλέον, για τα μεγαλύτερα από 10kb θραύσματα DNA, πραγματοποιείται πριν την αποδιάταξη τεμαχισμός με μερική αποπουρίνωση σε ασθενές όξινο περιβάλλον [Sambrook *et al*, 1989].

Σύμφωνα με τα παραπάνω, διαχωρισμένα θραύσματα μορίων DNA φωτογραφίζονται σε υπεριώδες φωτισμό δίπλα σε υποδεκάμετρο και στη συνέχεια αποδιάσσονται με επώαση του πηκτώματος σε διάλυμα αποδιάταξης (1.5M NaCl και 0.5M NaOH) και ήπια ανακίνηση δύο φορές για 40 min. Ακολουθεί εξουδετέρωση σε διάλυμα 1M Tris-Cl pH 8.0 και 1.5M NaCl, στις ίδιες συνθήκες. Στις περιπτώσεις που το μέγεθος των μελετούμενων θραυσμάτων DNA είναι μεγαλύτερο από 10Kb πραγματοποιείται, πριν την αποδιάταξη, αποπουρίνωση σε διάλυμα 0.25N HCl για 5 min. Τα νουκλεϊκά οξέα μεταφέρονται σε νάυλον μεμβράνες (HybondTM-NX, Amersham) χρησιμοποιώντας διάλυμα μεταφοράς 2x SSC (0.3M NaCl, 30mM κιτρικό νάτριο). Τα δεσμευμένα στη μεμβράνη μόρια DNA μονιμοποιούνται με έκθεση σε υπεριώδη ακτινοβολία (*UV cross-linking*) για 20 sec και με θέρμανση της μεμβράνης στους 80°C για 2 hrs.

η. Αλυσιδωτή αντίδραση πολυμεράσης (Polymerase Chain Reaction, PCR)

Η μέθοδος εφαρμόστηκε τόσο για την ανάκτηση τεμαχίων DNA με σκοπό την κλωνοποίησή τους, όσο και για τη ραδιοσήμανση τεμαχίων DNA. Για κάθε αντίδραση, χρησιμοποιήθηκαν 50-100ng γονιδιωματικού DNA ποντικού, 5-10ng κλωνοποιημένου DNA σε πλασμίδιο ή σε βακτηριοφάγο λ, ή 50-100ng DNA βιβλιοθήκης πλασμιδικών κλώνων cDNA ως μήτρα. Οι εκκινητές PCR είναι συνθετικά ολιγονουκλεοτίδια και χρησιμοποιήθηκαν σε συγκεντρώσεις 100ng ανά αντίδραση PCR. Ο όγκος των αντιδράσεων ήταν 15 – 50 μl και τυπικά περιελάμβανε: μίγμα dNTPs (250μM από κάθε νουκλεοτίδιο), 1x ρυθμιστικό διάλυμα *Taq* DNA πολυμεράσης (10mM Tris-Cl pH 8.5, 50mM KCl, 0.1% Triton X-100, 1.5mM MgCl₂), 1 ενζυμική μονάδα *Taq* DNA πολυμεράσης (3u/μl, Minotech) και 100μg/ml BSA (*bovine serum albumin, fraction V, SIGMA*). Οι αντιδράσεις πραγματοποιήθηκαν σε αυτόματο θερμοανακυκλωτή (Biometra, Perkin Elmer, Techne), που τυπικά ρυθμίζεται να εκτελέσει ένα αρχικό στάδιο αποδιάταξης του DNA στους 95°C για 5 min και ένα στάδιο θερμικής ανακύκλωσης που επαναλαμβάνεται 35 φορές και περιλαμβάνει: i) αποδιάταξη του DNA (95°C για 45sec), ii) υβριδοποίηση των εκκινητών στις συμπληρωματικές αλληλουχίες στόχους για 45sec, σε θερμοκρασία που υπολογίζεται με βάση το ποσοστό των βάσεων GC στην αλληλουχία των εκκινητών και εμπειρικά προκύπτει από τη σχέση: $\theta = T_m - (3-5)^\circ\text{C}$, όπου T_m η θερμοκρασία στην οποία το 50% των μορίων του εκκινητή έχουν υβριδοποιηθεί στις αλληλουχίες-στόχους [$T_m = 2 \times (\text{ποσοστό βάσεων AT}\%) + 4 \times (\text{ποσοστό βάσεων GC}\%)$], και iii) επέκταση των εκκινητών και σύνθεση της θυγατρικής αλυσίδας DNA σε θερμοκρασία όπου το ένζυμο παρουσιάζει την υψηλότερη ενεργότητα (72°C για την *Taq* DNA πολυμεράση της εταιρίας Minotech, 68-72°C για ένζυμα άλλων εταιριών) για χρόνο που κυμαίνεται από 30sec για προϊόντα PCR μεγέθους μικρότερου των 200bp έως και 90sec για προϊόντα PCR μεγέθους 1-2Kb. Ακολουθεί ένα στάδιο τελικής επέκτασης για 2-5 min. Στις περιπτώσεις που οι εκκινητές PCR φέρουν θέσεις αναγνώρισης από περιοριστικές ενδονουκλεάσες, πραγματοποιούνται αρχικά 5 κύκλοι σε θερμοκρασία υβριδοποίησης των εκκινητών ελαττωμένη κατά 2-5°C. Επιπλέον, για την ελαχιστοποίηση της πιθανότητας λάθους στην αλληλουχία του DNA, η οποία για τις συμβατικές *Taq* DNA πολυμεράσες είναι περίπου 1/1000nt, σε περιπτώσεις που το προϊόν της αντίδρασης επρόκειτο να κλωνοποιηθεί σε φορέα έκφρασης χρησιμοποιήθηκαν βελτιωμένα ένζυμα με ικανότητα επιδιόρθωσης (*proofreading*) των σφαλμάτων (Vent_r DNA Polymerase της εταιρίας New England Biolabs και Expand Long Template PCR System από την εταιρία Roche) και πραγματοποιήθηκε μικρότερος αριθμός κύκλων (25-30).

θ. Προσδιορισμός νουκλεοτιδικής αλληλουχίας μορίων DNA

Για τον σκοπό αυτό ακολουθήθηκαν διάφορες προσεγγίσεις που αποτελούν τροποποιήσεις της αρχικής μεθόδου, η οποία στηρίζεται στον τερματισμό της επιμήκυνσης της αλυσίδας του DNA μέσω της ενσωμάτωσης τριφωσφορικών διδεοξυνουκλεοτιδίων (*dideoxy chain termination, Sanger et al, 1977*). Χρησιμοποιήθηκαν ενθέματα υπερελεγκμένων πλασμιδίων, γραμμοποιημένα μόρια DNA ή προϊόντα PCR.

Προσδιορισμός νουκλεοτιδικής αλληλουχίας με PCR (cycle-sequencing)

Εφαρμόστηκε η μέθοδος της PCR, όπου ένας εκκινητής χρησιμοποιείται για την επιμήκυνση της νεοσυντιθέμενης αλυσίδας του DNA από την *Taq* DNA πολυμεράση παρουσία τριφωσφορικών διδεοξυνουκλεοτιδίων. Συγκεκριμένα, στο μίγμα της αντίδρασης προσθέτονται: ποσότητα DNA ανάλογη του μεγέθους του (200ng ανά Kb), 25ng εκκινητή, 5 ενζυμικές μονάδες τροποποιημένης *Taq* DNA πολυμεράσης (Sequencing Grade *Taq* DNA Polymerase, Promega) και 5μl 5x ρυθμιστικού διαλύματος (250mM Tris-Cl pH 9.0, 10mM MgCl₂), σε τελικό όγκο 17μl. Στη συνέχεια, 4μl αυτού του μίγματος αναμιγνύονται ξεχωριστά με 2μl από κάθε διάλυμα δεοξυνουκλεοτιδίων (dNTPs) και διδεοξυνουκλεοτιδίων (ddNTPs) τερματισμού. Τα διαλύματα τερματισμού A, G, C και T που χρησιμοποιούνται έχουν 20μM από κάθε dNTP και, επιπλέον, 350μM ddATP στο διάλυμα A, 30μM ddGTP στο διάλυμα G, 200μM ddCTP στο διάλυμα C και 600μM ddTTP στο διάλυμα T. Ακολουθεί αποδιάταξη στους 95°C (5 min) και 30 κύκλοι PCR σε αυτόματο θερμοανακυκλωτή. Η σήμανση των προϊόντων της αντίδρασης πραγματοποιείται είτε προσθέτοντας στο μίγμα 5μCi ραδιοϊσοτόπου α[³²P]-dATP ή α[³²P]-dCTP, είτε χρησιμοποιώντας εκκινητή που έχει προηγουμένως ραδιοσημανθεί στο 5'

άκρο του με γ [^{32}P]-dATP. Τα προϊόντα των αντιδράσεων αναμιγνύονται με 4μl διαλύματος 95% φορμαμίδης, 20mM EDTA pH 8.0, 0.05% *bromophenol blue* και 0.05% *xylene cyanol*, αποδιατάσσονται με θέρμανση στους 80°C για 2 min και ηλεκτροφορούνται σε αποδιατακτικό πήκτωμα ουρίας-πολυακρυλαμίδης [6% ακρυλαμίδη:δισ ακρυλαμίδη (38:2), 8M ουρία, 1x long run TBE] υπό την επίδραση ηλεκτρικού πεδίου σταθερής ισχύος 80Watt και χρησιμοποιώντας διάλυμα ηλεκτροφόρησης 1x long run TBE (133.8mM Tris-Cl, 44.5mM βορικό οξύ, 2.5mM EDTA, pH 8.3-8.7 στους 50°C). Το πήκτωμα μετά την ηλεκτροφόρηση μεταφέρεται σε χαρτί διήθησης (Whatman 3mm), στεγνώνει με θέρμανση σε κενό αέρος (BioRad Gel Dryer, Model 586) και ακολουθεί αυτοραδιογραφία και ανάγνωση της νουκλεοτιδικής σύστασης των μορίων DNA.

Αυτοματοποιημένη διαδικασία προσδιορισμού νουκλεοτιδικής αλληλουχίας με PCR

Ο προσδιορισμός της νουκλεοτιδικής αλληλουχίας πραγματοποιείται με PCR, με τη διαφορά ότι τα προϊόντα δε σημαίνονται ραδιενεργά, αλλά με φθορίζουσες ουσίες. Για το σκοπό αυτό χρησιμοποιούνται εκκινητές σημασμένοι με τη φθορίζουσα ουσία IRD-800 (phosphoramidite, $\text{C}_{59}\text{H}_{75}\text{N}_4\text{O}_6\text{PS}$, LI-COR 4000-33), οι οποίοι παρασκευάζονται από την εταιρία MWG Biotech. Ακολουθεί ηλεκτροφόρηση της αντίδρασης, ταυτόχρονη σάρωση από ακτίνα *laser* και παράλληλη ανάλυση της νουκλεοτιδικής αλληλουχίας με κατάλληλο υπολογιστικό πρόγραμμα. Οι αντιδράσεις πραγματοποιούνται με τη χρήση του SequiTherm EXCEL™ II Long-Read™ DNA Sequencing Kit-LC (Epicentre Technologies, Cat. No SE7701LC), ηλεκτροφορούνται σε αυτόματο αναλυτή τύπου GENE READIR 4200 (LI-COR) και αναλύονται με τη χρήση του προγράμματος BaseImageIR,v4.1, σύμφωνα με τις οδηγίες της κατασκευάστριας εταιρίας LI-COR.

14. Μέθοδοι ραδιοσήμανσης μορίων DNA

Ανιχνευτικά μόρια DNA με μεγέθη που κυμαίνονται από 0.5 έως 1.5Kb ραδιοσημαίνονται με τη μέθοδο της επέκτασης τυχαιών εκκινητών (*random priming*), ενώ μικρότερα μόρια σημαίνονται με τη χρήση της αντίδρασης PCR. Τέλος, ολιγονουκλεοτίδια (18-25nt) ραδιοσημαίνονται στο 5' άκρο τους για να χρησιμοποιηθούν ως ανιχνευτικά μόρια σε υβριδοποιήσεις ή ως εκκινητές σε αντιδράσεις προσδιορισμού της νουκλεοτιδικής αλληλουχίας μορίων DNA. Οι αντιδράσεις ραδιοσήμανσης που πραγματοποιήθηκαν περιγράφονται συνοπτικά παρακάτω:

α. Επέκταση τυχαιών εκκινητών

Η μέθοδος στηρίζεται στην ενεργότητα 5'→3' πολυμεράσης DNA του κλάσματος Klenow της DNA πολυμεράσης I, με την οποία ραδιοσημασμένα νουκλεοτίδια ενσωματώνονται στη νεοσυντιθέμενη αλυσίδα DNA, χρησιμοποιώντας ως εκκινητές έναν συνδυασμό τυχαιών εξανουκλεοτιδίων (N_6) που επαναδιατάσσεται στο DNA-μήτρα. Συγκεκριμένα, σε μια τυπική αντίδραση σήμανσης χρησιμοποιούνται 100ng γραμμικού τεμαχίου DNA (μεγέθους 0.5-1.5kb) σε όγκο 7.5μl, 5 μονάδες ενζύμου Klenow, 11.5μl διαλύματος 2x LS, 20μCi από κάθε ραδιενεργό ισότοπο [α - ^{32}P]dATP και [α - ^{32}P]dCTP (10μCi/μl, Amersham) και 10μg BSA σε τελικό όγκο αντίδρασης 25μl. Το διάλυμα 2x LS προκύπτει αναμιγνύοντας τρία διαλύματα : i) ρυθμιστικό διάλυμα 1M Hepes pH 6.6, ii) διάλυμα DTM που περιέχει 0.1mM dGTP και 0.1mM dTTP σε TM (250mM Tris-Cl pH 8.0, 5mM MgCl_2 , 50mM β-μερκαπτοαιθανόλη) και iii) διάλυμα OL (1mM Tris-Cl pH 7.5, 1mM EDTA pH 7.5, 90 units/ml τυχαιών εξανουκλεοτιδίων), σε αναλογία 25 (Hepes) : 25 (DTM) : 7 (OL). Αρχικά το DNA αποδιατάσσεται με βρασμό για 5 min και μεταφέρεται αμέσως σε πάγο όπου προσθέτονται τα υπόλοιπα αντιδραστήρια και η αντίδραση επώάζεται για 2-3 hrs στους 37°C ή για 12-14 hrs στους 25°C. Η ειδική ενεργότητα του ραδιοσημασμένου μορίου DNA που προκύπτει είναι 5×10^8 - 10^9 cpm/μg DNA.

β. Σήμανση 5' άκρου

Η μέθοδος εφαρμόζεται για τη σήμανση ολιγονουκλεοτιδίων (18-25nt), ή προϊόντων PCR μικρού μεγέθους (100-500bp), χρησιμοποιώντας την ενζυμική δράση της πολυνουκλεοτιδικής κινάσης του βακτηριοφάγου T4 η οποία, στις συνθήκες της αντίδρασης, καταλύει τη μεταφορά μιας φωσφορικής ομάδας από τη θέση γ της τριφωσφορικής αδενοσίνης (ATP) στο 5' άκρο πολυ- ή ολιγονουκλεοτιδίων (δίκλωνου και μονόκλωνου DNA και RNA) των οποίων έχει αφαιρεθεί η 5' φωσφορική ομάδα ενζυμικά (με δράση φωσφατάσης ή κινάσης παρουσία περίσσειας ADP) ή κατά την *in vitro* σύνθεσή τους, όπως συμβαίνει για τα συνθετικά ολιγονουκλεοτιδία [Sambrook *et al*, 1989]. Συγκεκριμένα, για τη σήμανση 100 ng συνθετικού ολιγονουκλεοτιδίου ή προϊόντος PCR, χρησιμοποιούνται 5 μονάδες ενζύμου κινάσης (New England Biolabs, 5u/μl), 1x ρυθμιστικό διάλυμα κινάσης (70 mM Tris-Cl pH 7.6, 10 mM MgCl₂, 5 mM DTT), 20μCi ραδιοϊσότοπο [γ -³²P]ATP (10μCi/μl, Amersham) σε συνολικό όγκο 20μl και η αντίδραση επωάζεται στους 37°C για 45 min. Η τυπική απόδοση της μεθόδου είναι 5x10⁷-10⁸ cpm/μg ολιγονουκλεοτιδίου ή προϊόντος PCR.

γ. Σήμανση με PCR

Η αλυσιδωτή αντίδραση πολυμεράσης ως μέθοδος σήμανσης, επιλέγεται σε περιπτώσεις που το μέγεθος των τεμαχίων DNA που πρόκειται να ραδιοσημανθούν είναι μικρότερο του 0.5Kb (ιδανικά 100-150bp). Αρχικά, πραγματοποιείται PCR (χωρίς ραδιοσήμανση) της οποίας το προϊόν (ποσότητα DNA 30-200ng) ηλεκτροφορείται, εξάγεται από το πήκτωμα αγαρόζης και εκλούεται σε 200μl απιονισμένου-αποστειρωμένου H₂O, παραμένοντας στους 4°C για 6-12 hrs. Περίπου 2μl από το διάλυμα που προκύπτει μετά την έκλυση χρησιμοποιούνται ως μήτρα σε μια δεύτερη αντίδραση PCR στην οποία προσθέτονται 10μCi (~20pM) ραδιοϊσότοπου [α -³²P]dATP (Amersham, 10μCi/μl, 370MBq/ml, ~3000Ci/mmol) και 12.5μM από κάθε νουκλεοτίδιο dCTP, dGTP και dTTP. Λόγω της περιορισμένης ποσότητας ραδιοσημασμένου νουκλεοτιδίου [α -³²P]dATP που χρησιμοποιείται σε αυτή την αντίδραση, η ποσότητα του τελικού προϊόντος είναι θεωρητικά της τάξεως του 1ng, όμως η ειδική ενεργότητα του ραδιοσημασμένου μορίου που προκύπτει είναι της τάξεως του 10⁹ cpm/μg προϊόντος PCR. Οι αντιδράσεις σήμανσης με PCR πραγματοποιούνται με τις τυπικές συγκεντρώσεις ενζύμου και ρυθμιστικού διαλύματος, 35 κύκλους PCR και συνθήκες που ρυθμίζονται με βάση την αλληλουχία των εκκινητών.

δ. Μέτρηση της ειδικής ενεργότητας ραδιοσημασμένου μορίου DNA

Η ειδική ενεργότητα εκφράζει το ποσοστό των ραδιενεργών φωσφορικών ομάδων ³²P που έχουν ενσωματωθεί στο μόριο DNA που ραδιοσημαίνεται με τις παραπάνω μεθόδους. Για τη μέτρησή της, μια ποσότητα αντίστοιχη περίπου του ενός εκατοστού των ραδιοσημασμένων προϊόντων των παραπάνω αντιδράσεων κατακρημνίζεται παρουσία παγωμένου 10% τριχλωροοξικού οξέος (TCA) και τα μη ενσωματωμένα ραδιενεργά νουκλεοτίδια απομακρύνονται με παγωμένη ακετόνη στις συνθήκες που περιγράφονται αναλυτικά από τους Sambrook *et al*. Στη συνέχεια, τα δείγματα μετρώνται σε σπινθηρομετρητή (Beckman LS1701).

15. Ανίχνευση τεμαχίων DNA με υβριδοποίηση

Αποδιαταγμένα μόρια DNA που έχουν μεταφερθεί και μονιμοποιηθεί σε νάυλον μεμβράνη μπορούν, σε κατάλληλες συνθήκες, να υβριδοποιηθούν με συμπληρωματικά ραδιοσημασμένα ανιχνευτικά μόρια DNA και στη συνέχεια να ανιχνευθούν με αυτοραδιογραφία. Η σύσταση του διαλύματος υβριδοποίησης, η διάρκεια και η θερμοκρασία της επώασης είναι συνθήκες που ρυθμίζονται σύμφωνα με τα σχετικά πρωτόκολλα [Sambrook *et al*, 1989]. Τα πειράματα υβριδοποίησης που πραγματοποιήθηκαν περιγράφονται συνοπτικά παρακάτω:

α. Υβριδοποίηση σε γονιδιωματικό DNA ποντικού

Μετά τον τεμαχισμό με περιοριστικά ένζυμα και την ανάλυση κατά Southern γονιδιωματικού DNA ποντικού στελέχους C57BL, η μεμβράνη προϋβριδοποιείται με συνεχή ανακίνηση για 6hrs στους 65°C σε 10ml διαλύματος υβριδοποίησης (100mM NaH₂PO₄, 50mM Na₄P₂O₇, 10mM EDTA, 7% SDS, 1% BSA, pH 7.0-7.5), παρουσία αποδιαταγμένου και μηχανικά τεμαχισμένου DNA σπέρματος σολομού ή ρέγγας (*carrier* DNA), σε συγκέντρωση 100 µg/ml. Ακολουθεί αντικατάσταση του διαλύματος με 10ml φρέσκου διαλύματος υβριδοποίησης, στο οποίο προσθέτονται *carrier* DNA ίδιας συγκέντρωσης και περίπου 10⁷cpm αποδιαταγμένου με βρασμό ραδιοσημασμένου ανιχνευτή. Ακολουθεί επώαση με συνεχή ανακίνηση στους 65°C για 12-16 hrs. Στη συνέχεια, η περίσσεια της ραδιενέργειας απομακρύνεται με πλύσιμο της μεμβράνης σε διάλυμα 2xSSC (0.3M NaCl, 30mM κιτρικό νάτριο) αρχικά στους 25°C και στη συνέχεια στους 65°C για 20-30 min και 3-4 επαναλήψεις, ελέγχοντας τις κρούσεις που παρέμεναν στη μεμβράνη με μετρητή Geiger. Ακολουθεί αυτοραδιογραφία με χρόνους έκθεσης που κυμαίνονται από 12 έως και 48 hrs, ανάλογα με την ένταση του σήματος.

β. Υβριδοποίηση σε κλωνοποιημένα τεμάχια DNA

Υβριδοποιήσεις σε κλωνοποιημένα τεμάχια DNA πραγματοποιούνται α) σε μεμβράνες που φέρουν DNA βακτηρίων μετασχηματισμένων με ανασυνδυασμένα πλασμίδια, για να ελεγχθεί η παρουσία ενθέματος και β) σε μεμβράνες που φέρουν ηλεκτροφορητικά διαχωρισμένα θραύσματα ανασυνδυασμένων πλασμιδίων ή λ βακτηριοφάγων. Ακολουθήθηκε η ίδια διαδικασία με τη διαφορά ότι, λόγω της πολύ ελαττωμένης πολυπλοκότητας του μονιμοποιημένου στη μεμβράνη DNA, η προϋβριδοποίηση και η υβριδοποίηση πραγματοποιήθηκαν χωρίς προσθήκη *carrier* DNA και σε μικρότερους χρόνους. Συγκεκριμένα, η διάρκεια της προϋβριδοποίησης περιορίζεται στα 30 min, και της υβριδοποίησης στις 3-5 hrs. Ο χρόνος έκθεσης μετά την αυτοραδιογραφία ήταν συνήθως μικρότερος από 5 min.

16. Κλωνοποίηση τεμαχίων DNA σε πλασμίδια

Χρησιμοποιήθηκαν κλασσικές μέθοδοι κλωνοποίησης θραυσμάτων DNA, μετά από τον ενζυμικό τεμαχισμό μεγαλύτερων μορίων, διαχωρισμό με ηλεκτροφόρηση, απομόνωση και (μερικές φορές) κατάλληλη τροποποίηση των άκρων τους, ώστε αυτά να είναι συμβατά για σύνδεση με τα άκρα (συμπληρωματικά ή τυφλά) του γραμμοποιημένου πλασμιδιακού φορέα. Η πλασμιδιακή κατασκευή αναπτύσσεται μετασχηματίζοντας βακτηριακά κύτταρα (*E.coli* DH5a) ή, σε ορισμένες περιπτώσεις, χρησιμοποιήθηκε για το μετασχηματισμό κυττάρων σακχαρομύκητα του στελέχους PJ69-4a.

α. Απομόνωση θραύσματος DNA από πήκτωμα αγαρόζης

Μετά την ηλεκτροφόρηση, το πήκτωμα παρατηρείται στο υπεριώδες φως και αποκόπτεται το τεμάχιο της αγαρόζης που περιέχει το συγκεκριμένο θραύσμα DNA. Η αγαρόζη πολτοποιείται με τη βοήθεια νυστεριού και μεταφέρεται σε δοκιμαστικό σωληνάριο Eppendorf 0.5ml, το οποίο έχει τρυπηθεί στη βάση του με βελόνα σύρριγγας 26g και περιέχει υαλοβάμβακα μέχρι το 1/3 του όγκου του. Το σωληνάριο με την αγαρόζη τοποθετείται σε Eppendorf 1.5ml και φυγοκεντρείται στις 8.000 rpm για 10 min. Το διάλυμα συλλέγεται από το σωληνάριο 1.5ml και η φυγοκέντρηση επαναλαμβάνεται όσες φορές χρειάζεται. Ίχνη αγαρόζης απομακρύνονται εκχυλίζοντας το διάλυμα με χλωροφόρμιο. Το DNA κατακρημνίζεται με προσθήκη 0.1 όγκου 3M CH₃COONa pH 5.2 και 2 όγκων παγωμένης απόλυτης αιθανόλης, μετά από παραμονή στους -80°C για 15 min και φυγοκέντρηση (12.000 rpm για 30 min στους 4°C). Το ίζημα ξεπλένεται με 70% αιθανόλη και επαναδιαλύεται στον επιθυμητό όγκο TE (10mM Tris-Cl, 1mM EDTA, pH 8.0).

β. Κατασκευή τυφλών άκρων σε γραμμικά μόρια DNA με προεξέχοντα 5' άκρα

Στις περιπτώσεις που τα γραμμοποιημένα μόρια του πλασμιδιακού φορέα και του ενθέματος που επρόκειτο να κλωνοποιηθεί είχαν προεξέχοντα μη συμπληρωματικά 5' άκρα, εφαρμόστηκε η μέθοδος κατασκευής τυφλών (*blunt*)

άκρων, ώστε να είναι δυνατή η αντίδραση σύνδεσης (*ligation*). Χρησιμοποιήθηκε το κλάσμα Klenow της DNA πολυμεράσης I, και συγκεκριμένα η 5'→3' ενεργότητα πολυμερισμού (προσθήκη νουκλεοτιδίων στο υπολειπόμενο 3' άκρο, με μήτρα το προεξέχον 5' άκρο), όπως περιγράφεται από τους Sambrook *et al* (1989). Συνοπτικά, για μια τυπική αντίδραση με 1.5μg γραμμοποιημένου DNA με 5' προεξέχοντα άκρα, χρησιμοποιήθηκαν 5 μονάδες ενζύμου Klenow (New England Biolabs), 1x ρυθμιστικού διαλύματος του ενζύμου Klenow (New England Biolabs), 50μM μίγματος dNTPs (12.5μM για κάθε dNTP), σε τελικό όγκο 40μl, επώαση στους 25°C για 30 min και τερματισμό με προσθήκη 1μl διαλύματος 0.5M EDTA και θέρμανση στους 72°C για 10 min.

γ. Σύνδεση γραμμικών μορίων DNA (*ligation*)

Για τη σύνδεση γραμμικών μορίων DNA χρησιμοποιείται το ένζυμο λιγάση του DNA του βακτηριοφάγου T4 (T4 DNA ligase), το οποίο καταλύει τη δημιουργία φωσφοδιεστερικού δεσμού ανάμεσα στην υδροξυλική ομάδα του 3' άκρου και τη φωσφορική ομάδα του παρακείμενου 5' άκρου δίκλωνων μορίων DNA ή RNA. Τα άκρα των μορίων DNA που συνδέονται είναι είτε τυφλά (*blunt*) είτε προεξέχοντα και συμβατά ή συμπληρωματικά, ενώ το τελικό προϊόν της αντίδρασης είναι επανακυκλοποιημένο γραμμικό μόριο, ανασυνδυασμένο πλασμιδίο ή πολυμερές (γραμμικό ή κυκλικό) μόριο. Ο σχεδιασμός κάθε αντίδρασης εξαρτάται από το είδος των άκρων, τις σχετικές τους συγκεντρώσεις και το επιθυμητό προϊόν της αντίδρασης [Sambrook *et al*, 1989]. Οι αντιδράσεις που πραγματοποιήθηκαν περιγράφονται παρακάτω:

Επανακυκλοποίηση πλασμιδιακού φορέα

Πλασμιδιακοί φορείς ή γραμμοποιημένα ανασυνδυασμένα πλασμιδία μεγέθους 3–15Kb και σε συγκέντρωση 1–2ng/μl, επανακυκλοποιούνται χρησιμοποιώντας 6 ενζυμικές μονάδες (Weiss units) λιγάσης (T4 DNA ligase, 6 Weiss units/μl, New England Biolabs), 1x ρυθμιστικό διάλυμα λιγάσης της ίδιας εταιρίας (50mM Tris-Cl pH 7.5, 10mM MgCl₂, 10mM DTT, 1mM ATP, 25μg/ml BSA) σε τελικό όγκο αντίδρασης 20μl. Οι αντιδράσεις πραγματοποιούνται στους 16°C για 12–16 hrs.

Κατασκευή ανασυνδυασμένων πλασμιδίων

Για την αντίδραση κατασκευής ανασυνδυασμένων πλασμιδίων χρησιμοποιήθηκε η ίδια συγκέντρωση ρυθμιστικού διαλύματος και ενζύμου με τις αντιδράσεις επανακυκλοποίησης, με τη διαφορά ότι ο τελικός όγκος της αντίδρασης ήταν μικρότερος (10–15μl) και η συνολική ποσότητα DNA μεγαλύτερη (10–100 ng/μl). Η ποσότητα DNA του γραμμοποιημένου φορέα διατηρείται σταθερή στα 30ng ανά αντίδραση, ενώ η ποσότητα του ενθέματος ήταν τέτοια ώστε η αναλογία μορίων φορέα : ενθέματος να είναι μεταξύ 1:5 (για προεξέχοντα άκρα και προσανατολισμένη κλωνοποίηση) και 1:20 (για τυφλά άκρα και μη προσανατολισμένη κλωνοποίηση). Οι αντιδράσεις έγιναν στους 16°C για 12–16 hrs.

δ. Μετασχηματισμός βακτηρίων με υπερελικωμένο πλασμιδιακό DNA

Οι μέθοδοι μετασχηματισμού βακτηρίων με μόρια υπερελικωμένων πλασμιδίων βασίζονται στην παρατήρηση ότι η κατεργασία βακτηριακών κυττάρων με παγωμένο διάλυμα χλωριούχου ασβεστίου που συνοδεύεται με σύντομη θέρμανση αυξάνει σημαντικά την ικανότητά τους να προσλαμβάνουν DNA πλασμιδίου [Cohen *et al*, 1972]. Τροποποιήσεις της αρχικής μεθόδου με τη χρήση βελτιωμένων στελεχών *E. coli* και διάφορων συνδυασμών δισθενών κατιόντων ή άλλων χημικών ενώσεων (DMSO, αναγωγικά μέσα), οδήγησαν σε απόδοση μετασχηματισμού της τάξης των 10⁸–10⁹ μετασχηματισμένων αποικιών ανά μg πλασμιδίου.

Υψηλότερη όμως απόδοση (10⁹–10¹⁰ αποικίες / μg πλασμιδίου) επιτυγχάνεται μετασχηματίζοντας τα βακτήρια με ηλεκτρική εκκένωση. Η τελευταία μέθοδος [Dower *et al*, 1988], αν και προκαλεί θάνατο στο 50–75% των βακτηρίων, μπορεί να επιτύχει υψηλές συχνότητες μετασχηματισμού ρυθμίζοντας την ισχύ του ηλεκτρικού πεδίου, τη διάρκεια του

παλμού και τη συγκέντρωση του DNA. Επιπλέον, η μέθοδος μετασχηματισμού με θερμικό σοκ φαίνεται να επηρεάζεται περισσότερο από το μέγεθος του πλασμιδίου (η απόδοση μειώνεται 10-100 φορές για πλασμίδια 10-20Kb) σε σχέση με τη μέθοδο της ηλεκτρικής εκκένωσης, για την οποία πλασμίδια μεγέθους 25-130Kb έχουν αναφερθεί ότι μετασχηματίζουν βακτήρια με συγκρίσιμες αποδόσεις [Dower *et al*, 1988].

Σύμφωνα με τα παραπάνω, για τον μετασχηματισμό βακτηριακών κυττάρων με DNA πλασμιδίου, αρχικά παρασκευάζονταν βακτηριακά κύτταρα δεκτικά (*competent*) για μετασχηματισμό και χρησιμοποιούνταν και οι δύο μέθοδοι, ανάλογα με τις απαιτήσεις του εκάστοτε πειράματος.

➤ **Μετασχηματισμός βακτηρίων με τη μέθοδο της ηλεκτρικής εκκένωσης (*electroporation*)**

Η παράμετρος που λαμβάνεται υπόψη στην παρασκευή δεκτικών κυττάρων είναι η όσο το δυνατόν χαμηλότερη περιεκτικότητα του βακτηριακού εναιωρήματος σε άλατα, τα οποία μπορούν να αυξήσουν την ιοντική ισχύ προκαλώντας αυξημένο κυτταρικό θάνατο κατά την ηλεκτρική εκκένωση.

*i) Παρασκευή δεκτικών (*electrocompetent*) βακτηρίων*

Καλλιέργεια κυττάρων *Escherichia coli* κατάλληλου στελέχους (DH5a, KC8) αναπτύχθηκε σε 50ml του θρεπτικού μέσου TYM, σχετικά φτωχού σε άλατα, με ολονύκτια επώαση στους 37°C και ανακίνηση στις 200 στροφές / min. Στη συνέχεια, 500μl μεταφέρονται σε 500ml LB και η καλλιέργεια συνεχίζεται σε κωνική φιάλη όγκου 2 λίτρων στις ίδιες συνθήκες μέχρι να φτάσει σε πυκνότητα αντίστοιχη της οπτικής απορρόφησης $OD_{600}=0.65-0.70$, ακολουθεί επώαση σε πάγο για 30 min και φυγοκέντρηση στις 3000 στροφές/ min για 15 min στους 4°C. Το ίζημα των κυττάρων επαναιωρείται αρχικά σε 500ml παγωμένου, αποστειρωμένου και αποιονισμένου H₂O, φυγοκεντρείται στις ίδιες συνθήκες και επαναιωρείται διαδοχικά σε H₂O όγκου 250ml, 50ml και 10ml. Τέλος, επαναιωρείται σε 1.5ml παγωμένου υδατικού διαλύματος γλυκερόλης 10% και τα δεκτικά για μετασχηματισμό βακτηριακά κύτταρα αποθηκεύονται στους -80°C σε όγκους 50μl και χρησιμοποιούνται σε χρονικό διάστημα περίπου ενός έτους με απόδοση μετασχηματισμού $5 \times 10^8 - 5 \times 10^9$ αποικίες ανά μg υπερελικωμένου πλασμιδίου.

ii) Διαδικασία μετασχηματισμού

Αποθηκευμένα στους -80°C δεκτικά για μετασχηματισμό βακτηριακά κύτταρα όγκου 50μl μεταφέρονται σε πάγο, αναμιγνύονται με 0.001 – 1ng υπερελικωμένου πλασμιδιακού DNA ή 3μl αντίδρασης σύνδεσης (*ligation*) μορίων DNA και παραμένουν στον πάγο για 15 min. Στη συνέχεια, μεταφέρονται σε παγωμένες κυψελίδες τύπου BTX P/N 610, διαστάσεων 1cm x 1cm, με κενό διάστημα (*chamber gap*) πάχους 1mm, και υποβάλλονται σε ηλεκτρική εκκένωση τάσης 1600V για 4-6msec στην ειδική συσκευή (Electro Cell Manipulator 600). Τα μετασχηματισμένα βακτηριακά κύτταρα μεταφέρονται σε 1ml παγωμένου θρεπτικού μέσου LB, επωάζονται στους 37°C για 2-3 hrs με ήπια ανακίνηση (100 - 150 στροφές/min) και τέλος αναπτύσσονται σε στερεό θρεπτικό μέσο LB για 16-20 hrs στους 37°C, παρουσία του κατάλληλου αντιβιοτικού.

➤ **Μετασχηματισμός βακτηρίων με τη μέθοδο του θερμικού σοκ**

Για τη χημική κατεργασία των βακτηρίων πριν το θερμικό σοκ χρησιμοποιείται συνδυασμός δισθενών κατιόντων Rb⁺², Mn⁺² και Ca⁺² σε συγκεντρώσεις και χρόνους επώασης που προτείνονται σε σχετικά πρωτόκολλα [Sambrook *et al*, 1989].

i) Παρασκευή βακτηριακών κυττάρων δεκτικών για μετασχηματισμό με θερμικό σοκ

Καλλιέργεια κυττάρων *Escherichia coli* κατάλληλου στελέχους αναπτύσσεται σε 50ml θρεπτικού μέσου LB με ολονύκτια επώαση στους 37°C και ανακίνηση στις 200 στροφές / min. Στη συνέχεια, 500μl μεταφέρονται σε 100ml LB και η καλλιέργεια συνεχίζεται σε κωνική φιάλη όγκου 2 λίτρων στις ίδιες συνθήκες μέχρι να φτάσει σε λογαριθμική φάση ανάπτυξης ($OD_{595}=0.4-0.5$), ακολουθεί επώαση σε πάγο για 15 min και φυγοκέντρηση στις 2000 στροφές / min για 10

min στους 4°C. Το ίζημα των κυττάρων επαναιωρείται αρχικά σε 40ml διαλύματος TFB1 (100mM RbCl₂, 50mM MnCl₂, 30mM CH₃COOK, 10mM CaCl₂, 15% γλυκερόλη, pH 5.8), επάζεται στον πάγο για 20 min και στη συνέχεια σε 5ml διαλύματος TFB2 (10mM MOPS, 10mM RbCl₂, 80mM CaCl₂, 15% γλυκερόλη, pH 7.0). Τέλος, τα δεκτικά για μετασχηματισμό βακτηριακά κύτταρα αποθηκεύονται στους -80°C σε όγκους 200μl και χρησιμοποιούνται σε χρονικό διάστημα περίπου ενός έτους με απόδοση μετασχηματισμού 5x10⁷-5x10⁸ αποικίες μετασχηματισμένων βακτηρίων ανά μg υπερελικωμένου πλασμιδιακού DNA.

ii) Διαδικασία μετασχηματισμού

Τα δεκτικά για μετασχηματισμό βακτηριακά κύτταρα όγκου 200μl μεταφέρονται από τους -80°C, όπου αποθηκεύονται, σε πάγο 15 min πριν την προσθήκη του πλασμιδιακού υπερελικωμένου DNA. Η παραμονή στον πάγο συνεχίζεται για 20 min και ακολουθεί απότομη μεταφορά των κυττάρων σε υδατόλουτρο θερμοκρασίας 42-43°C για 90sec, επαναφορά στον πάγο για 2min, προσθήκη 800μl θρεπτικού μέσου LB και επώαση στους 37°C για 1-3 hrs με ήπια ή και χωρίς ανακίνηση. Τέλος, τα μετασχηματισμένα βακτήρια αναπτύσσονται σε στερεό θρεπτικό μέσο LB για 16-20 hrs στους 37°C, παρουσία του κατάλληλου αντιβιοτικού.

Σημειώνεται ότι ο χρόνος επώασης σε LB μετά το θερμικό ή ηλεκτρικό σοκ, εξαρτάται από το γονίδιο ανθεκτικότητας σε αντιβιοτικό που φέρει το πλασμίδιο. Για ανθεκτικότητα σε αμπικιλίνη ή χλωραμφαινικόλη, ο χρόνος επώασης ήταν 45-60 min, ενώ για ανθεκτικότητα σε καναμυκίνη ο χρόνος ήταν 2-3 hrs. Μετά τον μετασχηματισμό των βακτηρίων έγιναν πειράματα ταυτοποίησης των κλώνων που έφεραν ανασυνδυασμένα πλασμίδια.

ε. Ταυτοποίηση ανασυνδυασμένων βακτηριακών κλώνων

Οι βακτηριακές αποικίες που φέρουν ανασυνδυασμένα πλασμίδια επιλέγονται με τις μεθόδους που περιγράφονται παρακάτω, και σε κάθε περίπτωση ακολουθεί απομόνωση του πλασμιδιακού DNA και ανάλυση με ενδονουκλεάσες περιορισμού.

i) Επιλογή βάσει χρώματος

Βασίζεται στην παρουσία του γονιδίου *LacZ* το οποίο κωδικοποιεί την β-γαλακτοσιδάση και βρίσκεται σε ορισμένους φορείς κλωνοποίησης (π.χ pBluescript). Συγκεκριμένα, ο πλασμιδιακός φορέας περιέχει το γονίδιο, του οποίου το ανοιχτό πλαίσιο ανάγνωσης διακόπτεται, χωρίς να καταστρέφεται, από τις πολλαπλές θέσεις κλωνοποίησης ενθεμάτων. Στα ανασυνδυασμένα πλασμίδια το γονίδιο *LacZ* δεν κωδικοποιεί τη β-γαλακτοσιδάση και στις αντίστοιχες βακτηριακές αποικίες δεν πραγματοποιείται αντίδραση χρώσης σε θρεπτικό μέσο που περιέχει το υπόστρωμα X-gal, με αποτέλεσμα αυτές οι αποικίες να εμφανίζονται λευκές.

ii) Επιλογή με υβριδοποίηση των βακτηριακών αποικιών σε μεμβράνες

Για τη μεταφορά και μονιμοποίηση βακτηριακών κλώνων σε μεμβράνες υβριδοποίησης εφαρμόζονταν η μέθοδος των Grunstein and Hogness (1975). Μεμβράνες υβριδοποίησης (Hybond-N⁺, Amersham) τοποθετούνται στα τρυβλία για 30sec. Ακολουθεί αποδιάταξη του DNA σε διάλυμα 1.5M NaCl και 0.5M NaOH για 5 min, εξουδετέρωση σε διάλυμα 1 M Tris-Cl pH 7.4 και 1.5 M NaCl, για 7 min και πλύσιμο σε 2x SSC για 5 min. Το δεσμευμένο DNA μονιμοποιείται με έκθεση σε υπεριώδη ακτινοβολία (*UV cross-linking*) και θέρμανση των μεμβρανών στους 80°C για 2 hrs. Ακολουθούν πειράματα υβριδοποίησης στα οποία ως ανιχνευτής χρησιμοποιείται ραδιοσημασμένο το επιθυμητό ένθεμα, ώστε να επιλεχθούν οι αντίστοιχοι κλώνοι.

iii) Επιλογή με αλυσιδωτή αντίδραση πολυμεράσης (PCR)

Στις περιπτώσεις που η επιλογή των ανασυνδυασμένων πλασμιδίων πραγματοποιείται με PCR, χρησιμοποιούνται εκκινητές που υβριδοποιούνται εκατέρωθεν της θέσης κλωνοποίησης. Συγκεκριμένα, κάθε βακτηριακή αποικία υπό εξέταση, μεταφέρεται σε 15μl αποστειρωμένου και απιονισμένου H₂O και 5μl αυτού του εναιωρήματος

χρησιμοποιείται ως 'μήτρα' σε αντιδράσεις PCR, ενώ με το υπόλοιπο εμβολιάζεται μια υγρή καλλιέργεια. Οι βακτηριακοί κλώνοι που φέρουν ανασυνδυασμένα πλασμίδια δίνουν προϊόν PCR μεγέθους μεγαλύτερου κατά το μέγεθος του ενθέματος, σε σχέση με αυτούς που φέρουν μη ανασυνδυασμένα πλασμίδια.

στ. Μετασηματισμός κυττάρων σακχαρομύκητα ρJ69-4a με πλασμιδιακό DNA

Για τον μετασηματισμό κυττάρων σακχαρομύκητα με πλασμιδιακό DNA ακολουθήθηκε το πρωτόκολλο των Agatep *et al* (1998), με ορισμένες τροποποιήσεις. Συγκεκριμένα, 50ml θρεπτικού μέσου YEP, το οποίο περιείχε και 0.25mM αδενίνης, εμβολιάζεται με μία αποικία κυττάρων σακχαρομύκητα στελέχους ρJ69-4a και επώαζεται για 16-20 hrs στους 30°C με ανακίνηση (250 rpm). Στη συνέχεια, η καλλιέργεια αραιώνεται κατάλληλα με θρεπτικό μέσο (μέχρι $OD_{550} = 0.2-0.4$) και η επώαση συνεχίζεται μέχρι την $OD_{550} = 0.6-0.7$ (συνολικός αριθμός κυττάρων: $\sim 10^9$). Ακολουθεί φυγοκέντρηση, επαναιώρηση του κυτταρικού ιζήματος σε 25ml αποστειρωμένου H_2O και δεύτερη επαναιώρηση σε 1ml φρέσκου διαλύματος οξικού λιθίου συγκέντρωσης 100mM. Το δείγμα φυγοκεντρείται (4000 rpm / 5-10 min), απομακρύνεται το υπερκείμενο και ακολουθεί νέα επαναιώρηση των κυττάρων σε 500μl διαλύματος οξικού λιθίου 100mM. Για κάθε μετασηματισμό, αναμιγνύονται: 0.5-1 μg πλασμιδιακού DNA και 500μg *carrier* DNA σε συνολικό όγκο 20μl και στη συνέχεια προσθέτονται 240μl διαλύματος 50% πολυαιθυλενικής γλυκόλης μοριακού βάρους 3350 (PEG 3350). Τέλος, προσθέτονται 100μl του κυτταρικού εναιωρήματος σε 100mM οξικού λιθίου, το δείγμα αναδεύεται ήπια και ακολουθεί προσθήκη 36μl πυκνού διαλύματος οξικού λιθίου (1M), ήπια ανάδευση (σε *vortex*) και επώαση για 30 min στους 30°C. Ακολουθεί θερμικό σοκ στους 42°C για 25 min, φυγοκέντρηση (3000rpm/10min), απομάκρυνση του υπερκείμενου, επαναιώρηση του ιζήματος σε 100μl θρεπτικού μέσου SC και άπλωμα των κυττάρων σε κατάλληλο θρεπτικό μέσο. Η επιλογή των μετασηματισμένων αποικιών πραγματοποιείται λόγω των αυξοτροφιών του στελέχους ρJ69-4a, που συμπληρώνονται από γονίδια των πλασμιδίων.

ζ. Διαχωρισμός πλασμιδίου ρVP16 από το πλασμίδιο ρGBT9

Η μέθοδος εφαρμόζεται με σκοπό να απομονωθεί πλασμιδιακό DNA από τους κλώνους – 'θηράματα' (*prey*) της βιβλιοθήκης cDNA (σε φορέα ρVP16), οι οποίοι επιλέγονται με βάση την αλληλεπίδραση που εμφανίζουν τα κωδικοποιούμενα από αυτούς πεπτίδια με την πρωτεΐνη 'δόλωμα' (*bait*) που κωδικοποιείται από το πλασμίδιο *cDNA-bait/pGBT9* στο σύστημα δύο υβριδίων σακχαρομύκητα. Επειδή κάθε επιλεγμένη αποικία κυττάρων σακχαρομύκητα περιέχει και τα δύο πλασμίδια (το '*prey*' cDNA σε ρVP16 και το '*bait*' cDNA σε ρGBT9), για την επιλογή μόνο του άγνωστου cDNA (*cDNA-prey/pVP16*) πραγματοποιείται, μετά την απομόνωση από σακχαρομύκητα, μετασηματισμός σε βακτήρια *E.coli* του στελέχους KC8, το οποίο παρουσιάζει αυξοτροφία σε ουρακίλη, τρυπτοφάνη, ιστιδίνη, αδενίνη και λευκίνη. Μετά τον μετασηματισμό, τα βακτήρια αναπτύσσονται σε θρεπτικό μέσο M9 που περιέχει ουρακίλη, τρυπτοφάνη, ιστιδίνη, αδενίνη (στις συγκεντρώσεις που αναφέρονται στην §4γ) και δεν περιέχει λευκίνη. Η αυξοτροφία της λευκίνης συμπληρώνεται από την παρουσία στα βακτήρια του πλασμιδίου *cDNA-prey/pVP16*, το οποίο με αυτόν τον τρόπο 'διασώζεται', σε αντίθεση με το πλασμίδιο *cDNA-bait/pGBT9* το οποίο, λόγω της παρουσίας τρυπτοφάνης στο θρεπτικό μέσο, δεν είναι απαραίτητο και αποβάλλεται από τα βακτήρια. Ακολουθεί απομόνωση πλασμιδιακού DNA σε μικρή κλίμακα. Επειδή το πλασμιδιακό DNA που απομονώνεται από το στέλεχος KC8 δεν είναι κατάλληλο για να χρησιμοποιηθεί σε αντιδράσεις αλληλούχησης, μετασηματίζονται με αυτό βακτήρια *E.coli* του στελέχους DH5a, από τα οποία απομονώνεται τελικά πλασμιδιακό DNA κατάλληλο για χρήση σε αντιδράσεις προσδιορισμού της νουκλεοτιδικής του σύστασης.

17. Σύστημα δύο υβριδίων σακχαρομύκητα (*yeast two-hybrid system*)

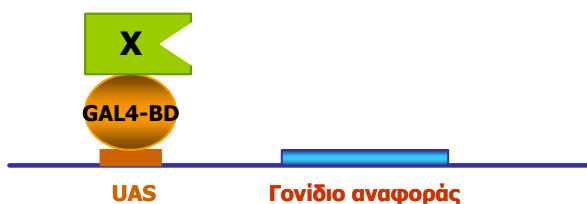
Πρόκειται για μια ευρέως διαδεδομένη πειραματική προσέγγιση, με σκοπό την ανίχνευση άμεσων πρωτεϊνικών αλληλεπιδράσεων σε ένα *in vivo* σύστημα. Χρησιμοποιείται τόσο για τη μελέτη της αλληλεπίδρασης δύο πρωτεϊνών, όσο και για την ταυτοποίηση πεπτιδίων που εμφανίζουν αλληλεπίδραση με μία υπό μελέτη πρωτεΐνη που χρησιμοποιείται σαν 'δόλωμα' (*bait*). Η αρχή της μεθόδου παρουσιάζεται διαγραμματικά στην Εικ.30. Η μέθοδος στηρίζεται στο ότι η μεταγραφική ενεργοποίηση ενός γονιδίου αναφοράς, η οποία λαμβάνει χώρα στον πυρήνα κυττάρων σακχαρομύκητα, εξαρτάται μόνο από την αλληλεπίδραση της υπό μελέτη πρωτεΐνης ('δόλωμα', '*bait*') με την πρωτεΐνη που προέκυψε από την επιλογή ('θήραμα', '*prey*').

Η παραπάνω προσέγγιση χρησιμοποιήθηκε για να ταυτοποιηθούν πρωτεΐνες-'θηράματα' (*prey*), που αλληλεπιδρούν με την υπό μελέτη πρωτεΐνη (*bait*), ελέγχοντας μια βιβλιοθήκη κλώνων cDNA εμβρύου ποντικού. Για την πραγματοποίηση του πειράματος ήταν απαραίτητη η κατασκευή υβριδικής πρωτεΐνης '*bait*', συντηγμένης με την υπομονάδα πρόσδεσης στο DNA του μεταγραφικού παράγοντα GAL4 (GAL4-BD), και η χρησιμοποίηση μιας βιβλιοθήκης κλώνων cDNA, οι οποίοι κωδικοποιούσαν υβριδικά πεπτίδια μέσου μήκους 100-250 αμινοξέων, συντηγμένα με την υπομονάδα ενεργοποίησης της μεταγραφής του παράγοντα VP16 (VP16-AD). Τα πρωτεϊνικά προϊόντα και των δύο παραπάνω κατασκευών οδηγούνταν στον πυρήνα κυττάρων στελέχους σακχαρομύκητα (PJ69-4a), λόγω σημάτων πυρηνικού εντοπισμού στην αμινοξική σύσταση των υπομονάδων GAL4-BD και VP16-AD. Επειδή το στέλεχος του σακχαρομύκητα που χρησιμοποιήθηκε στερείται ενδογενούς μεταγραφικού παράγοντα GAL4, η αλληλεπίδραση της πρωτεΐνης '*bait*' και ενός πεπτιδίου '*prey*' ήταν αναγκαία και ικανή να οδηγήσει σε ενεργοποίηση των γονιδίων αναφοράς (*ADE2*, *HIS3*, *LacZ*), που επάγονταν από GAL4 – ρυθμιζόμενους υποκινητές (GAL2, GAL1, GAL7).

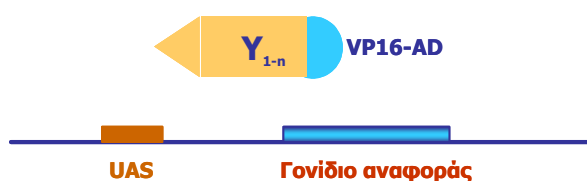
Η πειραματική διαδικασία που ακολουθήθηκε περιλαμβάνει:

1. Κλωνοποίηση του cDNA (ολόκληρου ή τμήματος) της πρωτεΐνης '*bait*' στον πλασμιδιακό φορέα έκφρασης pGBT9, στο αναγνωστικό πλαίσιο 3' της υπομονάδας πρόσδεσης στο DNA του μεταγραφικού παράγοντα GAL4 (GAL4-BD). Μετασχηματισμός κυττάρων σακχαρομύκητα του στελέχους PJ69-4a και επιλογή των ανασυνδυασμένων αποικιών.
2. Μετασχηματισμό του παραπάνω στελέχους (PJ69-4a: *cDNA-bait/pGBT9*) με τους κλώνους μιας βιβλιοθήκης cDNA εμβρύου ποντικού E9.5, κλωνοποιημένης στον πλασμιδιακό φορέα pVP16, στο 3' άκρο της υπομονάδας ενεργοποίησης της μεταγραφής του μεταγραφικού ενεργοποιητή VP16 (VP16-AD). Ανάπτυξη των διπλά μετασχηματισμένων κλώνων (PJ69-4a: *cDNA-bait/pGBT9*, *cDNA-prey/pVP16*) σε επιλεκτικό μέσο.
3. Επιλογή των αποικιών, οι οποίες εμφανίζουν ενεργοποίηση των γονιδίων αναφοράς (*ADE2*, *HIS3*, *LacZ*), εξαρτώμενη από την ταυτόχρονη παρουσία σ'αυτές των πλασμιδίων *cDNA-bait/pGBT9* και *cDNA-prey/pVP16*.
4. Απομόνωση πλασμιδιακού DNA από τους επιλεγμένους κλώνους σακχαρομύκητα και μετασχηματισμός βακτηρίων *E.coli* KC8. Ανάπτυξη των μετασχηματισμένων βακτηρίων σε συνθήκες επιλογής για την απώλεια του πλασμιδίου *cDNA-bait/pGBT9* και τη διατήρηση του *cDNA-prey/pVP16*.
5. Μετασχηματισμό βακτηρίων *E.coli* DH5a με πλασμιδιακό DNA *cDNA-prey/pVP16*, με σκοπό την απομόνωση πλασμιδιακού DNA καλής ποιότητας.
6. Προσδιορισμό της νουκλεοτιδικής αλληλουχίας των επιλεγμένων κλώνων *cDNA-prey/pVP16*.

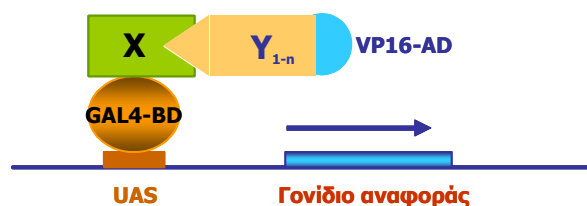
Το υβρίδιο GAL4-BD-X προσδένεται στο DNA



Τα υβρίδια VP16-AD-Y_{1-n} κωδικοποιούνται από μια βιβλιοθήκη cDNA



Αλληλεπίδραση των πρωτεϊνών X και Y προκαλεί μεταγραφική ενεργοποίηση



Εικ.30 Η αρχή της μεθόδου των δύο υβριδίων. GAL4-BD= GAL4 DNA binding domain, VP16-AD= VP16 activation domain, UAS= Upstream Activating Sequence, X=πρωτεΐνη - 'δόλωμα' (*bait*), Y=πρωτεΐνη - 'θήραμα' (*prey*).

α. Ανάπτυξη του στελέχους σακχαρομύκητα PJ69-4a

Το στέλεχος σακχαρομύκητα PJ69-4a αναπτύσσεται στους 30°C ή στους 25°C σε υγρό ή στερεό θρεπτικό μέσο YEP ή SC. Η σύσταση των θρεπτικών μέσων και οι συγκεντρώσεις των αμινοξέων και νουκλεοτιδίων που προσθέτονται σ'αυτό αναφέρονται παραπάνω (§4γ). Το είδος των αμινοξέων ή νουκλεοτιδίων που προσθέτονται σε κάθε περίπτωση, καθορίζονται με βάση το γονότυπο του στελέχους.

- *Στέλεχος αγρίου τύπου*: Το στέλεχος PJ69-4a εμφανίζει αυξοτροφία για λευκίνη, ουρακίλη, ιστιδίνη, τρυπτοφάνη και αδενίνη. Αναπτύσσεται είτε σε πλούσιο θρεπτικό μέσο (YEP), παρουσία αδενίνης, είτε στο συνθετικό μέσο SC, παρουσία Leu, Ura, His, Trp και Ade.

- *Στέλεχος PJ69-4a μετασχηματισμένο με το πλασμίδιο rGBT9*: Το πλασμίδιο rGBT9 συμπληρώνει την αυξοτροφία του PJ69-4a για τρυπτοφάνη. Επομένως, το μετασχηματισμένο με rGBT9 στέλεχος αναπτύσσεται σε SC παρουσία Leu, Ura, His και Ade.

- *Στέλεχος PJ69-4a μετασχηματισμένο με το πλασμίδιο rVP16*: Το πλασμίδιο rVP16 συμπληρώνει την αυξοτροφία του PJ69-4a για λευκίνη. Επομένως, το μετασχηματισμένο με rVP16 στέλεχος αναπτύσσεται σε SC παρουσία Ura, His, Trp και Ade.

- Διπλά μετασχηματισμένο στέλεχος PJ69-4a (με τα πλασμίδια *pGBT9* και *pVP16*): Στο διπλά μετασχηματισμένο στέλεχος συμπληρώνονται και οι δύο αυξοτροφίες (για τρυπτοφάνη και για λευκίνη) από τα πλασμίδια *pGBT9* και *pVP16*, αντίστοιχα. Επομένως, αυτό το στέλεχος αναπτύσσεται σε SC παρουσία Ura, His και Ade.

β. Επιλογή αλληλεπιδρώντων

Η επιλογή των πεπτιδίων που αλληλεπιδρούν με την πρωτεΐνη 'bait', πραγματοποιείται στο διπλά μετασχηματισμένο στέλεχος σακχαρομύκητα PJ69-4a, το οποίο φέρει τα πλασμίδια *bait/pGBT9* και *prey/pVP16*. Το στέλεχος αυτό αναπτύσσεται σε θρεπτικό μέσο SC παρουσία Ura, His και Ade. Για να ταυτοποιηθούν οι κλώνοι cDNA της βιβλιοθήκης οι οποίοι εμφανίζουν αλληλεπίδραση με την πρωτεΐνη 'bait', ελέγχεται η ικανότητά τους να αναπτύσσονται σε θρεπτικά μέσα επιλεκτικά για τη μεταγραφική ενεργοποίηση καθενός από τα τρία γονίδια αναφοράς του στελέχους PJ69-4a (*HIS3*, *ADE2*, και *LacZ*). Τα θρεπτικά μέσα που χρησιμοποιούνται σε κάθε περίπτωση ελέγχου μεταγραφικής ενεργοποίησης ήταν:

Θρεπτικό μέσο UA: SC με Ura, Ade και 3mM ή 10mM 3-aminotriazol (3-AT), για τον έλεγχο της ενεργοποίησης του γονιδίου αναφοράς *HIS3*. Η 3-αμινοτριάζολη (AT) είναι τοξική για τα κύτταρα που δεν παράγουν ιστιδίνη ή παράγουν χαμηλά επίπεδα ιστιδίνης, λόγω μη ειδικής ενεργοποίησης του υποκινητή GAL1.

Θρεπτικό μέσο UH: SC με Ura και His για τον έλεγχο ενεργοποίησης του γονιδίου αναφοράς *ADE2*.

Θρεπτικό μέσο UAH/X-gal: SC με Ura, Ade, His και X-gal, για τον έλεγχο της ενεργοποίησης του γονιδίου αναφοράς *LacZ*.

Η διαδικασία επιλογής που ακολουθήθηκε διακρίνεται σε τρία στάδια:

1^ο στάδιο επιλογής (διερεύνηση βιβλιοθήκης)

Περίπου 8×10^6 κλώνοι του διπλά μετασχηματισμένου στελέχους απλώνονται σε 4 τρυβλία διαστάσεων 23x23 cm σε στερεό θρεπτικό μέσο UA, το οποίο περιέχει μικρή συγκέντρωση αδενίνης (1.25mM), 20mM ουρακίλης, 5mM 3-AT και δεν περιέχει ιστιδίνη. Οι κλώνοι που αναπτύσσονται σ' αυτό το στάδιο επιλογής μεταφέρονται σε πρότυπα τρυβλία (*master plates*) με μη επιλεκτικό θρεπτικό μέσο (SC με Ura, Ade, His), όπου τοποθετούνται διατεταγμένοι.

2^ο στάδιο επιλογής (έλεγχος θετικών κλώνων 1^{ου} σταδίου)

Οι θετικοί κλώνοι του 1^{ου} σταδίου, διατεταγμένοι στα πρότυπα τρυβλία, μεταφέρονται με αποτύπωση (*replica*, Sambrook *et al*, 1989), σε καθένα από τα παραπάνω επιλεκτικά θρεπτικά μέσα (UA, UH και UAH/X-gal), καθώς και σε μη επιλεκτικό θρεπτικό μέσο UAH (SC με Ura, Ade, His), για πείραμα θετικού ελέγχου. Ως 'θετικοί' επιλέγονται: α) κλώνοι που εμφανίζουν ενεργοποίηση και των τριών γονιδίων αναφοράς και β) κλώνοι που εμφανίζουν ενεργοποίηση στα γονίδια *ADE2* και *HIS3* και δεν ενεργοποιούν το γονίδιο *LacZ*, επειδή η ενεργοποίηση του υποκινητή GAL7 απαιτεί περίπου τρεις φορές πιο ισχυρή αλληλεπίδραση σε σχέση με τους άλλους υποκινητές. Στη συνέχεια, από τους θετικούς κλώνους αυτού του σταδίου απομονώθηκε DNA του πλασμιδίου *cDNA-prey/pVP16*, που χρησιμοποιήθηκε σε αντίδραση προσδιορισμού της νουκλεοτιδικής του σύστασης. Ακολουθούσε σύγκριση της νουκλεοτιδικής αλληλουχίας των θετικών κλώνων τόσο μεταξύ τους, όσο και με βάσεις δεδομένων (GenBank, dbEST) ώστε να ταυτοποιηθούν οι θετικοί κλώνοι που αντιστοιχούν σε γνωστά γονίδια.

3^ο στάδιο επιλογής

Χρησιμοποιήθηκε πλασμιδιακό DNA από τους διαφορετικούς θετικούς κλώνους του 2^{ου} σταδίου, για να πραγματοποιηθεί με αυτό νέος μετασχηματισμός του στελέχους σακχαρομύκητα PJ69-4a, παρουσία ή όχι του πλασμιδίου *cDNA-bait/pGBT9*. Στη συνέχεια, οι μετασχηματισμένες αποικίες ελέγχθηκαν όπως περιγράφεται για το 2^ο στάδιο της επιλογής. Σκοπός αυτού του σταδίου ήταν να αποδειχθεί ότι ο κάθε επιλεγμένος κλώνος cDNA (*prey*) από το 2^ο στάδιο,

ήταν αυτός που κωδικοποιεί το πεπτιδίο που αλληλεπιδρά με την πρωτεΐνη 'bait' και να αποκλειστεί το ενδεχόμενο της παρουσίας περισσότερων του ενός *cDNA-prey/pVP16* πλασμιδίων στις επιλεγμένες αποικίες του 1^{ου} και 2^{ου} σταδίου.

γ. Πειράματα αρνητικού ελέγχου

Τα πειράματα αρνητικού ελέγχου είναι απαραίτητα στο πείραμα των δύο υβριδίων σακχαρομύκητα, ώστε να αποκλειστεί το ενδεχόμενο οι υποκινητές των γονιδίων αναφοράς να ενεργοποιούνται μόνο από την πρωτεΐνη που χρησιμοποιείται σαν 'δόλωμα' (*bait*) ή μόνο από την πρωτεΐνη – 'θήραμα' που επιλέγεται από τη βιβλιοθήκη (*prey*). Τα πειράματα αρνητικού ελέγχου που πραγματοποιήθηκαν ήταν:

- α) μετασχηματισμός του στελέχους σακχαρομύκητα PJ69-4a με το πλασμιδίο *cDNA-bait/pGBT9*,
- β) μετασχηματισμός του PJ69-4a με τον πλασμιδιακό φορέα pGBT9,
- γ) μετασχηματισμός του PJ69-4a με τον πλασμιδιακό φορέα pVP16.

Ακολούθησε έλεγχος ενεργοποίησης των τριών γονιδίων αναφοράς στις συνθήκες που περιγράφονται για το 2^ο στάδιο ελέγχου. Τα παραπάνω πειράματα αρνητικού ελέγχου πραγματοποιήθηκαν πριν τη διερεύνηση της βιβλιοθήκης, για να εξασφαλιστεί η αξιοπιστία των αποτελεσμάτων.

δ) μετασχηματισμός του PJ69-4a με καθέναν από τους επιλεγμένους κλώνους *cDNA-prey/pVP16* και έλεγχος ενεργοποίησης των τριών γονιδίων αναφοράς, με σκοπό να αποκλειστεί το ενδεχόμενο ενεργοποίησης των γονιδίων αναφοράς αποκλειστικά και μόνο λόγω ικανότητας πρόσδεσης του κωδικοποιούμενου πεπτιδίου-θήραματος' (*prey*) στις θέσεις αναγνώρισης του μεταγραφικού παράγοντα GAL4, στους υποκινητές των τριών γονιδίων αναφοράς.

18. Έκφραση, ανάλυση και ανίχνευση πρωτεϊνών

Χρησιμοποιήθηκαν διάφοροι μέθοδοι έκφρασης γονιδίων σε βακτήρια και σε κύτταρα θηλαστικών. Η έκφραση σε βακτήρια πραγματοποιήθηκε χρησιμοποιώντας το επαγόμενο σύστημα έκφρασης του *tac* υποκινητή και του *lacI* καταστολέα, ενώ στα κύτταρα των θηλαστικών ετερόλογα πεπτιδία εκφράζονταν μετά από παροδική διαμόλυνση των κυττάρων COS7 με τις αντίστοιχες πλασμιδιακές κατασκευές που έφεραν περιοχή έναρξης της αντιγραφής SV40 και υποκινητή CMV (Cytomegalovirus). Ακολουθούσε α) απομόνωση του πρωτεϊνικού εκχυλίσματος, β) καθαρισμός της υπερεκφρασμένης πρωτεΐνης ή των συμπλόκων που αυτή σχηματίζει, γ) ηλεκτροφόρηση πρωτεϊνικών μορίων και ανίχνευσή τους.

α. Παροδική διαμόλυνση κυτταρικών σειρών με πλασμιδιακές κατασκευές

Η παροδική παραγωγή ανασυνδυασμένων πεπτιδίων σε κύτταρα θηλαστικών επιτυγχάνεται κλωνοποιώντας τα αντίστοιχα cDNAs σε κατάλληλους πλασμιδιακούς φορείς με ρυθμιστικά στοιχεία που επάγονται ισχυρά στο κυτταρικό περιβάλλον, όπου εισάγονται οι πλασμιδιακές κατασκευές. Οι πρωτεΐνες που παράγονται μπορούν να ανιχνευθούν σε πειράματα υποκυτταρικού εντοπισμού, ή να αναλυθούν και να απομονωθούν σε βιοχημικά πειράματα (ανάλυση Western, ανοσοκατακρήμνιση κλπ).

Η παροδική διαμόλυνση των κυτταρικών σειρών πραγματοποιήθηκε σύμφωνα με τη μέθοδο των Graham και Van der Eb (1973), η οποία βασίζεται στην συγκατακρήμνιση του DNA με κρυστάλλους φωσφορικού ασβεστίου ($\text{Ca}_3(\text{PO}_4)_2$). Αναλυτικότερα, υπερεκκωμένο πλασμιδιακό DNA αναμιγνύεται με CaCl_2 τελικής συγκέντρωσης 125mM και ρυθμιστικό διάλυμα φωσφορικών αλάτων HEBS 1x (140mM NaCl, 0.75mM Na_2HPO_4 , 25mM HEPES pH 7.0), ώστε να

σχηματιστεί λεπτόκοκκο κρυσταλλικό ίζημα DNA-Ca₃(PO₄)₂. Το εναιώρημα των κρυστάλλων προστίθεται στα κύτταρα που βρίσκονται σε εκθετική φάση ανάπτυξης και οι κρύσταλλοι εισέρχονται με κυτταροφαγία στο εσωτερικό τους, κατά τη διάρκεια επώασης για 7 έως 12 hrs σε θερμοκρασία 37°C και ατμόσφαιρα με 5% CO₂. Η περίσσεια των κρυστάλλων απομακρύνεται ξεπλένοντας τα κύτταρα 2-3 φορές με 1xPBS ή με 1x HBSS (*Hank's Balanced Salt Solution*, Gibco BRL). Τέλος, με την προσθήκη νέου θρεπτικού μέσου, συνεχίζεται η ανάπτυξη των κυττάρων στις ίδιες συνθήκες για 12-36 hrs. Σύμφωνα με τα παραπάνω, τα πειράματα παροδικής διαμόλυνσης που παραγματοποιήθηκαν περιγράφονται συνοπτικά παρακάτω:

i) Παροδική διαμόλυνση με σκοπό την ανοσοανίχνευση της παραγόμενης πρωτεΐνης στο εσωτερικό των κυττάρων (με in situ ανοσοφθορισμό)

Τα κύτταρα αναπτύσσονται σε τρυβλία 6 θέσεων (διαμέτρου 3.5cm/θέση), στην επιφάνεια αντικειμενοφόρου με 3ml θρεπτικού μέσου DMEM, 10% FCS), αναναωμένο 2 περίπου hrs πριν τη διαμόλυνση. Για κάθε πείραμα χρησιμοποιήθηκε 0.25–1μg DNA υπερελικωμένου ανασυνδυασμένου πλασμιδίου και κατάλληλη ποσότητα από *carrier* DNA (μηχανικά τεμαχισμένο DNA σολομού), ώστε η συνολική ποσότητα του DNA να είναι 8μg. Ο συνολικός όγκος του διαλύματος ρυθμίζεται με απιονισμένο H₂O στα 135μl και στη συνέχεια προσθέτονται 15μl διαλύματος CaCl₂ 2.5M. Το μίγμα αναδεύεται ισχυρά (σε συσκευή Vortex) και σ'αυτό προσθέτονται σταγόνα-σταγόνα ίσος όγκος (150μl) διαλύματος 2x HEBS, με παράλληλη ανάδευση, ώστε να σχηματιστούν όσο το δυνατόν μικρότεροι κρύσταλλοι, οι οποίοι εισέρχονται ευκολότερα στα κύτταρα. Το λεπτόκοκκο ίζημα προστίθεται μετά από 15min σε θερμοκρασία δωματίου στα κύτταρα τα οποία επωάζονται στη συνέχεια (37°C, 5% CO₂) για 7-12 hrs. Τα πλυσίματα των κυττάρων γίνονται με 3ml 1x PBS ή 1x HBSS, και η επώαση συνεχίζεται στις ίδιες συνθήκες για 24 hrs με 5ml φρέσκου θρεπτικού μέσου DMEM. Ακολουθούν πειράματα ανοσοεντοπισμού της παραγόμενης πρωτεΐνης στα διαμολυσμένα κύτταρα, που είναι συνήθως το 10-20% του συνόλου.

ii) Παροδική διαμόλυνση με σκοπό την απομόνωση πρωτεϊνικού εκχυλίσματος

Τα κύτταρα αναπτύσσονται στην ίδια συνθήκες σε μεγαλύτερα τρυβλία (διαμέτρου 10cm ή 20cm), σε 10ml ή 20ml θρεπτικού μέσου, αντίστοιχα. Για κάθε διαμόλυνση χρησιμοποιούνται 24μg πλασμιδιακού DNA (για τρυβλία διαμέτρου 10cm) ή 40μg (για τρυβλία διαμέτρου 20cm), χωρίς να προστεθεί *carrier* DNA. Τα υπόλοιπα συστατικά προσθέτονται στις ίδιες συγκεντρώσεις και με τη σειρά που περιγράφεται παραπάνω. Ο τελικός όγκος του μίγματος είναι 1ml. Μετά την ολοκλήρωση της επώασης ακολουθεί απομόνωση ολικού πρωτεϊνικού εκχυλίσματος.

β. Ανίχνευση πρωτεΐνης σε κύτταρα θηλαστικών με in situ ανοσοφθορισμό

Τα πειράματα ανοσοεντοπισμού πραγματοποιήθηκαν ως επί το πλείστον σε κύτταρα (COS7) που αναπτύσσονται προσκολλημένα σε καλυπτρίδα τοποθετημένη μέσα σε τρυβλίο κυτταροκαλλιέργειας, χρησιμοποιώντας πρώτο αντίσωμα ειδικό για την υπό ανίχνευση πρωτεΐνη και δεύτερο αντίσωμα συζευγμένο με φθορίζουσα ουσία, ικανό να αλληλεπιδράσει με το πρώτο. Αναλυτικότερα, τα κύτταρα αναπτύσσονται σε μονοστιβάδες, που καλύπτουν την επιφάνεια καλυπτρίδων κατά 80-90%. Στη συνέχεια, πλένονται 3 φορές με 1x PBS και μονιμοποιούνται με 4% PFA (παραφορμαλδεΐδη) για 5 min. Ακολουθούν 3 πλυσίματα με 1x PBS και προσθήκη διαλύματος 20mM γλυκίνης για 10 min. Η γλυκίνη απομακρύνεται με 1x PBS και ακολουθεί επώαση σε διάλυμα 0.5% Triton-X και 1x PBS για 20 min, το οποίο καθιστά τις μεμβράνες των κυττάρων διαπερατές. Στη συνέχεια, προστίθεται παρεμποδιστικό διάλυμα (2mM MgCl₂, 5% BSA σε 1x PBS) για 10 min. Τα κύτταρα επωάζονται με το πρώτο αντίσωμα, αραιωμένο κατάλληλα σε παρεμποδιστικό διάλυμα (Πίνακας 11). Ο χρόνος επώασης κυμαίνεται από 45 min έως και 2 hrs, ανάλογα με το αντίσωμα (Πίνακας 11). Στη συνέχεια, τα κύτταρα ξεπλένονται τρεις φορές με παρεμποδιστικό διάλυμα χωρίς BSA και επωάζονται με το δεύτερο (φθορίζον) αντίσωμα στο σκοτάδι για 45 min. Ακολουθεί πλύσιμο των κυττάρων τρεις φορές με παρεμποδιστικό διάλυμα (χωρίς BSA) και μία φορά με 1x PBS και μεταφορά της καλυπτρίδας με την πλευρά των κυττάρων πάνω σε μια σταγόνα (~10μl) διαλύματος 1% n-propyl-gallate και 80% γλυκερόλης σε 1x PBS. Η καλυπτρίδα

σφραγίζεται αεροστεγώς και το δείγμα φυλάσσεται στους -20°C μέχρι να παρατηρηθεί σε απλό ή συνεστιακό (Leica TCS MP) μικροσκόπιο φθορισμού.

Στις περιπτώσεις πειράματος ανοσοεντοπισμού σε κύτταρα που δεν προσκολλώνται στο υπόστρωμα, ακολουθείται η ίδια διαδικασία με τη διαφορά ότι μετά το αρχικό πλύσιμο με 1x PBS, τα κύτταρα μεταφέρονται στην επιφάνεια αντικειμενοφόρων με επικάλυψη πολυλυσίνης (poly-D, O.Kindler GmbH & Co.) όπου παραμένουν για περίπου 1hr, μέχρι να προσκολληθούν στην αντικειμενοφόρο, λόγω του θετικού φορτίου της επίστρωσης poly-D.

Χρώση DNA με DAPI: Στις περιπτώσεις που υπήρξε ανάγκη να παρατηρηθούν οι πυρήνες των κυττάρων, πραγματοποιήθηκε χρώση του DNA με DAPI (*4',6-diamidino-2-phenylindole, dihydrochloride*). Συγκεκριμένα, μετά την επώαση των κυττάρων με το φθορίζον αντίσωμα, ακολουθεί πλύσιμο με 1x PBS και επώαση με διάλυμα DAPI 1μg/ml σε 1x PBS για 7 min. Η χρωστική απομακρύνεται με τρεις διαδοχικές πλύσεις με 1x PBS, τα κύτταρα τοποθετούνται σε αντικειμενοφόρο, όπως προηγουμένως, και η παρατήρηση γίνεται σε μικροσκόπιο φθορισμού.

γ. Απομόνωση ολικού πρωτεϊνικού εκχυλίσματος από καλλιέργεια κυττάρων θηλαστικών

Για την απομόνωση πρωτεϊνικού εκχυλίσματος, μονοστιβάδες κυττάρων αναπτύχθηκαν σε τρυβλία καλύπτοντας το υπόστρωμα κατά 90%. Μετά από έκπλυση από το θρεπτικό υλικό με παγωμένο 1x PBS που περιέχει και 1mM EDTA pH 8.0, τα κύτταρα αποκολλούνται από την επιφάνεια του τρυβλίου με τη βοήθεια πλαστικής ξύστρας και μεταφέρονται σε σωληνάριο φυγοκέντρου. Ακολουθεί φυγοκέντρηση ($1200\text{rpm}/4^{\circ}\text{C}/20\text{min}$), και επαναιώρηση των κυττάρων σε 5-6 όγκους διαλύματος λύσης (Πίνακας 13). Τα κύτταρα επωάζονται αρχικά για 20 min (4°C) και στη συνέχεια σε περιστρεφόμενο αναδευτήρα στους 4°C , μέχρι να παρατηρηθεί σχηματισμός νηματοειδούς ιζήματος από μεμβρανικά υπολείμματα και DNA. Ακολουθεί τεμαχισμός των μεγάλου μεγέθους συσσωματωμάτων και του DNA, μηχανικά, περνώντας το παρασκεύασμα από σύριγγα 26g περίπου 20 φορές. Τέλος, το πρωτεϊνικό εκχύλισμα φυγοκεντρείται ($10.000\text{rpm} / 4^{\circ}\text{C} / 5\text{min}$) και το υπερκείμενο φυλάσσεται στους -80°C με 10% γλυκερόλη.

δ. Παραγωγή πρωτεΐνης σε βακτήρια με επαγωγή

Για την παραγωγή πολυπεπτιδίων σε βακτήρια στελέχους *E. coli* BL21-Codon Plus(DE3)-RIL, τα cDNAs κλωνοποιήθηκαν σε πλασμιδιακούς φορείς έκφρασης (pGEX) στο καρβοξυτελικό άκρο του ενζύμου θειοτρανσφεράση της γλουταθειόνης (*Glutathione-S-Transferase*, GST). Η έκφραση των χιμαιρικών γονιδίων πραγματοποιήθηκε με ενεργοποίηση του υποκινητή *tac* παρουσία IPTG (*isopropyl-beta-D-1-thiogalactopyranoside*, $\text{C}_9\text{H}_{18}\text{O}_5\text{S}$) τελικής συγκέντρωσης 0.5mM στους 30°C για 4hrs. Η διαδικασία που ακολουθήθηκε περιγράφεται αναλυτικά σε διάφορα εργαστηριακά πρωτόκολλα [Sambrook *et al*, 1989/Cur. Prot. Molec. Biol., unit 16.7]. Συγκεκριμένα, 5-10 αποικίες ανασυνδυασμένων βακτηρίων αναπτύσσονται για 16-20 hrs στους 30°C , σε καλλιέργεια όγκου 50ml (LB, 100μg/ml αμπικιλίνη, 25μg/ml χλωραμφαινικόλη) και στη συνέχεια 2ml βακτηρίων μεταφέρονται σε 200ml θρεπτικού υλικού για χρόνο επώασης μέχρι τη λογαριθμική φάση ανάπτυξης ($\text{OD}_{600} = 0.65$). Ακολουθεί επαγωγή με προσθήκη 100μl διαλύματος IPTG 1M για 4 επιπλέον hrs και, εν συνεχεία, λύση των κυττάρων, απομόνωση του πρωτεϊνικού εκχυλίσματος (βλ. παρακάτω) και διαγνωστική ηλεκτροφόρηση σε πήκτωμα πολυακρυλαμίδης.

Μετά τη φυγοκέντρηση της βακτηριακής καλλιέργειας (2500rpm , 20min, 4°C) και την επαναιώρηση του ιζήματος σε 10ml διαλύματος λύσης (50mM Tris-Cl pH 8.0, 1mM EDTA, 100mM NaCl, 0.5% NP40, 0.5mg/ml BSA, 1mM PMSF), ακολουθεί επώαση σε πάγο για 10 min και λύση των κυττάρων με επίδραση υπερήχων (3 επαναλήψεις διάρκειας 1min, με ενδιάμεση επώαση 1min σε πάγο). Το δείγμα φυγοκεντρείται στις ίδιες συνθήκες και διαχωρίζεται το ίζημα (κυτταρικό τοίχωμα, μεμβράνη και DNA των βακτηρίων, έγκλειστα σωμάτια) από το υπερκείμενο (διαλυτές πρωτεΐνες).

Πίνακας 13. Διαλύματα λύσης κυττάρων θηλαστικών

<u>Διάλυμα 1 (RIPA)</u> Χρόνος επώασης : 30min	150mM NaCl 1% NP40, 0.5% DOC, 0.1% SDS, 50mM Tris-Cl pH 8.0 1mM PMSF
<u>Διάλυμα 2</u> Χρόνος επώασης : 1-3 hr	250mM NaCl, 1mM EDTA, 1% Triton X-100, 50mM Tris-Cl pH 7.5, 1mM PMSF
<u>Διάλυμα 3</u> Χρόνος επώασης: 1-3 hr	150mM NaCl, 1% Triton X-100, 0.05% SDS, 0.05% DOC, 100mM Tris-Cl pH7.5 1mM PMSF
<u>Διάλυμα 4</u> Χρόνος επώασης: 10 hr	150mM NaCl, 0.1% NP40, 50mM Hepes pH 7.0, 1mM PMSF

Σημ: Τα παραπάνω διαλύματα (1-4) χρησιμοποιήθηκαν εναλλακτικά, με βασικό κριτήριο τις ιδιότητες της πρωτεΐνης που πρόκειται να ανιχνευθεί στο κυτταρικό εκχύλισμα. Η λύση πρέπει να είναι αποτελεσματική ώστε να επιτρέπει ικανή ποσότητα της υπό μελέτη πρωτεΐνης στην υδατική φάση του εκχυλίσματος, πράγμα που εξασφαλίζεται από την παρουσία απορρυπαντικών στο διάλυμα. Σε πειράματα όπου απομονώνονται πρωτεϊνικά σύμπλοκα είναι συνήθως επιθυμητή η χρήση διαλυμάτων με μικρότερη συγκέντρωση απορρυπαντικών, τα οποία επιτρέπουν την απομόνωση όσο το δυνατό πιο άθικτων συμπλόκων. Στον Πίνακα φαίνεται η σύσταση των τεσσάρων διαφορετικών διαλυμάτων λύσης που χρησιμοποιήθηκαν. Ο απαιτούμενος χρόνος λύσης (επώαση στους 40°C) των κυττάρων, εξαρτάται από τη συγκέντρωση απορρυπαντικών στο διάλυμα λύσης.

Στις περιπτώσεις που σημαντική ποσότητα της υπερεκφρασμένης πρωτεΐνης είχε παγιδευτεί στο ίζημα, εφαρμόστηκε το πρωτόκολλο των Frangioni και Neel (1993). Συγκεκριμένα, το ίζημα της βακτηριακής καλλιέργειας μετά την επαγωγή επαναιωρείται σε 10ml διαλύματος STE (10mM Tris-Cl pH 8.0, 150mM NaCl, 1mM EDTA), στο οποίο προστίθεται λυσοζύμη. Μετά από επώαση για 15min στον πάγο, προστίθεται DTT τελικής συγκέντρωσης 5mM και μετά από 10min πραγματοποιείται λύση με 1.5% sarcosyl (N-laurylsarcosine, SIGMA), από αρχικό διάλυμα 10%. Ακολουθεί επώαση για 30min σε πάγο και επίδραση υπερήχων για ~3min, μέχρι την πλήρη ρευστοποίηση του εκχυλίσματος. Το δείγμα φυγοκεντρείται δύο φορές (12000rpm/4°C/15min) και μέρος του υπερεκφρασμένου διαλύματος ηλεκτροφορείται για την ποσοτικοποίηση της υπερεκφρασμένης πρωτεΐνης. Για τα πρωτεϊνικά εκχυλίσματα που υποβλήθηκαν σε χρωματογραφία συγγένειας με σφαιρίδια γλουταθειόνης για περαιτέρω καθαρισμό της πρωτεΐνης, στο υπερεκφρασμένο μετά τη φυγοκέντρηση προστέθηκε Triton-X σε τελική συγκέντρωση 2%. Τα πρωτεϊνικά εκχυλίσματα αποθηκεύονται στους -80°C, παρουσία 10% γλυκερόλης.

ε. Ηλεκτροφορητικός διαχωρισμός πρωτεϊνικού εκχυλίσματος

Ακολουθήθηκε η μέθοδος ανάλυσης πρωτεϊνών με ηλεκτροφόρηση σε πήκτωμα πολυακρυλαμίδης και SDS, όπως περιγράφεται αναλυτικά από τους Sambrook *et al.* Τα πήκτωμα πολυακρυλαμίδης περιλαμβάνουν δύο μέρη: το πρώτο για την ευθυγράμμιση των πρωτεϊνών σε ενιαίο μέτωπο (*stacking gel*) και το δεύτερο για τον διαχωρισμό τους (*separating gel*). Η σύσταση και οι διαστάσεις των πηκτωμάτων που χρησιμοποιήθηκαν φαίνονται στον Πίνακα 14. Το διάλυμα ηλεκτροφόρησης ήταν 25mM Tris-Cl, 192mM γλυκίνη, 1% SDS, pH 7.5. Τα δείγματα πριν την ηλεκτροφόρηση

αποδιατάσσονται με βρασμό για 5 min σε 1x *Laemli buffer* (50mM Tris-Cl pH6.8, 2% SDS, 0.1% *bromophenol blue*, 10% γλυκερόλη, 5% β-μερκαπτοαιθανόλη). Ακολουθεί χρώση των πρωτεϊνών με χρωστική *Coomasie blue* ή μεταφορά τους σε μεμβράνη (*Western blot*) και ανοσοανίχνευση τους.

Χρώση πρωτεϊνών με *Coomasie blue*

Μετά την ηλεκτροφόρηση, το πήκτωμα επωάζεται για περίπου 1 hr, με ήπια ανακίνηση, σε διάλυμα χρώσης που περιέχει 0.16%(w/v) *Coomasie blue*, 50% μεθανόλη και 12% οξικό οξύ και στη συνέχεια σε διάλυμα αποχρωματισμού (30% μεθανόλη, 10% οξικό οξύ, 3% γλυκερόλη), για 2 ή και περισσότερες ώρες.

στ. Ανάλυση 'Western blot'

Η μεταφορά των πρωτεϊνών από το πήκτωμα σε μεμβράνη νιτροκυτταρίνης (Protran, 0.2mm, Schleicher & Schuell) έγινε όπως περιγράφεται από τους Sambrook *et al*, με την πειραματική διάταξη που περιλαμβάνει η συσκευή Mini-PROTEAN 3 (Biorad). Για τη μεταφορά των πρωτεϊνών στη μεμβράνη χρησιμοποιήθηκε διάλυμα ίδιας σύστασης με αυτό της ηλεκτροφόρησης (βλ. παραπάνω), παρουσία 20% μεθανόλης, εφαρμόζοντας 110mA για 12–16 hrs ή 110Volts για 2-3 hrs.

Πίνακας 14. Σύσταση ηκτώματων SDS-PAGE

Σύσταση ηκτώματος 'ευθυγράμμισης' (<i>stacking gel</i>)	Σύσταση ηκτώματος 'ανάλυσης' (<i>separating gel</i>)
4.5% πολυακρυλαμίδη	8-20% πολυακρυλαμίδη
125mM Tris-Cl, pH 6.8	375mM Tris-Cl, pH 8.8
0.1% SDS	0.1% SDS
0.1% APS	0.1% APS
0.2% TEMED	0.05% TEMED

Σημ: Η συγκέντρωση της πολυακρυλαμίδης στο αναλυτικό πήκτωμα κυμαίνονταν από 8% έως 20%, ανάλογα με το μέγεθος των πρωτεϊνών που διαχωρίζονταν σε κάθε πείραμα [Sambrook *et al*, 1989]. Χρησιμοποιήθηκαν ηλεκτροφορητικές συσκευές Biorad (Mini-PROTEAN 3 Electrophoresis system), και ηκτώματα διαστάσεων 8.0cm x 7.3cm x 0.15cm (πλάτος x ύψος x πάχος). Το διάλυμα πολυακρυλαμίδης προκύπτει με ανάμιξη ακρυλαμίδης και δις-ακρυλαμίδης (*bis acrylamide*), σε αναλογία 29:1. APS = *ammonium persulfate*, TEMED = *N,N,N',N'* - tetramethylethylene diamine.

ζ. Ανοσοανίχνευση πρωτεϊνών μονιμοποιημένων σε μεμβράνες

Μετά την επώαση της μεμβράνης σε παρεμποδιστικό διάλυμα TBSTM [5% άπαχο γάλα σε 1x TBST (20mM Tris-Cl pH 7.6, 137mM NaCl, 0.1% Tween 20)] για 1-2 hrs σε θερμοκρασία δωματίου, ακολουθεί επώαση με το κατάλληλο αντίσωμα. Η συγκέντρωση του αντισώματος, η ιδανική θερμοκρασία και η διάρκεια της επώασης καθορίζεται εμπειρικά σε κάθε περίπτωση (βλ. Πίνακα 11). Στη συνέχεια, η μεμβράνη ξεπλένεται σε 1x TBST τρεις φορές για 10-20 min, με σκοπό την απομάκρυνση του αντισώματος που τυχόν είχε προσδεθεί μη ειδικά στη μεμβράνη. Ακολουθεί ανίχνευση του πρώτου αντισώματος με κατάλληλο δεύτερο αντίσωμα συζευγμένο με το ένζυμο HRP (*horse radish peroxidase*) σε διάλυμα TBSTM και σε συνθήκες που καθορίζονται για κάθε αντίσωμα (Πίνακας 11). Στη συνέχεια, η μεμβράνη ξεπλένεται με διάλυμα TBST τρεις φορές για 10-20 min και τέλος επωάζεται της μεμβράνης με μίγμα των αντιδραστηρίων

Supersignal WestPico Stable Peroxide και *Luminol/Enhancer* (PIERCE), τα οποία αποτελούν υπόστρωμα του ενζύμου HRP, σε αναλογία 1:1 για 5 min σε θερμοκρασία δωματίου. Το προϊόν της ενζυμικής αντίδρασης ανιχνεύεται με αυτοραδιογραφία.

η. Καθαρισμός της υπερεκφρασμένης πρωτεΐνης με χρωματογραφία συγγένειας

Τα χιμαιρικά πεπτίδια, υπερεκφρασμένα από τους πλασμιδιακούς φορείς της οικογένειας pGEX, καθαρίζονται μερικώς, με βάση την ιδιότητα της πρωτεΐνης GST (*glutathione-S-transferase*) να προσδένεται ισχυρά στο υπόστρωμά της, την γλουταθειόνη. Για το σκοπό αυτό χρησιμοποιήθηκαν ειδικά σφαιρίδια σεφαρόζης, καλυμμένα εξωτερικά με γλουταθειόνη (Glutathione Sepharose 4B, Amersham Biosciences Cat.no. 17-0756-01). Τα σφαιρίδια αρχικά εξισορροπούνται με 3 πλύσιμα σε παγωμένο διάλυμα πρόσδεσης (50mM Tris-Cl pH 8.0, 1mM EDTA, 100mM NaCl, 0.5% NP40, 0.5mg/ml BSA, 1mM PMSF) και στη συνέχεια επωάζονται με το πρωτεϊνικό εκχύλισμα στους 4°C για 2 hrs περίπου. Χρησιμοποιήθηκαν περίπου 100μl σφαιριδίων σε εναιώρημα αιθανόλης 20% για 200-500μg υπερεκφρασμένης πρωτεΐνης (ποσότητα που συνήθως αντιστοιχεί σε 0.5-1ml βακτηριακού εκχυλίσματος). Ακολουθεί διαχωρισμός των σφαιριδίων από το υπερκείμενο με φυγοκέντρηση (3000rpm/4°C /30sec), πλύσιμο των σφαιριδίων 2-3 φορές με 1ml διαλύματος πρόσδεσης χωρίς BSA, επαναιώρηση σε 1-2 όγκους του ίδιου διαλύματος και φύλαξη στους 4°C παρουσία 10% γλυκερόλης. Σημειώνεται ότι στις παραπάνω συνθήκες, το μεγαλύτερο μέρος (70-80%) της συντηγμένης με GST πρωτεΐνης προσδένεται στα σφαιρίδια. Στις περιπτώσεις που χρειάστηκε να γίνει έκλυση της πρωτεΐνης από τα σφαιρίδια, ακολούθησε επώαση των σφαιριδίων με ήπια ανακίνηση για 20 min σε θερμοκρασία δωματίου παρουσία 50-100mM γλουταθειόνης, σε διάλυμα πρόσδεσης χωρίς BSA. Η γλουταθειόνη απομακρύνεται από το διάλυμα της εκλουσμένης πρωτεΐνης με διαπίδυση (*dialysis*) μέσα από ημιπερατή μεμβράνη (Visking Dialysis Membrane, size 3, Medicell Int. Ltd), σε 200-300ml διαλύματος πρόσδεσης χωρίς BSA, παρουσία 10% γλυκερόλης και αναστολέα πρωτεασών (1mM PMSF) στους 4°C, με ήπια ανακίνηση της μεμβράνης για 12-16 hrs.

θ. *In vitro* δοκιμή αλληλεπίδρασης πρωτεΐνης με πρωτεΐνη συντηγμένη με GST, σε χρωματογραφία συγγένειας (GST-pulldown)

Η μέθοδος στηρίζεται στην ιδιότητα των χιμαιρικών πρωτεϊνών που φέρουν την πρωτεΐνη GST (θειοτρανσφεράση γλουταθειόνης) στο αμινοτελικό τους άκρο να καθαρίζονται από ένα πρωτεϊνικό εκχύλισμα, λόγω της συγγένειάς τους με τη γλουταθειόνη, με τη βοήθεια σφαιριδίων με επικάλυψη γλουταθειόνης (Glutathione Sepharose 4B). Η μέθοδος εφαρμόζεται και στη μελέτη πρωτεϊνικών αλληλεπιδράσεων, ανιχνεύοντας πρωτεΐνες που συνκατακρημνίζονται με αυτές που φέρουν GST και, επομένως, αλληλεπιδρούν μ'αυτές *in vitro*. Συγκεκριμένα, ελέγχθηκε η αλληλεπίδραση των συντηγμένων με GST πρωτεϊνών με πρωτεΐνες που είχαν συνθεθεί *in vitro*, χρησιμοποιώντας το σύστημα συζευγμένης μεταγραφής/μετάφρασης (T_NT Quick Coupled Transcription/Translation System) της εταιρίας Promega (βλ. παρακάτω), παρουσία ραδιενεργής μεθειονίνης (³⁵S-Met). Αναλυτικότερα, για κάθε πεπτίδιο του οποίου ελέγχθηκε η αλληλεπίδραση με μια χιμαιρική GST - πρωτεΐνη, κλωνοποιήθηκε το αντίστοιχο cDNA στον πλασμιδιακό φορέα έκφρασης pCDNA3.1 ή pCDNA3.1mycHis, ο οποίος φέρει στο 5'άκρο της θέσης κλωνοποίησης τον υποκινητή T7 και θέση έναρξης της μετάφρασης. Περίπου 1μg DNA πλασμιδίου χρησιμοποιείται για κάθε αντίδραση *in vitro* μεταγραφής/μετάφρασης, σύμφωνα με τις οδηγίες της Promega. Για κάθε δοκιμή αλληλεπίδρασης, χρησιμοποιήθηκαν 8μl αντίδρασης *in vitro* μετάφρασης (συνολικού όγκου 50μl) και 30μl σφαιριδίων γλουταθειόνης, στα οποία είχαν προηγουμένως προσδεθεί 5-10μg GST-πρωτεΐνης. Η αλληλεπίδραση πραγματοποιήθηκε για 5-12 hrs στους 4°C, σε περιστρεφόμενο αναδευτήρα και σε διάλυμα αλληλεπίδρασης (50mM Tris-Cl pH 8.0, 1mM EDTA, 100mM NaCl, 0.5% NP40, 0.5mg/ml BSA, 10% γλυκερόλη, 4mM PMSF) συνολικού όγκου 300μl. Ακολουθούσε φυγοκέντρηση (3000rpm/ 4°C/ 30sec), απομάκρυνση του υπερκείμενου και πλύσιμο των σφαιριδίων 3 φορές σε διάλυμα αλληλεπίδρασης, χωρίς BSA και γλυκερόλη, παρουσία PMSF τελικής συγκέντρωσης 1mM, στους 4°C.

Στη συνέχεια, πραγματοποιείται ηλεκτροφόρηση σε πήκτωμα πολυακρυλαμίδης. Τα δείγματα των πρωτεϊνών εκλύονται από τα σφαιρίδια με βρασμό των δειγμάτων για 5-10 min σε διάλυμα 1x *Laemli* (50mM Tris-Cl pH6.8, 2% SDS, 0.1% *bromophenol blue*, 10% γλυκερόλη, 5% β-μερκαπτοαιθανόλη) και ηλεκτροφορούνται. Ακολουθεί σταθεροποίηση του ηλεκτώματος σε διάλυμα 30% μεθανόλης, 10% οξικού οξέος και 3% γλυκερόλης για 30 min, ενίσχυση του σήματος με επώαση του ηλεκτώματος σε διάλυμα 1M σαλικλικού νατρίου pH 6.6 για 1 hr, ξήρανση στους 80°C υπό κενό και έκθεση σε φιλμ αυτοραδιογραφίας για 12-48 hrs στους -80°C. Η ανίχνευση της ραδιοσημασμένης πρωτεΐνης στο ίζημα των σφαιριδίων σημαίνει ότι αυτή αλληλεπιδρά με την συντηγμένη με GST πρωτεΐνη στο παραπάνω *in vitro* σύστημα. Για να ελεγχθεί η ειδικότητα της αλληλεπίδρασης, πραγματοποιείται παράλληλα και πείραμα αρνητικού ελέγχου, στο οποίο δοκιμάζεται η αλληλεπίδραση της ραδιοσημασμένης πρωτεΐνης με την πρωτεΐνη GST. Αναλυτικότερα:

i) Πρόσδεση της GST - πρωτεΐνης στα σφαιρίδια γλουταθειόνης

Σφαιρίδια γλουταθειόνης (Glutathione Sepharose 4B, Pharmacia) όγκου περίπου 30μl εξισορροπούνται με τρία πλυσίματα σε 1ml διαλύματος πρόσδεσης (50mM Tris-Cl pH 8.0, 1mM EDTA, 100mM NaCl, 0.5% NP40, 0.5mg/ml BSA, 1mM PMSF) στους 4°C και στη συνέχεια αναμιγνύονται με 100–500μg ολικού πρωτεϊνικού εκχυλίσματος βακτηρίων (συγκέντρωση ολικής πρωτεΐνης: ~2μg/μl), που περιέχει περίπου 5-50μg υπερεκφρασμένης πρωτεΐνης συντηγμένης με GST. Τα δείγματα επωάζονται στους 4°C, σε περιστρεφόμενο αναδευτήρα, για 2-5 hrs και στη συνέχεια φυγοκεντρώνονται (3000rpm/ 4°C/ 30sec), απομακρύνεται το υπερκείμενο και τα σφαιρίδια ξεπλένονται 3 φορές με διάλυμα πρόσδεσης στις ίδιες συνθήκες. Ακολουθεί ηλεκτροφόρηση όπως παραπάνω (1/20–1/10 του συνολικού όγκου), χρώση με *Coomassie* και ποσοτικοποίηση της προσδεμένης στα σφαιρίδια πρωτεΐνης, συγκριτικά με πρωτεΐνη γνωστής ποσότητας που έχει ηλεκτροφορηθεί.

ii) Δοκιμή 'GST pull-down' σε πρωτεϊνικό εκχύλισμα από κύτταρα θηλαστικών

Μια τροποποίηση της μεθόδου εφαρμόστηκε για να δοκιμαστεί η αλληλεπίδραση μεταξύ ενός πεπτιδίου συντηγμένου με GST και μιας ανασυνδυασμένης πρωτεΐνης που παράγεται σε κύτταρα θηλαστικών (COS7) μετά την παροδική τους διαμόλυνση με την αντίστοιχη πλασμιδιακή κατασκευή. Συγκεκριμένα, η λύση των κυττάρων COS7 πραγματοποιήθηκε σε διάλυμα λύσης: 50mM Tris-Cl pH 8.0, 1mM EDTA pH 8.0, 100mM NaCl, 0.5% NP40 και 1mM PMSF (~1ml διαλύματος/τρυβλίο 10cm). Ακολούθησε επώαση 500μg πρωτεϊνικού εκχυλίσματος για 12 hrs, σε περιστρεφόμενο αναδευτήρα στους 4°C, με 30μl σφαιρίδια γλουταθειόνης τα οποία είχαν προηγουμένως υποστεί χρωματογραφία συγγένειας με ~50μg GST–πρωτεΐνης, και εξισορροπηθεί στο ίδιο διάλυμα λύσης. Ακολούθησε η ίδια διαδικασία που περιγράφεται για τα βακτηριακά εκχυλίσματα και τέλος πραγματοποιήθηκε ανάλυση με *Western blot* και ανοσοανίχνευση της ανασυνδυασμένης πρωτεΐνης για την οποία δοκιμάστηκε η αλληλεπίδραση.

i. Συν-ανοσοκατακρήμνιση (*co-immunoprecipitation*)

Η μέθοδος εφαρμόστηκε για την ανίχνευση πρωτεϊνικών συμπλόκων που υπάρχουν *in vivo*, σε κύτταρα θηλαστικών. Για παράδειγμα, για την απομόνωση του συμπλόκου των πρωτεϊνών X και Y, σε 500-900μg πρωτεϊνικού εκχυλίσματος καλλιέργειας κυττάρων θηλαστικών, προστίθεται περίπου 0.5-2μg αντισώματος ειδικού για την πρωτεΐνη X και το δείγμα επωάζεται για 2-3 hrs στους 4°C, σε περιστρεφόμενο αναδευτήρα. Στη συνέχεια, προσθέτονται 8-20μl σφαιριδίων αгарόζης συζευγμένων με μόρια πρωτεΐνης A και πρωτεΐνης G (Protein A/G Plus Agarose, Santa Cruz Biotechnology Cat.no. sc-2003) και εξισορροπημένων σε διάλυμα λύσης των κυττάρων (Πίνακας 13), και η επώαση συνεχίζεται στις ίδιες συνθήκες για 5-12 hrs. Στο διάστημα αυτό, οι πρωτεΐνες A και G αλληλεπιδρούν με το Fc τμήμα του αντισώματος. Ακολουθεί φυγοκέντρωση του δείγματος (3000rpm, 30sec, 4°C), απομάκρυνση του υπερκείμενου και πλύσιμο των σφαιριδίων 3 φορές με το ίδιο διάλυμα. Τέλος, τα σφαιρίδια επαναιωρούνται σε μικρό όγκο (~50μl) διαλύματος και ακολουθεί ηλεκτροφόρηση σε αποδιατακτικό πήκτωμα (βλ. παραπάνω), ανάλυση με *Western blot* και ανοσοανίχνευση με αντίσωμα ειδικό για την πρωτεΐνη Y.

19. Ανάλυση αλληλουχιών DNA και πρωτεϊνών *in silico*

Παρακάτω αναφέρονται τα προγράμματα και οι βάσεις δεδομένων που χρησιμοποιήθηκαν για την ανάλυση αλληλουχιών DNA και πρωτεϊνών:

α. Ανάκτηση αλληλουχιών DNA ή πρωτεϊνών από τη βάση δεδομένων GenBank

Η GenBank είναι μια βάση δεδομένων με ελεύθερη πρόσβαση, η οποία περιέχει αλληλουχίες DNA και πρωτεϊνών που είτε κατατίθενται από ανεξάρτητες ερευνητικές ομάδες είτε καταχωρούνται μαζικά ως αποτέλεσμα των διεθνών προγραμμάτων αλληλούχησης γονιδιωμάτων και κλώνων cDNA [Benson *et al*, 2003]. Η πρόσβαση στην GenBank γίνεται μέσω του συστήματος Entrez του NCBI (*National Center for Biotechnology Information*), το οποίο είναι διαθέσιμο στον δικτυακό τόπο www.ncbi.nlm.nih.gov. Για την ανάκτηση μιας αλληλουχίας DNA ή πρωτεΐνης χρησιμοποιείται ο αντίστοιχος αριθμός πρόσβασης της αλληλουχίας ή κατάλληλη λέξη-κλειδί.

β. Πηγές δεδομένων αλληλούχησης γονιδιωμάτων (με ελεύθερη πρόσβαση)

Για την ανάκτηση αλληλουχιών γονιδιωμάτων DNA ή χρωμοσωμικών χαρτών του ανθρώπου ή του ποντικού, χρησιμοποιήθηκε η εφαρμογή MapView του συστήματος Entrez για τον άνθρωπο ή τον ποντικό, αντίστοιχα (www.ncbi.nlm.nih.gov/mapview) ή τα δεδομένα της βάσης Ensembl [Hubbard *et al*, 2002] του EBI (*European Bioinformatics Institute*) και του *Sanger Centre* (www.ensembl.org).

γ. Σύγκριση αλληλουχιών DNA ή πρωτεϊνών με τις αλληλουχίες της GenBank

Οι συγκρίσεις με τις νουκλεοτιδικές ή αμινοξικές αλληλουχίες της GenBank πραγματοποιήθηκαν με χρήση των προγραμμάτων της οικογένειας BLAST (*Basic Local Alignment Search Tool*) [Altschul *et al*, 1990], που είναι διαθέσιμα στον δικτυακό τόπο του NCBI (www.ncbi.nlm.nih.gov/BLAST). Τα προγράμματα αυτά χρησιμοποιούνται με σκοπό:

- τη σύγκριση μιας νουκλεοτιδικής αλληλουχίας με τις αλληλουχίες της βάσης δεδομένων (BLASTN),
- τη σύγκριση μιας αμινοξικής αλληλουχίας με τις αλληλουχίες της βάσης δεδομένων (BLASTP),
- τη μετάφραση μιας αλληλουχίας DNA και τη σύγκριση με τις πρωτεΐνες της βάσης δεδομένων (BLASTX), ή
- τη μετάφραση των αλληλουχιών DNA της βάσης δεδομένων και τη σύγκριση με μια αμινοξική αλληλουχία (TBLASTN).

Το αποτέλεσμα της σύγκρισης απεικονίζεται από το λογισμικό BLAST με τη μορφή νουκλεοτιδικών ή αμινοξικών αλληλουχιών (BLAST *hits*) στοιχημένων με την αλληλουχία που αποτελεί το ερώτημα (*Query*) της σύγκρισης. Επιπλέον, αναφέρεται το ποσοστό ταυτότητας μεταξύ του κάθε BLAST *hit* και της αλληλουχίας *Query*, καθώς και η τιμή E, η οποία εκφράζει τη στατιστική σπουδαιότητα κάθε στοίχισης, δηλαδή την πιθανότητα η στοίχιση να είναι τυχαία. Για όλες τις συγκρίσεις χρησιμοποιήθηκε εξ ορισμού ως όριο η τιμή $E=10$ και οι μικρότερες από αυτήν τιμές θεωρήθηκαν από το πρόγραμμα στατιστικά σημαντικές. Οι αποδεκτές στοίχισεις είχαν σχεδόν μηδενική τιμή E.

Για την ταυτοποίηση των συντηρημένων περιοχών μιας πρωτεΐνης πραγματοποιήθηκε σύγκριση της αμινοξικής της αλληλουχίας με όλα τα γνωστά συντηρημένα πρωτεϊνικά μοτίβα, με χρήση της εφαρμογής CDS (*Conserved Domain Search*) του BLAST. Επιπλέον, για την ταυτοποίηση των πρωτεϊνών με ίδια ή παρόμοια σύνθεση συντηρημένων περιοχών, χρησιμοποιήθηκε η εφαρμογή CDART (*Conserved Domain Architecture Retrieval Tool*).

δ. Στοίχιση αλληλουχιών DNA ή πρωτεϊνών και κατασκευή εξελικτικών δέντρων

Για τη στοίχιση δύο αλληλουχιών χρησιμοποιείται το πρόγραμμα Blast2seq (www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html). Για τη στοίχιση περισσότερων από δύο νουκλεοτιδικών ή αμινοξικών αλληλουχιών χρησιμοποιήθηκε το πρόγραμμα ClustalW στον δικτυακό τόπο www.ebi.ac.uk/clustalw. Για την κατασκευή φυλογενετικών δέντρων με βάση τη στοίχιση αμινοξικών αλληλουχιών, χρησιμοποιήθηκε το πρόγραμμα MEGA (*Molecular Evolutionary Genetics Analysis, version 2.1*) [Kumar *et al*, 2001] και συγκεκριμένα η μέθοδος των αποστάσεων UPGMA (*Unweighted Pair Group Method with Arithmetic Mean*), με βάση την οποία το μήκος των κλάδων του δέντρου

εκφράζουν τις αντικαταστάσεις αμινοξέων και οι εξελικτικές αποστάσεις που υπολογίζονται (*p-distances*) ανά ζεύγος συγκρινόμενων πρωτεϊνών εκφράζουν το ποσοστό των αμινοξέων που είναι διαφορετικά στις δύο πρωτεΐνες ($p=N_d / L$, όπου p είναι η εξελικτική απόσταση, N_d ο αριθμός των διαφορετικών αμινοξέων και L ο συνολικός αριθμός των αμινοξέων που συμμετέχουν στη στοίχιση). Το αποτέλεσμα της ανάλυσης αξιολογείται από τον ίδιο αλγόριθμο εφαρμόζοντας τη στατιστική δοκιμή '*Bootstrap test*' [Kumar *et al*, 2001]. Ως 'αξιόπιστη' θεωρείται η τοπολογία του δέντρου με 'τιμή *Bootstrap*' μεγαλύτερη από 90%.

ε. Μετάφραση αλληλουχιών DNA

Για να μεταφραστούν αλληλουχίες DNA *in silico*, χρησιμοποιήθηκε το πρόγραμμα DNAMAN v.4.0 (Lynnon Biosoftware).

στ. Ταυτοποίηση θέσεων αναγνώρισης περιοριστικών ενζύμων

Για να βρεθούν οι θέσεις αναγνώρισης περιοριστικών ενζύμων σε αλληλουχίες DNA, χρησιμοποιήθηκε το πρόγραμμα Webcutter v.2.0, στον δικτυακό τόπο www.firstmarket.com/cutter/cut2.html.

ζ. Ταυτοποίηση θέσεων SUMΟλίωσης

Για να ταυτοποιηθούν πιθανές θέσεις μετα-μεταφραστικής τροποποίησης πρωτεϊνών με τη σύζευξη σε αυτές του πεπτιδίου SUMO, χρησιμοποιήθηκε το πρόγραμμα SUMOplot (www.abgent.com/sumoplot.html).

ΠΛΑΣΜΙΔΙΑΚΕΣ ΚΑΤΑΣΚΕΥΕΣ

1. Υποκλωνοποιήσεις γονιδιωματικών κλώνων

Οι παρακάτω κλωνοποιήσεις τεμαχίων γονιδιωματικού DNA πραγματοποιήθηκαν στον φορέα pBluescript και χρησιμοποιήθηκαν για ανάλυση του προτύπου τεμαχισμού περιοριστικών ενζύμων και προσδιορισμό της νουκλεοτιδικής αλληλουχίας του γονιδίου *NeurI* του ποντικού. Αποτελούν υποκλώνους των βακτηριοφάγων που προέκυψαν από διερεύνηση γονιδιωματικής βιβλιοθήκης ποντικού.

α) 15.8kb EcoRI-NotI : Το DNA του φάγου Dph12 τεμαχίστηκε με EcoRI και NotI και υβριδοποιήθηκε με το ραδιοσημασμένο προϊόν PCR A-1. Η ζώνη 15.8Kb που έδινε θετικό σήμα στην υβριδοποίηση κλωνοποιήθηκε σε pBluescript / EcoRI-NotI. (Σημειώνεται ότι η θέση αναγνώρισης της ενδονουκλεάσης NotI ανήκει στον polylinker του φάγου)

β) 4.4kb BamHI-NotI : Είναι το 3' ακραίο κομμάτι της παραπάνω κατασκευής, προέκυψε τεμαχίζοντας το DNA του φάγου Dph12 με BamHI και NotI και υβριδοποιώντας με το προϊόν PCR A-1. Κλωνοποιήθηκε σε pBluescript / BamHI-NotI.

γ) 1.85kb BstYI : Προέκυψε με τεμαχισμό του DNA του φάγου Dph12 με BstYI και υβριδοποίηση με τον ανιχνευτή A-1. Αντιστοιχεί στο 5' ακραίο τμήμα της κατασκευής 4.4Kb BamHI-NotI, με άκρα BamHI-BstYI. Περιέχει περίπου 1.4kb γονιδιωματικής αλληλουχίας πάνω από το σημείο έναρξης της μετάφρασης, την κωδική αλληλουχία του πρώτου εξονίου (85bp) και 350bp από το 5' άκρο του πρώτου ιντρονίου. Κλωνοποιήθηκε σε pBluescript / BamHI.

δ) 1.4kb PstI : Προέκυψε με τεμαχισμό του DNA του φάγου Dph12 με PstI και υβριδοποίηση με τον ανιχνευτή A-1. Περιέχει περίπου 1.35kb γονιδιωματικής αλληλουχίας πάνω από το σημείο έναρξης της μετάφρασης μέχρι την PstI(35), που βρίσκεται στο εξόνιο E1. Κλωνοποιήθηκε σε pBluescript / PstI.

ε) 400bp SacII-PstI : Αντιστοιχεί στο 3' ακραίο κομμάτι της κατασκευής 1.4kb PstI και προέκυψε από τον τεμαχισμό αυτής με SacII και PstI. Κλωνοποιήθηκε σε pBluescript / SacII-PstI.

στ) 1.2kb BstYI : Προέκυψε με τεμαχισμό του DNA των φάγων Cph1 και Dph42 με BstYI και υβριδοποίηση με τον ανιχνευτή 1A. Περιέχει το 3' άκρο του πρώτου ιντρονίου, μέχρι την BstYI (218), η οποία βρίσκεται στο εξόνιο E2. Κλωνοποιήθηκε σε pBluescript/BamHI, με προσανατολισμό T7(5' – 3') T3.

ζ) 3.8kb PstI: Προέκυψε από τους φάγους Bph1 και Cph42, ύστερα από τεμαχισμό του DNA τους με PstI, ανάλυση κατά Southern και υβριδοποίηση με διάφορους ανιχνευτές. Περιέχει τα εξόνια E2 και E3. Κλωνοποιήθηκε σε pBluescript / PstI, με προσανατολισμό T3 (5'-3') T7.

η) 3.9kb BamHI: Προέκυψε από τους φάγους Bph1 και Cph42, ύστερα από πέψη του DNA τους με BamHI, ανάλυση κατά Southern και υβριδοποίηση με διάφορους ανιχνευτές. Περιέχει τα εξόνια E2 και E3, επικαλύπτεται με την κατασκευή 3.8kb PstI και κλωνοποιήθηκε σε pBluescript / BamHI, με προσανατολισμό T3 (5'-3') T7.

θ) 4.8kb PstI: Προέκυψε από τον φάγο Cph36, ύστερα από πέψη του DNA του με PstI. Περιέχει τα εξόνια E4, E5 και το εξόνιο E6, μέχρι τη θέση PstI (νουκλεοτιδίο 1734 του cDNA Y15160) που βρίσκεται 9 νουκλεοτιδία 3' του κωδικόνιου τερματισμού της μετάφρασης. Κλωνοποιήθηκε σε pBluescript/PstI με προσανατολισμό T7 (5'-3') T3.

ι) 300bp PstI-NotI: Προέκυψε από την κατασκευή 4.8kb PstI, ύστερα από πέψη με NotI. Περιέχει το 3' άκρο του τρίτου ιντρονίου, μέχρι την NotI(679) που βρίσκεται στο εξόνιο E4. Κλωνοποιήθηκε σε pBluescript/NotI, με προσανατολισμό T3 (5'-3') T7.

κ) 4.8kb PstI/NotI-recycled: Προέκυψε από την κατασκευή 4.8kb PstI, ύστερα από πέψη με NotI και επανακυκλοποίηση του πλασμιδίου.

2. Κλωνοποίηση του *Neurl* cDNA σε φορείς έκφρασης

Το cDNA του *Neurl* από τον κλώνο *Neurl* (2433)/pBluescript κλωνοποιήθηκε σε διάφορους φορείς με σκοπό την έκφρασή του σε κύτταρα θηλαστικών ή σακχαρομύκητα. Προέκυψαν οι εξής κατασκευές:

α) *Neurl/pCDNA3.1*: Από τον κλώνο *Neurl*(2433)/pRSETC αποκόπηκε θραύσμα μεγέθους 2082bp με τα ένζυμα BamHI και SpeI και συγκολλήθηκε με τον φορέα pCDNA3.1, ο οποίος είχε γραμμοποιηθεί με τα ένζυμα BamHI και XbaI. Η κατασκευή αυτή κωδικοποιεί την αλληλουχία των αμινοξέων 5-574 του *Neurl*. Η θέση έναρξης της μετάφρασης και εννέα αμινοξικά κατάλοιπα στο αμινοτελικό άκρο προέρχονται από την κατασκευή *Neurl/pRSETC*.

β) *Neurl/pCDNA3.1mycHisA*: Το ένθεμα (1769bp) προέκυψε με αντίδραση PCR, χρησιμοποιώντας ως DNA-μήτρα την κατασκευή *Neurl*(2433)/pRSETC και εκκινητές τους *BamHI-pRSETC-F* και *EcoRI-Neurl-R* (Πίνακας 8). Ο φορέας γραμμοποιήθηκε με τα ένζυμα BamHI και EcoRI. Η χιμαιρική πρωτεΐνη που κωδικοποιείται από αυτήν την κατασκευή περιλαμβάνει τα αμινοξέα 5-574 του *Neurl* καθώς και τους επιτόπους myc και 6xHis, στο καρβοξυτελικό άκρο. Η θέση έναρξης της μετάφρασης και εννέα αμινοξικά κατάλοιπα στο αμινοτελικό άκρο προέρχονται από την κατασκευή *Neurl/pRSETC*.

γ) *Neurl(5-390)/pCDNA3.1*: Το ένθεμα (1218bp) προέκυψε τεμαχίζοντας την κατασκευή *Neurl/pCDNA3.1* με τα ένζυμα BamHI (5' θέση κλωνοποίησης) και SmaI (μοναδική θέση στο νουκλεοτίδιο 1170 του cDNA Y15160). Η κλωνοποίηση έγινε σε φορέα pCDNA3.1 γραμμοποιημένο με BamHI και EcoRV. Από αυτήν την κατασκευή κωδικοποιούνται τα αμινοξέα 5-390 του *Neurl*.

δ) *Neurl(391-574)/pCDNA3.1HisB*: Το ένθεμα (578bp) προέκυψε από τον τεμαχισμό της κατασκευής *Neurl/pCDNA3.1mycHisA* με τα ένζυμα SmaI (μοναδική θέση στο νουκλεοτίδιο 1170 του cDNA Y15160) και NotI (μοναδική θέση στην περιοχή πολλαπλών θέσεων κλωνοποίησης του pCDNA3.1mycHisA). Η κλωνοποίηση έγινε σε φορέα pCDNA3.1HisB, ο οποίος είχε γραμμοποιηθεί με τα ίδια ένζυμα.

ε) *Neurl/pGBT9*: Το ένθεμα (1785bp) προέκυψε με αντίδραση PCR χρησιμοποιώντας ως μήτρα το ένθεμα του πλασμιδίου *Neurl(2433)/pBluescript* και εκκινητές τα συνθετικά ολιγονουκλεοτίδια *Neurl-2hyb-F* και *Neurl-2hyb-R* (Πίνακας 8), που στο 3' άκρο τους φέρουν αλληλουχία *Neurl* και στο 5' άκρο τους αλληλουχία συμπληρωματική με τις περιοχές 845-874bp και 903-932bp, αντίστοιχα, οι οποίες βρίσκονται εκατέρωθεν της πολλαπλής θέσης κλωνοποίησης του pGBT9 (Genbank ID: U07646, Cloning vector pGBT9, 5524bp). Χρησιμοποιήθηκε το σύστημα *Expand Long Template PCR* (System 3) της εταιρίας Roche, στο οποίο η αντίδραση PCR πραγματοποιείται από βελτιωμένο ένζυμο, με ικανότητα επιδιόρθωσης των σφαλμάτων (*proofreading*). Για κάθε αντίδραση, αναμιχθηκε το DNA-μήτρα (600ng) με 1mM από κάθε dNTP, 300ng από κάθε εκκινητή, 1x ρυθμιστικού διαλύματος (system 3), 10% φορμαμίδης, ενώ η συγκέντρωση του MgCl₂ ρυθμίστηκε σε 6mM. Η αντίδραση πραγματοποιήθηκε με 35 κύκλους θερμικής ανακύκλωσης, οι οποίοι περιελάμβαναν αποδιάταξη στους 92°C για 45sec, υβριδοποίηση των εκκινητών στους 58°C για 45sec και σύνθεση DNA στους 68°C για 90sec. Στη συνέχεια, το προϊόν της αντίδρασης ηλεκτροφορήθηκε, αποκόπηκε από το πήκτωμα αγαρόζης και χρησιμοποιήθηκε σε πείραμα μετασχηματισμού κυττάρων σακχαρομύκητα PJ69-4a, για την κλωνοποίησή του στον πλασμιδιακό φορέα pGBT9 με ομόλογο ανασυνδυασμό. Συγκεκριμένα, για τον μετασχηματισμό 10⁸-5x10⁹ κυττάρων σακχαρομύκητα χρησιμοποιήθηκαν 3μg του προϊόντος PCR και 400ng γραμμοποιημένου πλασμιδίου pGBT9.

Με αυτόν τον τρόπο, προστέθηκαν σε κάθε άκρο του ενθέματος πλευρικές αλληλουχίες μήκους 30bp, ταυτόσημες με αυτές του pGBT9, ώστε να είναι δυνατή η κλωνοποίηση με ομόλογο ανασυνδυασμό στα κύτταρα του σακχαρομύκητα. Σημειώνεται ότι επειδή από το DNA-μήτρα λείπουν τα πρώτα 12 νουκλεοτίδια του *Neurl* cDNA Y15160, αυτά προστέθηκαν στον εκκινητή *Neurl-2hyb-F*, ώστε η κατασκευή αυτή να κωδικοποιεί το πλήρες μόριο *Neurl* (574 αμινοξέα) σε σύντηξη στο αμινοτελικό του άκρο με την υπομονάδα πρόσδεσης στο DNA του μεταγραφικού παράγοντα GAL4.

στ) *Neurl (1-521)/pGBT9*: Ένα ελλειμματικό μόριο *Neurl* που περιλαμβάνει τα αμινοξικά κατάλοιπα 1-521 και κατασκευάστηκε ως προϊόν της αντίδρασης PCR με τους εκκινητές *EcoRI-Neurl-del-F* και *BamHI-Neurl-del-R*. Για την αντίδραση PCR χρησιμοποιήθηκε ένζυμο με ικανότητα επιδιόρθωσης (*proofreading*) των σφαλμάτων (Vent_R DNA Polymerase της εταιρίας New England Biolabs), στις συνθήκες που προτείνονται από την εταιρία, παρουσία MgSO₄ τελικής συγκέντρωσης 6mM και 10% φορμαμίδης. Πραγματοποιήθηκαν 5 κύκλοι

θερμικής ανακύκλωσης με θερμοκρασία υβριδοποίησης των εκκινητών στους 55°C και 30 κύκλοι με υβριδοποίηση στους 60°C. Κάθε κύκλος περιελάμβανε στάδιο αποδιάταξης στους 94 °C για 45sec και στάδιο σύνθεσης DNA στους 72 °C για 90sec.

ζ) *Neurl (515-574)/pGBT9* : Ένα ελλειματικό μόριο *Neurl* που περιλαμβάνει τα αμινοξικά κατάλοιπα 515-574 και κατασκευάστηκε ως προϊόν της αντίδρασης PCR με τους εκκινητές *EcoRI-RZD-F* και *BamHI-RZD-R*. Για την αντίδραση PCR χρησιμοποιήθηκε ένζυμο με ικανότητα επιδιόρθωσης (proofreading) των σφαλμάτων (*Vent_r* DNA Polymerase της εταιρίας New England Biolabs), στις συνθήκες που προτείνονται από την εταιρία. Πραγματοποιήθηκαν 2 κύκλοι θερμικής ανακύκλωσης με θερμοκρασία υβριδοποίησης των εκκινητών στους 57°C και 32 κύκλοι με υβριδοποίηση στους 67°C. Κάθε κύκλος περιελάμβανε στάδιο αποδιάταξης στους 94 °C για 45sec και στάδιο σύνθεσης DNA στους 72°C για 20sec.

Ως μήτρα για τις αντιδράσεις PCR των κατασκευών (6) και (7) χρησιμοποιήθηκε πλασμιδιακό DNA της κατασκευής πλήρους μήκους *Neurl/pGBT9*. Τα προϊόντα των αντιδράσεων PCR, ύστερα από κατεργασία των άκρων με τις περιοριστικές ενδονουκλεάσες *EcoRI* και *BamHI*, κλωνοποιήθηκαν στον πλασμιδιακό φορέα *pGBT9*, ο οποίος είχε προηγουμένως γραμμοποιηθεί με τα ίδια ένζυμα.

η) *Neurl (227-404)/pEGFPC1* : Η πλασμιδιακή κατασκευή *Neurl (227-404)/pEGFPC1* (βλ. Υλικά και Μεθόδους) τεμαχίστηκε με *EcoRI* και *HincII* και απομονώθηκε η ζώνη μήκους 530bp, που κλωνοποιήθηκε στον πλασμιδιακό φορέα *pEGFPC1*, ο οποίος είχε γραμμοποιηθεί με *EcoRI* και *SmaI*. Η πλασμιδιακή κατασκευή που προέκυψε κωδικοποιεί την ανασυνδυσμένη πρωτεΐνη GFP-*Neurl(227-404)*.

θ) *Neurl (227-488)/pEGFPC1* : Η πλασμιδιακή κατασκευή *Neurl (227-404)/pEGFPC1* (βλ. Υλικά και Μεθόδους) τεμαχίστηκε με *AraI* και απομονώθηκε η ζώνη μήκους 5.5Kb, η οποία περιέχει τον φορέα *pEGFPC1* (4.7Kb) και τα 1466 νουκλεοτίδια από το 5' του ενθέματος. Από την επανακυκλοποίηση αυτής της ζώνης προέκυψε η πλασμιδιακή κατασκευή που κωδικοποιεί την ανασυνδυσμένη πρωτεΐνη GFP-*Neurl(227-488)*.

ι) *Neurl (404-574)/pEGFPC1* : Η πλασμιδιακή κατασκευή *Neurl (227-404)/pEGFPC1* (βλ. Υλικά και Μεθόδους) τεμαχίστηκε με *EcoRI* και *HincII* και απομονώθηκε η ζώνη μήκους 1Kb, που κλωνοποιήθηκε στον πλασμιδιακό φορέα *pEGFPC1*, ο οποίος είχε γραμμοποιηθεί με *HindIII* και μετατροπή των άκρων του σε τυφλά με *Klenow*. Η πλασμιδιακή κατασκευή που προέκυψε κωδικοποιεί την ανασυνδυσμένη πρωτεΐνη GFP-*Neurl(404-574)*.

3. Κλωνοποίηση των cDNA-‘θηραμάτων’ που προέκυψαν από το πείραμα δύο υβριδίων σε φορείς έκφρασης

Για την έκφραση των cDNAs που απομονώθηκαν στο πείραμα δύο υβριδίων με ‘δόλωμα’ την πρωτεΐνη NeurI σε βακτήρια και σε κύτταρα θηλαστικών, αυτά κλωνοποιήθηκαν στους φορείς έκφρασης pGEX-3x και pCDNA3.1HisA, αντίστοιχα.

α) cDNA-prey / pGEX-3x : Τα ‘prey’ cDNAs αποκόπηκαν από τον πλασμιδιακό φορέα pVP16 με BamHI και EcoRI και κλωνοποιήθηκαν στις ίδιες θέσεις του φορέα pGEX-3x, στο αναγνωστικό πλαίσιο της πρωτεΐνης GST. Προέκυψαν οι κατασκευές:

- 12(*Lek1*) / pGEX-3x, η οποία κωδικοποιεί τη χιμαιρική πρωτεΐνη GST-Lek1(1349-1616)
- 42(*Ubc9*) / pGEX-3x, η οποία κωδικοποιεί τη χιμαιρική πρωτεΐνη GST-Ubc9(1-155)
- 57(*Numa1*) / pGEX-3x, η οποία κωδικοποιεί τη χιμαιρική πρωτεΐνη GST- Numa1(905-1057)
- 26(*Dlc-1*) / pGEX-3x, η οποία κωδικοποιεί τη χιμαιρική πρωτεΐνη GST- Dlc-1(51-89)

β) cDNA-prey / pCDNA3.1His. Τα ακόλουθα ‘prey’ cDNAs αποκόπηκαν από τον πλασμιδιακό φορέα pVP16 με BamHI και EcoRI και κλωνοποιήθηκαν στις ίδιες θέσεις του φορέα pCDNA3.1HisA, ο οποίος κωδικοποιεί πεπτιδία σημασμένα στο αμινοτελικό άκρο με τον επίτοπο Xpress. Προέκυψαν οι κατασκευές:

- 57(*Numa1*) / pCDNA3.1HisA, η οποία κωδικοποιεί τη χιμαιρική πρωτεΐνη Xpress-Numa1(905-1057)
- 25(*Numa1*) / pCDNA3.1HisA, η οποία κωδικοποιεί τη χιμαιρική πρωτεΐνη Xpress-Numa1(984-1123)
- 12(*Lek1*) / pCDNA3.1HisA, η οποία κωδικοποιεί τη χιμαιρική πρωτεΐνη Xpress-Lek1(1349-1616)
- 42(*Ubc9*) / pCDNA3.1HisA, η οποία κωδικοποιεί τη χιμαιρική πρωτεΐνη Xpress-Ubc9(1-155)
- 26(*Dlc-1*) / pCDNA3.1HisA, η οποία κωδικοποιεί τη χιμαιρική πρωτεΐνη Xpress-Dlc-1(51-89)
- 36(*Peg3*) / pCDNA3.1HisA, η οποία κωδικοποιεί τη χιμαιρική πρωτεΐνη Xpress-Peg3(15-146)
- 20(*Gas8*) / pCDNA3.1HisA, η οποία κωδικοποιεί τη χιμαιρική πρωτεΐνη Xpress-Gas8(125-235)
- 41(*Dnmt3a*) / pCDNA3.1HisA, η οποία κωδικοποιεί τη χιμαιρική πρωτεΐνη Xpress-Dnmt3a(406-539)
- 45(*Evpl*) / pCDNA3.1HisA, η οποία κωδικοποιεί τη χιμαιρική πρωτεΐνη Xpress-Evpl(270-372)
- 55(*Ankycorbin*) / pCDNA3.1HisA, η οποία κωδικοποιεί τη χιμαιρική πρωτεΐνη Xpress-Ankycorbin(461-620).

Τέλος, το ένθεμα του κλώνου #34(*Lek1*), αποκόπηκε από τον pVP16 με NotI και κλωνοποιήθηκε στην ίδια θέση του φορέα pCDNA3.1HisC, ώστε να προκύψει η κατασκευή 34(*Lek1*)/ pCDNA3.1HisC, η οποία κωδικοποιεί τη χιμαιρική πρωτεΐνη Xpress-Lek1(492-607).

4. Κλωνοποίηση του cDNA του γονιδίου *Sumo-1*

α) *Sumo-1/pGEM-T Easy* : Μέρος της κωδικής περιοχής του γονιδίου *Sumo-1* του ποντικού (Genbank ID: NM_009460, νουκλεοτίδια 120-409) παρήχθη με αντίδραση PCR, στην οποία χρησιμοποιήθηκαν ως DNA-μήτρα πλασμιδιακά DNAs (50ng) της cDNA βιβλιοθήκης εμβρύου ποντικού E9.5 και οι εκκινητές *BamHI-Sumo-F* και *ScaI-Sumo-R* (Πίνακας 9). Για την αντίδραση χρησιμοποιήθηκε *Taq* πολυμεράση με ικανότητα επιδιόρθωσης σφαλμάτων (Expand High Fidelity PCR system, Roche, Cat.no 1732641). Πραγματοποιήθηκαν 35 κύκλοι θερμικής αποδιάταξης στους 95°C για 45sec, υβριδοποίησης των εκκινητών στους 62°C για 45sec και ενζυμικού πολυμερισμού στους 72°C για 30sec. Το προϊόν PCR (300bp) κλωνοποιήθηκε στον φορέα pGEM-T Easy, χρησιμοποιώντας το pGEM-T Vector System I (Promega Cat.no. A3600).

β) *Sumo-1(1-97)/pGEX-3x* : Το ένθεμα προέκυψε τεμαχίζοντας την κατασκευή *Sumo-1/pGEM-T Easy* με τα ένζυμα *BamHI* και *ScaI* και απομονώνοντας τη ζώνη μεγέθους 300bp, η οποία στη συνέχεια κλωνοποιήθηκε σε φορέα pGEX-3x, γραμμοποιημένο με τα ένζυμα *BamHI* και *SmaI*. Η κατασκευή αυτή κωδικοποιεί την ενεργή μορφή του μορίου *Sumo-1* (αμινοξέα 1-97) συντηγμένη στο αμινοτελικό της άκρο με την πρωτεΐνη GST.

γ) *Sumo-1(1-97)/pEGFP-C1* : Το ένθεμα προέκυψε τεμαχίζοντας την κατασκευή *Sumo-1/pGEM-T Easy* αρχικά με *BamHI*, κατασκευή τυφλών άκρων με *Klenow*, τεμαχισμό με *ScaI*, απομόνωση της ζώνης μεγέθους 300bp και κλωνοποίησή της στον φορέα pEGFP-C1, ο οποίος είχε υποστεί γραμμοποίηση με *BglIII* και μετατροπή των άκρων του σε τυφλά με *Klenow*. Η κατασκευή αυτή κωδικοποιεί την ενεργή μορφή του μορίου *Sumo-1* (αμινοξέα 1-97) συντηγμένη στο αμινοτελικό της άκρο με την πρωτεΐνη GFP.

5. Κλωνοποίηση του cDNA του γονιδίου *Dlc-1*

α) *Dlc-1/pGEM-T Easy* : Η κωδική περιοχή του γονιδίου *Dlc-1* του ποντικού (Genbank ID: BC008106, νουκλεοτίδια 100-368) παρήχθη με αντίδραση PCR, στην οποία χρησιμοποιήθηκαν ως DNA-μήτρα πλασμιδιακά DNAs (50ng) της cDNA βιβλιοθήκης εμβρύου ποντικού E9.5 και οι εκκινητές *BamHI-DLC1-F* και *EcoRII-DLC1-R* (Πίνακας 9). Για την αντίδραση χρησιμοποιήθηκε *Taq* πολυμεράση με ικανότητα επιδιόρθωσης σφαλμάτων (Expand High Fidelity PCR system, Roche, Cat.no 1732641). Πραγματοποιήθηκαν 5 κύκλοι θερμικής ανακύκλωσης με υβριδοποίηση των εκκινητών στους 50°C για 45sec και 30 κύκλοι με υβριδοποίηση στους 62°C. Κάθε κύκλος περιελάμβανε στάδιο θερμικής αποδιάταξης στους 95°C για 45sec και σύνθεσης DNA στους 72°C για 30sec. Το προϊόν PCR (280bp) κλωνοποιήθηκε στον φορέα pGEM-T Easy, χρησιμοποιώντας το pGEM-T Vector System I (Promega Cat.no. A3600).

β) *Dlc-1/pGEX-3x* : Το ένθεμα προέκυψε από τον τεμαχισμό της κατασκευής *Dlc-1/pGEM-T Easy* με τα ένζυμα *BamHI* και *EcoRI*, απομόνωση του θραύσματος μεγέθους 280bp και κλωνοποίησή του στον φορέα

pGEX-3x, οποίος είχε γραμμοποιηθεί με τα ίδια ένζυμα. Η κατασκευή αυτή κωδικοποιεί την πρωτεΐνη Dlc-1 (Dynein light chain 1, 8KD), συντηγμένη στο αμινοτελικό της άκρο με την πρωτεΐνη GST.

γ) *Dlc-1/pCDNA3.1HisA* : Το ένθεμα προέκυψε από τον τεμαχισμό της κατασκευής *Dlc-1/pGEM-T Easy* με τα ένζυμα BamHI και EcoRI, απομόνωση του θραύσματος μεγέθους 280bp και κλωνοποίησή του στον φορέα pCDNA3.1HisA, ο οποίος είχε γραμμοποιηθεί με τα ίδια ένζυμα. Η κατασκευή αυτή κωδικοποιεί την πρωτεΐνη Dlc-1 (89 αμινοξέα), συντηγμένη στο αμινοτελικό της άκρο με τον συνθετικό επίτοπο Xpress.

ΑΠΟΤΕΛΕΣΜΑΤΑ

1. Συμβολή στη δημιουργία του φυσικού χάρτη του χρωμοσώματος 10 του ανθρώπου

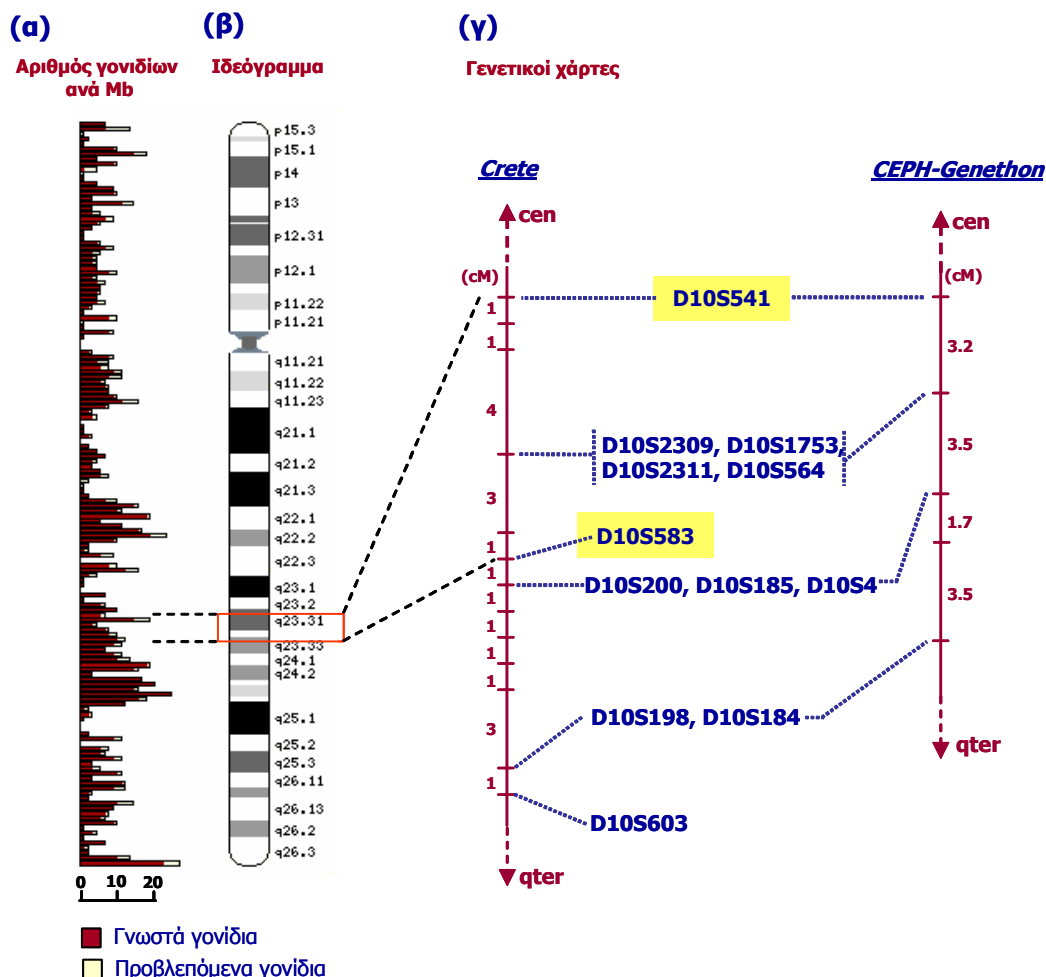
Για την αλληλούχηση του χρωμοσώματος 10 από τα συνεργαζόμενα εργαστήρια του διεθνούς προγράμματος ανάλυσης του ανθρώπινου γονιδιώματος, ήταν απαραίτητη η δημιουργία ενός φυσικού χάρτη διατεταγμένων κλώνων γονιδιωματικού DNA (*contig map*). Οι κλώνοι αυτοί αποτελούν την πρώτη ύλη για τα πειράματα προσδιορισμού της νουκλεοτιδικής αλληλουχίας του χρωμοσώματος. Οι κλώνοι τεχνητών χρωμοσωμάτων βακτηριοφάγου P1 (PACs) και βακτηρίων (BACs) είναι ιδανικοί για τον σκοπό αυτό, επειδή συνδυάζουν το μεγάλο μήκος ενθέματος (100-200 Kb) με αυξημένη σταθερότητα [Bentley *et al*, 2001].

Στα πλαίσια της δημιουργίας του φυσικού χάρτη του χρωμοσώματος 10, συλλέχθηκαν κλώνοι ανθρώπινου γονιδιωματικού DNA από βιβλιοθήκες PAC (RPCI-1) και BAC (RPCI-11), για την περιοχή του χρωμοσώματος ανάμεσα στους γενετικούς δείκτες D10S541 και D10S583, η οποία χαρτογραφείται στη ζώνη 10q23.31–q23.33. Επιλέξαμε αυτή την περιοχή επειδή εμφανίζει υψηλή πυκνότητα σε γονίδια (Εικ.31). Η συλλογή κλώνων PAC και BAC στηρίχθηκε: α) σε κλώνους YAC (*Yeast Artificial Chromosomes*) της περιοχής του γενετικού δείκτη D10S2311, χαρτογραφημένους από προηγούμενες μελέτες μας της αρχικής φάσης του προγράμματος φυσικής χαρτογράφησης του χρωμοσώματος 10 [Marzella, Kokkinaki *et al*, 1997], και β) σε αλληλουχίες EST (*Expressed Sequence Tags*), χαρτογραφημένες με την τεχνολογία των ακτινοβολημένων κυτταρικών υβριδίων.

α. Επιλογή κλώνων PAC

Η βιβλιοθήκη RPCI-1 υβριδοποιήθηκε με ραδιοσημασμένα προϊόντα αντιδράσεων Alu-PCR. Για τη σύνθεση των ανιχνευτών χρησιμοποιήθηκαν εκκινητές που αναγνωρίζουν τις χαρακτηριστικές για το ανθρώπινο DNA επαναλήψεις Alu, σε συνδυασμό με DNA κλώνων YAC (8EH2, 10BA1, 32GE7 και 37IF9) που χαρτογραφούνταν στην περιοχή του γενετικού δείκτη D10S2311. Τα YACs αυτά είχαν απομονωθεί προηγουμένως διερευνώντας με PCR τη βιβλιοθήκη ICI των Anand *et al* (1990) για τον γενετικό δείκτη D10S2311. Στη συνέχεια, διαπιστώθηκε ότι στα ίδια YACs περιέχονται και οι δείκτες D10S1753 και D10S564, οι οποίοι σύμφωνα με τον γενετικό χάρτη *CEPH-Genethon* (www.cephb.fr/ceph-genethon-map.html), καθώς και δεδομένα γενετικής χαρτογράφησης του εργαστηρίου μας [Marzella, Kokkinaki *et al*, 1997], εμφανίζουν μηδενική γενετική απόσταση από τον D10S2311 (Εικ.31). Για καθένα από τα YACs, πραγματοποιήθηκαν αντιδράσεις Alu-PCR με τους εκκινητές CL1, CL2, AluS, AluJ, Alu-end και Alu-4723 (Υλικά και Μέθοδοι). Τα προϊόντα των αντιδράσεων είχαν μέγεθος που κυμαίνονταν από 0.1 έως και 5Kb.

Ανάλογα με το μέγεθός τους, ραδιοσημάνθηκαν είτε στο 5' άκρο τους, είτε με τη μέθοδο των τυχαίων εκκινήτων και στη συνέχεια χρησιμοποιήθηκαν για τη διερεύνηση της βιβλιοθήκης RPCI-1. Από τα πειράματα αυτά προέκυψαν 22 κλώνοι PAC (Πίνακας 15). Ακολούθησαν πειράματα χαρτογράφησης των κλώνων.



Εικ.31 Η περιοχή του ανθρώπινου χρωμοσώματος 10 μεταξύ των δεικτών D10S541 και D10S583. Παριστάνονται διαγραμματικά: **α)** η κατανομή αφθονίας γνωστών ή προβλεπόμενων γονιδίων (γονιδιακή πυκνότητα), με βάση τα πρόσφατα στοιχεία της βάσης δεδομένων *Ensembl* (www.ensembl.org/Homo_sapiens/mapview?chr=10), **β)** το ιδεόγραμμα του χρωμοσώματος, με βάση το πρότυπο ζώνωσης μετά από χρώση *Giemsa*, **γ)** Οι γενετικοί χάρτες της περιοχής, από το εργαστήριό μας (*Crete*, [Marzella, Kokkinaki *et al*, 1997]) και από τα κέντρα γονιδιωματικής ανάλυσης *CEPH* και *Genethon* (www.cephb.fr/ceph-genethon-map.html). Φαίνονται οι αποστάσεις σε *centiMorgan* (cM) ανάμεσα στους γενετικούς δείκτες.

Χαρτογράφηση των κλώνων PAC με βάση το περιεχόμενό τους σε γενετικούς δείκτες

Για να χαρτογραφηθούν οι επιλεγμένοι κλώνοι PAC με βάση το περιεχόμενό τους σε γνωστούς γενετικούς δείκτες, απομονώθηκε από αυτούς πλασμιδιακό DNA, στο οποίο ανιχνεύθηκαν με αντίδραση PCR οι γενετικοί δείκτες που περιείχονταν και στα YACs της περιοχής (D10S1753, D10S564 και D10S2311). Συγκεκριμένα, ο γενετικός δείκτης D10S1753 ανιχνεύθηκε σε δύο κλώνους PAC (P15 και P22), ο D10S564 σε οκτώ (P1, P7, P8, P11, P12, P14, P17 και P21), ενώ ο D10S2311 σε πέντε (P7, P13, P14, P17 και P18). Οι γενετικοί δείκτες D10S564 και D10S2311 ανιχνεύθηκαν ταυτόχρονα σε τρεις κλώνους (P7, P14 και P17), γεγονός που επιβεβαιώνει τη μικρή φυσική τους απόσταση. Οι υπόλοιποι δέκα κλώνοι δεν ήταν δυνατόν να

χαρτογραφηθούν με αυτή την προσέγγιση επειδή δεν περιείχαν κανέναν από τους παραπάνω γενετικούς δείκτες αλλά ούτε και κάποιον από τους γειτονικούς (D10S2309, D10S1735).

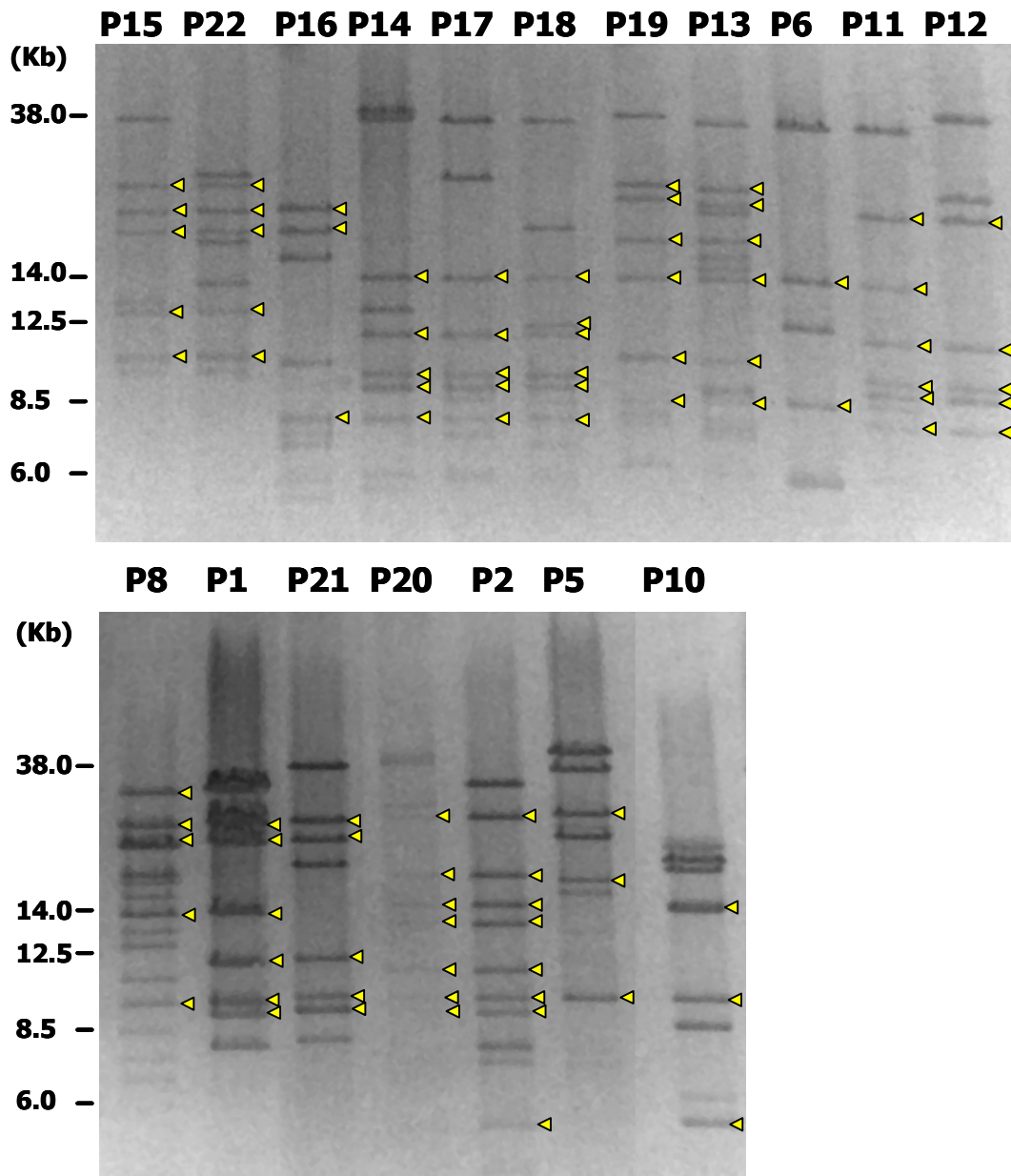
Πίνακας 15. Κλώνι PAC της περιοχής του γενετικού δείκτη D10S2311

#	ID	(Kb)	#	ID	(Kb)
P1	216P14	200	P12	74P12	120
P2	219B14	100	P13	63O24	220
P3	226F15	110	P14	27H17	340
P4	264C24	100	P15	23G1	120
P5	241M10	100	P16	37J13	100
P6	241N10	100	P17	156B16	120
P7	241O11	100	P18	172M13	140
P8	308A19	280	P19	183E18	200
P9	300I22	70	P20	173J4	100
P10	32O16	100	P21	105N10	170
P11	82C17	160	P22	135E5	200

Σημ: Για καθέναν από τους επιλεγμένους κλώνους PAC της βιβλιοθήκης RPCI-1, αναφέρεται ο αύξοντας αριθμός που του δόθηκε (#), το κωδικό του όνομα (ID) και το μέγεθός του (Kb), όπως υπολογίστηκε από την ανάλυση με ένζυμα περιορισμού.

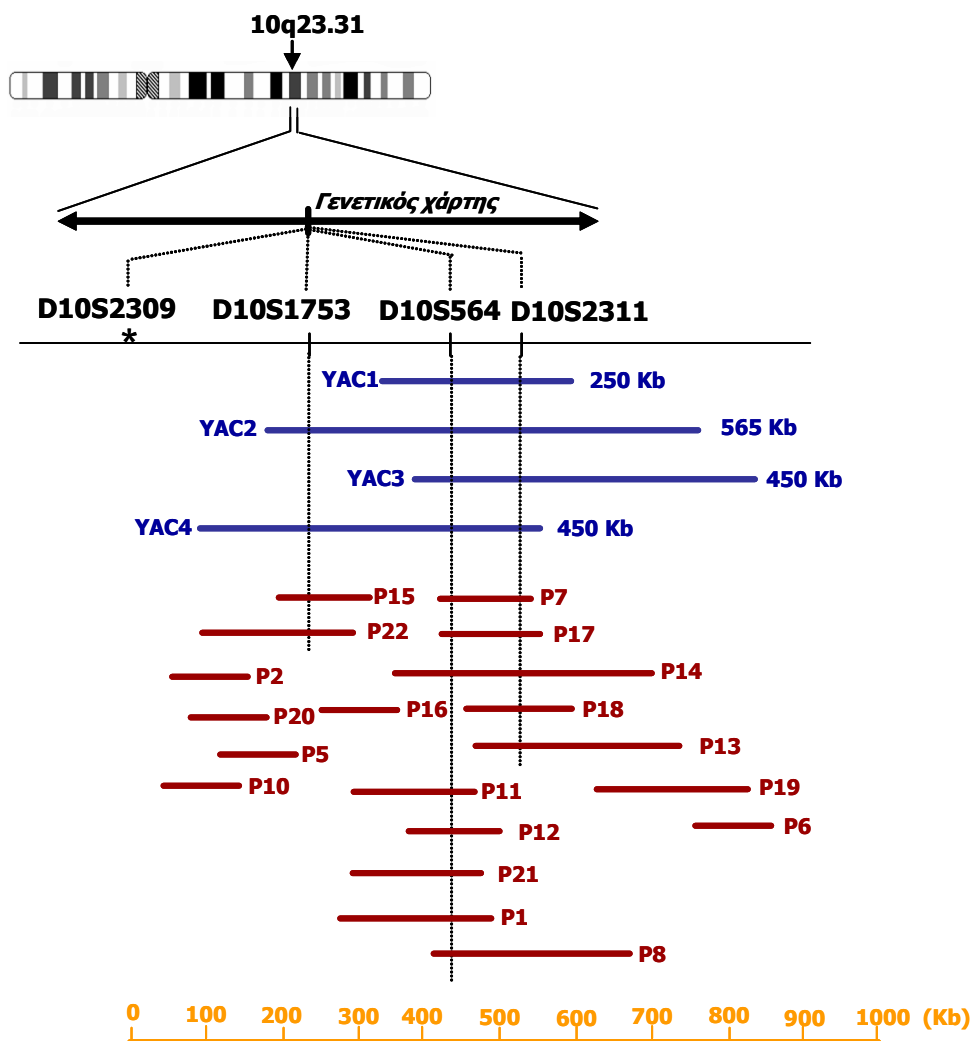
Χαρτογράφηση των κλώνων PAC με βάση τις ομοιότητες στο πρότυπο ενζυμικού τεμαχισμού

Για να χαρτογραφηθούν οι επιλεγμένοι κλώνοι PAC με βάση τις μεταξύ τους ομοιότητες στο πρότυπο τεμαχισμού τους από ένζυμα περιορισμού, το DNA κάθε κλώνου τεμαχίστηκε με την περιοριστική ενδονουκλεάση BamHI. Επειδή τα ενθέματα των κλώνων κατά την κατασκευή της συγκεκριμένης βιβλιοθήκης είχαν προκύψει από μερικό τεμαχισμό ανθρώπινου γονιδιωματικού DNA με την περιοριστική ενδονουκλεάση MboI και είχαν κλωνοποιηθεί στη συμβατή μοναδική θέση BamHI του φορέα pCYPAC2N [Ιωαννου & de Jong, 1996], η BamHI αναμένονταν να αναγνωρίσει αρκετές θέσεις στο εσωτερικό των ενθεμάτων, ώστε να μας επιτρέψει τέτοιου είδους ανάλυση. Τα τεμαχισμένα DNAs αναλύθηκαν σε ηλεκτροφόρηση παλλόμενου ηλεκτρικού πεδίου (PFGE), σε συνθήκες που ευνοούν τον διαχωρισμό μορίων DNA μεγέθους 5-50 Kb (Εικ.32). Στη συνέχεια, προσδιορίστηκαν οι ομάδες των κλώνων που εμφάνιζαν επικαλύψεις στο πρότυπο των ζωνών (Εικ.32,33). Η ηλεκτροφόρηση επαναλήφθηκε με κατάλληλη σειρά των δειγμάτων, ώστε να επιβεβαιωθεί η παρουσία κοινών ζωνών σε αυτά. Από την ανάλυση αυτή επιβεβαιώθηκε η διάταξη των κλώνων που είχαν χαρτογραφηθεί με βάση το περιεχόμενό τους σε γενετικούς δείκτες και επιπλέον χαρτογραφήθηκαν 7 κλώνοι από τους 10 που δεν περιείχαν γνωστούς γενετικούς δείκτες (Εικ.33). Συγκεκριμένα, οι κλώνοι P2, P20, P10 και P5 εμφάνιζαν επικαλύψεις μεταξύ τους αλλά και με τον κλώνο P22, ο οποίος περιείχε τον γενετικό δείκτη D10S1735. Επιπλέον, οι κλώνοι P6 και P19 επικαλύπτονταν με τον κλώνο P13, ο οποίος περιείχε τον γενετικό δείκτη D10S2311, ενώ ο κλώνος P16 επικαλύπτονταν με τον P14, αλλά και με τους P15 και P22 (Εικ.33).



Εικ.32 Πρότυπο τεμαχισμού 18 κλώνων PAC με την περιοριστική ενδονουκλεάση **BamHI**. Πλασμιδιακό DNA (~2μg) από κάθε κλώνο επώαστηκε με 20units ενζύμου (BamHI, Minotech) για 3 hrs στους 37°C, παρουσία σπερμιδίνης 1mM και κατάλληλου ρυθμιστικού διαλύματος (Minotech), σε τελικό όγκο 30μl. Ακολούθησαν εκχυλίσεις με φαινόλη και χλωροφόρμιο, κατακρήμνιση με αιθανόλη, επαναδιάλυση σε 20μl TE (10mM Tris-Cl, 1mM EDTA, pH 8.0), ηλεκτροφόρηση σε παλλόμενο ηλεκτρικό πεδίο (PFGE) και χρώση με βρωμιούχο αιθίδιο. Τέλος, το πήκτωμα φωτογραφήθηκε σε υπεριώδη φωτισμό (φαίνονται οι αντίστοιχες αρνητικές εικόνες). Τα κίτρινα βέλη υποδεικνύουν τις κοινές ζώνες μεταξύ των δειγμάτων, που χρησιμοποιήθηκαν για την ανάλυση των επικαλύψεων των κλώνων PAC.

Με τις παραπάνω προσεγγίσεις, οι 19 από τους 22 κλώνους PAC τοποθετήθηκαν σε *contig* (Εικ.33). Από τον ενζυματικό τεμαχισμό των 22 κλώνων με BamHI ή με NotI υπολογίστηκαν και τα μεγέθη των ενθεμάτων, αθροίζοντας τις ζώνες κάθε πλασμιδίου και αφαιρώντας 16Kb που αντιστοιχούν στον πλασμιδιακό φορέα (Πίνακας 15).



Εικ.33 Χάρτης επικαλυπτόμενων κλώνων YAC και PAC (*contig*) της περιοχής του γενετικού δείκτη D10S2311. Οι κατακόρυφες γραμμές υποδεικνύουν τη θέση των αντίστοιχων γενετικών δεικτών στους γονιδιωματικούς κλώνους YAC (μπλε) και PAC (κόκκινο). Οι κλώνοι PAC τοποθετούνται με βάση τις επικαλύψεις στο πρότυπο ενζυμικού τεμαχισμού τους (βλ. Εικ.). Η ακριβής θέση του D10S2309 (αστερίσκος), ο οποίος δεν περιέχεται σε κανέναν από τους κλώνους, προσδιορίστηκε με βάση τα πρόσφατα στοιχεία αλληλούχησης της περιοχής (GenBank ID: NT_030059).

β. Φυσικός χάρτης των κλώνων YAC και PAC στην περιοχή του D10S2311

Με βάση το μέγεθος των κλώνων YAC και PAC, το περιεχόμενό τους σε γενετικούς δείκτες αλλά και πειράματα μελέτης των μεταξύ τους επικαλύψεων, διαπιστώθηκε ότι οι κλώνοι καλύπτουν συνολικά ένα διάστημα περίπου 800Kb στην περιοχή 10q23.31, περιέχοντας τους γενετικούς δείκτες D10S1753, D10S564 και D10S2311 (Εικ.33). Λόγω της μηδενικής γενετικής απόστασης των συγκεκριμένων δεικτών (Εικ.31) δεν ήταν δυνατόν να προσδιοριστεί ο προσανατολισμός των κλώνων ως προς το κεντρομερές και το τελομερές του μεγάλου βραχίονα του χρωμοσώματος 10. Όμως, διαπιστώθηκε η διάταξη των τριών γενετικών δεικτών ως εξής: D10S1753 – D10S564 – D10S2311. Οι φυσικές αποστάσεις μεταξύ των δεικτών προσδιορίστηκαν με βάση τα μεγέθη και τις επικαλύψεις των κλώνων PAC. Συγκεκριμένα, η απόσταση ανάμεσα στους δείκτες D10S1753 και D10S564 υπολογίζεται περίπου 200Kb, με βάση το μέγεθος και τη θέση του κλώνου P8, ο οποίος περιέχει τον D10S564, αλλά όχι τον D10S1753 (Εικ.33). Αντίστοιχα, από το

μέγεθος του μικρότερου από τους κλώνους που ταυτόχρονα περιέχουν τους D10S564 και D10S2311 (κλώνος P7, Εικ.33), προκύπτει ότι η μεταξύ τους απόσταση είναι μικρότερη των 100Kb. Με βάση τα παραπάνω, δεν ήταν δυνατή η ακριβής τοποθέτηση στον φυσικό χάρτη του γενετικού δείκτη D10S2309, ο οποίος εμφανίζει μηδενική απόσταση από τους παραπάνω δείκτες στον γενετικό χάρτη (Εικ.31). Όμως, επειδή δεν ανιχνεύεται στους κλώνους YAC και PAC που εξετάστηκαν, υποθέτουμε ότι τοποθετείται είτε σε απόσταση μεγαλύτερη από 200Kb από την πλευρά του δείκτη D10S1753, ή σε απόσταση μεγαλύτερη των 400Kb από την πλευρά του D10S2311.

Τα παραπάνω συμπεράσματα επιβεβαιώνονται με βάση τα πρόσφατα στοιχεία από την ανάλυση του ανθρώπινου γονιδιώματος (βλ. νουκλεοτιδική αλληλουχία με GenBank ID NT_030059). Πράγματι, ισχύει η προβλεπόμενη διάταξη των τριών γενετικών δεικτών και οι ακριβείς μεταξύ τους φυσικές αποστάσεις είναι 186,7 Kb (ανάμεσα στους D10S1753 και D10S564) και 94.6 Kb (ανάμεσα στους D10S564 και D10S2311). Επιπλέον, ο γενετικός δείκτης D10S2309 τοποθετείται πριν τον D10S1753, σε απόσταση 231.6 Kb, ενώ η διάταξη των τεσσάρων γενετικών δεικτών ως προς το κεντρομερές (cen) είναι: cen – D10S2309 - D10S1753 – D10S564 – D10S2311 (Εικ.33).

γ. Επιλογή κλώνων BAC

Στη συνέχεια επιλέχθηκαν κλώνοι BAC από τη βιβλιοθήκη RPCI-11, οι οποίοι περιείχαν γονίδια της περιοχής μεταξύ των γενετικών δεικτών D10S541 και D10S583. Η ιδιαίτερα υψηλή συχνότητα χαρτογραφημένων αλληλουχιών EST στην περιοχή (Εικ.31) μας επέτρεπε τη συλλογή μεγάλου αριθμού κλώνων, οι οποίοι θα χρησιμοποιούνταν για την αλληλούχηση και τη λεπτομερέστερη χαρτογράφηση αυτών των γονιδίων. Για την επιλογή των κλώνων BAC, η βιβλιοθήκη RPCI-11 υβριδοποιήθηκε με ομάδες ραδιοσημασμένων αλληλουχιών EST που προέκυψαν από ανθρώπινο γονιδιωμικό DNA ως προϊόντα αντιδράσεων PCR με ειδικούς εκκινητές (Υλικά και Μέθοδοι). Η επιλογή των ESTs πραγματοποιήθηκε με βάση τη θέση τους στους χάρτες ακτινοβολημένων κυτταρικών υβριδίων *Genemap'99* (www.ncbi.nlm.nih.gov/genemap99) και 'NCBI framework map' (www.ncbi.nlm.nih.gov/genome/guide). Τα ESTs επιλέγονταν ώστε να ανήκουν σε διαφορετικά γονίδια, με βάση την ομαδοποίησή τους σε μεταγραφικές μονάδες από τη βάση δεδομένων *Unigene* (www.ncbi.nlm.nih.gov/UniGene). Βέβαια, με την παράλληλη πρόοδο στην αλληλούχηση κωδικών αλληλουχιών, για ορισμένα από τα ESTs που αρχικά κατατάσσονταν σε διαφορετικές ομάδες *Unigene*, στη συνέχεια διαπιστώσαμε ότι αντιπροσωπεύουν το ίδιο γονίδιο. Το γεγονός αυτό δεν ήταν αναγκαστικά πλεονασμός στην πειραματική διαδικασία, εφόσον σε πολλά ανθρώπινα γονίδια έχει παρατηρηθεί μέγεθος που ξεπερνά τα 100Kb, ώστε να απαιτούνται δύο κλώνοι BAC για την πλήρη τους κάλυψη [Lander *et al*, 2001]. Στον Πίνακα 16 φαίνονται τα 32 ESTs που χρησιμοποιήθηκαν ως ανιχνευτές και οι ομάδες *Unigene* τις οποίες εκπροσωπούν. Με βάση σημερινά δεδομένα, τα 32 αυτά ESTs αντιπροσωπεύουν 26 διαφορετικά γονίδια, για 14 από τα οποία υπάρχουν στοιχεία για τη λειτουργία τους (βλ. Συζήτηση).

Πίνακας 16. Ανιχνευτές ESTs για τη βιβλιοθήκη RPCI-11

<i>Sanger ID</i>	<i>EMBL ID</i>	<i>Unigene</i>	Γνωστά γονίδια
ΟΜΑΔΑ I			
Stsg41319	D12036	Hs.116122	pantothenate kinase 1 (PANK1)
Stsg26781	D20012	Hs.364441	
Stsg41741	H14604	Hs.116122	pantothenate kinase 1 (PANK1)
Stsg32258	NP_065850	Hs.16229	
Stsg30644	AA169811	Hs.274413	
Stsg16023	R97513	Hs.116122	pantothenate kinase 1 (PANK1)
Stsg55729	A004R39	Hs.252839	interferon-induced protein with tetratricopeptide repeats 5 (IFIT5)
Stsg55566	L16782	Hs.240	M phase phosphoprotein 1 (MPP1)
Stsg2721	Z40249	Hs.95907	multiple inositol polyphosphate histidine phosphatase, 1 (MINPP1)
Stsg21644	N51865	Hs.177162	
ΟΜΑΔΑ II			
Stsg21645	N51873	Hs.47357	cholesterol 25-hydroxylase
Stsg29905	AA143609	Hs.169274	interferon-induced protein with tetratricopeptide repeats 2 (IFIT2)
Stsg22113	N63988	Hs.169274	interferon-induced protein with tetratricopeptide repeats 2 (IFIT2)
Stsg3666	H01317	Hs.274230	3'-phosphoadenosine 5'-phosphosulfate synthase 2 (PAPSS2)
Stsg15752	R78938	Hs.208641	actin, alpha 2 (ACTA2)
Stsg2551	Z39336	Hs.406224	
Stsg32004	N67037	Hs.355697	glutamate dehydrogenase 1 (GLUD1)
Stsg12880	H57760	Hs.37497	
Stsg8438	H64542	Hs.16369	
Stsg39120	L21195	Hs.73739	5-hydroxytryptamine (serotonin) receptor 7 (adenylate cyclase-coupled) (HTR7)
Stsg55704	H10009	Hs.73739	5-hydroxytryptamine (serotonin) receptor 7 (adenylate cyclase-coupled) (HTR7)
ΟΜΑΔΑ III			
Stsg55850	SHGC-35401	Hs.448589	ankyrin repeat domain 1 (ANKRD1)
Stsg31377	AA115498	Hs.246914	
Stsg15491	R63338	Hs.280776	tankyrase, TRF1-interacting ankyrin-related ADP-ribose polymerase 2 (TNKS2)
Stsg8384	H60064	Hs.37716	
Stsg55807	A006A33	Hs.100960	
Stsg35189	R49668	Hs.100960	
Stsg30011	AA149620	Hs.71999	
Stsg31710	AA194926	Hs.500529	
Stsg2803	R36999	Hs.131945	cytoplasmic polyadenylation element binding protein 3 (CPEB3)
Stsg12559	H04015	Hs.180930	B-TFIID transcription factor-associated (BTAF1), 170kD
Stsg35957	N68018	Hs.180930	B-TFIID transcription factor-associated (BTAF1), 170kD

Σημ: Τα ESTs επιλέχθηκαν με βάση τη χαρτογράφησή τους στη συγκεκριμένη περιοχή στους διαθέσιμους χάρτες ακτινοβολημένων κυτταρικών υβριδίων. Για κάθε EST αναφέρεται ο κωδικός *Sanger* και *EMBL* και η ομάδα *Unigene* στην οποία ανήκει. Για κάθε γνωστό γονίδιο αναφέρεται το όνομα της πρωτεΐνης που κωδικοποιεί. Οι ομάδες EST I, II και III αντιστοιχούν σε αυτές του Πίνακα 5.

Τα 32 ESTs τοποθετούνταν στους χάρτες ακτινοβολημένων κυτταρικών υβριδίων 'Genemap'99' και 'NCBI framework map' στο διάστημα μεταξύ των γενετικών δεικτών D10S541 και D10S583. Οι δείκτες

αυτοί περιέχονταν στους διαθέσιμους γενετικούς χάρτες του χρωμοσώματος 10 και η μεταξύ τους γενετική απόσταση υπολογίζονταν στα 10cM (Εικ.31). Όπως διαπιστώθηκε και αργότερα με την πλήρη αλληλούχηση της περιοχής [Lander *et al*, 2001], η φυσική απόσταση των δύο δεικτών δεν ξεπερνά τις 5Mb, γεγονός που υποδηλώνει αυξημένο ποσοστό μειωτικών ανασυνδυασμών στο τμήμα αυτό του χρωμοσώματος.

Πραγματοποιήθηκαν τρία πειράματα υβριδοποίησης της βιβλιοθήκης RPCI-11 με τις ομάδες ανιχνευτών I, II και III (Πίνακας 16). Αντίστοιχα, από κάθε υβριδοποίηση προέκυψαν 108, 134 και 74 θετικά σήματα και ταυτοποιήθηκαν οι αντίστοιχοι κλώνοι BAC, με βάση την ονοματολογία των κλώνων της συγκεκριμένης βιβλιοθήκης. Επειδή σε κάθε ομάδα ανιχνευτών χρησιμοποιήθηκαν ESTs από διαφορετικές ομάδες *Unigene*, δεν παρατηρήθηκαν κοινοί κλώνοι μεταξύ των διαφορετικών υβριδοποιήσεων. Χρησιμοποιώντας ως ανιχνευτές 26 διαφορετικά γονίδια, συλλέχθηκαν συνολικά 316 κλώνοι BAC (Πίνακας 17).

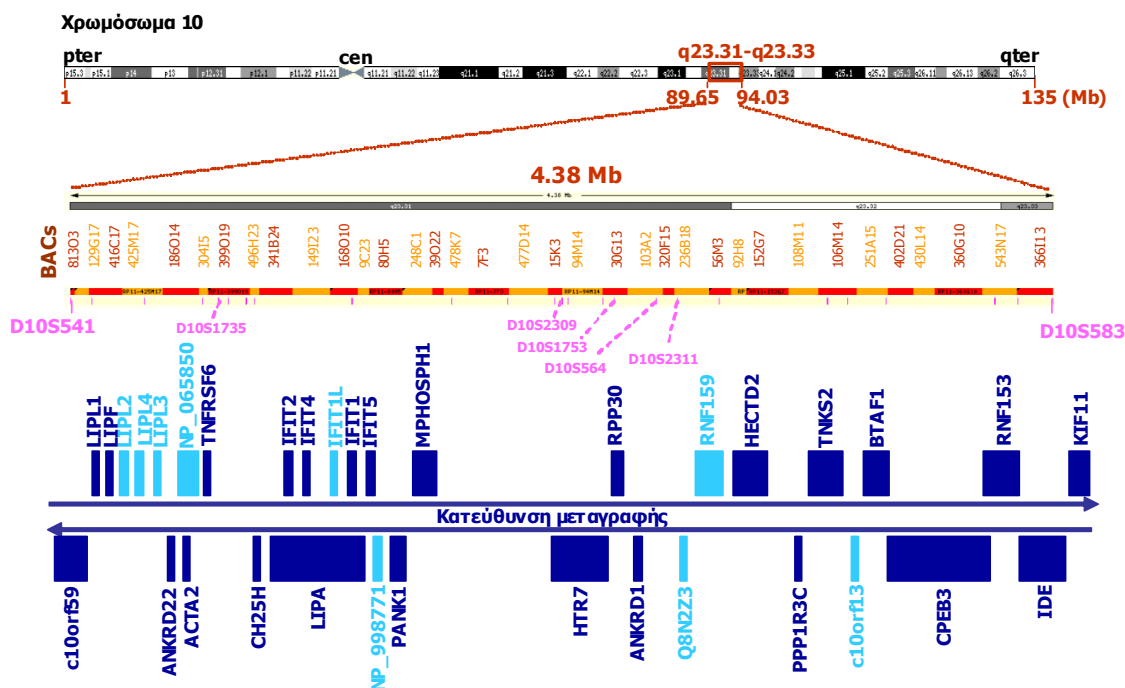
δ. Φυσικός χάρτης κλώνων BAC και αλληλούχηση της περιοχής D10S541-D10S583

Στη συνέχεια, οι κλώνοι PAC και BAC που επιλέχθηκαν από τη συγκεκριμένη χρωμοσωμική περιοχή, ενσωματώθηκαν σε μεγαλύτερες ομάδες επικαλυπτόμενων κλώνων (*contigs*), με στόχο την πλήρη κάλυψη του χρωμοσώματος 10 του ανθρώπου. Η κατασκευή των *contigs* στηρίχθηκε σε δεδομένα σύγκρισης του προτύπου τεμαχισμού των κλώνων με περιοριστικές ενδονουκλεάσες (*fingerprinting*), αλλά και σε επιπλέον υβριδοποιήσεις, στις οποίες χρησιμοποιούνταν ως ανιχνευτές ακραίες αλληλουχίες κλώνων BAC, μικροδορυφορικοί γενετικοί δείκτες καθώς και αλληλουχίες STS (*Sequenced Tag Sites*) ειδικές για το χρωμόσωμα 10. Τέλος, επιλέχθηκαν 35 κλώνοι BAC που κάλυπταν την περιοχή σχηματίζοντας το ελάχιστο 'μονοπάτι' (*tiling path*), οι οποίοι χρησιμοποιήθηκαν σε πειράματα τυχαίας (*shotgun*) αλληλούχησης. Τα πειράματα κατασκευής των *contigs* και η αλληλούχηση των γονιδιωματικών κλώνων πραγματοποιήθηκαν από το συνεργαζόμενο Ινστιτούτο *Sanger* (*Hinxton*, UK). Σύμφωνα με τον φυσικό χάρτη που προέκυψε, η περιοχή του χρωμοσώματος ανάμεσα στους γενετικούς δείκτες D10S541 και D10S583 έχει μήκος 4.38Mb. Η επεξεργασία της νουκλεοτιδικής αλληλουχίας από τα προγράμματα γονιδιακής πρόβλεψης της βάσης δεδομένων *Ensembl* κατέληξε στην ταυτοποίηση 51 γονιδίων *in silico*, 23 από τα οποία είχαν χρησιμοποιηθεί ως ανιχνευτές για την ταυτοποίηση των κλώνων BAC. Αξίζει να σημειωθεί ότι για τρεις από τους 26 διαφορετικούς ανιχνευτές (τα γονίδια *MINPP1*, *PAPSS2* και *GLUD1*, βλ. Πίνακα 16), ενώ με βάση τους χάρτες των 'radiation hybrids' χαρτογραφούνταν ανάμεσα στους δείκτες D10S541 και D10S583, η πλήρης αλληλούχηση της περιοχής έδειξε ότι τοποθετούνται σε απόσταση μικρότερη των 2 Mb πριν τον D10S541, προς το κεντρομερές. Στην Εικ. 34 φαίνεται η θέση των 34 προβλεπόμενων γονιδίων της περιοχής που κωδικοποιούν πρωτεΐνη, από τα οποία 25 αντιστοιχούν σε γνωστές πρωτεΐνες και 9 αντιστοιχούν σε άγνωστες μέχρι τώρα κωδικές αλληλουχίες που χαρτογραφήθηκαν από προγράμματα πρόβλεψης γονιδίων, λόγω της ομοιότητάς τους με κλώνους cDNA. Επιπλέον, στην περιοχή χαρτογραφήθηκαν και 10 γονίδια που κωδικοποιούν RNA, 7 θεωρητικά (*putative*) γονίδια και 17 ψευδογονίδια (Deloukas *et al*, 2004/ βλ. Συζήτηση).

Πίνακας 17. Επιλεγμένοι κλώνοι BAC για την περιοχή D10S541-D10S583

ESTs	Κλώνοι BAC
Ομάδα I Stsg41319 Stsg26781 Stsg41741 Stsg32258 Stsg30644 Stsg16023 Stsg55729 Stsg55566 Stsg2721 Stsg21644	108 BACs: 9C23, 21I24, 35C6, 40O16, 40N12, 47K24, 56N18, 57C13, 62J15, 64K2, 66J10, 71O3, 72B17, 73A16, 80H5, 85C24, 115J15, 115N17, 116F16, 119F19, 122D20, 125J8, 129G17, 135H15, 136P13, 138O9, 148D8, 159F21, 163J10, 163M19, 168O10, 172I22, 175I5, 176F8, 181I8, 188H10, 192D22, 199A4, 204L12, 218E3, 219J20, 220I11, 222G3, 233H8, 228A3, 229A6, 234L12, 248C1, 256D19, 257B22, 262D21, 262L20, 263D23, 263J2, 266F3, 266P5, 269E3, 276C24, 276D24, 300F14, 318P7, 319E5, 319G4, 322M19, 329L12, 332A11, 338H11, 360H16, 368C20, 375N21, 377E16, 380L16, 384G14, 388M4, 399L7, 400M3, 404P11, 408O9, 410H2, 410L2, 416B5, 425C21, 426L18, 433G8, 435J1, 443N19, 446L8, 454J2, 457H8, 458G8, 458K23, 462K11, 465M9, 469C6, 488G5, 489D4, 491K2, 498K17, 513B1, 518M4, 522E21, 525K6, 527H8, 528M22, 542J7, 545J24, 556N15, 575C13
Ομάδα II Stsg21645 Stsg29905 Stsg22113 Stsg3666 Stsg15752 Stsg2551 Stsg32004 Stsg12880 Stsg8438 Stsg39120 Stsg55704	134 BACs: 10B9, 11K21, 11O21, 15N11, 16A13, 18P4, 21I24, 30C6, 34D1, 35C6, 38O2, 41C12, 41M11, 57C13, 64K2, 64O14, 66I10, 66O15, 67N3, 77F13, 92E19, 95A18, 96C23, 96E23, 98K15, 98K17, 103A2, 106M22, 113H4, 115J15, 115N17, 119F19, 121L14, 122D20, 130D9, 131D2, 136P13, 146I16, 148D8, 149I23, 149J19, 155M6, 155M14, 159F21, 163J10, 163M19, 167H16, 168H1, 168M11, 175I5, 186P7, 187C5, 187H10, 201M14, 219J20, 220B16, 220I11, 222G3, 223O16, 224L16, 228A3, 231I18, 234L12, 235B19, 245A22, 245D15, 245J6, 249G24, 253L6, 256F2, 260B2, 262D21, 264A1, 264D5, 264E2, 264I7, 266F19, 267E5, 276C24, 276D24, 287K24, 288J8, 289A23, 289F23, 293L10, 303A10, 309B15, 319B14, 337H16, 338H11, 360H20, 362O4, 363B10, 364F9, 368C20, 377D9, 377E16, 392P7, 399O19, 401P6, 416C1, 421C21, 435J15, 442F10, 443N19, 448L14, 454J2, 460M16, 460N21, 463K12, 474H4, 474I15, 477D14, 478D6, 481N24, 488C24, 489D4, 491K20, 492L19, 494D18, 496G4, 498A22, 498B22, 498H16, 498I6, 513B1, 517O8, 518M4, 527H8, 528M22, 560G4, 567B5, 570L24
Ομάδα III Stsg55850 Stsg31377 Stsg15491 Stsg8384 Stsg55807 Stsg35189 Stsg30011 Stsg31710 Stsg2803 Stsg12559 Stsg35957	74 BACs: 11P23, 21B1, 27I3, 30E16, 43D10, 67L13, 70C15, 71F8, 81N7, 88A22, 90J1, 99P7, 112H11, 119K6, 126H3, 127J3, 160G7, 149I15, 150M17, 158J11, 162B9, 164M14, 180I12, 180I13, 192O5, 231H4, 231M20, 251A15, 254N8, 265L24, 267A16, 275G24, 291E10, 298H8, 303D7, 304L15, 305E9, 307A13, 320F15, 325B17, 329B8, 349I10, 368O6, 394H20, 402D21, 408A11, 412I14, 415L13, 420E3, 424D3, 430F14, 430L14, 439A24, 441G1, 441G9, 453J9, 459I18, 481N24, 484K9, 489M10, 492L22, 505K23, 508I18, 510B24, 530C19, 532L6, 534O5, 534O13, 539D21, 541K9, 558K13, 567B5, 567L6, 568I23

Σημ: Οι κλώνοι της βιβλιοθήκης RPCI-11 που επιλέχθηκαν, υβριδοποιώντας με τα ESTs της πρώτης στήλης, χωρισμένα σε τρεις ομάδες (I, II, III).



Εικ.34 Φυσικός χάρτης της περιοχής D10S541-D10S583. Η περιοχή αυτή τοποθετείται μεταξύ των χρωμοσωμικών ζωνών q23.31 και q23.33 του χρωμοσώματος 10 και έχει μήκος 4.38 Mb. Για την πλήρη αλληλούχηση της από το Ινστιτούτο *Sanger*, χρησιμοποιήθηκαν 35 ελάχιστα επικαλυπτόμενοι κλώνοι BAC της βιβλιοθήκης RPCI-11, των οποίων φαίνονται τα ονόματα. Επιπλέον, παρουσιάζονται τα γονίδια της περιοχής, όπως προβλέπονται από το πρόγραμμα *Ensembl*. Συνολικά υπάρχουν 34 γονίδια που κωδικοποιούν πρωτεΐνη στην περιοχή, από τα οποία 25 αντιστοιχούν σε γνωστές πρωτεΐνες (μπλε) και 9 αντιπροσωπεύουν νέα μεταγράφα (γαλάζιο). Στοιχεία από βάση δεδομένων *Ensembl* και Deloukas *et al*, 2004.

2. Ανάλυση γονιδιωματικής οργάνωσης του γονιδίου *Neuralized* (*Neurl*) του ποντικού

Για τη μελέτη της οργάνωσης του *Neurl* στο γονιδίωμα του ποντικού αναλύθηκαν τμήματα γονιδιωματικού DNA κλωνοποιημένα σε βακτηριοφάγους Lambda FIXII, τα οποία απομονώθηκαν από φαγική βιβλιοθήκη, καθώς και ολικό γονιδιωματικό DNA ποντικού στελέχους C57BL. Επιπλέον, μελετήθηκε η γονιδιωματική οργάνωση *in silico*, αξιοποιώντας δεδομένα από προγράμματα αλληλούχησης κλώνων cDNA ή γονιδιωματικού DNA του ποντικού, τα οποία έχουν καταχωρηθεί στις βάσεις dbEST, *Ensembl* και *Genbank*.

α. Χαρακτηρισμός των επιλεγμένων γονιδιωματικών κλώνων

Στα πειράματα αρχικής διερεύνησης της βιβλιοθήκης καθώς και σε ακόλουθα επιβεβαιωτικά πειράματα χρησιμοποιήθηκαν ως ανιχνευτές το ένθεμα του κλώνου cDNA *Neurl*(2433bp)/*pBluescript*, καθώς και 10 επικαλυπτόμενα τμήματά του, τα οποία παρασκευάστηκαν είτε με ενζυματικό τεμαχισμό του αρχικού κλώνου, είτε ως προϊόντα PCR με κατάλληλους εκκινητές (Υλικά και Μέθοδοι). Συνολικά πραγματοποιήθηκαν τέσσερα πειράματα διερεύνησης της φαγικής βιβλιοθήκης (A, B, C και D) σε καθένα από τα οποία ελέγχθηκαν 6x10⁵ βακτηριοφάγοι με διαφορετικούς κάθε φορά ανιχνευτές, ανάλογα με την περιοχή του γονι-

δίου για την οποία αναζητούσαμε γονιδιωματικούς κλώνους. Από τα πειράματα αυτά προέκυψαν 17 βακτηριοφάγοι (Πίνακας 18).

Πίνακας 18. Επιλεγμένοι βακτηριοφάγοι

Πείραμα	Ανιχνευτής	Επιλεγμένοι γονιδιωματικοί κλώνοι
A	2433	Aph1 , Aph3 , Aph5, Aph6, Aph8, Aph11
B	1A	Bph1 , Bph2, Bph3 , Bph4, Bph5, Bph7, Bph8 , Bph9
C	5, 9	Cph36 , Cph42
D	7	Dph12

Σημ: Οι βακτηριοφάγοι που προέκυψαν από κάθε πείραμα διερεύνησης της γονιδιωματικής βιβλιοθήκης ποντικού. Συνολικά προέκυψαν 21 φάγοι από τα τέσσερα πειράματα. Με έντονα γράμματα φαίνονται οι κλώνοι που στη συνέχεια βρέθηκε ότι αρκούν για να καλύψουν την αντίστοιχη γονιδιωματική περιοχή (οι υπόλοιποι είναι όμοιοι ή σημαντικά επικαλυπτόμενοι).

Οι επιλεγμένοι κλώνοι αναλύθηκαν λεπτομερέστερα κατά *Southern*, με ενζυματικό τεμαχισμό του DNA τους και υβριδοποιήσεις με κατάλληλους ανιχνευτές. Τα αποτελέσματα αυτής της ανάλυσης παρουσιάζονται στον Πίνακα 19. Το μέγεθος του ενθέματος κάθε βακτηριοφάγου (Πίνακας 19) προσδιορίστηκε με ενζυματικό τεμαχισμό DNA με NotI και ηλεκτροφόρηση σε πήκτωμα αγαρόζης, παρουσία κατάλληλου δείκτη μοριακών μεγεθών. Από τον Πίνακα 19 είναι φανερό ότι ολόκληρη η κωδική αλληλουχία του *Neurl* (1725 νουκλεοτίδια) αντιπροσωπεύεται στους επιλεγμένους γονιδιωματικούς κλώνους. Οι κλώνοι που επιλέχθηκαν για την υποκλωνοποίηση γονιδιωματικών υποπεριοχών με σκοπό την ανάλυση μέρους της νουκλεοτιδικής αλληλουχίας και τη μελέτη της οργάνωσης εξονίων – ιντρονίων του γονιδίου ήταν:

- Ο φάγος Cph42, οποίος περιέχει κωδική αλληλουχία από τον ανιχνευτή 1A (νουκλεοτίδια 13-218) μέχρι τον εκκινητή 8 (νουκλεοτίδιο 1382). Από αυτόν υποκλωνοποιήθηκαν τα θραύσματα 3.8Kb PstI και 3.9Kb BamHI.
- Ο φάγος Cph36, ο οποίος δεν υβριδοποιείται με τους 5' ανιχνευτές (A-1, 1, 1A και 5-10) αλλά επεκτείνεται καθοδικά του φάγου Cph42, εφόσον δίνει προϊόν PCR με το ζεύγος εκκινητών 8-12. Από αυτόν υποκλωνοποιήθηκε ένα θραύσμα PstI μεγέθους 4.8Kb.
- Ο φάγος Dph12, ο οποίος είναι ο μοναδικός που δίνει προϊόν PCR με το ζεύγος εκκινητών A-1, περιέχει τα πρώτα 85 νουκλεοτίδια κωδικής αλληλουχίας (θέση εκκινητή 1) και επεκτείνεται κυρίως προς την 5' πλευρική περιοχή του γονιδίου. Από αυτόν υποκλωνοποιήθηκε ένα θραύσμα EcoRI-NotI μήκους 15.8Kb καθώς και μικρότερα (1.4Kb PstI, 1.85Kb BstYI).

- Ο φάγος *Aph3*, από τον οποίο κλωνοποιήθηκε ένα θραύσμα *XbaI* μεγέθους 6Kb, που χρησιμοποιήθηκε για την αλληλούχηση τμήματος της 3' μη μεταφραζόμενη περιοχή του γονιδίου.

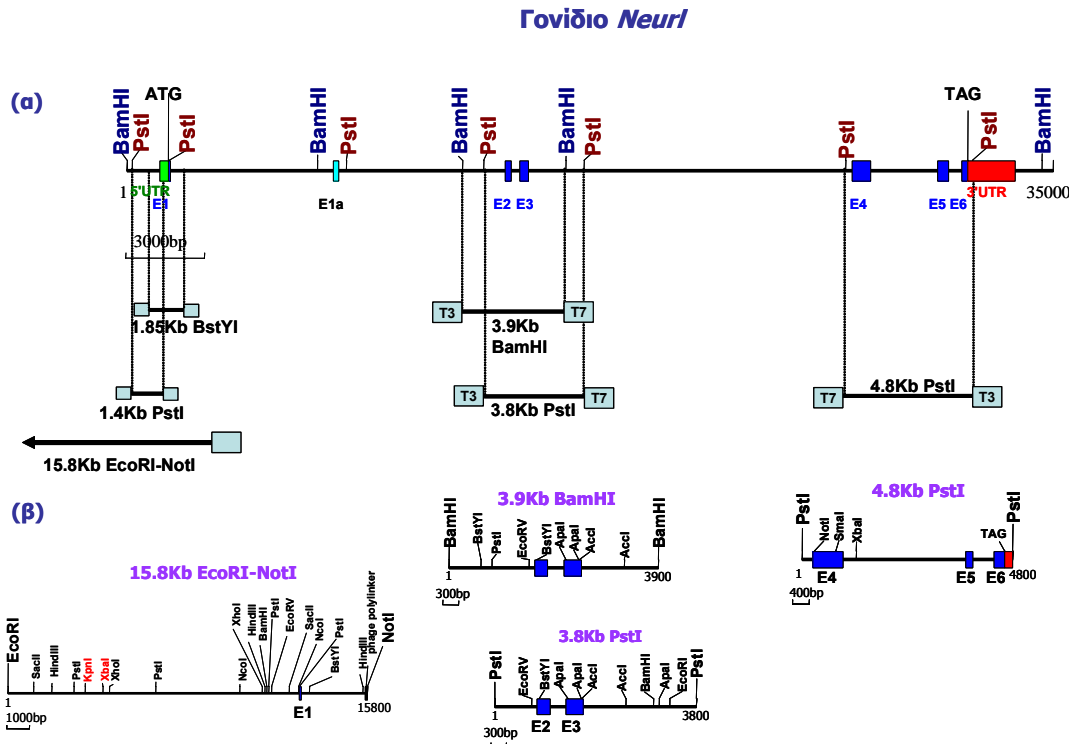
Πίνακας 19. Χαρακτηρισμός των γονιδιωματικών κλώνων

Ιχνηθέτης	A p h 1	A p h 3	A p h 5	A p h 6	A p h 8	A p h 11	B p h 1	B p h 2	B p h 3	B p h 4	B p h 5	B p h 7	B p h 8	B p h 9	C p h 36	C p h 42	D p h 12
2433	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
1	+	-	+	-	-	-	+	+	+	+	+	+	+	+	-	+	+
1A	-	-	+	-	-	-	+	+	+	+	+	+	+	+	-	+	+
2	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	-
3	+	+	-	+	+	+	-	-	-	-	-	-	-	-	+	+	-
4	+	-	-	+	-	+	+	+	+	+	+	+	+	+	+	+	-
5	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	-
6 (PCR A-1)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
7 (PCR 5-10)	±	-	+	-	-	-	+	+	+	+	+	+	+	+	-	+	-
8 (PCR 3-11)	±	-	±	-	-	-	±	±	±	±	±	±	±	±	±	±	-
9 (PCR 8-12)	+	+	-	+	-	+	±	±	±	±	±	±	±	±	+	±	-
Μέγεθος ενθέματος (Kb)	18	11	16	7.5	15	7.5	16	16	18	19	23	18	18	20	16	19	18

Σημ: Το αποτέλεσμα της ανάλυσης των 17 γονιδιωματικών κλώνων με υβριδοποιήσεις ή αντιδράσεις PCR, για τους αντίστοιχους ανιχνευτές. Με + συμβολίζεται η παρουσία σήματος στην υβριδοποίηση ή προϊόντος PCR και με - η απουσία. Για τα προϊόντα PCR που χρησιμοποιήθηκαν ως ανιχνευτές σε υβριδοποιήσεις, συμβολίζεται με ± η ταυτόχρονη παρουσία σήματος στην υβριδοποίηση αλλά απουσία προϊόντος PCR στους γονιδιωματικούς κλώνους, με τους αντίστοιχους εκκινητές. Το μέγεθος του ενθέματος κάθε γονιδιωματικού κλώνου σε Kb φαίνεται στην τελευταία σειρά του πίνακα. Οι ανιχνευτές που χρησιμοποιήθηκαν περιγράφονται στο Κεφάλαιο 'Υλικά και Μέθοδοι' (Πίνακας 10).

Τα θραύσματα γονιδιωματικού DNA που επιλέχθηκαν από τον ενζυματικό τεμαχισμό των βακτηριοφάγων, κλωνοποιήθηκαν στον πλασμιδιακό φορέα *pBluescript* και στη συνέχεια χρησιμοποιήθηκαν ως DNA-μήτρα σε αντιδράσεις προσδιορισμού της νουκλεοτιδικής αλληλουχίας με εκκινητές σχεδιασμένους από την αλληλουχία του *Neur1* cDNA ή με τους 'εξωτερικούς' εκκινητές T3 και T7 του πλασμιδίου *pBluescript*. Από αυτή την ανάλυση προσδιορίστηκαν τα όρια εξονίων-ιντρονίων του γονιδίου. Επιπλέον, σε αυτούς τους υπόκλωνους μελετήθηκε λεπτομερώς το πρότυπο των θέσεων αναγνώρισης μιας σειράς περιοριστικών ενζύμων, ώστε να χαρτογραφηθούν με ακρίβεια τα εξόνια του γονιδίου και να υπολογιστούν τα μεγέθη των ιντρονίων. Το αποτέλεσμα της χαρτογράφησης των θέσεων αναγνώρισης περιοριστικών ενδο-

νοκλεασών για τους υπόκλωνους 15.8Kb EcoRI-NotI, 3.9Kb BamHI, 3.8Kb PstI και 4.8Kb PstI φαίνεται στην Εικ.35.

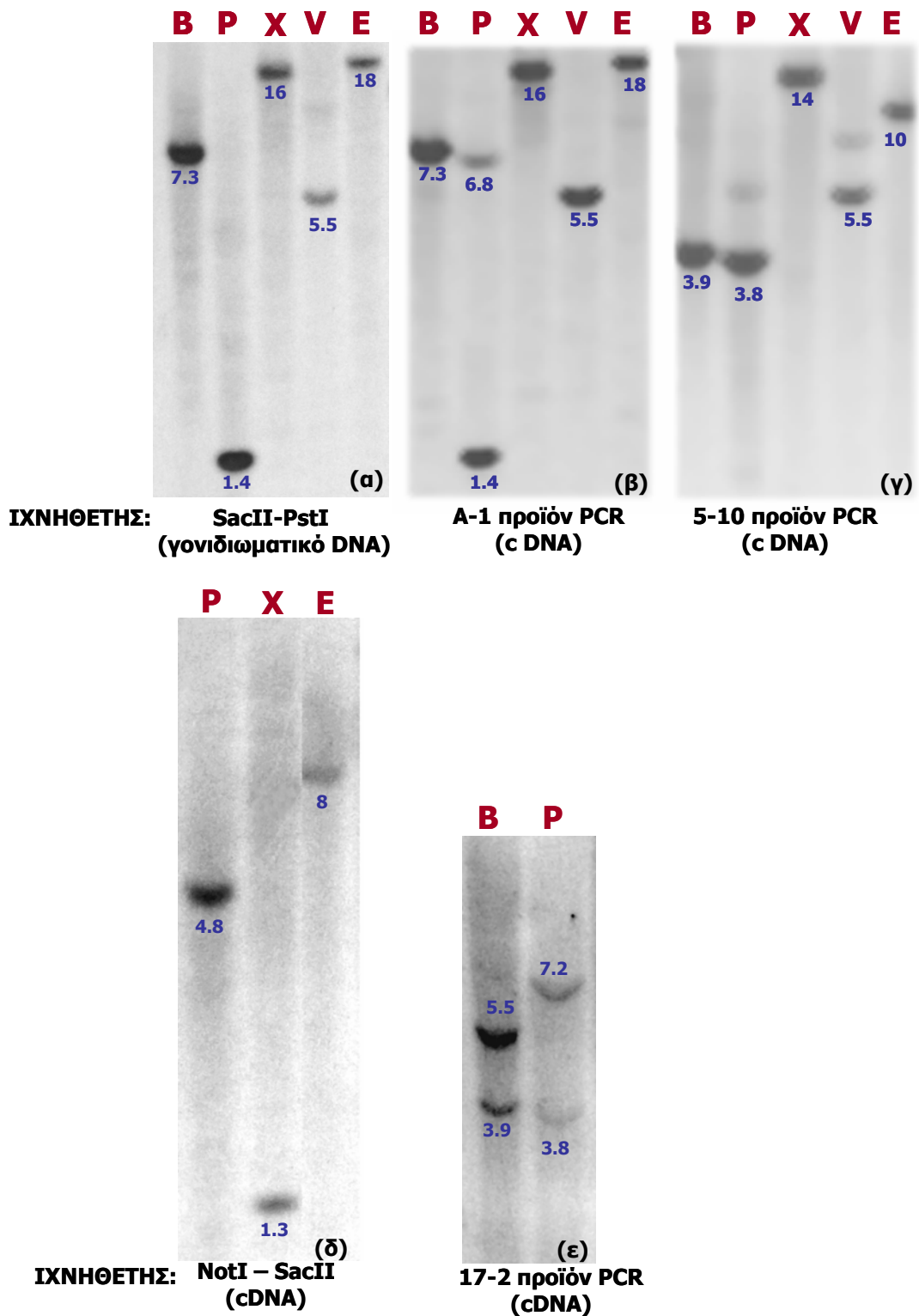


Εικ.35 Ανάλυση των γονιδιωματικών υποκλώνων με ένζυμα περιορισμού. (α) Οι υπόκλωνοι 1.85Kb BstYI, 1.4Kb PstI και 15.8Kb EcoRI-NotI περιέχουν το πρώτο εξόνιο (E1) του γονιδίου *Neur1*, οι υπόκλωνοι 3.9Kb BamHI και 3.8Kb PstI περιέχουν τα εξόνια E2 και E3, ενώ ο υπόκλωνος 4.8Kb PstI περιέχει τα εξόνια E4, E5 και E6. Η κλωνοποίηση κάθε υπόκλωνου περιγράφεται στο Κεφάλαιο 'Πλασμιδιακές Κατασκευές'. T3 και T7 είναι εκκινητές του πλασμιδιακού φορέα rBluescript. (β) Χαρτογράφηση των θέσεων αναγνώρισης περιοριστικών ενζύμων στους υπόκλωνους 15.8Kb EcoRI-NotI, 3.9Kb BamHI, 3.8Kb PstI και 4.8Kb PstI.

β. Ανάλυση γονιδιωματικού DNA

Ολικό γονιδιωματικό DNA ποντικού στελέχους C57BL αναλύθηκε κατά *Southern* και υβριδοποιήθηκε με κατάλληλους ραδιοσημασμένους ανιχνευτές. Σκοπός μας ήταν να διαπιστωθεί σε αυτό η παρουσία των ίδιων θραυσμάτων DNA που είχαν υποκλωνοποιηθεί από τους βακτηριοφάγους, ώστε να αποκλειστεί το ενδεχόμενο ανασυνδυασμού στο ένθεμα των επιλεγμένων φάγων. Στην Εικ.36 παρουσιάζεται το αποτέλεσμα της ανάλυσης με τους εξής ανιχνευτές: α) το τμήμα του γονιδιωματικού κλώνου Dph12 SacII- PstI, μήκους 400bp, που αντιστοιχεί στο 3' άκρο του υπόκλωνου 1.4Kb PstI (Εικ.35), β) Το προϊόν PCR A-1, γ) το προϊόν PCR 5-10, δ) το τμήμα του cDNA NotI(νουκλεοτίδιο 679) - SacII(νουκλεοτίδιο 1103), που αντιστοιχεί στο 4ο εξόνιο (ανιχνευτής 4, Πίνακας 10) και ε) Το προϊόν RT-PCR 17-2, το οποίο αντιπροσωπεύει τα εξόνια E1a¹) και E2.

¹ Ο εκκινητής 17 έχει σχεδιαστεί από τη διαθέσιμη αλληλουχία (EST AA667491) ενός εναλλακτικού εξονίου του *Neur1*, του E1a, το οποίο εντοπίστηκε *in silico* και χαρτογραφήθηκε στο πρώτο ιντρόνιο του γονιδίου. (βλ. παρακάτω)



Εικ.36 Το αποτέλεσμα της ανάλυσης ολικού γονιδιωματικού DNA ποντικού στελέχους C57BL με BamHI (B), PstI (P), XbaI (X), EcoRV (V) και EcoRI (E) και υβριδοποίησης με τους ανιχνευτές που αναφέρονται κάτω από κάθε εικόνα. Κάτω από κάθε ζώνη φαίνεται το μέγεθός της σε Kb.

Ο ανιχνευτής A-1 ανιχνεύει δύο θραύσματα γονιδιωματικού DNA τεμαχισμένου με PstI (Εικ.36β), μήκους 6.8 Kb και 1.4 Kb. Η μικρότερη ζώνη αντιστοιχεί στον 1.4 Kb PstI υπόκλωνο του φάγου Dph12, που αντιπροσωπεύει τη γονιδιωματική περιοχή 5' της αρχής μετάφρασης, ενώ η μεγαλύτερη αποτελεί μέρος του

πρώτου ιντρονίου. Η περιοχή SacII-PstI αποτελεί μέρος του 1.4Kb PstI υπόκλωνου και γι' αυτό το πρότυπο υβριδοποίησης με αυτόν τον ανιχνευτή επικαλύπτεται σημαντικά και επιβεβαιώνει το αποτέλεσμα του ανιχνευτή cDNA A-1 (Εικ.36α). Στην υβριδοποίηση με τον ανιχνευτή 5-10 (εξόνια 2 και 3) επιβεβαιώνεται η παρουσία, στο γονιδιωματικό DNA, των θραυσμάτων 3.9Kb BamHI και 3.8Kb PstI (υπόκλωνοι του φάγου Cph42) (Εικ.36γ). Στην υβριδοποίηση με μέρος του εξονίου 4 (Εικ.36δ), επιβεβαιώνεται η παρουσία του θραύσματος 4.8Kb PstI, (υπόκλωνος του φάγου Cph36). Τέλος, στην υβριδοποίηση με τον ανιχνευτή 17-2 (Εικ.36ε), αποκαλύπτεται ένα καινούριο θραύσμα BamHI μήκους 5.5Kb, το οποίο περιέχει το εξόνιο E1a. Από το τεμαχισμένο με PstI γονιδιωματικό DNA ανιχνεύεται το αναμενόμενο θραύσμα 3.9Kb καθώς και ένα θραύσμα μήκους 7.2Kb.

γ. Προσδιορισμός ορίων εξονίων - ιντρονίων

i) Με βάση το προϊόν PCR εκκινητών του Neurl cDNA (2^ο και 5^ο ιντρόνιο)

Στους γονιδιωματικούς κλώνους και σε ολικό γονιδιωματικό DNA ποντικού δοκιμάστηκε η αντίδραση PCR με τα ζεύγη των εκκινητών A-1, A-2, 5-10, 3-11 και 8-12. Από αυτά, τα ζεύγη 5-10 και 8-12 δίνουν προϊόντα μεγαλύτερα σε σχέση με αυτά που δίνουν στο cDNA, το ζεύγος A-1 δίνει προϊόν μεγέθους ίσου με το προβλεπόμενο από το cDNA μέγεθος (~140bp), ενώ για τα A-2 και 3-11 η αντίδραση δεν δίνει προϊόν. Επιπλέον, σχεδιάστηκε ο εκκινητής 17 από τη διαθέσιμη νουκλεοτιδική αλληλουχία του εναλλακτικού εξονίου E1a (βλ. παρακάτω) και δοκιμάστηκε η αντίδραση PCR σε γονιδιωματικούς κλώνους, σε ολικό γονιδιωματικό DNA καθώς και σε προϊόν αντίδρασης αντίστροφης μεταγραφής από ολικό RNA εγκεφάλου ποντικού με τον εκκινητή 5. Η μόνη περίπτωση στην οποία η αντίδραση PCR έδωσε προϊόν (~130bp) ήταν με μήτρα το προϊόν RT. Στη συνέχεια, προσδιορίστηκε η νουκλεοτιδική αλληλουχία των προϊόντων PCR 5-10 και 8-12 και συγκρίθηκε με την αντίστοιχη αλληλουχία cDNA. Στα σημεία συρραφής εξονίων-ιντρονίων αναγνωρίστηκαν τα χαρακτηριστικά συντηρημένα νουκλεοτιδία GT (5' άκρο ιντρονίου) και AG (3' άκρο ιντρονίου).

2^ο ιντρόνιο: Η θέση και το μέγεθός του προσδιορίστηκε από τη νουκλεοτιδική αλληλουχία των προϊόντων της αντίδρασης PCR με τους εκκινητές 5 και 10 σε μεμονωμένους γονιδιωματικούς κλώνους και σε ολικό γονιδιωματικό DNA. Οι εκκινητές αυτοί δίνουν σε γονιδιωματικό DNA προϊόν μεγαλύτερο κατά 300 περίπου ζεύγη βάσεων σε σχέση με αυτό που δίνουν στο cDNA. Η θέση του ιντρονίου εντοπίζεται μεταξύ των νουκλεοτιδίων 327 και 328 του cDNA (Εικ.37).

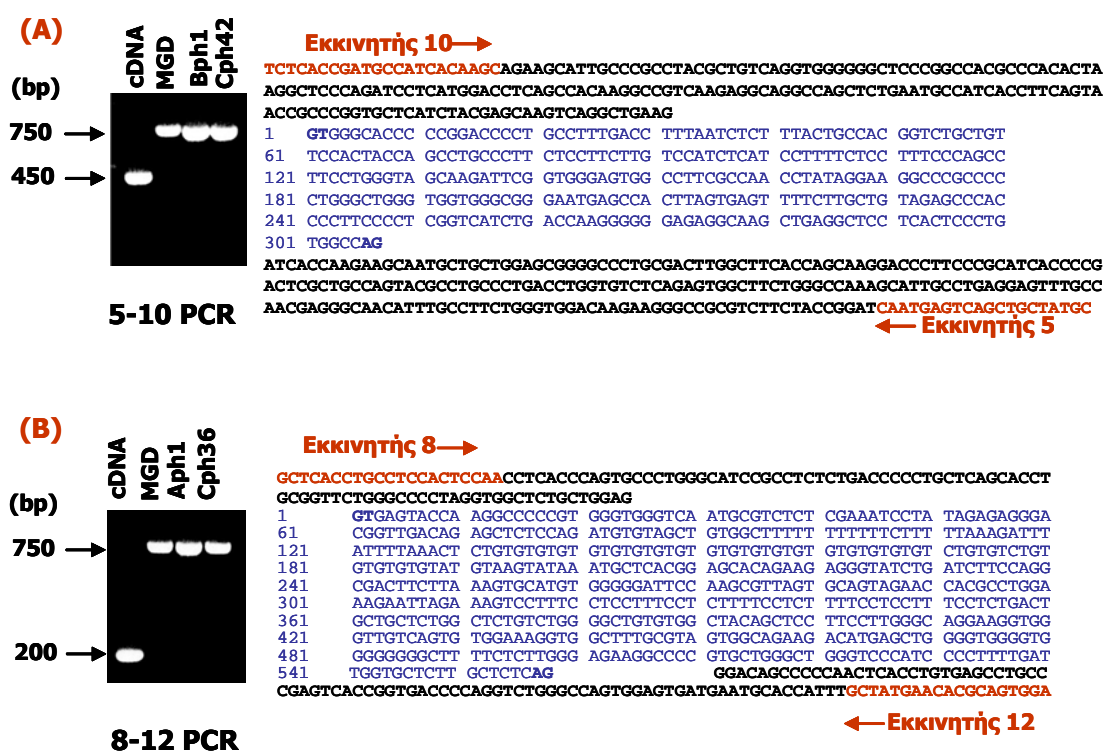
5^ο ιντρόνιο: Η θέση και το μέγεθός του προσδιορίστηκε από τη νουκλεοτιδική αλληλουχία των προϊόντων της αντίδρασης PCR με τους εκκινητές 8 και 12 σε μεμονωμένους γονιδιωματικούς κλώνους και σε ολικό γονιδιωματικό DNA. Οι εκκινητές αυτοί δίνουν σε γονιδιωματικό DNA προϊόν μεγαλύτερο κατά 500 περίπου ζεύγη βάσεων σε σχέση με αυτό που δίνουν στο cDNA. Η θέση του ιντρονίου εντοπίζεται μεταξύ των νουκλεοτιδίων 1486 και 1487 του cDNA (Εικ.37).

ii) Με βάση την νουκλεοτιδική αλληλουχία υποκλωνοποιημένων τμημάτων των φάγων

Για τα ιντρόνια που δεν ήταν δυνατή η αντίδραση PCR, πραγματοποιήθηκε υποκλωνοποίηση των τμημάτων που περιέχουν τα άκρα τους σε pBluescript και προσδιορισμός της νουκλεοτιδικής τους αλληλου-

χίας με εκκινητές του *Neur/cDNA* ή και 'εξωτερικούς' εκκινητές του pBluescript. Το μέγεθός τους μπορούσε να εκτιμηθεί λαμβάνοντας υπόψην το πρότυπο τεμαχισμού και υβριδοποίησης των φάγων και του ολικού γονιδιωματικού DNA.

1^ο ιντρόνιο: Το 5' άκρο του προσδιορίστηκε με τον εκκινητή A στον υπόκλωνο 1.85Kb BstYI του φάγου Dph12. Το 3' άκρο του προσδιορίστηκε με τον εκκινητή 2 στον υπόκλωνο 3.8Kb PstI του φάγου Cph42. Η θέση του προσδιορίστηκε μεταξύ των νουκλεοτιδίων 85 και 86 του cDNA. Το μέγεθος του 1ου ιντρονίου δεν ήταν δυνατό να προσδιοριστεί μόνο με βάση τη χαρτογράφηση των φάγων, γιατί η περιοχή αυτή δεν καλύπτονταν πλήρως από γονιδιωματικούς κλώνους. Μια εκτίμηση του μεγέθους αυτού του ιντρονίου έγινε με βάση την ανάλυση ολικού γονιδιωματικού DNA με τους δύο ανεξάρτητους τρόπους που περιγράφονται παρακάτω (Εικ. 38).

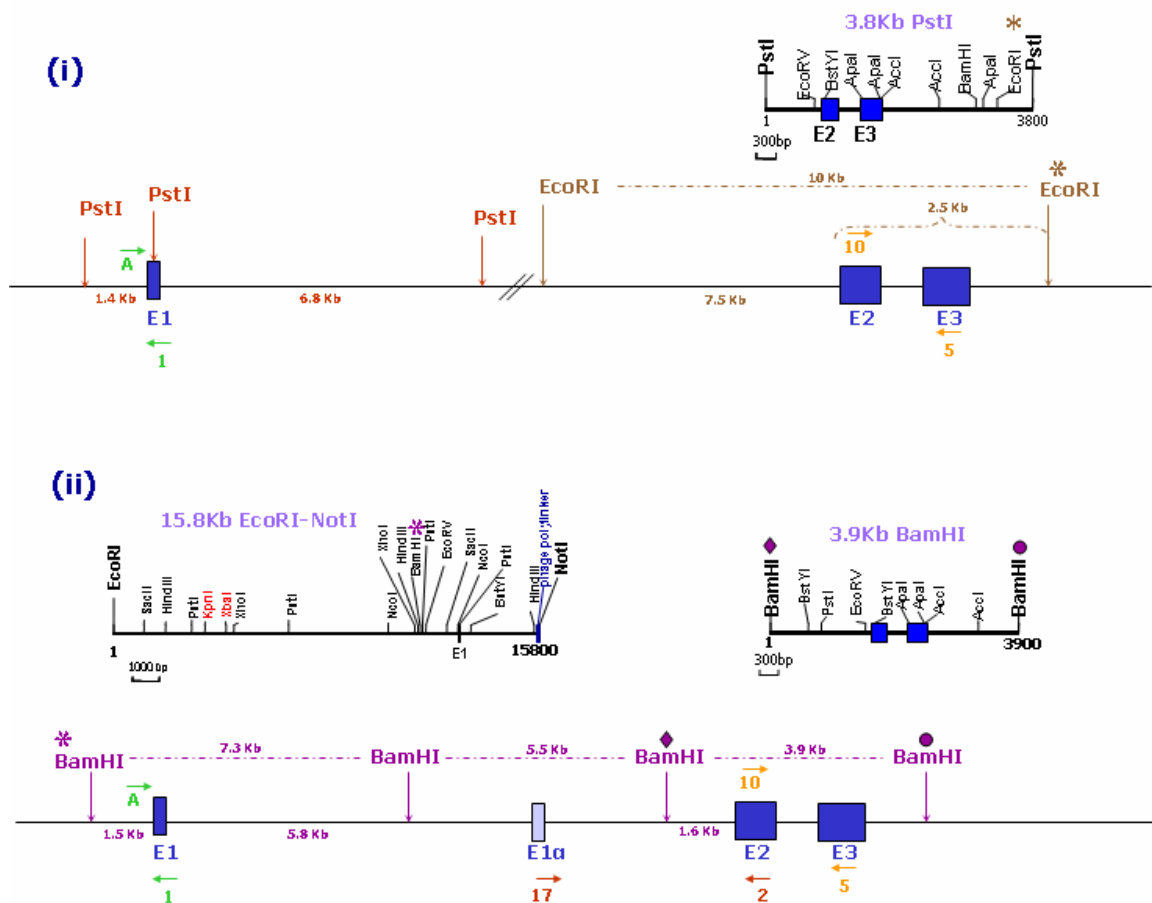


Εικ.37 Προσδιορισμός της θέσης και του μεγέθους των ιντρονίων 2 (A) και 5 (B), όπως αυτά προκύπτουν αναλύοντας τα προϊόντα PCR με τα ζεύγη εκκινητών 5-10 και 8-12, αντίστοιχα. Bph1 και Cph42 είναι δύο από τους φάγους που δίνουν προϊόν PCR με το ζεύγος εκκινητών 5-10. Aph1 και Cph36 είναι δύο από τους φάγους που δίνουν προϊόν PCR με το ζεύγος εκκινητών 8-12. MGD και cDNA είναι τα προϊόντα PCR με μήτρα 50ng γονιδιωματικού DNA ποτικού ή 5ng πλασμιδίου *Neur(2433)/pBluescript*, αντίστοιχα.

- Ο ανιχνευτής A-1 (Εικ.36β) αναγνωρίζει ένα θραύσμα PstI 6.8Kb (5' άκρο του 1^{ου} ιντρονίου), το οποίο περιέχει μόνο 50bp κωδικής αλληλουχίας (εξόνιο 1). Αντίστοιχα, για το 3' άκρο του ιντρονίου, ο ανιχνευτής 5-10 αναγνωρίζει στο γονιδιωματικό DNA ένα θραύσμα EcoRI μεγέθους 10Kb (Εικ.36γ). Από τη χαρτογράφηση του υπόκλωνου 3.8Kb PstI προκύπτει ότι υπάρχει μία θέση αναγνώρισης EcoRI που απέχει 2.5Kb από το 3' άκρο του εξονίου 2. Επομένως, 7.5Kb του γονιδιωματικού θραύσματος EcoRI ανήκουν στο 1^ο ιντρόνιο. Με την υπόθεση ότι τα δύο αυτά γονιδιωματικά θραύσματα δεν είναι επικαλυπτόμενα, το μέγεθος του 1^{ου} ιντρονίου μπορεί να υπολογιστεί στα (7.5 + 6.8) 14.4 Kb τουλάχιστον.

- Ο ανιχνευτής 17-2 (προϊόν PCR από τα εξόνια E1a και E2) αναγνωρίζει δύο γονιδιωματικά θραύσματα BamHI (Εικ.36ε): το 5.5Kb που περιέχει το εξόνιο E1a και το 3.9Kb που περιέχει τα εξόνια 2 και 3 (3.9Kb BamHI, υπόκλωνος φάγου Cph42). Από το τελευταίο, 1.6Kb ανήκουν στο 1^ο ιντρόνιο. Επιπλέον, ο ανιχνευτής A-1 αναγνωρίζει το BamHI 7.3Kb θραύσμα, από το οποίο τμήμα μήκους 5.8Kb ανήκει στο 1^ο ιντρόνιο. Με την υπόθεση ότι τα παραπάνω θραύσματα BamHI είναι συγγραμικά, το εκτιμώμενο μέγεθος του 1^{ου} ιντρονίου είναι (5.5 + 1.6 + 5.8) 12.9Kb τουλάχιστον.

Με βάση τα παραπάνω, το 1^ο ιντρόνιο του γονιδίου πρέπει να είναι μεγαλύτερο από 13-15Kb.



Εικ.38 Εκτίμηση του μεγέθους του πρώτου ιντρονίου με δύο τρόπους (i και ii), από τους οποίους προκύπτει μέγεθος μεγαλύτερο από 13-15Kb.

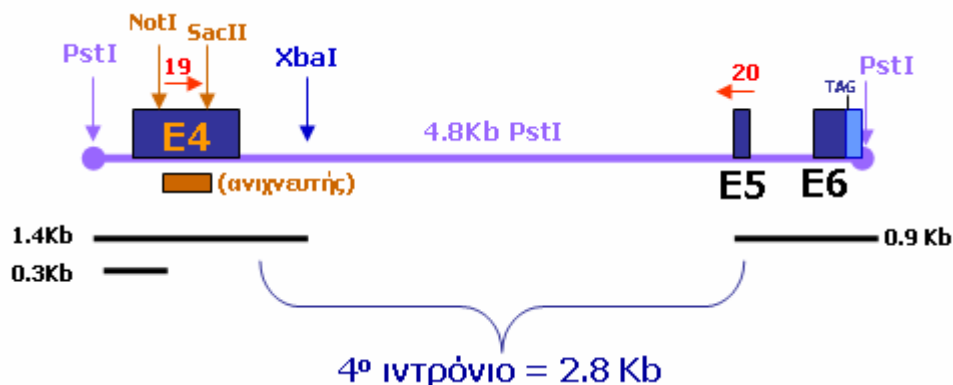
3^ο ιντρόνιο: Το 5' άκρο του προσδιορίστηκε με τον εκκινητή 3 στον υπόκλωνο 3.8Kb PstI του φάγου Cph42. Το 3' άκρο του προσδιορίστηκε με τον εξωτερικό εκκινητή T3 του υπόκλωνου 300bp PstI-NotI, ο οποίος προέκυψε από το τμήμα PstI 4.8Kb του φάγου Cph36 ύστερα από τεμαχισμό με NotI, η οποία αναγνωρίζει το νουκλεοτιδίο 679 του cDNA (εξόνιο 4) και μια θέση στην T7 πλευρική περιοχή της PstI. Η θέση του ιντρονίου προσδιορίστηκε μεταξύ των νουκλεοτιδίων 649 και 650 του cDNA. Το μέγεθός του προσδιορίστηκε στα 12Kb (Εικ.39).

4^ο ιντρόνιο: Το 5' άκρο του προσδιορίστηκε με τον εκκινητή 19 στον υπόκλωνο PstI 4.8Kb του φάγου Cph36, ενώ το 3' άκρο του προσδιορίστηκε με τον εκκινητή 20 στον ίδιο υπόκλωνο. Η θέση του 4^{ου} ιντρο-



Εικ.39 Προσδιορισμός μεγέθους του 3^{ου} ιντρονίου, με βάση το μέγεθος του βακτηριοφάγου Cph42, το οποίο είχε υπολογιστεί από ανάλυση ενζυμικού περιορισμού (19Kb). Το 5' ακραίο τμήμα του φάγου, μέχρι και το εξόνιο E3, έχει μήκος 3.5 Kb, όπως προκύπτει από τη χαρτογράφηση του υπόκλωνου 3.8Kb PstI. Το 3' ακραίο τμήμα του περιέχει τα εξόνια E4 και E5, το ιντρόνιο 4 και μέρος του ιντρονίου 5 (Πίνακας 19). Το μέγεθος του 3' ακραίου τμήματος είναι 3.5Kb, με βάση τη χαρτογράφηση του υπόκλωνου PstI 4.8Kb.

νίου προσδιορίστηκε μεταξύ των νουκλεοτιδίων 1339 και 1340 του cDNA, ενώ το μέγεθός του υπολογίστηκε στα 2.8Kb (Εικ.40).



Εικ. 40 Προσδιορισμός μεγέθους του 4^{ου} ιντρονίου. Το μέγεθος του 4^{ου} ιντρονίου υπολογίστηκε με βάση τη χαρτογράφηση του υπόκλωνου 4.8Kb PstI. Σ' αυτόν τον υπόκλωνο χαρτογραφείται θέση XbaI σε απόσταση 1.4Kb από τη θέση κλωνοποίησης PstI. Το θραύσμα PstI-XbaI μήκους 1.4Kb περιλαμβάνει ολόκληρο το εξόνιο E4, όπως φαίνεται από την υβριδοποίηση με τον ανιχνευτή cDNA NotI-SacII. Η θέση NotI του εξονίου E4 απέχει 0.3Kb από τη θέση κλωνοποίησης PstI. Το 3' άκρο του κλώνου αντιστοιχεί στη θέση PstI που βρίσκεται 9bp 3' της θέσης λήξης της μετάφρασης, η οποία απέχει 0.9Kb από το 3' άκρο του 4^{ου} ιντρονίου. Με βάση τα παραπάνω, το μέγεθος του 4^{ου} ιντρονίου υπολογίζεται στα 2.8Kb. Η θέση του προσδιορίστηκε με αλληλούχηση αυτού του κλώνου χρησιμοποιώντας τους εκκινητές 19 και 20, που φαίνονται με κόκκινα βέλη.

δ. Αλληλούχηση των περιοχών 5'UTR και 3'UTR

Η αλληλούχηση στην περιοχή 5'UTR του γονιδίου πραγματοποιήθηκε στον υπόκλωνο 1.4Kb PstI του φάγου Dph12 και στο θραύσμα του PstI-SacII μήκους 1Kb, με εκκινητές T3 και T7 του ρBluescript. Με αυτόν τον τρόπο προσδιορίστηκε αλληλουχία μήκους 786bp από την 5' πλευρική περιοχή του γονιδίου. Σε αυτήν περιέχεται και η περιοχή 647bp που αντιστοιχεί στο 5'UTR, σύμφωνα με το EST CA328532. Η αλληλούχηση της περιοχής 3'UTR δεν ήταν πλήρης. Πραγματοποιήθηκε με τον εκκινητή 7 στον υπόκλωνο 6Kb XbaI του φάγου Aph3. Προσδιορίστηκαν τα πρώτα 134 νουκλεοτιδία της 3' μη μεταφραζόμενης περιοχής του γονιδίου.

Το συνολικό αποτέλεσμα του προσδιορισμού νουκλεοτιδικής αλληλουχίας από τους γονιδιωματικούς κλώνους και η γονιδιωματική οργάνωση του *NeurI* φαίνεται στην Εικ. 41.

1	AGGAAAGGGA	GACGCTCAAC	CTCCGAAGCG	CCCATCTACC	CCTGACGTCA	GGTCCGGAGC	
61	AGTGCCGCCC	CCAAACCTGG	GTCCCGCGCG	CCTTCGAGAA	CTTATAGCCA	CGGGGAACGC	
121	GGCGCGCCCC	GAGACCACA	TCTGGCAGAG	CGGGCGCGCG	CGTGTCACTA	AGACGCTCAT	
181	TAACCTGGCC	AGCTGTGGCC	ATAACAACCG	GAGCCCCGGA	GTCTGGCGGA	CGCTAGGGCG	
241	GGGGTTGGGG	GCGGCGGCGT	CGGGAGGGAC	CTGGCGGGAG	CGGGGGCGGT	CCACACGAGC	
301	CGGGCGCTGC	AAGCTCACCC	GCGGTGGGA	GGCGCCACGC	ACGTGCGCTC	GGCGGTCCCC	
361	GCGGGCGGGG	ACCGAGCCCG	CCTACGAGCC	TCCGAGCAGC	CGCGGCGCGA	GCCCCGACTC	
421	TATGGCCCTC	GGGGAGCGCG	CCGCCCTGCG	CCGCCACCC	GCAGCCGGAA	TCCGAGCTCC	
481	AGCACCAGGC	AGAGCGGCAG	GCTCGCGCGG	CGGGCCGGGG	CCCAGGGGCC	GGGCGGCKCC	
541	GGGCGCGGGG	CATGGGAGCC	CGGTAGCCGG	GTGGCGCGAG	GACCCCCACG	GCGGGCAGGA	
601	CCCGCCGAGC	GAAGACCGCC	AAGCCAGGG	AGAGGGGGAG	ACAGCAAACC	GCCCCGCCCC	5' UTR
661	GGTCGCGACC	CCGCCAGGGC	CCTCCCTTGG	CACGCGCACC	TGCGCGCGCA	CACACTTGCA	
721	CACCTTTCCT	CCGCTCCTGC	CCAGCCTTGC	CTGGCTCCCG	CCGGAGCACC	CCCCGGGGGC	
781	GATGCCATGG	GTAATAAATT	CTCCAGTGTG	TCCTCTCTGC	AGCGAGGAAA	CCCAGCCCGC	E1
841	CGGTCCGCGG	GCCACCCCCA	GAACCTCAA	GGTAGGCTTC	GCTGACCCGG	GAGCAAATTT	
901	GCAGGGAGGG	TTGCGGCGCA	GGTCACTGT	
1	AGTCACGACC	CACAGGTGGA	GAACCAGTTG	GGTGGCAGTG	GTTGGGGACA	AAGTGAAGG	I1
61	TGGCAGGAAG	GCTGTGGTTC	ACAGCTGGCT	GTGCTGGTGA	AGGAGTAAAT	GCCGAGCCGT	
121	GGGTGGTCCC	GCTTCTTCCC	TGGACCCCTC	CCTAGGAGAC	CCCAATCTGT	CTCTTGGTCC	
181	TATTAGATT	TCTGACGCTA	TGGCCTGTTT	CTATTAGATT	CTCTGACGCT	ATGGCCTGTT	
241	CCTTTGCAGA	CTCCATCGGG	GGCTCCTTCC	CGGTGCCCTC	TCACCAGATG	CATCACAAGC	
301	AGAAGCATTG	CCCGCTACG	CTGTCAAGTG	GGGGGCTCCC	GGCCACGCCG	CTGCTCTTCC	E2
361	ACCCCCACAC	TAAGGGCTCC	CAGATCCTCA	TGGACCTCAG	CCACAAGGCC	GTCAAGAGGC	
421	AGGCCAGCTT	CTGCAATGCC	ATCACCTTCA	GTAACCGCCC	GGTGTCTCAT	TACGAGCAA	
481	TCAGGCTGAA	GGTGGGCACC	CCCGGACCCC	TGCCTTTGAC	CTTTAATCTC	TTTACTGCCA	
541	CGGTCTGCTG	TTCCACTACC	AGCCTGCCCT	TCTCCTTCTT	GTCCATCTCA	TCCTTTTCTC	
601	CTTTCCGAGC	CTTCTGGGT	AGCAAGATTC	GGTGGGAGTG	GCCTTCGCCA	ACCTATAGGA	I2
661	AGGCCCGCCC	CCTGGGCTGG	GTGGTGGGCG	GGAATGAGCC	ACTTAGTGAG	TTTTCTTGCT	
721	GTAGAGCCCA	CCCCTTCCCC	TCCGTCATCT	GACCAAGGGG	GGAGAGGCAA	GCTGAGGCTC	
781	CTCACTCCCT	GTGGCCAGAT	CACCAAGAAG	CAATGCTGCT	GGAGCGGGGC	CCTGCGACTT	
841	GGCTTACCA	GCAAGGACCC	TTCCCGCATC	CACCCGACTC	CGCTGCCCAA	GTACGCGCTG	E3
901	CCTGACCTGG	TGTCTCAGAG	TGGCTTCTGG	GCCAAAGCAT	TGCCTGAGGA	GTTTGCCAAC	
961	GAGGGCAACA	TCATTGCCTT	CTGGGTGGAC	AAGAAAGGGC	CGCTCTTCTA	CCGATCAAT	
1021	GAGTCAGCTG	CTATGCTTTT	CTTCAGTGGG	GTCCGGACGG	CGGACCGGCT	CTGGGCCCTG	
1081	GTGGACGCTT	ACGGCCTCAC	GCGGGGTGTC	CAGCTGCTAG	GTGAGTGGCT	TCTTTCTGGT	I3
1141	CCGTTAGTTC	AGGTCTGGGA	CTAACCTTCA	TTCCCATGAG	TCGTGGTGGC	TCTGGTGG...	
1	TTAGTGCTCA	GGTGGCGCTT	CTGCGACGAC	CCCTGTGGGA	CGGTGACAGT	GCCCTGTCT	
61	CTGCGCCTCC	CCGTGCGAGA	CAGCGAGCTG	GTGCTGCCCG	ACTGCCTGCG	GCCGCGCTCC	
121	TTCAACGCGC	TGCGCGCGCC	GTGCTGCGGG	TGCGAGGCGG	ATGAGGCGCG	CCTGTCCGTG	
181	AGCCTGTGCG	ACCTCAAGCT	GCCGGGAGCC	GACGGCGACG	ACGGCGCACC	GCCTGCCGGC	
241	TGCCCGATCC	CGCAGAATC	GCTCAATTCT	CAGCACAGCC	GCGCGCTGCC	GGCGCAGCTC	
301	GAGCGCGACC	TGCGCTTCCA	CGCGCTTCCG	CGCCCGCGCG	ACGTCCGCAT	CCTGGACGAG	
361	CAGACGCTGG	CGCGCTGGA	GCACGGGCGC	GACGAGCGCG	CGCTCGTCTT	CACCAGCCGG	
421	CCGTTGCGCG	TGGCCGAGAC	CATCTTCACT	AAGGTCACGC	GCTCGGGTGG	GCGCGGGCCG	E4
481	GGCGCGCTGT	CGTTGCGGGT	CACCAAGTGT	GACCCCGGCA	CGCTCGGCC	CGCGGACCTG	
541	CCCTTCAAGC	CCGAGGCCCT	GGTGGACCCG	AAGGAGTTCT	GGGCGGTGTG	TCCGCTGCC	
601	GGGCTCTGTC	ACAGCGGCGA	CATCTGGGC	CTGGTGGTCA	ACGCGGACGG	AGAGCTGCAC	
661	CTGAGTCACA	ACGGCGCGGC	GGCCGGCATG	CAGCTGTGCG	TGGATCCCTC	GCAGCCCTCC	
721	TGATGCTCT	TCAGCCTGCA	TGGCGCCATC	ACGACGGTCC	GCATCCTCGG	TGAGTCCGG	
781	GTGCTATCCC	CCAGTCCATG	ATGCTTTTTC	TAGGCTTAGG	TGAGTCTGCG	GTAGGGTGG	
841	AGCATCTGTG	TCCACGGATG	CGGGAAGAGA	GGAATGAACC	TCAGAAGAAC	CCAAAATGCA	
901	CACAAATGGA	CCAGGCTTAC	ACCCAGATGG	TTGTATTAAG	GGGCGAGGGC	AGTGACTTTC	
961	TGGCTTGTTC	CAGGCTGCTT	GTTGGCCTAG	GAATAAGTA	GCTGT.....	I4
1	CCTACAAAAC	TAGTAGAGCA	ATGGACCGGG	GACTCCATTC	CTGAAGGCCA	AAAGGATCCT	
61	AGGGAAAAGT	TGGCTTGGAA	ATTGCGGTGA	TATAGAACAT	GATACTACCT	CCTGACCCTA	
121	ATAGGTGGGC	TTCAGAGCAT	GGGCGGAGAT	GAACCAGCCT	GTGGGGATCC	TGCCCTCAGC	
181	TGTGGCTGTC	CTCTGCCCTC	AGCCTGTGGC	TGCCCTCTGG	GCAGGTGTTG	AGCCCTGTGG	
241	GGCTGGAAC	GAAACTTTCC	TCCCTTAGG	CTAATGGCAA	GTGTGACCAT	CCACTTTCTC	
301	AGGCTCCACC	ATCATGACTG	AACGGGGTGG	CCCATCTCTC	CCCTGCTCAC	CTGCCCTCC	E5
361	TCCAACTCCA	CCCAGTGCCC	TGGGCATCCG	CCTCTCTGAC	CCCCTGCTCA	GCACCTGCGG	
421	TTCTGGGCC	CTAGGTGGCT	CTGCTGGAGG	TGAGTACCAA	GGCCCCGTG	GGTGGTCAA	
481	TGCGTCTCTC	GAAATCCTAT	AGAGAGGGAC	GGTTGACAGA	GCTCTCCAGA	TGTGTAGCTG	
541	TGGCTTTTTT	TTTTTCTTTT	TAAAGATTTA	TTTTAAACTC	TGTGTGTGTG	TGTGTGTGTG	
601	TGTGTGTGTG	TGTGTGTGTG	TGTGTGTGTG	TGTGTGTGTG	TAAGTATAAA	TGCTCACGGA	
661	GCACAGAAGA	GGGTATCTGA	TCTTCCAGGC	GACTTCTTAA	AGTGCATGTG	GGGGATTCCA	
721	AGCGTTAGTG	CAGTAGAACC	ACGCTGGGAA	AGAATTAGAA	AGTCTTTTCC	TCCTTTCTCC	I5
781	TTTTCTCTTT	TTCTCCTTTT	CCTCTGACTG	CTGCTCTGGC	TCTGTCTGGG	GCTGTGTGGC	
841	TACAGCTCCT	TCCTTGGGCA	GGAAGGTGGG	TTGTCACTGT	GGAAGGTGG	CTTTGCGTAG	
901	TGGCAGAAGA	CATGAGCTGG	GGTGGGGTGG	GGGGGGCTTT	TCTCTTGGGA	GAAGGCCCCG	
961	TGCTGGGCTG	GGTCCCATCC	CCTTTTGATT	GGTGTCTTGG	CTCTCAGGGA	CAGCCCCCAA	
1021	CTCACCTGTG	AGCCTGCCCG	AGTCACCGGT	GACCCAGGT	CTGGGCCAGT	GGAGTGTATG	E6
1081	ATGCACCAAT	TGCTATGAAC	ACGCAGTGG	TACAGTCACT	TACAGTGTG	GCCACATGTG	
1141	CCCTGTCTAC	TCTGTGGGCG	TGCGCCCTCA	GAAAGCCCTG	CACGCTGTCT	GCCCCATCTG	
1201	CCGTCGCCCC	ATCAAGGACA	TCAATCAAGC	CTACCCGACG	TCCAGCCCA	CTGCAGAGCC	
1261	GCCCTGCAC	GCCCACTTCT	GCAGACTGCA	GCCCAAGCCG	CACCAGCTGA	GGGCAAGGCC	
1321	AACAGGGTCC	CTTCTTCTA	CAGTTTGGAA	ACTCTTCTC	CTTTTCCATT	AAACGTGGGA	3' UTR

Εικ.41 Η νουκλεοτιδική αλληλουχία του γονιδίου *Neur1*, όπως προσδιορίστηκε από τα υποκλωνοποιημένα τμήματα και τα προϊόντα PCR των γονιδιωματικών κλώνων. Τα τμήματα της νουκλεοτιδικής αλληλουχίας που αντιστοιχούν σε κωδική αλληλουχία (εξόνια E1 – E6) φαίνονται με κίτρινο. Η αλληλουχία του 5'UTR (σύμφωνα με το EST CA328532) παρουσιάζεται με μπλε, ενώ το τμήμα του 3'UTR που προσδιορίστηκε από αυτή την μελέτη φαίνεται με πράσινο.

ε. Ανάλυση της γονιδιωματικής οργάνωσης του *Neurl* *in silico*

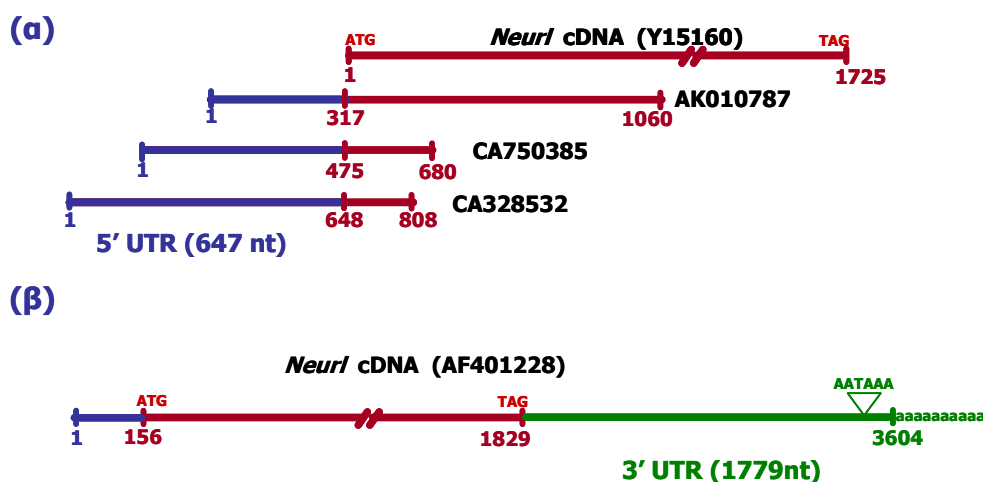
Η οργάνωση του γονιδίου *Neurl* στο χρωμόσωμα 19 του ποντικού μελετήθηκε και *in silico*, χρησιμοποιώντας τις διαθέσιμες αλληλουχίες EST και cDNA, καθώς και τη γονιδιωματική αλληλουχία που προέκυψε από το πρόγραμμα ανάλυσης του γονιδιώματος του ποντικού.

ι) Σύγκριση του *Neurl* cDNA (Y15160) με άλλα cDNAs και ESTs του ποντικού

Το *Neurl* cDNA Y15160 συγκρίθηκε με τα ESTs του ποντικού που έχουν καταχωρηθεί στη βάση δεδομένων dbEST καθώς και με πλήρη μόρια cDNA της βάσης δεδομένων Genbank. Από τις συγκρίσεις αυτές προσδιορίστηκε μέρος της 5' μη μεταφραζόμενης περιοχής, ολόκληρη η 3' μη μεταφραζόμενη περιοχή, ενώ διαπιστώθηκε η ύπαρξη ενός εναλλακτικού μεταγράφου του γονιδίου (βλ. παρακάτω).

Προσδιορισμός της 5' μη μεταφραζόμενης περιοχής (5'UTR)

Συγκρίνοντας την κωδική αλληλουχία του *Neurl* (1725bp, Y15160) με τα ESTs του ποντικού στη βάση δεδομένων dbEST προκύπτουν μεταξύ άλλων και 3 ESTs (AK010787, CA750385 και CA328532), τα οποία φέρουν επιπλέον αλληλουχία στο 5' άκρο (Εικ.42). Το EST CA328532 εκτείνεται περισσότερο προς το 5' UTR και με βάση αυτό, το μήκος της 5' μη μεταφραζόμενης περιοχής του γονιδίου πρέπει να είναι τουλάχιστον 647bp.



Εικ.42 Προσδιορισμός των περιοχών 5'UTR και 3'UTR του *Neurl* *in silico*, από τα ESTs του ποντικού στη βάση δεδομένων dbEST. (α) Με βάση το EST CA328532, το μήκος της περιοχής 5'UTR είναι τουλάχιστον 647nt. (β) Με βάση τον κλώνο cDNA AF401228, το μήκος της περιοχής 3' UTR είναι 1779nt. Στα διαγράμματα των ESTs και κλώνων cDNA, με κόκκινο φαίνεται η κωδική αλληλουχία, με μπλε η περιοχή 5' UTR και με πράσινο η περιοχή 3' UTR.

Προσδιορισμός της 3' μη μεταφραζόμενης περιοχής (3'UTR)

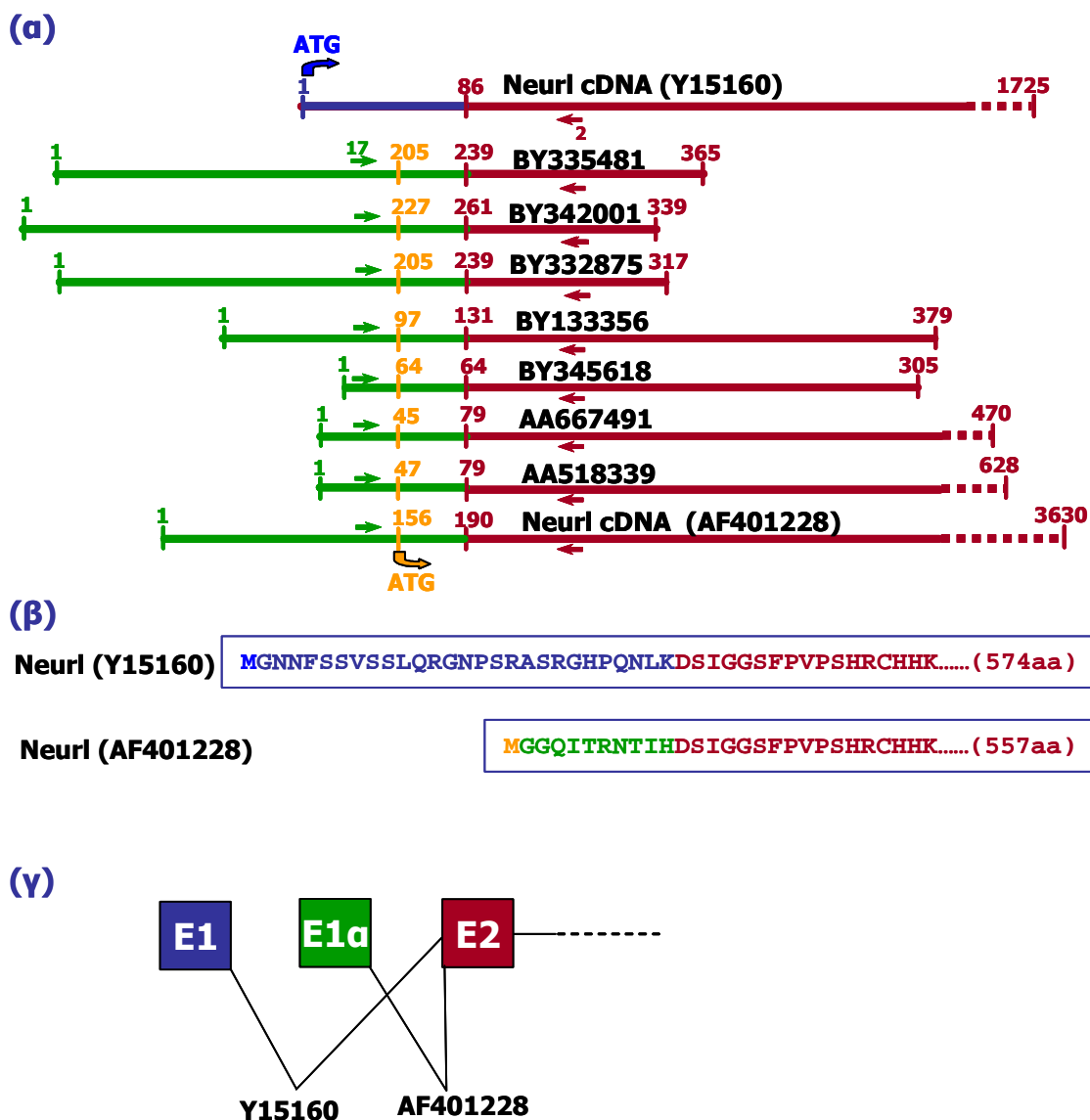
Η χιμαιρική αλληλουχία του *Neurl* cDNA Y15160 (βλ. Υλικά και Μέθοδοι), μήκους 2445bp, περιέχει 1749 νουκλεοτίδια του *Neurl*, από τα οποία τα 25 τελευταία ανήκουν στην περιοχή 3'UTR. Η υπόλοιπη 3' μη μεταφραζόμενη περιοχή προέκυψε *in silico* από τον κλώνο cDNA AF401228, μήκους 3630bp, ο οποίος περιλαμβάνει 1779 νουκλεοτίδια 3'UTR, σήμα πολυαδενυλίωσης και ουρά poly-A (Εικ.42).

Προσδιορισμός του εναλλακτικού εξονίου E1a

Από τη σύγκριση της κωδικής αλληλουχίας του *Neurl* (1725bp, Y15160) με ESTs του ποντικού προέκυψαν μεταξύ άλλων και 7 ESTs (τα BY335481 και BY332875 από ινοβλάστες συνδετικού ιστού αρθρώσεων, τα BY342001 και BY345618 από συνολικό ιστό αρθρώσεων, τα AA667491 και AA518339 από καλλιέργεια μυϊκών ινών και το BY133356 από έμβρυο E17.5), για τα οποία το σημείο έναρξης της ταυτότητάς τους με την αλληλουχία Y15160 ήταν πάντα το νουκλεοτίδιο 86 της αλληλουχίας Y15160, που συμπίπτει με το 5' άκρο του δεύτερου εξονίου. Η αλληλουχία που δεν εμφάνιζε ομοιότητα με την Y15160 ήταν επικαλυπτόμενη στα ESTs (Εικ.43). Η αλληλουχία αυτή περιέχεται και στο πλήρους μήκους μετάγραφο του *Neurl* AF401228, το οποίο εκφράζεται στους σκελετικούς μυες των ενηλίκων [Vollrath *et al*, 2001] και κωδικοποιεί μια πρωτεΐνη *Neurl* 557 αμινοξέων με διαφορετικό αμινοτελικό άκρο από αυτό του πεπτιδίου *Neurl* Y15160 (574 αμινοξέα), όπως φαίνεται στην Εικ.43. Επειδή στις βάσεις δεδομένων που ελέγχθηκαν δεν βρέθηκε κανένα EST ή πλήρες cDNA που να φέρει αλληλουχία του εξονίου E1 (νουκλεοτίδια 1-85 του Y15160) σε συνδυασμό με το εξόνιο E1a και επειδή στην αλληλουχία του E1a υπάρχει κωδικόνιο έναρξης της μετάφρασης και ανοικτό πλαίσιο ανάγνωσης 34 νουκλεοτιδίων, θεωρητικά υπάρχουν δύο μετάγραφα *Neurl*, της μορφής E1-E2-.... (~4.4Kb) και E1a-E2-.... (~3.7Kb), σύμφωνα και με την ανάλυση *Northern* των Vollrath *et al* (2001). Το μέγεθος της περιοχής 5'UTR_a του εναλλακτικού αυτού μεταγράφου είναι, με βάση το EST BY342001, τουλάχιστον 226bp (Εικ.43).

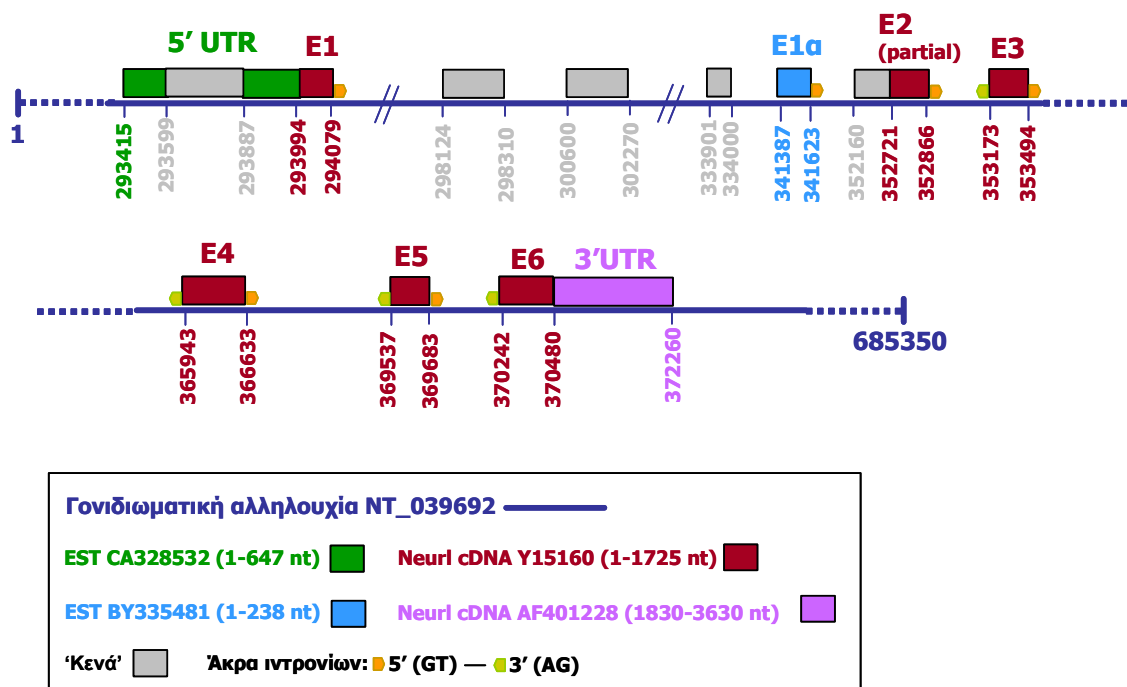
ii) Σύγκριση των *Neurl* cDNAs με τη γονιδιωματική αλληλουχία του ποντικού

Αλληλουχίες cDNA του *Neurl* συγκρίθηκαν με τη γονιδιωματική αλληλουχία του ποντικού. Συγκεκριμένα, οι αλληλουχίες cDNA που χρησιμοποιήθηκαν για τη σύγκριση ήταν: α) η κωδική αλληλουχία του *Neurl* (νουκλεοτίδια 1-1725 του cDNA Y15160), β) η περιοχή 5'UTR (νουκλεοτίδια 1-647 του EST CA328532), γ) Η περιοχή 3' UTR (νουκλεοτίδια 1830-3630 του cDNA AF401228), δ) το εναλλακτικό εξόνιο E1a (νουκλεοτίδια 1-238 του EST BY335481). Και στις τέσσερις παραπάνω συγκρίσεις, οι αλληλουχίες που αναγνωρίστηκαν ανήκαν στη γονιδιωματική αλληλουχία NW_000148 μήκους 2.17Mb, η οποία έχει προκύψει από συνδυασμό μικρότερων γονιδιωματικών αλληλουχιών και περιέχει κενά γνωστού μήκους (0.2–3Kb). Συγκεκριμένα, το γονίδιο *Neurl* εντοπίζεται στην αλληλουχία NT_039692 μήκους 684.35 Kb, η οποία αποτελεί μέρος της προηγούμενης. Το αποτέλεσμα της σύγκρισης φαίνεται στην Εικ.44. Το πλήρες μετάγραφο Y15160 (1725bp) αναγνωρίζει έξι περιοχές γονιδιωματικού DNA, οι οποίες αντιστοιχούν στα εξόνια E1– E6. Στα άκρα των ιντρονίων αναγνωρίστηκαν τα συντηρημένα νουκλεοτίδια GT (5') και AG (3'). Τα μεγέθη των εξονίων και των ιντρονίων είναι τα αναμενόμενα από την ανάλυση των γονιδιωματικών κλώνων. Με διαθέσιμη τη νουκλεοτιδική αλληλουχία ήταν δυνατή και η ακριβής περιγραφή του πρώτου ιντρονίου, του οποίου το μέγεθος προσδιορίστηκε στα 58.6Kb. Σε αυτό το ιντρόνιο και σε απόσταση 47.3Kb από το E1, βρίσκεται το εναλλακτικό εξόνιο E1a. Έλεγχος στις βάσεις δεδομένων dbEST και Genbank για ESTs που επεκτείνονται μέχρι και 500bp από το εξόνιο E1a δεν έδωσε κανένα αποτέλεσμα, πράγμα που σημαίνει ότι το μέγεθος αυτού του εξονίου που περιλαμβάνει και περιοχή 5'UTR είναι περίπου 240bp.



Εικ.43 *In silico* προσδιορισμός του εναλλακτικού εξονίου E1α. (α) Το cDNA Y15160 εμφανίζει ομοιότητα με τα ESTs BY335481, BY342001, BY332875, BY133356, BY345618, AA667491, AA518339 και με το πλήρες *Neur1* cDNA AF401228, μετά το νοκλεοτίδιο 85 (κόκκινο χρώμα). Τα πρώτα 85 νοκλεοτίδια (εξόνιο E1, μπλε χρώμα) αντικαθίστανται σε αυτά τα μετάγραφα από την αλληλουχία του εξονίου E1α (πράσινο). Τα πράσινα βέλη υποδεικνύουν τη θέση του εκκινήτη 17, ο οποίος σχεδιάστηκε για τον πειραματικό εντοπισμό του εξονίου E1α και τα κόκκινα βέλη τη θέση του εκκινήτη 2 που χρησιμοποιήθηκε για τον ίδιο σκοπό. Στο εξόνιο E1α υπάρχει κωδικόνιο έναρξης της μετάφρασης στο αναγνωστικό πλαίσιο του *Neur1*, από το οποίο κωδικοποιείται η ισομορφή AF401228. (β). Στο σχήμα (γ) παριστάνεται διαγραμματικά το εναλλακτικό *splicing* των εξονίων E1 και E1α με το εξόνιο E2 του *Neur1*. (Η σχετική θέση των εξονίων E1 και E1α προκύπτει από δεδομένα του προγράμματος αλληλούχησης του γονιδιώματος του ποντικού).

Το EST CA328532 αντιπροσωπεύει την 5' μη μεταφραζόμενη περιοχή του κλώνου *Neur1* Y15160, το τμήμα της γονιδιωματικής αλληλουχίας NT_039692 που αντιστοιχεί στην περιοχή 5' του εξονίου E1, το οποίο όμως περιλαμβάνει και ένα 'κενό' 350 νοκλεοτιδίων. Η περιοχή αυτή βρέθηκε στη βάση δεδομένων των επιμέρους γονιδιωματικών αλληλουχιών (*Trace archive, www.ncbi.nlm.nih.gov/Traces/trace.cgi*), με κωδικό 85864620, ύστερα από σύγκριση με το EST CA328532 και με την αντίστοιχη περιοχή που είχε αλληλουχηθεί από τους γονιδιωματικούς κλώνους του γονιδίου *Neur1* (βλ. παραπάνω).



Εικ.44 Σύγκριση γονιδιωματικής αλληλουχίας ποντικού με cDNAs του *Neur1*. Για τον *in silico* προσδιορισμό της γονιδιωματικής οργάνωσης του *Neur1*, συγκρίθηκαν τα ESTs CA328532, BY335481 και τα πλήρους μήκους μετάγραφα Y15160 και AF401228 με την γονιδιωματική αλληλουχία NT_039692 του χρωμοσώματος 19 του ποντικού. Στο σχήμα φαίνονται διαγραμματικά οι περιοχές ομοιότητας για κάθε EST και πλήρες μετάγραφο, που αναπαριστούν την οργάνωση του γονιδίου *Neur1*, σύμφωνα με τον χρωματικό κώδικα που φαίνεται στο πλαίσιο.

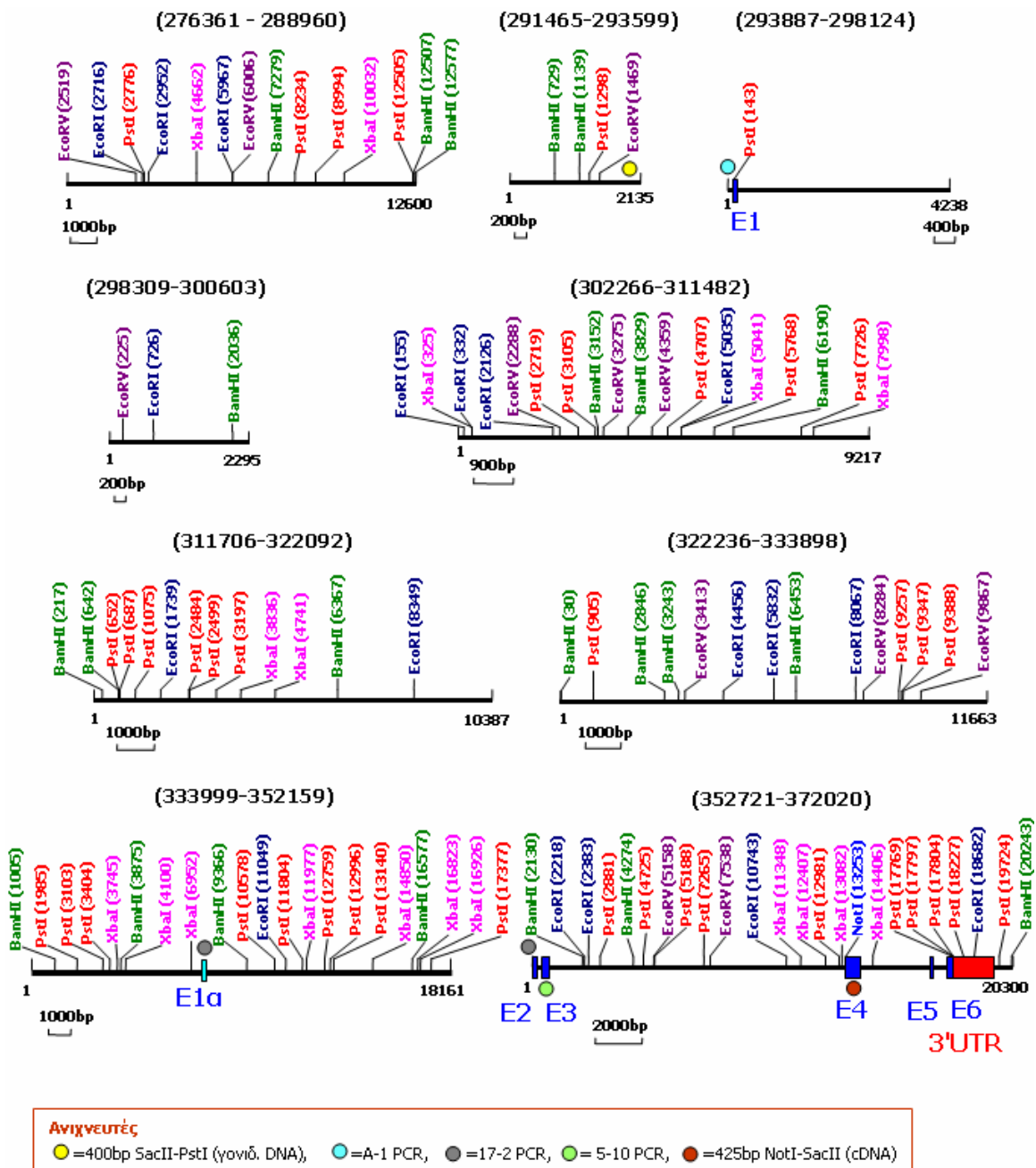
Περίπου 1Kb γονιδιωματικής αλληλουχίας ανοδικά της 5' πλευρικής περιοχής του *Neur1* διερευνήθηκε για το αν αντιπροσωπεύεται σε ESTs ή πλήρη μόρια cDNA των βάσεων δεδομένων dbEST και Genbank και το αποτέλεσμα αυτού του ελέγχου ήταν αρνητικό, γεγονός που υποδηλώνει ότι το μήκος της 5'UTR περιοχής του *Neur1* Y15160 είναι περίπου 647bp, όπως προκύπτει και από το EST CA328532. Από τη σύγκριση της γονιδιωματικής αλληλουχίας με τα νουκλεοτίδια 1830-3630 του cDNA AF401228, προσδιορίστηκε η περιοχή 3'UTR, μήκους 1779bp, η οποία περιλαμβάνει χαρακτηριστικό σήμα πολυαδενυλίωσης. Τέλος, σε περιοχή γονιδιωματικού DNA μήκους περίπου 97Kb, η οποία περιλαμβάνει το γονίδιο *Neur1*, εντοπίστηκαν οι θέσεις αναγνώρισης για τις περιοριστικές ενδονουκλεάσες BamHI, EcoRI, PstI, XbaI, EcoRV και NotI και ο χάρτης που προέκυψε (Εικ.45) συγκρίθηκε με αυτόν που είχε προκύψει από τις πειραματικές προσεγγίσεις (χαρακτηρισμός βακτηριοφάγων και ανάλυση ολικού γονιδιωματικού DNA).

Τα προβλεπόμενα μεγέθη των θραυσμάτων με βάση την *in silico* ανάλυση προσεγγίζουν τα μεγέθη των ζωνών που προέκυψαν από τις πειραματικές προσεγγίσεις, για τους ανιχνευτές που χρησιμοποιήθηκαν (Πίνακας 20). Στις περιπτώσεις που παρατηρούνται διαφορές, αυτές πιθανώς οφείλονται στις περιοχές της γονιδιωματικής αλληλουχίας που δεν έχουν ακόμα αλληλουχηθεί (κενά).

Πίνακας 20. Σύγκριση πειραματικής και *in silico* μεθόδου ανάλυσης

Ιχνηθέτης	Περιοριστικό ένζυμο	<i>In silico</i>	Πείραμα
SacII - PstI	BamHI	7,743	7,3
	PstI	1,268	1,4
	XbaI	16,202	16,0
	EcoRV	5,602	5,5
	EcoRI	16,710	18,0
A-1	BamHI	7,743	7,3
	PstI	1,268 / 11388	1,4 / 6,8
	XbaI	16,202	16,0
	EcoRV	5,602	5,5
	EcoRI	16,710	18,0
5-10	BamHI	4,276	3,9
	PstI	4,227	3,8
	XbaI	13,145	14,0
	EcoRV	<i>25,778</i>	5,5
	EcoRI	9,892	10,0
NotI-SacII	PstI	4,788	4,8
	XbaI	1,324	1,3
	EcoRI	7,939	8,0
17-2	BamHI	4,276 / 5,491	3,9 / 5,5
	PstI	4,227 / 7,174	3,8 / 7,2

Σημ: Σύγκριση των ζωνών που προβλέπονται από την *in silico* ανάλυση της γονιδιωματικής αλληλουχίας στην περιοχή του γονιδίου *NeurI* (στήλη '*In silico*') με τα μεγέθη των ζωνών που προέκυψαν από την πειραματική ανάλυση γονιδιωματικού DNA (στήλη 'Πείραμα') για τους συγκεκριμένους ανιχνευτές και τις περιοριστικές ενδονουκλεάσες που αναφέρονται. Τα μεγέθη των προβλεπόμενων από την *in silico* ανάλυση ζωνών που εμφανίζουν μεγάλη διαφορά από τις πειραματικές τιμές φαίνονται με πλάγια γράμματα. Όλα τα μεγέθη είναι σε Kb.



Εικ.45 Θέσεις αναγνώρισης περιοριστικών ενζύμων στην περιοχή του γονιδίου *Neur1*. Τα διαγράμματα παρουσιάζουν εννέα διαστήματα της γονιδιωματικής αλληλουχίας του ποντικού NT_039692, που περιέχουν το γονίδιο *Neur1*. Οι αριθμοί στις παρενθέσεις αντιστοιχούν σε νουκλεοτίδια DNA της αλληλουχίας NT_039692. Μεταξύ των εννέα διαδοχικών διαστημάτων υπάρχουν 'κενά' γνωστού μεγέθους (0.1-2.5Kb), λόγω ελλιπούς αλληλούχησης της περιοχής. E1-E6 είναι τα εξόνια του *Neur1*. Οι κύκλοι συμβολίζουν τους ανιχνευτές που χρησιμοποιήθηκαν στον πειραματικό προσδιορισμό της οργάνωσης του *Neur1*, σύμφωνα με τον χρωματικό κώδικα που φαίνεται στο πλαίσιο (Υλικά και Μέθοδοι, Πίνακας 10).

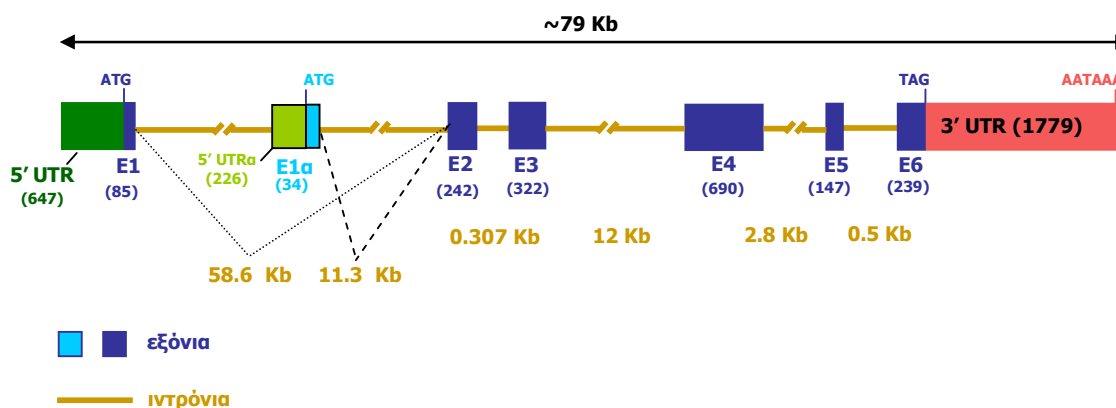
Λαμβάνοντας υπόψη τα παραπάνω, για την οργάνωση του γονιδίου *Neur1* του ποντικού προκύπτει η διάταξη της Εικόνας 46. Από αυτήν συμπεραίνουμε ότι:

Το γονίδιο *Neur1* εκτείνεται σε περιοχή μήκους περίπου 79Kb στο γονιδίωμα του ποντικού.

- Η κωδική του αλληλουχία (1725 νουκλεοτίδια, Υ15160) διακρίνεται σε 6 εξόνια με μεγέθη: 85, 242, 322, 690, 147 και 239 br, τα οποία διαχωρίζονται από 5 ιντρόνια με μεγέθη: ~60, 0.3, 12, 2.8 και 0.5 Kb.

Υπάρχει ένα τουλάχιστον επιπλέον εξόνιο του γονιδίου (το E1a), το οποίο δεν συμπεριλαμβάνεται στην κωδική αλληλουχία των 1725 νουκλεοτιδίων (Y15160), αλλά υπάρχει σε άλλους κλώνους cDNA και ESTs. Το εξόνιο αυτό συμμετέχει σε εναλλακτικό μετάγραφο του *Neurl* και χαρτογραφείται στο ιντρόνιο μεταξύ των εξονίων E1 και E2, σε απόσταση 47.3Kb από το E1.

- Η 5'UTR περιοχή του γονιδίου έχει μήκος περίπου 647bp , με βάση το EST CA328532.
- Η 3'UTR περιοχή του γονιδίου έχει μήκος 1779bp, με βάση το cDNA AF401228.



Εικ.46 Η οργάνωση του *Neurl* στο γονιδίωμα του ποντικού, όπως προέκυψε από τον χαρακτηρισμό των γονιδιωματικών κλώνων, την ανάλυση ολικού γονιδιωματικού DNA, και την *in silico* ανάλυση των διαθέσιμων αλληλουχιών cDNA και γονιδιωματικού DNA. Σε παρένθεση φαίνεται ο αριθμός των νουκλεοτιδίων κωδικής αλληλουχίας κάθε εξονίου.

στ. Σύγκριση γονιδιωματικής οργάνωσης των γονιδίων *Neurl* και *NEURL*

Όπως φαίνεται από μεγάλης κλίμακας γονιδιωματικές συγκρίσεις [Waterston *et al*, 2002a], τα ορθόλογα γονίδια ανθρώπου και ποντικού εμφανίζουν ομοιότητες στο μήκος και τον συνολικό αριθμό των εξονίων τους, στην αναλογία μεγέθους των ιντρονίων τους, καθώς και στις ρυθμιστικές τους περιοχές (βλ. Εισαγωγή). Για να διαπιστωθεί η κοινή καταγωγή των γονιδίων *Neurl* του ποντικού και *NEURL* του ανθρώπου, πραγματοποιήθηκε σύγκριση της γονιδιωματικής τους οργάνωσης. Η οργάνωση του γονιδίου *NEURL* προσδιορίστηκε *in silico*, συγκρίνοντας τη νουκλεοτιδική αλληλουχία του cDNA AF029729 (3634 nt) με τη γονιδιωματική αλληλουχία NT_030059 (~43 Mb) του ανθρώπινου χρωμοσώματος 10. Μέρος της περιοχής 5'UTR προσδιορίστηκε από το EST U87864, ενώ η παρουσία του εναλλακτικού εξονίου E1a διαπιστώθηκε με βάση 16 ανθρώπινα ESTs (CB992271, U69206, CB992681, CB992205, CB991269, CB991009, BX098217, CA948745, CA941769, CA772780, CA771909, BQ631750, BQ286290, BI712632, BG656185, CB068968) από βιβλιοθήκες εγκεφάλου ή παγκρεατικού αδένου, στα οποία το εξόνιο E1a ακολουθείται από το E2, όπως συμβαίνει και στο γονίδιο του ποντικού. Σύμφωνα με το EST CB992271 και την γονιδιωματική αλληλουχία (NT_030059), το εξόνιο E1a περιέχει 34 νουκλεοτίδια κωδικής αλληλουχίας και τουλάχιστον 152 νουκλεοτίδια μη μεταφραζόμενης περιοχής (5' UTRa, Εικ.47).

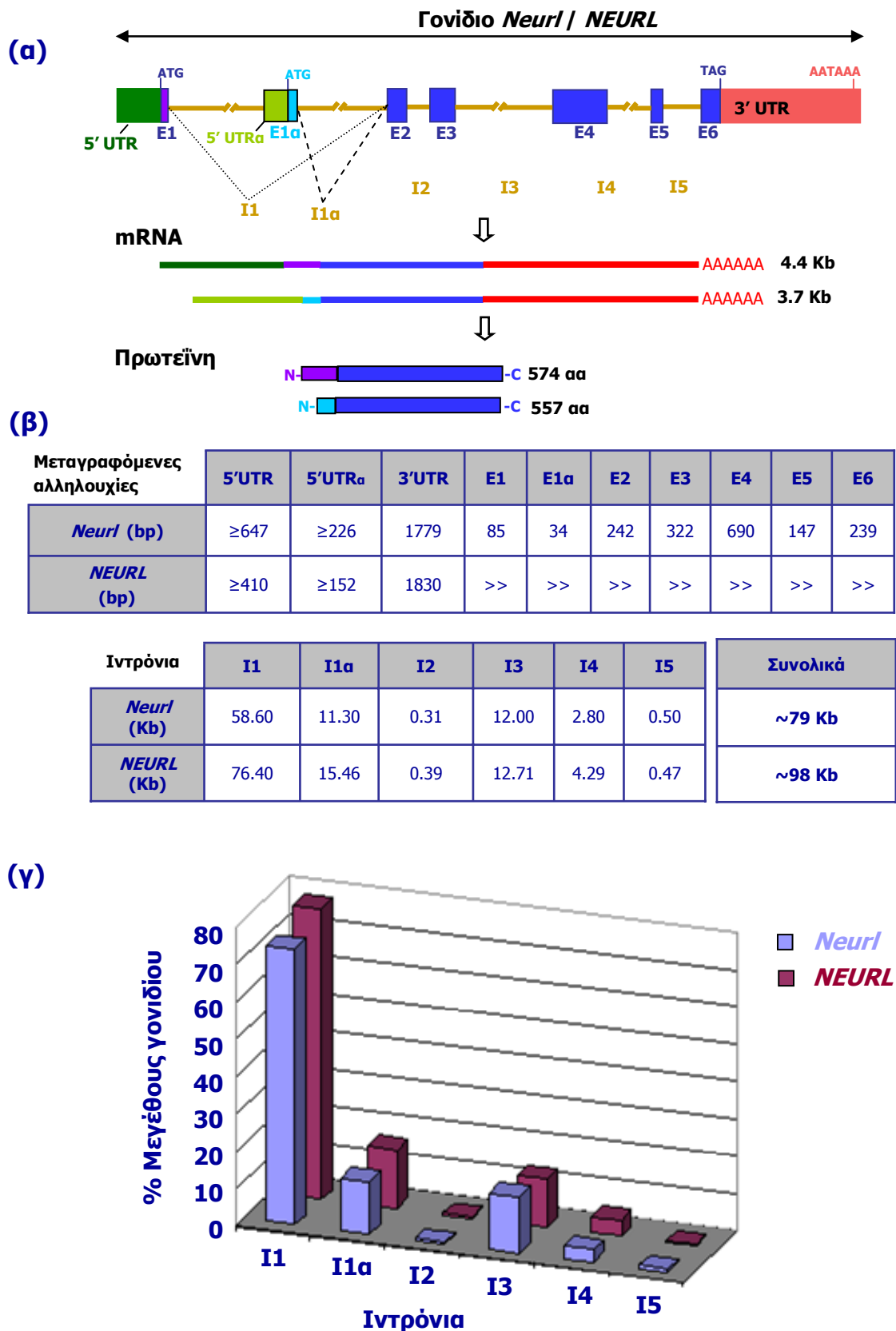
Από τη σύγκριση της γονιδιωματικής οργάνωσης των *NEURL* και *Neurl* φαίνεται ότι ο αριθμός των εξονίων και των κωδικών νουκλεοτιδίων/εξόνιο διατηρούνται όμοια στα δύο ορθόλογα γονίδια. Τα ιντρόνια του ανθρώπινου γονιδίου (ιδιαίτερα το 1ο και το 4ο) είναι μεγαλύτερα από τα αντίστοιχα στο γονίδιο του ποντικού, όμως διατηρείται η σχετική αναλογία μεγεθών. Συγκρίσιμα είναι και τα μεγέθη των περιοχών

3'UTR των δύο ορθόλογων γονιδίων (Εικ.47). Όσον αφορά το μέγεθος της περιοχής 5'UTR, αυτό δεν μπορεί να προσδιοριστεί με ακρίβεια *in silico*. Με βάση το ανθρώπινο EST U87864, η περιοχή 5'UTR του *NEURL* είναι τουλάχιστον 410bp. Σύμφωνα με τα παραπάνω, το ανθρώπινο γονίδιο *NEURL* (~99Kb) είναι μεγαλύτερο από αυτό του ποντικού (~79Kb), ενώ το μεγαλύτερο ποσοστό του μεγέθους και των δύο γονιδίων (~70-80%) αντιπροσωπεύεται από το πρώτο ιντρόνιο.

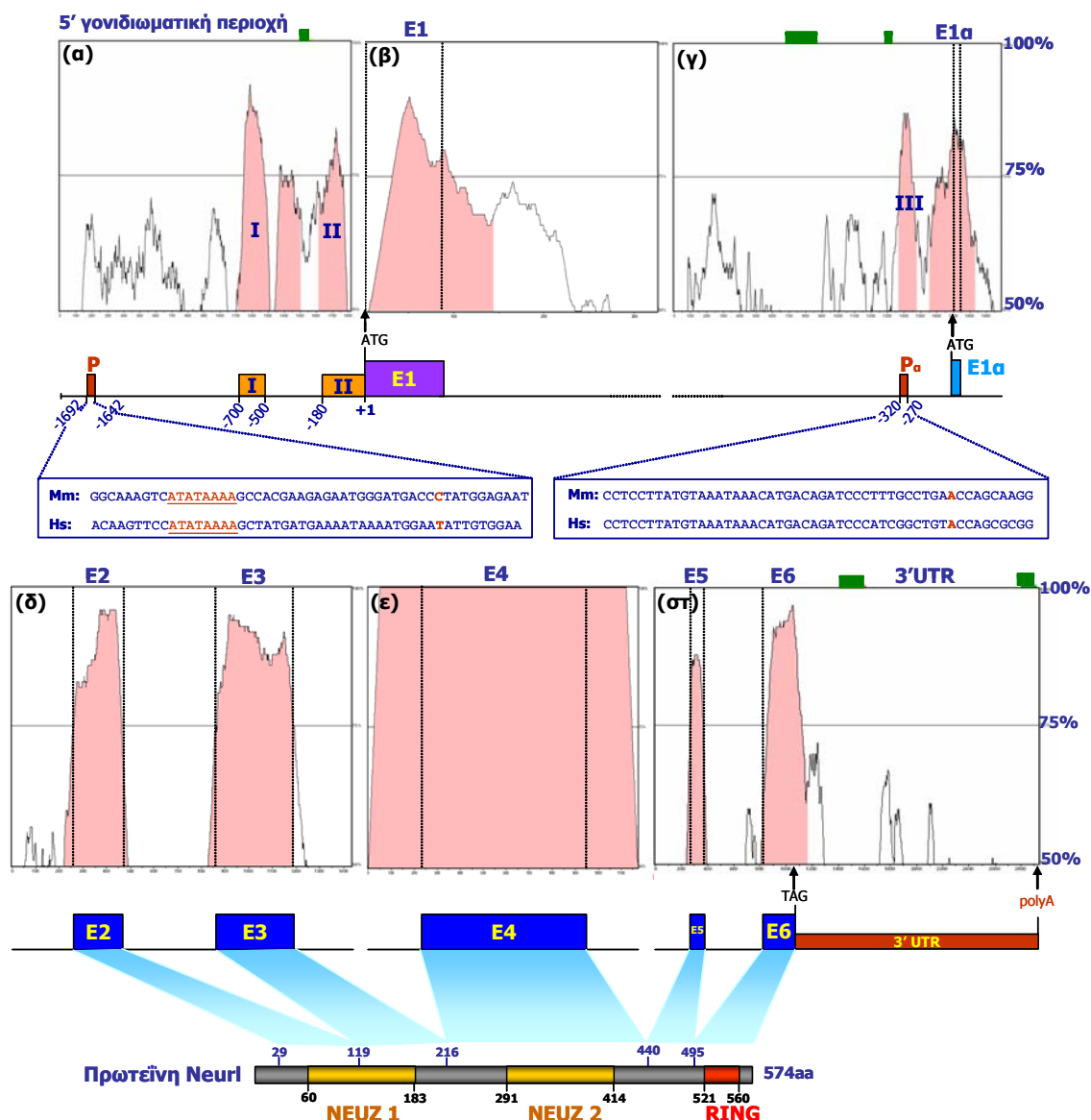
η. Σύγκριση γονιδιωματικής αλληλουχίας ανθρώπου και ποντικού στις περιοχές των γονιδίων *Neurl* και *NEURL*

Συγκρίσεις μεταξύ ορθόλογων γονιδίων του ανθρώπου και του ποντικού στο επίπεδο της νουκλεοτιδικής αλληλουχίας είναι χρήσιμες για την αποκάλυψη ρυθμιστικών περιοχών, λόγω της μεγαλύτερης συντήρησης που εμφανίζουν αυτές σε σχέση με την γενική νουκλεοτιδική ομοιότητα των δύο γονιδιωμάτων (60-70%). Με το δεδομένο ότι δεν υπάρχει πληροφορία για τη μεταγραφική ρύθμιση του γονιδίου *Neuralized* στα θηλαστικά, μια τέτοια σύγκριση είναι αναγκαία για την οριοθέτηση των πιθανών ρυθμιστικών περιοχών του. Στην ανάλυση χρησιμοποιήθηκαν τα εξής τμήματα των γονιδιωματικών αλληλουχιών NT_030059 (*H. sapiens*) και NT_039692 (*M. musculus*): α) 1812 νουκλεοτίδια από την περιοχή 5' του σημείου έναρξης της μετάφρασης από το εξόνιο E1, β) Η κωδική αλληλουχία του εξονίου E1 (85bp) και 240 νουκλεοτίδια από το 5' άκρο του ιντρονίου I1, γ) 1986 νουκλεοτίδια που περιλαμβάνουν την κωδική αλληλουχία του εναλλακτικού εξονίου E1a (34bp), μέρος της 5' πλευρικής περιοχής της (1712bp) και το 5' άκρο του ιντρονίου I1a (240bp) δ) 240 νουκλεοτίδια από το 3' άκρο του πρώτου ιντρονίου, το εξόνιο E2, το ιντρόνιο I2, το εξόνιο E3 και 240 νουκλεοτίδια από το 5' άκρο του ιντρονίου I3, ε) το εξόνιο E4 και 240 νουκλεοτίδια καθενός από τα ιντρόνια που το περιβάλλουν, στ) 240 νουκλεοτίδια από το 3' άκρο του ιντρονίου I4, το εξόνιο E5, το ιντρόνιο E5, το εξόνιο E6 και η περιοχή 3' UTR. Για τις συγκρίσεις χρησιμοποιήθηκε το πρόγραμμα mVISTA (www-gsd.lbl.gov/vista/index.shtml), στο οποίο νουκλεοτιδική ταυτότητα με ποσοστό μεγαλύτερο του 75% θεωρείται σημαντική.

Όπως ήταν αναμενόμενο και από προηγούμενες συγκρίσεις [Pavliopoulos *et al*, 2002], διαπιστώθηκε μεγάλη συντήρηση κατά μήκος των κωδικών περιοχών των δύο γονιδίων. Σημαντική είναι και η συντήρηση στα σημεία συναρμογής των εξονίων (*splice junctions*), ενώ η ομοιότητα των ιντρονίων μειώνεται σε απόσταση μικρότερη των 200 νουκλεοτιδίων από τα άκρα τους. Η περιοχή 3' UTR των γονιδίων εμφανίζει τη μικρότερη συντήρηση. Αντίθετα, στις 5' πλευρικές περιοχές των εξονίων E1 και E1a παρατηρούνται διαστήματα με σημαντικά ποσοστά νουκλεοτιδικής ταυτότητας (Εικ.48). Συγκεκριμένα, οι συντηρημένες περιοχές I και II (Εικ.48α) ανήκουν στην 5'UTR περιοχή, της οποίας το ελάχιστο μέγεθος εκτιμάται με βάση τα ESTs, (βλ. παραπάνω), ενώ η περιοχή III περιέχει έναν πιθανό εναλλακτικό υποκινητή (Pa), με βάση το πρόγραμμα πρόβλεψης που χρησιμοποιήθηκε. Σημειώνεται ότι αυτές οι περιοχές δεν ανήκουν στο μεταγγραφο RP11-225H22.5, το οποίο έχει εντοπιστεί στο πρώτο ιντρόνιο του *Neurl* από προγράμματα γονιδιακής πρόβλεψης (βλ. Συζήτηση).



Εικ.47 Σύγκριση της οργάνωσης των γονιδίων *Neur1* και *NEURL*. (α) Η διάταξη εξονίων και ιντρονίων είναι ίδια στα δύο γονίδια. Από το εναλλακτικό *splicing* των εξονίων E1 ή E1 α με το E2 προκύπτουν οι ισομορφές 574aa και 557aa που φαίνονται στο σχήμα. (β) Σύγκριση μεγεθών των μεταγραφόμενων και μη μεταγραφόμενων περιοχών. Τα εξόνια των δύο γονιδίων έχουν ίδιο μέγεθος, ενώ τα ιντρόνια του *NEURL* είναι συνολικά μεγαλύτερα. Τα μεγέθη των περιοχών 5' και 3' UTR είναι εκτιμήσεις με βάση τα διαθέσιμα ESTs. Τα ιντρόνια συμβολίζονται με I1-I5. (γ) Στα δύο γονίδια διατηρείται σταθερή η σχέση μεγεθών των ιντρονίων.



Εικ.48 Συγκριτική ανάλυση της γονιδιωματικής περιοχής των γονιδίων *Neur1* και *NEURL*. Στα διαγράμματα (α)-(ςτ) παρουσιάζεται το ποσοστό της νουκλεοτιδικής ταυτότητας (άξονας Y) κατά μήκος κάθε συγκρινόμενης περιοχής (άξονας X). Συγκρίθηκαν οι περιοχές των εξονίων E1, E1α, E2-E6, καθώς και μη μεταφραζόμενες περιοχές (βλ. κείμενο). Μεγαλύτερη ομοιότητα παρατηρείται στην κωδική αλληλουχία. P και Pa είναι οι περιοχές των πιο πιθανών υποκινητών, σύμφωνα με το πρόγραμμα πρόβλεψης NNPP (www.fruitfly.org/seq_tools/pro-moter.html). I, II και III είναι μη μεταφραζόμενες περιοχές με σημαντική ομολογία μεταξύ των δύο γονιδίων. Τα πράσινα ορθογώνια στα διαγράμματα (α), (γ) και (ςτ) συμβολίζουν περιοχές επαναλαμβανόμενου DNA. Όπως φαίνεται στο σχήμα της πρωτεΐνης Neur1, οι περιοχές DNA με υψηλότερη συντήρηση κωδικοποιούν τα *domains* NEUZ1, NEUZ2 και RING.

Από το αποτέλεσμα της συγκριτικής αυτής ανάλυσης (Εικ.48) είναι φανερό ότι τα εξόνια που κωδικοποιούν τα διαφορετικά *domains* της πρωτεΐνης Neur1 (E2, E3, E4 και E6) είναι περισσότερο συντηρημένα σε σχέση με τα υπόλοιπα που κωδικοποιούν το αμινοτελικό άκρο (E1, E1α) ή περιοχή μεταξύ των *domains* (E5).

3. Διερεύνηση πρωτεϊνικών αλληλεπιδράσεων του Neur1 με το σύστημα 'δύο υβριδίων' στον σακχαρομύκητα

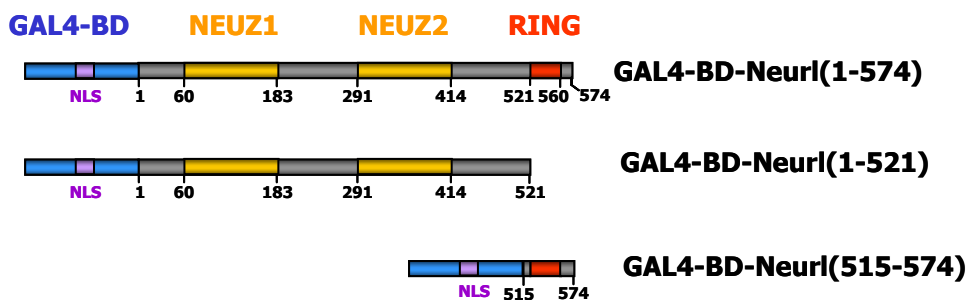
Στην προσπάθεια να μελετήσουμε τη βιολογική λειτουργία της πρωτεΐνης Neutralized των θηλαστικών (Neur1), ήταν αναγκαία η ταυτοποίηση των πρωτεϊνών με τις οποίες η Neur1 αλληλεπιδρά. Όπως ήδη έχει αναφερθεί, η πρωτεΐνη Neutralized ανήκει στην ετερογενή οικογένεια των πρωτεϊνών με domain RING, δομή που έχει επανειλημμένως συσχετιστεί με πρωτεϊνικές αλληλεπιδράσεις. Πρωτεΐνες με RING έχουν βρεθεί να συμμετέχουν σε μεγάλη ποικιλία βιολογικών λειτουργιών, ρυθμίζοντας πχ μηχανισμούς ογκογένεσης, απόπτωσης, ανάπτυξης και ιικής μόλυνσης. Μεγάλος αριθμός αναφορών στην πρόσφατη βιβλιογραφία περιγράφουν το ρόλο του RING στις τροποποιήσεις πρωτεϊνών από ουβικουΐτινη ή μόρια ομόλογα της ουβικουΐτινης, διαδικασίες που επιτελούνται από πολυπρωτεϊνικά ενζυμικά σύμπλοκα [Borden 2000 / Jackson *et al*, 2000 / Joazeiro & Weissman 2000]. Σε πρωτεΐνες που φέρουν RING έχει αποδοθεί ο ρόλος του 'μεσολαβητή' (*adaptor*), όσον αφορά τη σύνδεση πρωτεϊνικών υποστρωμάτων με τα παραπάνω ενζυμικά σύμπλοκα. Συνήθως, ο ρόλος αυτός επιτυγχάνεται και από επιπλέον *domains* του ίδιου μορίου, που αναγνωρίζουν τα διαφορετικά υποστρώματα, όπως συμβαίνει με την πρωτεΐνη c-Cbl, της οποίας η δράση ως E3 λιγάσης ουβικουΐτινης εξαρτάται εκτός από το RING και από το SH2 *domain*, που αλληλεπιδρά με τις πρωτεϊνικές κινάσες RPTK (*receptor protein tyrosine kinases*), καταλύοντας την ουβικουΐτινίωσή τους [Joazeiro *et al*, 1999]. Με βάση τα παραπάνω, η ταυτοποίηση των πρωτεϊνών που αλληλεπιδρούν με την πρωτεΐνη Neur1 είναι ιδιαίτερα σημαντική στην πορεία για την κατανόηση της βιολογικής της λειτουργίας.

Για τον σκοπό αυτό εφαρμόστηκε η μέθοδος του διυβριδικού συστήματος σακχαρομύκητα (*yeast two-hybrid system*), σύμφωνα με την οποία η πρωτεΐνη Neur1 χρησιμοποιήθηκε ως 'δόλωμα' (*bait*) για να ταυτοποιηθούν πρωτεΐνες - θηράματα (*prey*) που αλληλεπιδρούν άμεσα με αυτήν, ελέγχοντας μια βιβλιοθήκη κλώνων cDNA εμβρύου ποντικού E9.5. Για την πραγματοποίηση του πειράματος ήταν απαραίτητη η κατασκευή υβριδικής πρωτεΐνης Neur1, συντηγμένης με την υπομονάδα πρόσδεσης στο DNA του μεταγραφικού παράγοντα GAL4 (GAL4-BD) και η χρησιμοποίηση μιας βιβλιοθήκης κλώνων cDNA, οι οποίοι κωδικοποιούν υβριδικά πεπτίδια μέσου μήκους 100-250 αμινοξέων, συντηγμένα με την υπομονάδα ενεργοποίησης της μεταγραφής (VP16-AD) του παράγοντα VP16 (Υλικά και Μέθοδοι). Τα πρωτεϊνικά προϊόντα και των δύο παραπάνω κατασκευών οδηγούνται στον πυρήνα κυττάρων σακχαρομύκητα, λόγω σημάτων πυρηνικού εντοπισμού στην αμινοξική σύσταση των υπομονάδων GAL4-BD και VP16-AD. Επειδή το στέλεχος του σακχαρομύκητα που χρησιμοποιήθηκε (PJ69-4a) στερείται ενδογενούς μεταγραφικού παράγοντα GAL4, η αλληλεπίδραση της υβριδικής πρωτεΐνης Neur1 και ενός υβριδικού πεπτιδίου της βιβλιοθήκης είναι αναγκαία και ικανή να προκαλέσει ενεργοποίηση των γονιδίων αναφοράς (*ADE2*, *HIS3*, *LacZ*), που επάγονται από GAL4-ρυθμιζόμενους υποκινητές (*GAL2*, *GAL1*, *GAL7*).

α. Διερεύνηση της βιβλιοθήκης cDNA με 'δόλωμα' το πλήρες *Neur1*

1^ο στάδιο επιλογής: Η κωδική περιοχή του γονιδίου *Neur1* (1725bp) κλωνοποιήθηκε στον πλασμιδιακό φορέα pGBT9 με ομόλογο ανασυνδυασμό ενθέματος και φορέα, στα κύτταρα σακχαρομύκητα pJ69-4a (βλ. Κεφ. 'Πλασμιδιακές Κατασκευές', *Neur1/pGBT9*). Το προϊόν αυτής της κατασκευής είναι η χιμαιρική πρωτεΐνη GAL4-BD-*Neur1* (Εικ.49), η οποία, αφού πρώτα ελέγχθηκε ότι δεν είναι από μόνη της ικανή να ενεργοποιήσει τη μεταγραφή των γονιδίων αναφοράς, χρησιμοποιήθηκε ως δόλωμα για τη διερεύνηση της βιβλιοθήκης cDNA στα κύτταρα σακχαρομύκητα pJ69-4a. Ελέγχθηκαν 8x10⁶ διπλά μετασχηματισμένες αποικίες σακχαρομύκητα και μετά την πρώτη επιλογή συλλέχθηκαν 63 αποικίες, οι οποίες μεταφέρθηκαν διατεταγμένες σε δύο πρότυπα τρυβλία (*master plates*) με μη επιλεκτικό θρεπτικό μέσο (SC με Ura, Ade, His).

Πρωτεΐνες – 'δόλωμα' ('bait')



Εικ.49 Πρωτεΐνες-‘δόλωμα’ (*bait*) που χρησιμοποιήθηκαν στο πείραμα δύο υβριδίων. Η GAL4-BD-*Neur1*(1-574) είναι η πλήρους μήκους πρωτεΐνη Neutralized του ποντικού, συντηγημένη με την υπομονάδα GAL4-BD, που χρησιμοποιήθηκε στη διερεύνηση της βιβλιοθήκης. Τα τμήματά της, *Neur1*(1-521) και *Neur1*(515-574) δοκιμάστηκαν ως ‘δόλωμα’ στα επιλεγμένα ‘θηράματα’ (*preys*). Η κλωνοποίηση των αντίστοιχων τμημάτων cDNA του *Neur1* στον φορέα περιγράφεται στο Κεφ. ‘Πλασμιδιακές Κατασκευές’. (NLS= σήμα πυρηνικού εντοπισμού, GAL4-BD= υπομονάδα πρόσδεσης στο DNA του μεταγραφικού παράγοντα GAL4)

2^ο στάδιο επιλογής: Χρησιμοποιήθηκαν τα θρεπτικά μέσα UA, UH και UAH/X-gal (Υλικά και Μέθοδοι). Για το θρεπτικό μέσο UA δοκιμάστηκαν δύο διαφορετικές συγκεντρώσεις 3-αμινοτριαζόλης (3 και 10mM). Από τις 63 επιλεγμένες αποικίες, οι 57 αναπτύχθηκαν στα τρυβλία επιλογής. Στο τρυβλίο ελέγχου ενεργοποίησης του γονιδίου *LacZ*, 17 από τις αποικίες εμφάνισαν μπλε χρώμα (Εικ.50, Πίνακας 21). Στη συνέχεια, απομονώθηκε πλασμιδιακό DNA από τις 57 αποικίες και επιλέχθηκε από κάθε αποικία το πλασμίδιο *cDNA-prey/pVP16*, με μετασχηματισμό κυττάρων *E.coli* του στελέχους KC8. Ακολούθησε μετασχηματισμός κυττάρων *E.coli* του στελέχους DH5a για την απομόνωση πλασμιδιακού DNA καλής ποιότητας και τον προσδιορισμό της νουκλεοτιδικής σύστασης των 57 κλώνων *cDNA-prey/pVP16*.

Πίνακας 21. Αποτελέσματα 2^{ου} σταδίου επιλογής

#	UAH	UH	UA 3mM AT	UA 10mM AT	UAH Xgal
1	+	+	+	+	w
2	+	+	+	+	w
3	+	+	+	+	w
4	+	+	+	+	b
5	+	+	+	+	w
6	+	+	+	+	w
7	+	+	+	+	w
8	+	+	+	+	w
9	+	+	+	+	w
10	+	+	+	+	w
11	+	+	+	+	w
12	+	+	+	+	w
13	+	+	+	+	w
14	+	+	+	+	w
15	+	+	+	+	w
16	+	+	+	+	b
17	+	+	+	+	w
18	+	+	+	+	b
19	+	+	+	+	b
20	+	+	+	+	w
21	+	+	+	+	w
22	+	+	+	+	w
24	+	+	+	+	w
25	+	+	+	+	b
26	+	+	+	+	b
27	+	+	+	+	b
28	+	+	+	+	w
29	+	+	+	+	w
30	+	+	+	+	w
31	+	+	+	+	b
32	+	+	+	+	w
34	+	+	+	+	w
35	+	+	+	+	w
36	+	+	+	+	w
37	+	+	+	+	w
38	+	+	+	+	w
39	+	+	+	+	b
40	+	+	+	+	w
41	+	+	+	+	w
42	+	+	+	+	w
43	+	+	+	+	w
44	+	+	+	+	b
45	+	+	+	+	w
46	+	+	+	+	w
47	+	+	+	+	b
48	+	+	+	+	w
49	+	+	+	+	w
50	+	+	+	+	b
52	+	+	+	+	b
53	+	+	+	+	w
54	+	+	+	+	w
55	+	+	+	+	w
57	+	+	+	+	w
58	+	+	+	+	b
59	+	+	+	+	b
60	+	+	+	+	b
62	+	+	+	+	w

Σημ: Οι 57 αποικίες που αναπτύχθηκαν στο 2^ο στάδιο επιλογής. Με (+) συμβολίζεται η ανάπτυξη στα διαφορετικά θρεπτικά μέσα. Με 'b' ή 'w' (τελευταία στήλη) συμβολίζονται οι αποικίες που εμφανίζουν μπλε (b) ή λευκό (w) χρώμα παρουσία X-gal. UAH είναι μη επιλεκτικό θρεπτικό μέσο (πείραμα θετικού ελέγχου), U= ουρακίλη, A= αδενίνη, H= ιστιδίνη. Η ακριβής σύσταση των θρεπτικών μέσων αναφέρεται στο Κεφ. Ύλικά και Μέθοδοι.

γ. Ομάδες επικαλυπτόμενων κλώνων

- Ομάδα I: 22 επικαλυπτόμενοι κλώνοι, οι οποίοι αντιπροσωπεύουν τμήμα της κωδικής περιοχής του γονιδίου *Numa1* (*Nuclear Mitotic Apparatus protein 1*).
- Ομάδες IIα και IIβ: 6 και 4 κλώνοι, αντίστοιχα, οι οποίοι αντιπροσωπεύουν δύο διαφορετικές περιοχές του mRNA του γονιδίου *Lek1* (*Leucine, glutamic acid, lysine family 1*).
- Ομάδα III: 2 επικαλυπτόμενοι κλώνοι, που αντιπροσωπεύουν τμήμα της κωδικής περιοχής του γονιδίου *Peg3* (*Paternally expressed gene 3*).
- Ομάδα IV: 2 όμοιοι κλώνοι, που αντιπροσωπεύουν τμήμα του μεταγράφου του γονιδίου *Cab140* (*Calcium binding protein 140*).
- Ομάδα V: 6 κλώνοι που αντιπροσωπεύουν τμήμα της κωδικής περιοχής του γονιδίου *Lam2b* (*Laminin B2*), το οποίο κωδικοποιεί την Β2 υπομονάδα της λαμνίνης.

Για όλους τους παραπάνω κλώνους ελέγχθηκε ότι τα ενθέματα έχουν κλωνοποιηθεί στο σωστό αναγνωστικό πλαίσιο του φορέα rVP16, ώστε να κωδικοποιούν τα αντίστοιχα πεπτίδια.

δ. Μοναδιαίοι κλώνοι

Από τους 15 μοναδιαίους κλώνους που απομονώθηκαν, οι 13 αντιστοιχούν σε γνωστά mRNAs, από τα οποία 6 είναι κλωνοποιημένα στο σωστό αναγνωστικό πλαίσιο του πλασμιδιακού φορέα rVP16 (Πίνακας 23), 5 είναι κλωνοποιημένα σε λάθος αναγνωστικό πλαίσιο, ένας κλώνος (ο #14) ανήκει στην 3' μη μεταφραζόμενη περιοχή και, τέλος, ο κλώνος #46 είναι προϊόν μεταγραφής με κατεύθυνση αντίστροφη από αυτήν γνωστού γονιδίου του ποντικού. Συγκεκριμένα, οι έξι κλώνοι με γνωστή νουκλεοτιδική αλληλουχία και σωστό πλαίσιο μετάφρασης είναι:

- Ο κλώνος #20, ο οποίος αντιπροσωπεύει περιοχή του γονιδίου *Gas8* (ή *Gas11*, *G1 arrest – specific*).
- Ο κλώνος #26, ο οποίος αντιπροσωπεύει το γονίδιο *Dlc-1* (*Dynein, cytoplasmic, light chain 1*) ή *LC8* και κωδικοποιεί 39 αμινοξέα από το καρβοξυτελικό άκρο της ελαφριάς αλυσίδας (8KD) της κυτταροπλασματικής δυνεΐνης.
- Ο κλώνος #41, που κωδικοποιεί μέρος της μεθυλτρανσφεράσης 3α του DNA (γονίδιο *Dnmt3a*).
- Ο κλώνος #42, που αντιπροσωπεύει το γονίδιο *Ubc9* (*Ubiquitin conjugating enzyme 9*) και κωδικοποιεί 155 από τα 158 αμινοξέα του γονιδιακού προϊόντος.
- Ο κλώνος #45, που αντιστοιχεί στο γονίδιο *Enp1*, το οποίο κωδικοποιεί την πρωτεΐνη *Enoplakin*.
- Τέλος, ο κλώνος #55 αντιπροσωπεύει το γονίδιο του ποντικού που κωδικοποιεί την πρωτεΐνη *ankycorbin*.

Πίνακας 22. Ομάδες επικαλυπτόμενων κλώνων

<p>Ομάδα I</p> <p>(22 κλώνοι)</p>	<p>mRNA <i>Numa1</i> (BC049791, 7214nt) nt: 2870-3526</p> <p>Πρωτεΐνη <i>Numa1</i> (AAH49791, 2094aa) aa: 902-1123</p>
<p>Ομάδα II</p> <p>α) (6 κλώνοι)</p> <p>β) (4 κλώνοι)</p>	<p>mRNA: <i>Lek1</i> (XM_129658, 9409nt), Πρωτεΐνη: <i>Lek1</i> (XP_129658, >2898aa)</p> <p>α) 21 466 566 nt: 1397-1821 28 492 592 aa: 466-607 → 34, 50 492 607</p> <p>β) → 12 1349 1616 nt: 4046-4964 24 1381 1443 aa: 1349-1655 8 1461 1520 37 1463 1603 44 1458 1655 52 1462 1622</p> <p>30 aa</p>
<p>Ομάδα III</p> <p>(2 κλώνοι)</p>	<p>→ 36 15 146 38 61 166</p> <p>mRNA: <i>Peg3</i> (AF038939, 5560nt), nt: 54-450 Πρωτεΐνη: <i>Peg3</i> (AAB96922, 1571aa), aa: 15-166</p> <p>30 aa</p>
<p>Ομάδα IV</p> <p>(2 κλώνοι)</p>	<p>→ 48, 49 820 923</p> <p>mRNA: <i>Cab140</i> (BC050107, 4406nt), nt: 2474-2768 Πρωτεΐνη: <i>Cab140</i> (AAH50107, 999aa), aa: 820-923</p> <p>30 aa</p>
<p>Ομάδα V</p> <p>(6 κλώνοι)</p>	<p>→ 3,4 1009 1190 9 1022 1163 19 1009 1111 32 1008 1126 35 1001 1130</p> <p>mRNA <i>Lam2b</i> (J02930, 7554nt) nt: 3194-3763</p> <p>Πρωτεΐνη <i>Laminin B2</i> (AAA39408, 1605aa) aa: 1001-1190</p> <p>30 aa</p>

Σημ: Ομάδες επικαλυπτόμενων κλώνων cDNA, που προέκυψαν από τη διερεύνηση της βιβλιοθήκης με 'δόλωμα' το πλήρες μόριο *Neur1*. Με κόκκινους χαρακτήρες (2^η στήλη) αναφέρονται οι αριθμοί των κλώνων cDNA. Παρουσιάζεται διαγραμματικά η σχετική επικάλυψη των κλώνων κάθε ομάδας. Στα άκρα κάθε κλώνου (οριζόντιες γραμμές) φαίνεται η αρίθμηση των αμινοξέων που κωδικοποιούνται, με βάση τις αλληλουχίες αναφοράς. Τα κόκκινα βέλη υποδεικνύουν τους κλώνους κάθε ομάδας που επιλέχθηκαν για το 3^ο στάδιο επιλογής. Οι κλώνοι #39, 53, 54 και 62 της ομάδας I ήταν χιμαιρικοί στο 3' άκρο τους, το οποίο παρουσιάζεται με διαφορετικό χρώμα. Σε κάθε ομάδα, αναφέρονται τα μετάγραφα (mRNAs) που αντιπροσωπεύονται, καθώς και οι αντίστοιχες πρωτεΐνες. Σε παρένθεση (2^η στήλη) είναι οι αριθμοί πρόσβασης στη βάση δεδομένων Genbank και το μέγεθος κάθε νουκλεοτιδικής (nt) ή αμινοξικής (aa) αλληλουχίας. Σε κάθε περίπτωση, υποδεικνύεται το συνολικό διάστημα (με πράσινο) της αλληλουχίας αναφοράς, που αντιπροσωπεύεται από τους κλώνους cDNA.

Πίνακας 23. Μοναδιαίοι γνωστοί κλώνοι

# 20	mRNA: <i>Gas8</i> ή <i>Gas11</i> (Growth arrest – specific 8 gene 8 ή 11, NM_018855, 1702nt) nt: 585-915 Πρωτεΐνη: Gas8 ή Gas11 (NP_061343, 489aa) aa: 125-235
# 26	mRNA: <i>Dlc-1</i> ή <i>LC8</i> (Dynein, cytoplasmic, light chain 1, 8KD, BC008106, 657nt) nt: 247-588 Πρωτεΐνη: Dlc-1 ή LC8 (AAH08106, 89aa) aa: 51-89
# 41	mRNA: <i>Dnmt3a</i> (DNA methyltransferase 3A, NM_007872, 9637nt) nt: 1431-1833 Πρωτεΐνη: Dnmt3a (NP_031898, 908aa) aa: 406-539
# 42	mRNA: <i>Ubc9</i> (Ubiquitin conjugating enzyme 9, X99739, 780nt) nt: 47-550 Πρωτεΐνη: Ubc9 (CAA68072, 158aa) aa: 1-155
# 45	Γονίδιο: <i>Evpl</i> (envoplakin), εξόνια 8-10 (AJ319609 nt 620-727 και 868-963 – AJ319610 nt 466-570) Πρωτεΐνη: Evpl (Q9D952, 2035aa) aa: 270-372
# 55	mRNA: <i>Ankycorbin</i> mRNA (AF202315, 3024nt) nt: 1407-1889 Πρωτεΐνη: Ankycorbin (AAG24483, 979aa) aa: 461-620

Σημ: Οι μοναδιαίοι κλώνοι, οι οποίοι ήταν κλωνοποιημένοι σε σωστό αναγνωστικό πλαίσιο (*in frame*) και οι νουκλεοτιδικές και αμινοξικές αλληλουχίες, τις οποίες αντιπροσωπεύουν. Με πλάγιους χαρακτήρες αναφέρονται τα ονόματα γονιδίων και με κανονικούς χαρακτήρες τα ονόματα πρωτεϊνών. Στις παρενθέσεις φαίνονται οι αριθμοί πρόσβασης κάθε αλληλουχίας στη βάση δεδομένων Genbank και τα μεγέθη των αντίστοιχων αλληλουχιών. Για το γονίδιο *Evpl* του ποντικού δεν είναι γνωστή η πλήρους μήκους αλληλουχία του μεταγράφου του, γι'αυτό οι αλληλουχίες που αναφέρονται είναι από ανάλυση γονιδιωματικού DNA. (aa= αμινοξέα, nt= νουκλεοτίδια)

ε. 3^ο στάδιο επιλογής

Σ' αυτό το στάδιο, η επιλογή πραγματοποιήθηκε σε 13 κλώνους, αντιπροσωπευτικούς των επιλεγμένων του 2ου σταδίου (βλ. Πίνακες 22,23). Για να διαπιστωθεί ότι η μεταγραφική ενεργοποίηση των γονιδίων αναφοράς που παρατηρήθηκε στα προηγούμενα στάδια οφείλεται στην ταυτόχρονη παρουσία τόσο της υβριδικής πρωτεΐνης GAL4-BD-Neur1, όσο και κάθε μίας από τις VP16-AD-prey πρωτεΐνες, πραγματοποιήθηκαν εκ νέου μετασχηματισμοί κυττάρων σακχαρομύκητα ρJ694a με την κατασκευή *Neur1/pGBT9*, καθώς και με καθέναν από τους 13 επιλεγμένους *cDNA-prey/pVP16* κλώνους. Οι μετασχηματισμένες αποικίες ελέγχθηκαν για την ενεργοποίηση της μεταγραφής των τριών γονιδίων αναφοράς στα θρεπτικά μέσα I, II και III. Και οι 13 επιλεγμένοι κλώνοι ήταν ικανοί να ενεργοποιούν τη μεταγραφή των γονιδίων *ADE2* και *HIS3*, ενώ 3 από αυτούς [#4(Lek1), #25(Numa1) και #26(Dlc-1)], ενεργοποιούσαν και το *LacZ* γονίδιο. Γενικά, παρατηρήθηκε το ίδιο πρότυπο ενεργοποίησης των γονιδίων-μαρτύρων με αυτό των προηγούμενων σταδίων επιλογής (βλ. Πίνακα 21). Είναι σημαντικό να τονιστεί ότι η μεταγραφική ενεργοποίηση που παρατηρήθηκε, εξαρτάται απόλυτα από την παρουσία της υβριδικής πρωτεΐνης GAL4-BD-Neur1 στα κύτταρα του σακχαρομύκητα. Αυτό διαπιστώθηκε σε πειράματα αρνητικού ελέγχου, στα οποία πραγματοποιήθηκε μετασχηματισμός με τις κατασκευές *cDNA-prey/pVP16*, παρουσία μόνο του πλασμιδια-

κού φορέα pGBT9. Στις μετασχηματισμένες αποικίες που προέκυψαν, δεν παρατηρήθηκε ενεργοποίηση των γονιδίων-μαρτύρων.

στ. Έλεγχος αλληλεπίδρασης των επιλεγμένων πεπτιδίων με τα ελλειμματικά μόρια *Neur1(1-521)* και *Neur1(515-574)*

Οι 13 επιλεγμένοι κλώνοι των προηγούμενων σταδίων ελέγχθηκαν για το αν προκαλούν μεταγραφική ενεργοποίηση εξαρτώμενη από τον GAL4, παρουσία των ελλειμματικών μορίων *Neur1(1-521)* και *Neur1(515-574)*, που χαρακτηρίζονται από την απουσία της περιοχής του RING στο πρώτο μόριο και την απουσία των δύο NEUZ επαναλήψεων στο δεύτερο, αντίστοιχα (Εικ.49). Για κάθε κλώνο *cDNA-prey/pVP16*, πραγματοποιήθηκε μετασχηματισμός σε κύτταρα σακχαρομύκητα pJ694a, παρουσία είτε της κατασκευής *Neur1(1-521)/pGBT9*, είτε της *Neur1(515-574)/pGBT9* και οι μετασχηματισμένες αποικίες μεταφέρθηκαν στα θρεπτικά μέσα επιλογής για τον έλεγχο της ενεργοποίησης των γονιδίων αναφοράς (Πίνακας 24). Οι κλώνοι *cDNA-prey/pVP16* που αντιπροσωπεύουν τα γονίδια *Numa1*, *Lek1*, *Ubc9*, *Lam2b* και *Gas8* εμφάνισαν ίδιο πρότυπο ενεργοποίησης των γονιδίων αναφοράς παρουσία του υβριδίου GAL4-BD-*Neur1(1-521)*, με αυτό που εμφάνιζαν παρουσία του πλήρους μορίου GAL4-BD-*Neur1* (Πίνακας 21). Αντίθετα, οι κλώνοι των γονιδίων *Dlc-1*, *Peg3*, *Dnmt3A*, *Evpl*, *Cab140* και *Ankycorbin* απέτυχαν να ενεργοποιήσουν τη μεταγραφή των γονιδίων-μαρτύρων απουσία της υπομονάδας RING του *Neur1*. Όμως, όπως διαπιστώθηκε από τον έλεγχο της μεταγραφικής ενεργοποίησης παρουσία του μορίου GAL4-BD-*Neur1(515-574)*, η περιοχή του RING δεν είναι αρκετή για να αλληλεπιδράσει με κανένα από τα 'θηράματα' που δοκιμάστηκαν, ώστε να ενεργοποιήσει τα γονίδια-μάρτυρες. Βέβαια, δεν αποκλείουμε την πιθανότητα η έκφραση της περιοχής RING να είναι τοξική για τον σακχαρομύκητα, όπως έχει παρατηρηθεί σε πειράματα εκτοπικής υπερέκφρασης της αντίστοιχης περιοχής του δροσοφιλικού *neur* [Lai & Rubin, 2001b]. Τα αποτελέσματα αυτά οδηγούν στο συμπέρασμα ότι η παρουσία των επαναλήψεων NEUZ1 και NEUZ2 είναι αρκετή για την αλληλεπίδραση του *Neur1* με τα 'θηράματα' *Numa1*, *Lek1*, *Ubc9*, *Lam2b* και *Gas8*, αλλά όχι και για την αλληλεπίδρασή του με τις πρωτεΐνες *Dlc-1*, *Peg3*, *Dnmt3a*, *Evpl*, *Cab140* και *Ankycorbin*, για τις οποίες μόνο το πλήρες μόριο *Neur1* είναι ικανό προκαλέσει ενεργοποίηση των γονιδίων-μαρτύρων στον σακχαρομύκητα (Εικ.51).

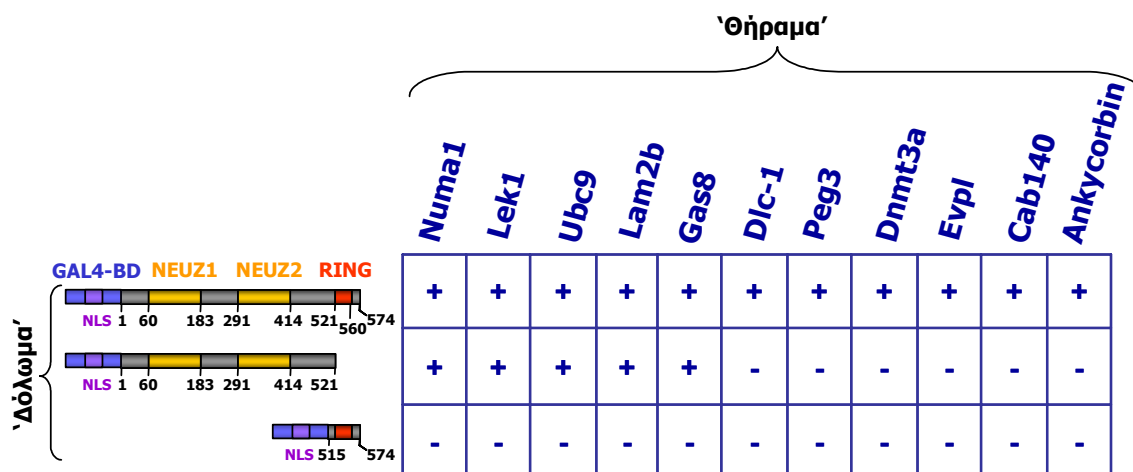
ζ. Επιλογή κλώνων για περαιτέρω ανάλυση

Το επόμενο βήμα μετά την ταυτοποίηση των πρωτεϊνών που αλληλεπιδρούν με την πρωτεΐνη *Neur1* στα κύτταρα του σακχαρομύκητα ήταν η επιβεβαίωση αυτών των αλληλεπιδράσεων σε *in vitro* συνθήκες (μέθοδος *GST-pulldown*) αλλά και στα κύτταρα των θηλαστικών (συν-ανοσοκατακρήμνιση). Οι μέθοδοι αυτές απαιτούν την κλωνοποίηση των υπό μελέτη cDNAs σε πλασμιδιακούς φορείς κατάλληλους για έκφραση των αντίστοιχων πρωτεϊνών σε βακτήρια και σε κύτταρα θηλαστικών (Υλικά και Μέθοδοι). Με στόχο τη συλλογή όσο το δυνατόν περισσότερων πληροφοριών για τη λειτουργία του *Neur1* και λόγω του μεγάλου αριθμού των πρωτεϊνών που ταυτοποιήθηκαν στο σύστημα δύο υβριδίων, επελέγησαν αρχικά ορισμένες μόνο από αυτές τις πρωτεΐνες για περαιτέρω ανάλυση. Τα κριτήρια της επιλογής ήταν ο αριθμός των αντιπροσωπευτικών κλώνων cDNA που είχαν επιλεγεί για κάθε πρωτεΐνη στο πείραμα των δύο υβριδίων, σε

Πίνακας 24. Πείραμα δύο υβριδίων με 'δόλωμα' ελλειμματικά μόρια Neur1

# κλώνου cDNA	Γονίδιο	A) Bait: GAL4-BD-Neur1 (1-521)					B) Bait: GAL4-BD-Neur1 (515-574)				
		UAH	UH	UA 3mM AT	UA 10mM AT	UAH Xgal	UAH	UH	UA 3mM AT	UA 10mM AT	UAH Xgal
25	<i>Numa1</i>	+	+	+	+	w	+	-	-	-	w
57	>>	+	+	+	+	w	+	-	-	-	w
34	<i>Lek1</i>	+	+	+	+	w	+	-	-	-	w
12	>>	+	+	+	+	b	+	-	-	-	W [®]
42	<i>Ubc9</i>	+	+	+	+	w	+	-	-	-	w
4	<i>Lam2b</i>	+	+	+	+	b	+	-	-	-	W [®]
20	<i>Gas8</i>	+	+	+	+	w	+	-	-	-	w
26	<i>Dlc-1</i>	+	-	-	-	W [®]	+	-	-	-	W [®]
36	<i>Peg3</i>	+	-	-	-	w	+	-	-	-	w
41	<i>Dnmt3a</i>	+	-	-	-	w	+	-	-	-	w
45	<i>Evpl</i>	+	-	-	-	w	+	-	-	-	w
49	<i>Cab140</i>	+	-	-	-	w	+	-	-	-	w
55	<i>Ankycorbin</i>	+	-	-	-	w	+	-	-	-	w

Σημ: Οι 13 επιλεγμένοι κλώνοι cDNA δοκιμάστηκαν σε πείραμα δύο υβριδίων, με δόλωμα το Neur1(1-521) (A) και το Neur1(515-574) (B). Με (+) συμβολίζεται η ανάπτυξη στα θρεπτικά μέσα. Με 'b' ή 'w' συμβολίζονται οι αποικίες που εμφανίζουν μπλε (b) ή λευκό (w) χρώμα παρουσία X-gal. Με W[®], οι αποικίες που ενώ εμφάνιζαν μπλε χρώμα στο πείραμα με το πλήρες μόριο Neur1 ή το Neur1(1-521), εμφανίζουν λευκό χρώμα με το μόριο Neur1(1-521) ή το Neur1(515-574), αντίστοιχα. UAH είναι μη επιλεκτικό θρεπτικό μέσο (πείραμα θετικού ελέγχου), U=ουρακίλη, A=αδερίνη, H=ιστιδίνη (βλ. 'Υλικά και Μέθοδοι')



Εικ.51 Συνολική απεικόνιση των αποτελεσμάτων του πειράματος των δύο υβριδίων. Για κάθε συνδυασμό πρωτεΐνης - 'δόλωμα' (Neur1) και πρωτεΐνης - 'θήραμα', απεικονίζεται η εμφάνιση (+) ή απουσία αλληλεπίδρασης (-), όπως αυτή ανιχνεύθηκε με βάση τη μεταγραφική ενεργοποίηση των γονιδίων-μαρτύρων στα κύτταρα σακχαρομύκητα rJ69-4a, στα οποία εκφράστηκαν οι παραπάνω συνδυασμοί πρωτεϊνών. (Οι πρωτεΐνες - 'θήραμα' εκφράστηκαν σε σύντηξη με την υπομονάδα ενεργοποίησης της μεταγραφής VP16-AD.)

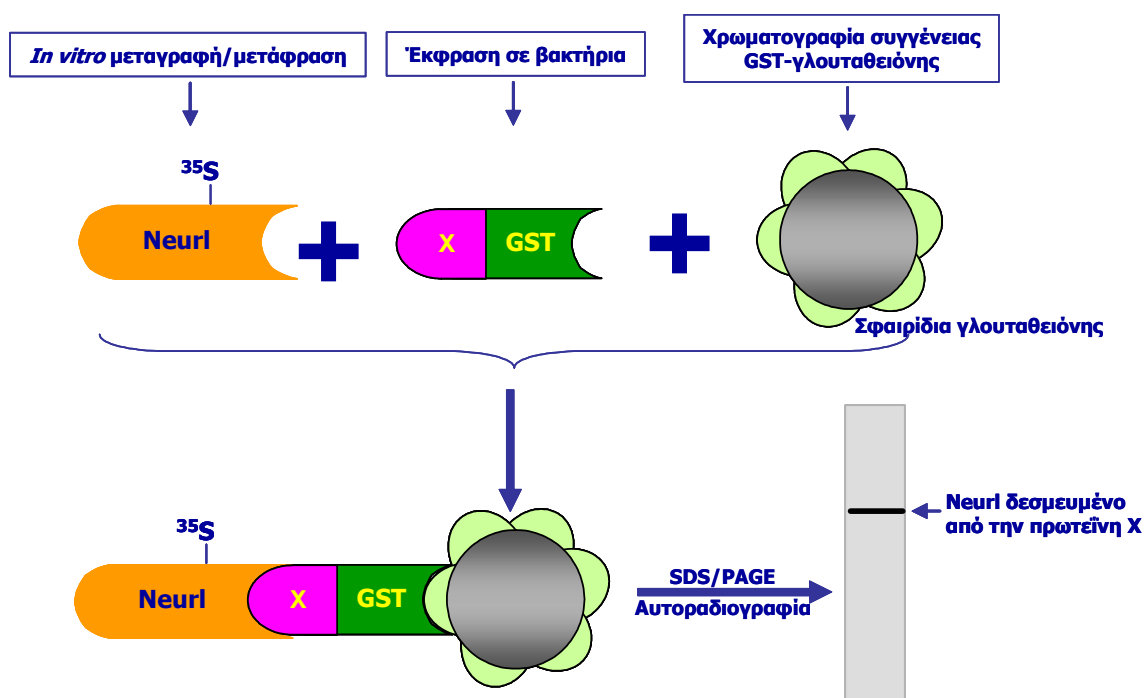
συνδυασμό και με τις υπάρχουσες πληροφορίες για τη λειτουργία της. Με βάση αυτά τα κριτήρια επιλέχθηκαν οι πρωτεΐνες *Numa1*, *Lek1*, *Dlc-1* και *Ubc9*.

Σύμφωνα με αποτελέσματα προηγούμενων μελετών μας για τη δομή, το πρότυπο έκφρασης και τον υποκυτταρικό εντοπισμό της πρωτεΐνης Neur1 του ποντικού [Pavlopoulos *et al*, 2002 και Εισαγωγή], τον

φαινότυπο της απενεργοποίησης του *Neur1* [Vollrath *et al*, 2001 / Ruan *et al*, 2001], αλλά και τις ενδείξεις από ορθόλογα μόρια Neutralized για τη βιοχημική του δράση ως E3 λιγάση [Yeh *et al*, 2001/ Lai *et al*, 2001/ Deblandre *et al*, 2001], είναι δυνατόν να δικαιολογηθούν οι αλληλεπιδράσεις της πρωτεΐνης *Neur1* με καθένα από τα παραπάνω μόρια (βλ. Συζήτηση). Είναι όμως σημαντικό να τονιστεί ότι κανένα από τα 11 'θηράματα' δεν έχει αναφερθεί ως 'ψευδώς θετικό' αποτέλεσμα σε πείραμα δύο υβριδίων [Hengen, 1997 και www.fccc.edu/research/labs/golemis/InteractionTrapInWork.htm]. Επομένως, η επιλογή των παραπάνω μορίων στηρίχθηκε στην πιθανή συσχέτιση της λειτουργίας τους με το *Neur1*, με βάση τα διαθέσιμα για κάθε πρωτεΐνη στοιχεία (βλ. Συζήτηση).

4. Έλεγχος αλληλεπιδράσεων της πρωτεΐνης *Neur1* *in vitro*

Από το πείραμα των δύο υβριδίων προέκυψαν οι υβριδικές πρωτεΐνες που αλληλεπιδρούν με την πρωτεΐνη *Neur1*, προκαλώντας μεταγραφική ενεργοποίηση των γονιδίων-μαρτύρων του σακχαρομύκητα. Στη συνέχεια, για 4 από αυτές (*Ubc9*, *Dlc-1*, *Lek1* και *Numa1*) επιβεβαιώνουμε την άμεση αλληλεπίδραση με την πρωτεΐνη *Neur1*, χρησιμοποιώντας την *in vitro* μέθοδο 'GST-pulldown' (Εικ.52). Συγκεκριμένα, ελέγχθη-



Εικ.52 Σχηματική αναπαράσταση της *in vitro* μεθόδου 'GST-pulldown', που χρησιμοποιήθηκε. Η πρωτεΐνη *Neur1* παράγεται ραδιοσημασμένη *in vitro*, παρουσία μεθειονίνης- ^{35}S . Οι συντηγμένες με GST πρωτεΐνες (X-GST) παράγονται σε βακτήρια. Τα μόρια της πρωτεΐνης *Neur1* που αλληλεπιδρούν με την X, δεσμεύονται στα σφαιρίδια με επικάλυψη γλουταθειόνης, λόγω της συγγένειας του ενζύμου GST (*Glutathione-S-transferase*) με το υπόστρωμά του (γλουταθειόνη), κατακρημνίζονται και ανιχνεύονται με SDS/PAGE και αυτοραδιογραφία (Υλικά και Μέθοδοι).

καν οι αλληλεπιδράσεις τόσο του πλήρους μήκους μορίου Neur1(5-574), όσο και των επιμέρους τμημάτων του Neur1(5-390) και Neur1(391-574) με τα πεπτίδια Ubc9, Dlc-1, Lek1 και Numa1. Στα παρακάτω πειράματα 'GST-pulldown', τα μόρια Neur1 έχουν προκύψει από *in vitro* σύνθεση και είναι ραδιοσημασμένα με Met-S35, ενώ τα πεπτίδια Ubc9, Dlc-1, Lek1 και Numa1 εκφράστηκαν σε βακτήρια σε σύντηξη με την *Glutathione-S-transferase* (GST).

α. Έλεγχος αλληλεπιδράσεων του Neur1 πλήρους μήκους

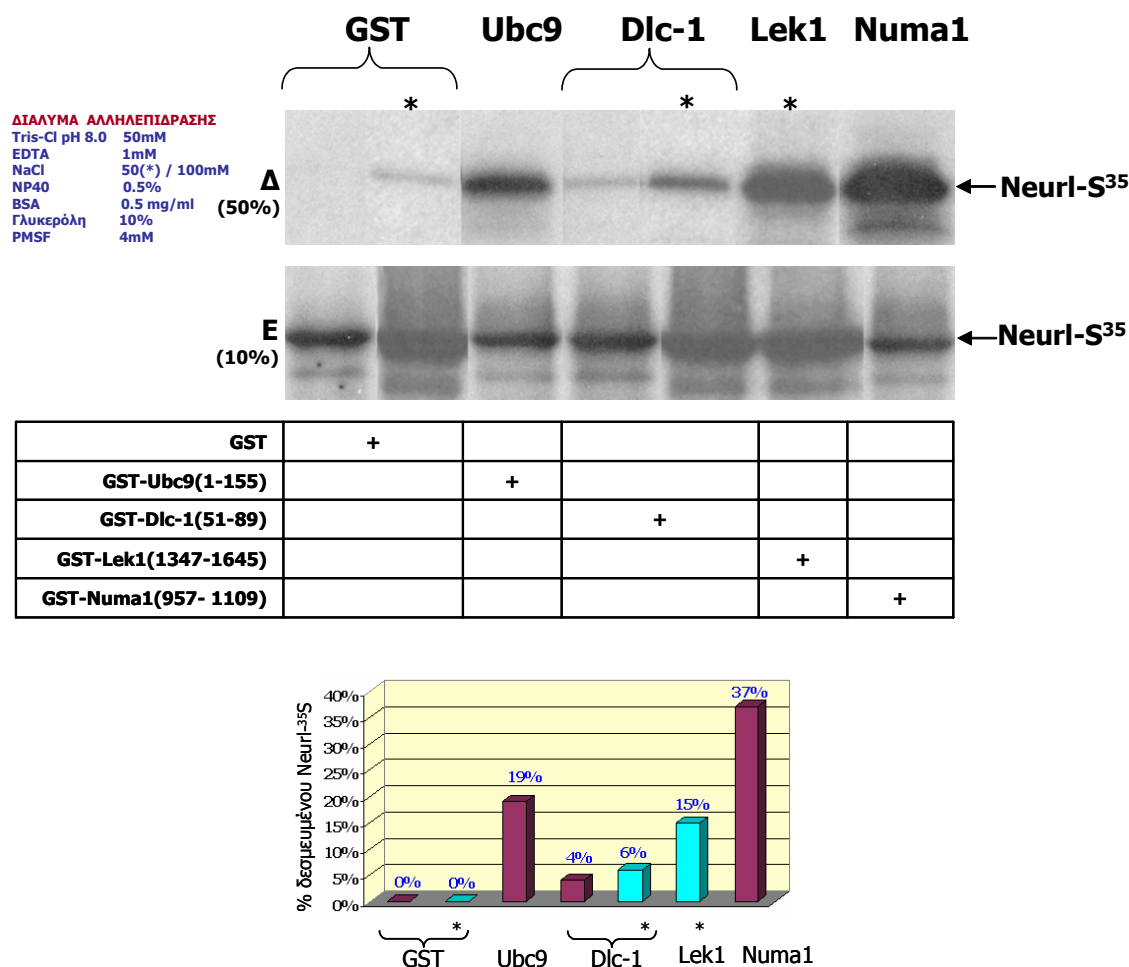
Για να ελεγχθεί η άμεση αλληλεπίδραση της πρωτεΐνης Neur1 με τα πεπτίδια Ubc9, Dlc-1, Lek1 και Numa1, βακτηριακά κύτταρα *E.coli* BL21-DE3-Codon Plus μετασχηματίστηκαν με κάθε μία από τις κατασκευές *42(Ubc9)/pGEX-3x*, *26(Dlc-1)/pGEX-3x*, *12(Lek1)/pGEX-3x* και *57(Numa1)/pGEX-3x* (βλ. Κεφ. 'Πλασμιδιακές Κατασκευές') και ακολούθησε επαγωγή με 0.5mM IPTG για 4 ώρες στους 30°C. Αφού ελέγχθηκε η παραγωγή των πεπτιδίων, ακολούθησε καθαρισμός τους από το βακτηριακό εκχύλισμα με χρωματογραφία συγγένειας σε σφαιρίδια γλουταθειόνης και στη συνέχεια πραγματοποιήθηκε η δοκιμή 'GST-pulldown', αναμιγνύοντας κάθε πεπτίδιο με ραδιοσημασμένη πρωτεΐνη Neur1 (προϊόν *in vitro* μεταγραφής-μετάφρασης της κατασκευής *Neur1/pCDNA3.1*, παρουσία μεθειονίνης-35S), στις συνθήκες που περιγράφονται στην Εικ.53. Από αυτά τα πειράματα, επιβεβαιώνεται η αλληλεπίδραση των πεπτιδίων Numa1, Lek1, Ubc9 και Dlc-1 με την πρωτεΐνη Neur1 *in vitro*. Συγκεκριμένα, στις συνθήκες αλληλεπίδρασης *in vitro*, παρατηρήθηκε ισχυρότερη δέσμευση του Neur1 στην πρωτεΐνη Numa1, σε σχέση με τις υπόλοιπες, πράγμα που επιβεβαιώνεται και από τα αποτελέσματα του πειράματος των δύο υβριδίων. Ακολουθούν οι πρωτεΐνες Lek1 και Ubc9, ενώ μικρότερη συγγένεια σύνδεσης, η οποία ενισχύεται σε διάλυμα ελαττωμένης αλατότητας (50mM NaCl), διαπιστώθηκε μεταξύ των πρωτεϊνών Neur1 και Dlc-1.

β. Έλεγχος αλληλεπίδρασης των ελλειμματικών μορίων Neur1

Με σκοπό τη χαρτογράφηση των περιοχών αλληλεπίδρασης στο μόριο Neur1, δοκιμάστηκαν οι αλληλεπιδράσεις των επιλεγμένων πεπτιδίων Ubc9, Dlc-1, Lek1 και Numa1 με δύο ελλειμματικά μόρια Neur1: α) το Neur1(5-390), που περιέχει την επανάληψη NEUZ1, την ενδιάμεση περιοχή NEUZ1-NEUZ2 και τα 100 N-τελικά αμινοξέα της NEUZ2 και β) το Neur1(391-574), που περιέχει το υπόλοιπο μόριο Neur1 (24 C-τελικά αμινοξέα της NEUZ2, την ενδιάμεση περιοχή NEUZ2-RING και το RING domain) (Εικ.54). Πραγματοποιήθηκαν πειράματα *in vitro* δοκιμής 'GST-pulldown', στα οποία τα χιμαιρικά πεπτίδια Ubc9, Dlc-1, Lek1 και Numa1 (προσδεδεμένα σε σφαιρίδια γλουταθειόνης) αναμίχθηκαν με: α) το προϊόν της *in vitro* μεταγραφής-μετάφρασης της κατασκευής *Neur1(5-390)/pCDNA3.1*, και β) το προϊόν της *in vitro* μεταγραφής-μετάφρασης της κατασκευής *Neur1(391-574)/pCDNA3.1HisB* (βλ. 'Πλασμιδιακές Κατασκευές').

Από αυτές τις δοκιμές αλληλεπίδρασης διαπιστώθηκε ελαττωμένη συγγένεια των ελλειμματικών μορίων Neur1 και με τα 4 πεπτίδια (Ubc9, Dlc-1, Lek1, Numa1), σε σχέση με το πλήρους μήκους Neur1. Όμως, φαίνεται ότι το αμινοτελικό τμήμα Neur1(5-391), που περιέχει τις επαναλήψεις NEUZ, διατηρεί σε σημαντικό βαθμό τη συγγένεια και με τα 4 πεπτίδια (Εικ.55). Αντίθετα, το συμπληρωματικό C-τελικό τμήμα Neur1(391-574), που περιέχει την ενδιάμεση περιοχή NEUZ2-RING και το RING, αλληλεπιδρά με ελαττω-

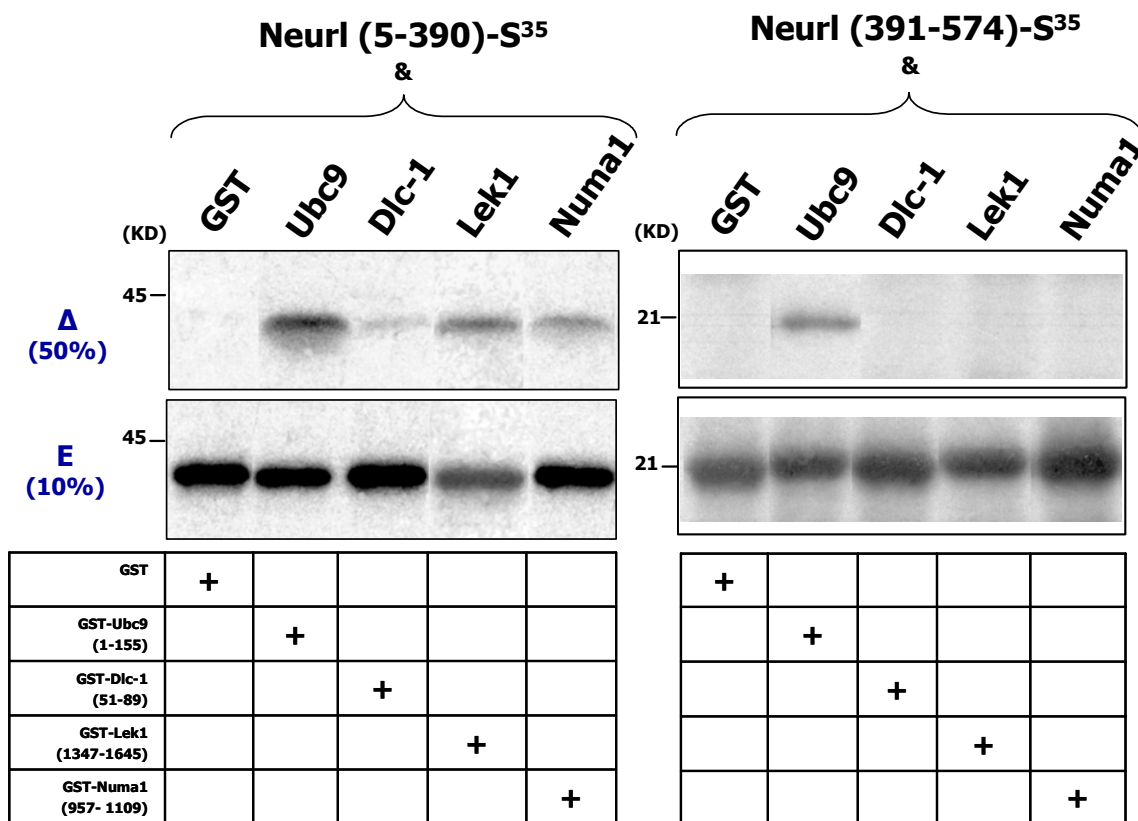
μένη συγγένεια μόνο με την πρωτεΐνη Ubc9, ενώ δεν αρκεί για την αλληλεπίδραση του Neur1 με τις πρωτεΐνες Dlc-1, Lek1 και Numa1 (Εικ.55).



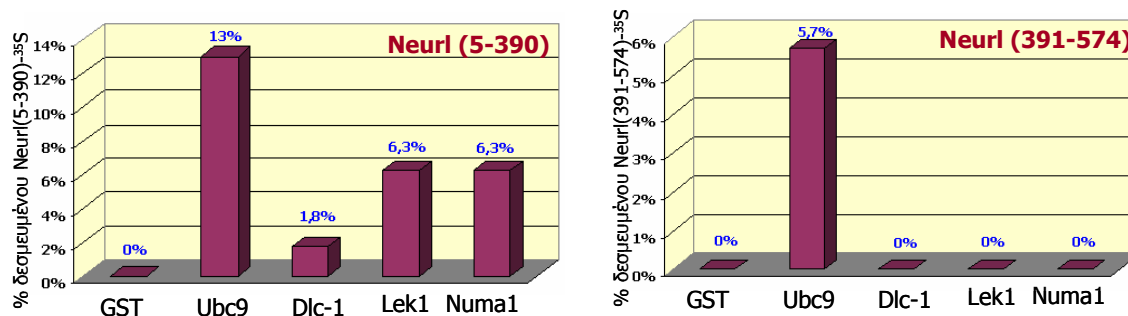
Εικ.53 *In vitro* δοκιμές αλληλεπίδρασης 'GST-pulldown' του Neur1(5-574) με τα πεπτίδια GST-Ubc9(1-155), GST-Dlc-1(51-89), GST-Lek1(1347-1645) και GST-Numa1(957-1109). Με (+) συμβολίζεται στον πίνακα η παρουσία της κάθε χημικικής πρωτεΐνης, δεσμευμένης στα σφαιρίδια γλουταθειόνης, στο πείραμα του οποίου το αποτέλεσμα εικονίζεται στην αντίστοιχη διαδρομή. Στην πρώτη σειρά φαίνεται το 50% των δειγμάτων του συμπλόκου Neur1/GST-πρωτεΐνης, στα οποία ανιχνεύεται το δεσμευμένο (Δ) Neur1. Στην δεύτερη σειρά φαίνεται το 10% των δειγμάτων του υπερκείμενου διαλύματος (E= ελεύθερο Neur1) των αντίστοιχων πειραμάτων αλληλεπίδρασης. Για κάθε πείραμα, 5μg πρωτεΐνης GST (πείραμα αρνητικού ελέγχου) ή 5μg από κάθε συζευγμένο με GST πεπτίδιο, δεσμεύτηκαν σε σφαιρίδια γλουταθειόνης και επώαστηκαν με 8μl αντίδρασης *in vitro* μεταγραφής/μετάφρασης της πλασμιδιακής κατασκευής *Neur1(5-574)/pCDNA3.1* σε διάλυμα αλληλεπίδρασης (50mM Tris-Cl pH 8.0, 1mM EDTA, 50 ή 100mM NaCl, 0.5% NP40, 0.5 mg/ml BSA, 10% γλυκερόλη, 4mM PMSF), συνολικού όγκου 300μl. Με αστερίσκους (*) και στίγες μπλε χρώματος συμβολίζονται τα πειράματα στα οποία χρησιμοποιήθηκε 50mM NaCl. Τα δείγματα επώαστηκαν για 12 ώρες στους 4°C, σε περιστρεφόμενο αναδευτήρα. Ακολούθησε φυγοκέντρηση και ηλεκτροφόρηση σε πήκτωμα SDS-πολυακρυλαμίδης 10%. Το πήκτωμα επώαστηκε για 30 λεπτά σε σταθεροποιητικό διάλυμα (30% μεθανόλη, 10% οξικό οξύ, 3% γλυκερόλη) και για 1 ώρα σε διάλυμα 1M σαλικλικού νατρίου pH 6.6. Ακολούθησε έκθεση σε φιλμ αυτοραδιογραφίας για 12 ώρες στους -80°C, στο οποίο ανιχνεύθηκε το ραδιοενεργό μόριο Neur1 (Neur1-S³⁵), μεγέθους 62KD. Τέλος, παριστάνεται διαγραμματικά το ποσοστό του δεσμευμένου Neur1 σε κάθε πείραμα.

Για τις πρωτεΐνες Numa1, Ubc9 και Lek1, τα αποτελέσματα αυτά είναι συγκρίσιμα με αυτά των πειραμάτων δύο υβριδίων στα οποία χρησιμοποιήθηκαν ως 'δόλωμα' τα ελλειμματικά μόρια Neur1(1-521) και Neur1(515-574). Και οι τρεις πρωτεΐνες αλληλεπίδρασαν με το Neur1(1-521) στα κύτταρα του σακχαρομύκητα, όπως και στο *in vitro* πείραμα αλληλεπίδρασης με το Neur1(5-390). Επομένως, φαίνεται ότι τα αμινοξέα 5-390 του Neur1 είναι αρκετά για την αλληλεπίδρασή του με τις πρωτεΐνες Numa1, Ubc9 και Lek1. Η πρωτεΐνη Ubc9 δεν εμφάνισε αλληλεπίδραση με το Neur1(515-574), που περιέχει το RING domain, στα κύτταρα

του σακχαρομύκητα, ενώ στο *in vitro* πείραμα αλληλεπίδρα με μικρή συγγένεια με το Neur1(391-574). Αυτό πιθανόν να σημαίνει ότι η αλληλεπίδραση του Ubc9 με το Neur1 δεν περιλαμβάνει το RING domain. Η Dlc-1 στο πείραμα των δύο υβριδίων δεν αλληλεπίδρασε με κανένα από τα ελλειμματικά μόρια Neur1, ενώ *in vitro* αλληλεπίδρα με το Neur1(5-390).

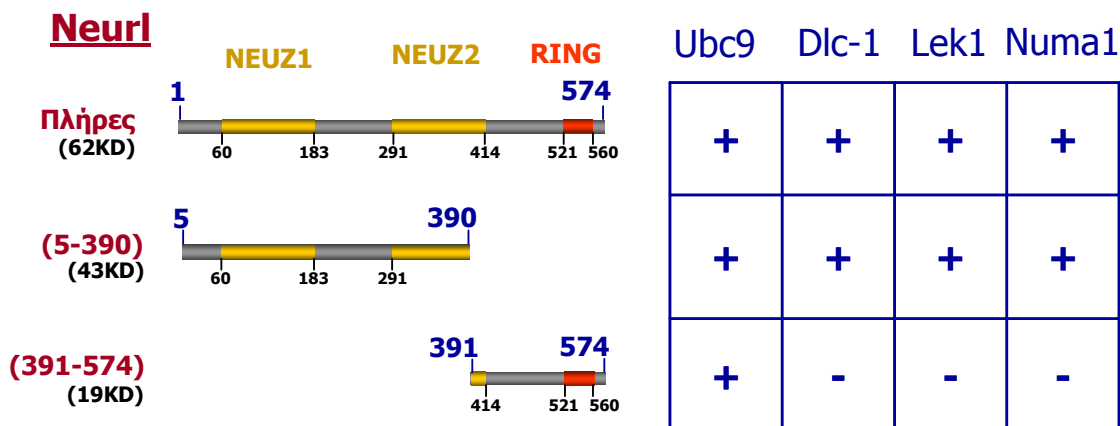


ΔΙΑΛΥΜΑ ΑΛΛΗΛΕΠΙΔΡΑΣΗΣ: 50mM Tris-Cl pH 8.0, 1mM EDTA, 100mM NaCl, 0.5% NP40, 0.5 mg/ml BSA, 10% γλυκερόλη και 4mM PMSF



Εικ. 55 *In vitro* δοκιμές αλληλεπίδρασης 'GST-pull-down' των ελλειμματικών μορίων Neur1(5-390) και Neur1(391-574) με τα χημικά πεπτίδια GST-Ubc9(1-155), GST-Dlc-1(51-89), GST-Lek1(1347-1645) και GST-Numa1(957-1109). Με (+) συμβολίζεται στους πίνακες η παρουσία της κάθε χημικής πρωτεΐνης, δεσμευμένης στα σφαιρίδια γλουταθειόνης, στο πείραμα του οποίου το αποτέλεσμα εικονίζεται στην αντίστοιχη διαδρομή. Οι δοκιμές 'GST-pull-down' πραγματοποιήθηκαν στις συνθήκες που περιγράφονται στην Εικ.53, με τις ραδιοσημασμένες πρωτεΐνες Neur1(5-390)-S³⁵ και Neur1(391-574)-S³⁵ και 5μg κάθε GST-πεπτιδίου, σε διάλυμα αλληλεπίδρασης συνολικού όγκου 300μl. Για κάθε *in vitro* δοκιμή αλληλεπίδρασης ηλεκτροφορήθηκε το 50% του δεσμευμένου (Δ) μορίου Neur1(5-390)-S³⁵ ή Neur1(391-574)-S³⁵ και το 10% του υπερκείμενου διαλύματος (E= ελεύθερα μόρια Neur1(5-390)-S³⁵ ή Neur1(391-574)-S³⁵) σε πήκτωμα SDS-πολυακρυλαμίδης 15%. Ακολούθησε ενίσχυση του σήματος με σαλικλικό νάτριο και αυτοραδιογραφία για 12 ώρες στους -80°C. Τα διαγράμματα παριστάνουν τα ποσοστά των δεσμευμένων μορίων Neur1(5-390)-S³⁵ και Neur1(391-574)-S³⁵ για κάθε πείραμα.

Η διαφορά αυτή πιθανόν οφείλεται στη διαφορετική φύση των δύο μεθόδων και ίσως να αντανακλά τη σχετικά χαμηλή συγγένεια σύνδεσης αυτής της πρωτεΐνης με το *Neur1*, όπως διαπιστώθηκε και από το πείραμα 'GST-pulldown' με ολόκληρο το *Neur1* cDNA.



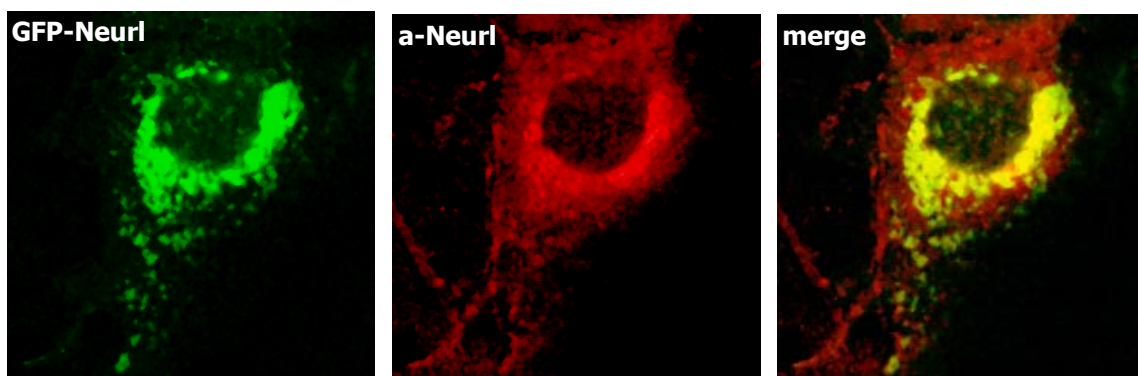
Εικ. 54 Χαρτογράφηση των αλληλεπιδράσεων στην πρωτεΐνη *Neur1*. Αποτελέσματα αλληλεπίδρασης των GST-συντηγμένων πεπτιδίων *Ubc9*, *Dlc-1*, *Lek1* και *Numa1* με το πλήρες μήκος *Neur1* ή τα ελλειμματικά μέρη *Neur1*(5-390) και *Neur1*(391-574), που εκφράστηκαν *in vitro*. Με (+) και (-) συμβολίζεται η εμφάνιση ή η απουσία αλληλεπίδρασης, αντίστοιχα. Η κλωνοποίηση των ελλειμματικών μορίων *Neur1* περιγράφεται στο Κεφάλαιο 'Πλασμιδιακές Κατασκευές'.

5. Διερεύνηση των πρωτεϊνικών αλληλεπιδράσεων του *Neur1* με οπτική μικροσκοπία

Μια καλή ένδειξη αλληλεπίδρασης δύο ενδοκυττάρων πρωτεϊνών είναι να παρουσιάζουν επικάλυψη στο κυτταρικό/ιστικό πρότυπο έκφρασης. Μετά τον προσδιορισμό των πρωτεϊνών που αλληλεπιδρούν με το *Neur1* στο σύστημα δύο υβριδίων του σακχαρομύκητα και την επιβεβαίωση αυτών των αλληλεπιδράσεων *in vitro*, ακολούθησαν πειράματα συγκριτικής μελέτης του υποκυτταρικού εντοπισμού τους, καθώς και συγκριτικής ανάλυσης του προτύπου έκφρασης των αντίστοιχων γονιδίων. Από την τελευταία μελέτη διαπιστώθηκε σημαντική επικάλυψη του προτύπου έκφρασης των γονιδίων *Ubc9*, *Numa1*, *Dlc-1* και *Lek1* με αυτήν του *Neur1* κατά την ανάπτυξη του ποντικού (Μπαζοπούλου ΜΔ, Διπλωματική Εργασία, Παν/μιο Κρήτης, 2002, βλ. και Συζήτηση). Εδώ μελετάται ο υποκυτταρικός συνεντοπισμός της πρωτεΐνης *Neur1* με τις αλληλεπιδρούσες πρωτεΐνες *Ubc9*, *Numa1* και *Dlc-1* σε κυτταρικές σειρές θηλαστικών και σε σπερματοζώαρια ενήλικου ποντικού. Επιπλέον, μετά την πρόσφατη διαπίστωση των Ling *et al* (2004) ότι η πρωτεΐνη *Dnmt3a* (το 'θήραμα' #41 του *Neur1*) αλληλεπιδρά με τη *Ubc9* και τροποποιείται από SUMO-1, διερευνήσαμε το συνεντοπισμό του *Neur1* και με αυτή την πρωτεΐνη στα κύτταρα θηλαστικών.

Για την ανίχνευση του ενδογενούς *Neur1* χρησιμοποιήθηκε το πολυκλωνικό αντίσωμα α-*Neur1* (Εικ.56). Για την ανίχνευση των υπόλοιπων ενδογενών πρωτεϊνών χρησιμοποιήθηκαν τα αντίστοιχα μονοκλωνικά αντισώματα και ακολούθησε ανοσοφθορισμός και παρατήρηση σε συνεστιακό ή σε οπτικό μικροσκόπιο φθορισμού ('Υλικά και Μέθοδοι'). Τα πειράματα αυτά πραγματοποιήθηκαν σε κυτταρικές σειρές

θηλαστικών (COS7, CV1, HeLa, Caco2). Η ανίχνευση της ενδογενούς πρωτεΐνης Numa1 ήταν δυνατή σε μία μόνο από τις κυτταρικές σειρές που δοκιμάστηκαν, γεγονός που πιθανόν σχετίζεται με τη σταδιακή απώλεια της συγκεκριμένης πρωτεΐνης στις κυτταρικές σειρές μετά από συγκεκριμένο αριθμό διαιρέσεων ή συνεχιζόμενη καλλιέργεια [Taimen *et al*, 2000]. Παρατηρώντας τον εντοπισμό του Neur1 σε διάφορες φάσεις της κυτταρικής διαίρεσης διαπιστώσαμε συνεντοπισμό με τους μικροσωληνίσκους, με την υπομονάδα της κυταροπλασματικής δυνεΐνης Dlc-1 και με τη Numa1. Τέλος, στα πειράματα που ακολουθούν ανιχνεύονται και τα ανασυνδυασμένα μόρια GFP-Neur1(5-574), Xpress-Ubc9, GFP-SUMO-1 και Flag-Dnmt3a. Ως μοριακοί δείκτες για ακριβέστερη περιγραφή του υποκυτταρικού εντοπισμού των υπό μελέτη πρωτεϊνών χρησιμοποιήθηκαν τα αντισώματα α -Lamin A/C και α - β -tubulin, που ανιχνεύουν τις λαμίνες A και C (πρωτεΐνες του εσωτερικού πυρηνικού φακέλου) και την β -tubulin (συστατικό των μικροσωληνίσκων), αντίστοιχα.

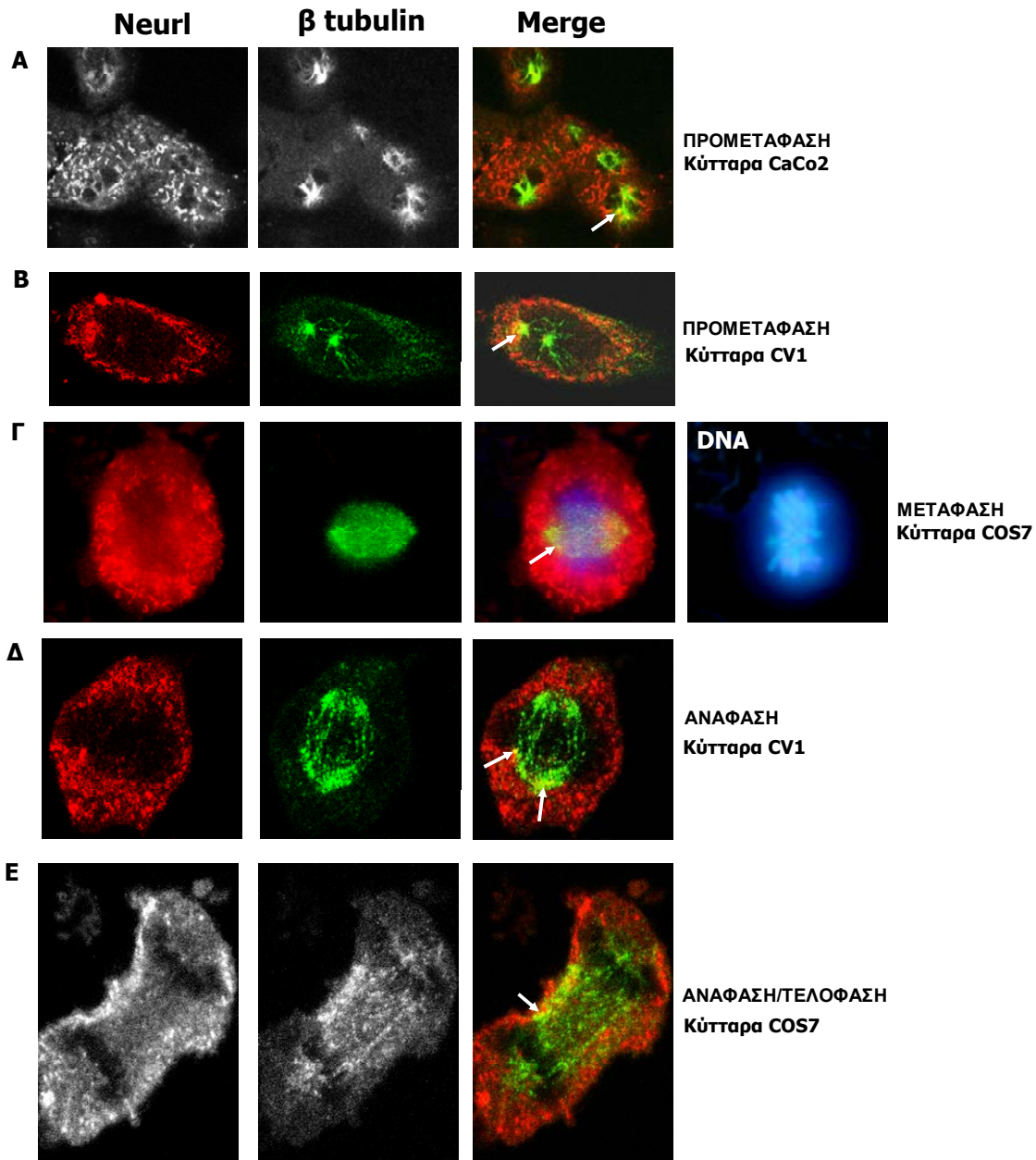


Εικ.56 Έλεγχος ειδικότητας του α -Neur1 σε κύτταρα θηλαστικών. Κύτταρα COS7 που εκφράζουν παροδικά την ανασυνδυασμένη πρωτεΐνη GFP-Neur1 (πράσινο) επώαστηκαν με το α -Neur1 και ανιχνεύθηκαν με ανοσοφθορισμό (κόκκινο). Από τον αναμενόμενο συνεντοπισμό (κίτρινο) ενδογενούς και ανασυνδυασμένης πρωτεΐνης προκύπτει ότι το πολυκλωνικό αντίσωμα α -Neur1 αναγνωρίζει ειδικά την πρωτεΐνη Neur1. [Ε. Κουτέλου, Διπλωματική Εργασία, Παν/μιο Κρήτης, 1999]

α. Εντοπισμός του Neur1 σε διάφορες φάσεις της κυτταρικής διαίρεσης

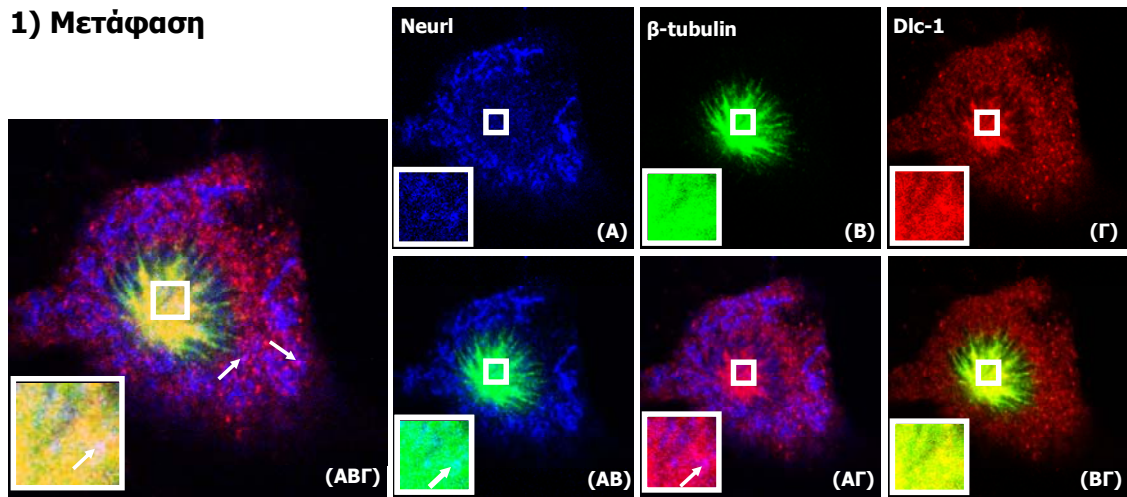
Από προηγούμενα πειράματα υποκυτταρικού εντοπισμού της πρωτεΐνης Neur1 σε κύτταρα θηλαστικών, με υπερέκφραση χιμαιρικών πολυπεπτιδίων Neur1 [Pavlopoulos, Kokkinaki *et al*, 2002] καθώς και ανοσοανίχνευση του ενδογενούς Neur1 [Ε. Κουτέλου, δημοσίευτα αποτελέσματα], διαπιστώθηκε ο εντοπισμός της πρωτεΐνης σε στικτές δομές, διάσπαρτες στο κυτταρόπλασμα, οι οποίες πολλές φορές σχημάτιζαν ένα δακτύλιο γύρω από τον πυρήνα, ενώ σε σημαντικό ποσοστό των κυττάρων παρατηρήθηκε και πυρηνικός εντοπισμός της πρωτεΐνης. Τα κύτταρα που παρουσιάζουν το παραπάνω πρότυπο έχουν πυρήνες με οργανωμένη πυρηνική μεμβράνη και βρίσκονται σε φάσεις S ή G1, όπως διαπιστώνεται με ταυτόχρονη χρώση του DNA με DAPI.

Επειδή σημαντικό ποσοστό των πρωτεϊνών-‘θηραμάτων’ του Neur1 συμμετέχουν στη διαδικασία της κυτταρικής διαίρεσης και ιδιαίτερα στον σχηματισμό της μιτωτικής ατράκτου (Numa1, Dlc-1) ή έχουν ρόλους που σχετίζονται με την κυτταρική διαίρεση (Lek1) [Goodwin *et al*, 1999/Ashe *et al*, 2004], μελετήσαμε τον εντοπισμό της πρωτεΐνης Neur1 στις διάφορες φάσεις της κυτταρικής διαίρεσης, χρησιμοποιώντας την β -tubulin ως μοριακό δείκτη. Ανιχνεύοντας ταυτόχρονα τις δύο πρωτεΐνες, διαπιστώνουμε συνεντοπισμό κατά την διαίρεση των κυττάρων. Ο συνεντοπισμός είναι μεγαλύτερος στα σημεία που

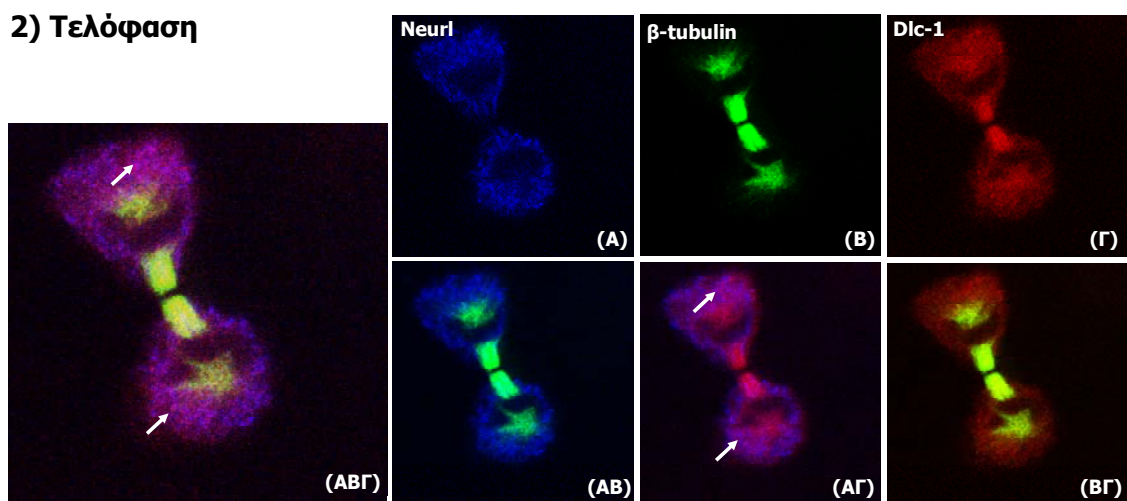


Εικ.57 Υποκυτταρικός εντοπισμός του Neur1 συγκριτικά με το δίκτυο των μικροσωληνίσκων (β tubulin) σε διάφορα στάδια της μίτωσης κυττάρων θηλαστικών. Σε κυτταρικές σειρές θηλαστικών (COS7, CV1, CaCo2) ανοσοανιχνεύθηκαν ταυτόχρονα οι πρωτεΐνες Neur1 (κόκκινο) και β -tubulin (πράσινο) και παρατηρήθηκε συνεντοπισμός (κίτρινο, βέλη) κατά την πρόωμη (A, B) και όψιμη (Γ) μετάφαση, την ανάφαση (Δ) και στα αρχικά στάδια της τελοφασης (Ε). Οπτικές τομές των κυττάρων παρατηρήθηκαν σε συνεστιακό μικροσκόπιο (A,B,Δ,E). Η εικόνα Γ προέκυψε από οπτικό μικροσκόπιο φθορισμού, με ταυτόχρονη παρατήρηση της χρώσης DAPI του DNA.

1) Μετάφαση



2) Τελόφαση

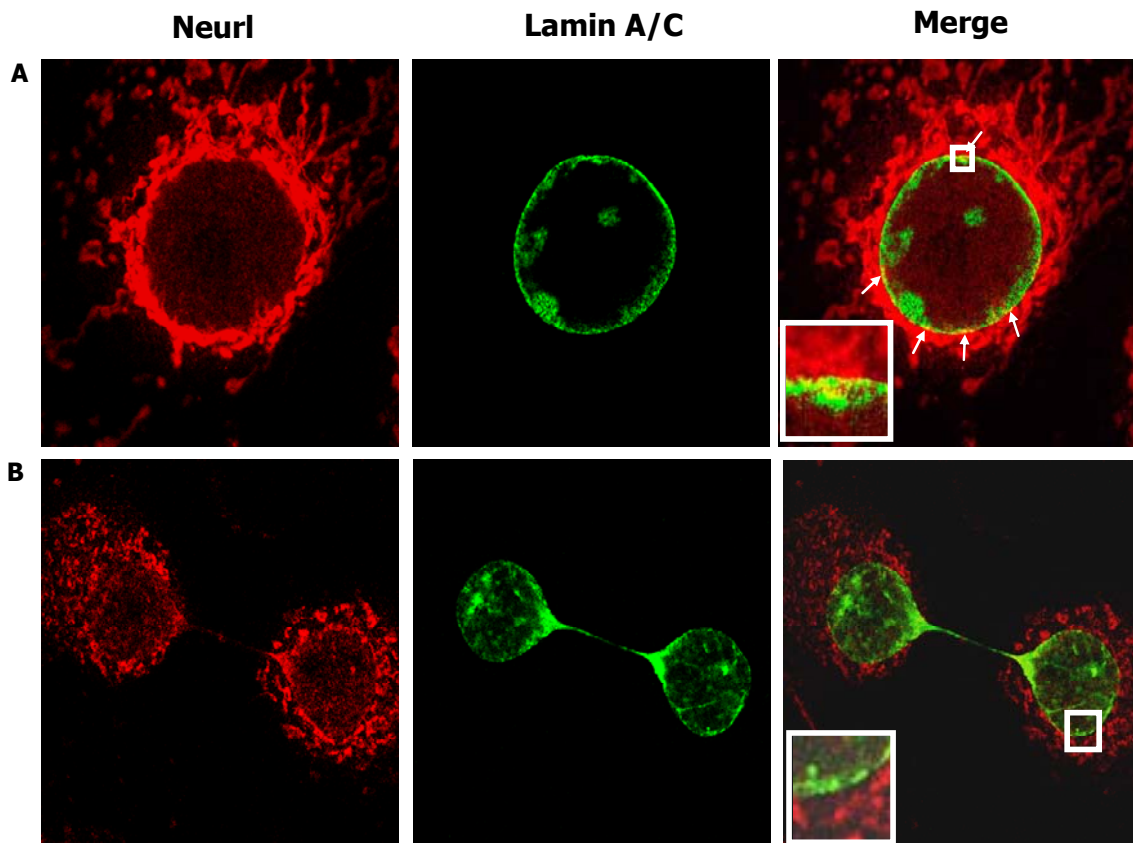


Εικ.58 Συνεντοπισμός Neu1 και Dlc-1 σε διαιρούμενα κύτταρα κατά τη μετάφαση (1) και τελόφαση (2). Οπτικές τομές κυττάρων COS7, μετά από ταυτόχρονη ανοσοανίχνευση των ενδογενών πρωτεϊνών Neu1 (A), β -tubulin (B) και Dlc-1 (Γ). **(1)**: Στη μετάφαση το Neu1 συνεντοπίζεται με την Dlc-1 κυρίως στην περιφέρεια της μιτωτικής ατράκτου (βλ. ιώδες στα AΓ και ABΓ, βέλη). Σε μεγαλύτερη μεγέθυνση (πλαίσιο) φαίνεται το μικρό ποσοστό μορίων Neu1 που συνεντοπίζεται με τη β -tubulin (γαλάζιο, AB) και με τη Dlc-1 (ιώδες, AΓ) στους πόλους της μιτωτικής ατράκτου. **(2)**: Κατά την τελόφαση, το Neu1 συνεντοπίζεται με την Dlc-1 κυρίως στην περιφέρεια των νεοσχηματιζόμενων πυρήνων (βλ. ιώδες στα AΓ και ABΓ, βέλη). Η Dlc-1 συνεντοπίζεται με την β -tubulin κατά μήκος των μικροσωληνίσκων της ατράκτου (κίτρινο, BΓ), ενώ δεν παρατηρείται σημαντικός συνεντοπισμός του Neu1 με την β -tubulin.

αρχίζουν να σχηματίζονται οι 'πολικοί αστέρες' των μικροσωληνίσκων και κοντά στους πόλους της ατράκτου στη μετάφαση και ανάφαση. Αντίθετα, κατά την τελόφαση δεν παρατηρήθηκε σημαντικός συνεντοπισμός της πρωτεΐνης Neur1 με την β -tubulin (Εικ.57). Τέλος, οι δύο πρωτεΐνες εντοπίζονται μαζί με την Dlc-1 στην περιοχή του κεντροσώματος (Εικ.58), όπου διαπιστώθηκε και η παρουσία της Numa-1 (βλ. παρακάτω).

β. Πυρηνικός εντοπισμός της πρωτεΐνης Neur1

Αρκετές από τις πρωτεΐνες-‘θηράματα’ του Neur1 εντοπίζονται και στον πυρήνα ή είναι αποκλειστικά πυρηνικές. Συγκεκριμένα, η Numa1 βρίσκεται στον πυρήνα ταχέως πολλαπλασιαζόμενων κυττάρων και στους πόλους της μιτωτικής ατράκτου κατά την κυτταρική διαίρεση [Taimen *et al*, 2000/Yang CH *et al*, 1992], η Lek1 έχει βρεθεί να αλληλεπιδρά με την πρωτεΐνη Rb στον πυρήνα κυττάρων θηλαστικών ρυθμίζοντας την είσοδο τους στη φάση της μίτωσης [Ashe *et al*, 2004], η Ubc9 (*SUMO conjugating enzyme*) έχει εντοπιστεί στο εσωτερικό του πυρήνα και αλληλεπιδρά με συστατικά του πυρηνικού φακέλου [Mingot *et al*, 2001/ Zhang *et al*, 2002], η Dnmt3a (*de novo DNA methyl transferase 3a*) εντοπίζεται στην ετεροχρωματίνη [Bachman *et al*, 2001], αλληλεπιδρά με απακετυλάσες ιστονών [Margot *et al*, 2001/Fuks *et al*, 2001] και προκαλεί μεταγραφική καταστολή, ενώ η Dlc-1 (υπομονάδα της κυτταροπλασματικής και αξονικής δυνεΐνης και της μωσίνης V) εντοπίζεται στον πυρήνα αλλά και στο κυτταρόπλασμα [Crepieux *et al*, 1997]. Όπως, λοιπόν, προκύπτει από το πρότυπο υποκυτταρικού εντοπισμού των πρωτεϊνών-‘θηραμάτων’ του Neur1, η αλληλεπίδραση των περισσότερων από αυτές (Numa1, Lek1, Ubc9, Dnmt3a) με την πρωτεΐνη Neur1 πρέπει είτε να συμβαίνει κατά την κυτταρική διαίρεση, είτε σε αυτήν να συμμετέχει ένα ποσοστό της πρωτεΐνης Neur1 που εντοπίζεται στο εσωτερικό του πυρήνα. Τέλος, δεν μπορεί να αποκλειστεί και το ενδεχόμενο αλληλεπίδρασης του Neur1 με αυτές τις πρωτεΐνες στις περιοχές των πυρηνικών πόρων, εφόσον τουλάχιστον για την Ubc9 είναι γνωστός ο εντοπισμός της στους πυρηνικούς πόρους και η μετακίνησή της από το κυτταρόπλασμα στον πυρήνα μέσω της importin 13 [Mingot JM *et al*, 2001], ενώ σύμφωνα με προηγούμενες παρατηρήσεις μας το Neur1 σχηματίζει έναν δακτύλιο γύρω από τον πυρήνα, λιγότερο ή περισσότερο συμπαγή [Pavlioroulos *et al*, 2002]. Για να διερευνήσουμε τον ενδεχόμενο εντοπισμό του Neur1 στον πυρήνα κυττάρων θηλαστικών με μεθόδους ανοσοεντοπισμού, χρησιμοποιήσαμε ως μοριακούς δείκτες τις λαμίνες A και C, πρωτεΐνες του εσωτερικού πυρηνικού φακέλου. Στο μεγαλύτερο ποσοστό των μεσοφασικών κυττάρων που παρατηρήθηκαν, ο δακτύλιος που σχημάτιζε το Neur1 γύρω από τον πυρήνα παρουσίαζε σε πολλά σημεία επικάλυψη με τον εσωτερικό πυρηνικό φάκελο, ενώ διαπιστώθηκε και χρώση του a-Neur1 στο εσωτερικό του πυρήνα, η οποία ήταν περισσότερο έντονη σε μικρό ποσοστό των κυττάρων (Εικ.59), ενδείξεις που υποδηλώνουν παρουσία μέρους της πρωτεΐνης Neur1 στον πυρήνα και μπορούν να δικαιολογήσουν τις αλληλεπιδράσεις της με τα συγκεκριμένα ‘θηράματα’. Σημειώνεται ότι οι παραπάνω παρατηρήσεις υποστηρίζονται και από βιοχημικά πειράματα κυτταρικής κλασμάτωσης, στα οποία σημαντικό μέρος του Neur1 εντοπίζεται στο πυρηνικό κλάσμα κυττάρων COS7 [E. Κουτέλου, αδημοσίευτα αποτελέσματα], καθώς και από τα πειράματα των Timmusk *et al* (2002), στα οποία η χρήση λεπτομυκίνης B (αναστολέας εξόδου από τον πυρήνα μέσω του συστήματος CRM1/exportin1) οδήγησε στον σχεδόν αποκλειστικά πυρηνικό εντοπισμό της ανασυνδυσμένης πρωτεΐνης Neur1 στα κύτταρα Neuro2a.



Εικ.59 Υποκυτταρικός εντοπισμός του Neur1 συγκριτικά με τις πρωτεΐνες του πυρηνικού φακέλου Lamin A/C. Οπτικές τομές κυττάρων COS7, ύστερα από ταυτόχρονη ανοσοανίχνευση των ενδογενών πρωτεϊνών Neur1 (κόκκινο) και Lamin A/C (πράσινο). Σε μη διαιρούμενα κύτταρα (Α), το Neur1 εμφανίζει επικάλυψη με τον εσωτερικό πυρηνικό φάκελο, στην περιφέρεια του πυρήνα. Αντίθετα, στους νέους πυρήνες που σχηματίζονται κατά την τελόφαση (Β), δεν υπάρχει σημαντική επικάλυψη του Neur1 με τον πυρηνικό φάκελο. Σε πλαίσιο φαίνονται ενδεικτικές περιοχές σε μεγέθυνση. Τα βέλη υποδεικνύουν πιθανά σημεία συνεντοπισμού Neur1 και Lamin A/C (κίτρινο).

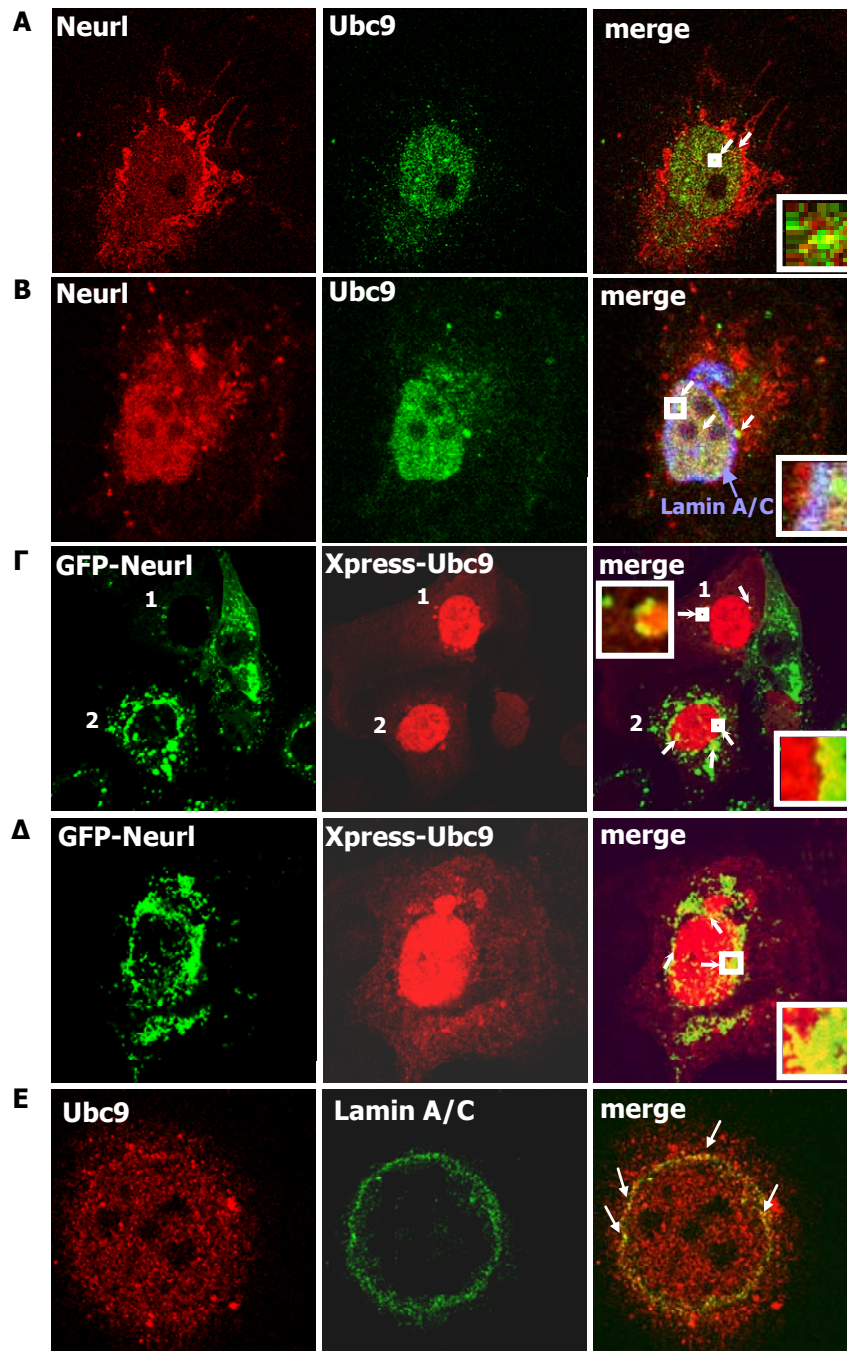
γ. Συνεντοπισμός των πρωτεϊνών *Neur1* και *Ubc9*

Για την πρωτεΐνη *Ubc9* είναι γνωστό ότι συσσωρεύεται στον πυρήνα κυττάρων θηλαστικών, στον οποίο εισέρχεται μέσω του μεταφορέα *Importin 13*, με μηχανισμό που ελέγχεται από την υδρόλυση του *RanGTP* [Mingot *et al*, 2001]. Επιπλέον, η *Ubc9* ανιχνεύεται σε μικρότερο βαθμό σε στικτές δομές στο κυτταρόπλασμα και στην περιοχή γύρω από τον πυρήνα. Ως βασικό συστατικό του μηχανισμού SUMΟλίωσης, η *Ubc9* αλληλεπιδρά με συστατικά των κυτταροπλασματικών και πυρηνοπλασματικών ινιδίων των πυρηνικών πόρων [Zhang *et al*, 2002]. Το πρότυπο αυτό συμφωνεί με αυτό που παρατηρούμε στα πειράματα χρώσης κυττάρων COS7 με τα αντισώματα α-*Ubc9* και α-*Lamin A/C* (Εικ.60Ε). Η *Ubc9* συσσωρεύεται κυρίως στον πυρήνα, σχηματίζοντας στικτές δομές. Διαπιστώνεται σημαντική επικάλυψη με τον πυρηνικό φάκελο. Μικρότερο ποσό της πρωτεΐνης *Ubc9* ανιχνεύεται σε στικτές δομές στο κυτταρόπλασμα (Εικ.60Ε). Ίδιο πρότυπο εντοπισμού παρατηρήθηκε και σε πειράματα ανοσοφθορισμού με το αντίσωμα α-*Xpress* σε κύτταρα που είχαν διαμολυνθεί παροδικά με την πλασμιδιακή κατασκευή *42(Ubc9)/pCDNA3.1HisA* (Εικ.60Γ,Δ). Ο υποκυτταρικός εντοπισμός της πρωτεΐνης *Ubc9* μελετήθηκε συγκριτικά με αυτόν του *Neur1* (Εικ.60Α-Δ). Ανιχνεύθηκαν είτε τα ενδογενή μόρια *Neur1* και *Ubc9*, είτε οι ανασυνδυασμένες πρωτεΐνες *GFP-Neur1(5-574)* και *Xpress-Ubc9*. Και στις δύο περιπτώσεις παρατηρήθηκε συνεντοπισμός σε διακριτά σημεία της περιπυρηνικής περιοχής ή και του εσωτερικού του πυρήνα.

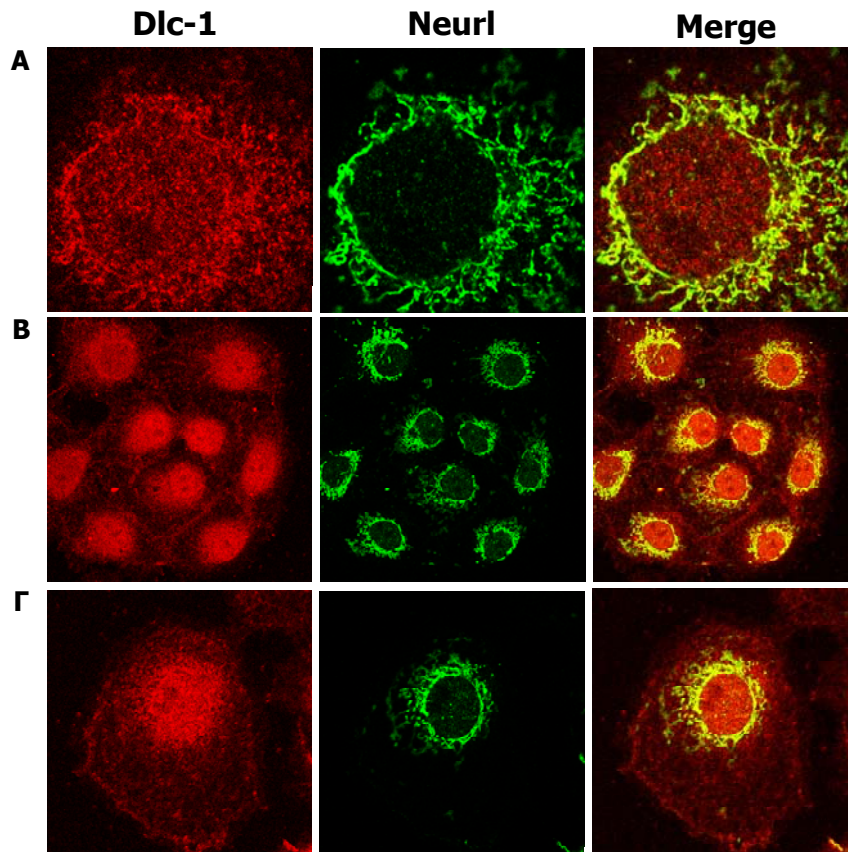
δ. Συνεντοπισμός των πρωτεϊνών *Neur1* και *Dlc-1*

Η πρωτεΐνη *Dlc-1* είναι υπομονάδα των πρωτεϊνικών συμπλόκων της κυτταροπλασματικής και αξονημικής δυνεΐνης και της μυοσίνης V, τα οποία αλληλεπιδρούν με τα ινίδια των μικροσωληνίσκων και της ακτίνης, αντίστοιχα. Κατά συνέπεια, το μονοκλωνικό αντίσωμα α-*Dlc-1* ανιχνεύει την πρωτεΐνη *Dlc-1* παντού στο κύτταρο, με εντονότερη χρώση στο εσωτερικό του πυρήνα και στην περιπυρηνική περιοχή και ασθενέστερη στην περιφέρεια του κυττάρου. Επιπλέον, διαπιστώθηκε η παρουσία της *Dlc-1* στην περιοχή του κεντροσώματος, στη μιτωτική άτρακτο και ιδιαίτερα στην περιοχή των πόλων της ατράκτου κατά την κυτταρική διαίρεση, σε συμφωνία και με προηγούμενες παρατηρήσεις [Alonso *et al*, 2001/ Mercedes *et al*, 2000/ Gaetz & Karoor, 2004]. Τέλος, όπως έχει αναφερθεί από βιοχημικά πειράματα και παρατηρήσεις ηλεκτρονικής μικροσκοπίας σε μαστιγοφόρους μονοκύτταρους ευκαρυωτικούς οργανισμούς του γένους *Chlamydomonas* [Yang *et al*, 2001], διαπιστώσαμε εντοπισμό της *Dlc-1* στο αξόνημα των μαστιγίων σπερματοζωαρίων ενήλικου ποντικού.

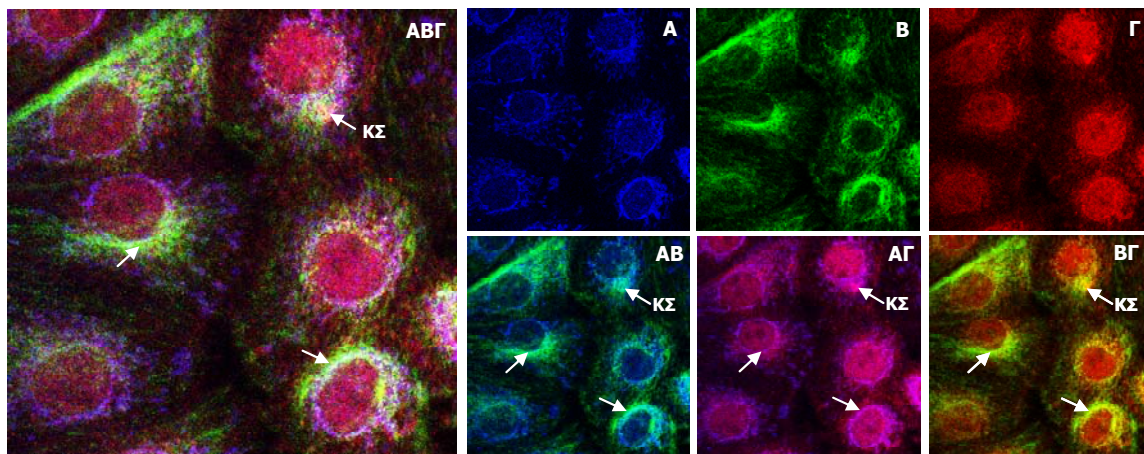
Στα πειράματα που παρουσιάζονται παρακάτω, ανιχνεύονται οι ενδογενείς πρωτεΐνες *Dlc-1* και *Neur1* σε κύτταρα COS7 και σε σπερματοζωάρια. Διαπιστώνουμε σημαντική επικάλυψη των δύο πρωτεϊνών στο κυτταρόπλασμα των κυττάρων COS7 και κυρίως στην περιοχή γύρω από τον πυρήνα, όπου εντοπίζεται το μεγαλύτερο ποσοστό του *Neur1* (Εικ.61). Επιπλέον, συνεντοπισμός των δύο πρωτεϊνών παρατηρήθηκε και στην περιοχή του κεντροσώματος, όπως φαίνεται από τα πειράματα τριπλής χρώσης κυττάρων COS7 με τα αντισώματα α-*Neur1*, α-*Dlc-1* και α-*β-tubulin* (Εικ.62). Ο ταυτόχρονος συνεντοπισμός του *Neur1* με την *Dlc-1* και τους μικροσωληνίσκους (*β-tubulin*) διατηρείται πιο περιορισμένος μέχρι και το στάδιο της μετάφρασης, όπου μικρό ποσοστό *Neur1* εντοπίζεται στην περιοχή των μικροσωληνίσκων της ατράκτου (Εικ.58),



Εικ.60 Συνεντοπισμός των πρωτεϊνών Neur1 και Ubc9 σε κύτταρα θηλαστικών. (A-E): Οπτικές τομές κυττάρων COS7. **(A,B):** Ταυτόχρονη ανοσοανίχνευση των ενδογενών πρωτεϊνών Neur1 (κόκκινο) και Ubc9 (πράσινο). Στο κύτταρο A, το ποσοστό του Neur1 στο εσωτερικό του πυρήνα είναι μικρότερο και ο συνεντοπισμός με Ubc9 είναι περιορισμένος σε σημεία της περιφέρειας ή του εσωτερικού του πυρήνα (βέλη). Αντίθετα, το κύτταρο B έχει μεγαλύτερο ποσοστό πυρηνικού Neur1, με αποτέλεσμα αυξημένο συνεντοπισμό με την πρωτεΐνη Ubc9. Τα όρια του πυρήνα στο κύτταρο B φαίνονται με ταυτόχρονη χρώση με το αντίσωμα α-Lamin A/C. **(Γ,Δ):** Ταυτόχρονη παροδική διαμόλυνση με τις πλασμιδιακές κατασκευές *Neur1/pEGFP1* και *42(Ubc9)/pCDNA3.1HisA* και ανίχνευση των παραγόμενων πρωτεϊνών GFP-Neur1 (πράσινο) και Xpress-Ubc9 (κόκκινο). Όπως και με τα ενδογενή μόρια (A,B), παρατηρούνται περιοχές συνεντοπισμού των ανασυνδυασμένων μορίων στην περιφέρεια και στο εσωτερικό του πυρήνα. Το κύτταρο 1 της εικόνας Γ εκφράζει χαμηλά επίπεδα GFP-Neur1, όμως εμφανίζει συνεντοπισμό των δύο πρωτεϊνών στην περιφέρεια του πυρήνα. Σε πλαίσιο φαίνονται χαρακτηριστικές περιοχές σε μεγέθυνση. **(Ε):** Πυρηνικός εντοπισμός της πρωτεΐνης Ubc9. Οπτική τομή κυττάρου COS7, ύστερα από ταυτόχρονη ανοσοανίχνευση της ενδογενούς πρωτεΐνης Ubc9 (κόκκινο) και των πρωτεϊνών του εσωτερικού πυρηνικού φακέλου Lamin A/C. Η Ubc9 εντοπίζεται κυρίως στο εσωτερικό του πυρήνα και σε μικρότερο ποσοστό στην περιφέρεια του, σχηματίζοντας στικτές δομές. Επιπλέον, σε αρκετά σημεία παρατηρείται συνεντοπισμός με τον πυρηνικό φάκελο (βέλη).



Εικ.61 Συνεντοπισμός Neu1 και Dlc-1 σε κύτταρα θηλαστικών. (Α-Γ): Οπτικές τομές κυττάρων COS7, ύστερα από ταυτόχρονη ανοσοανίχνευση των ενδογενών πρωτεϊνών Dlc-1 (κόκκινο) και Neu1 (πράσινο). Στην επικάλυψη των δύο τομών (merge) φαίνεται σημαντικός συνεντοπισμός (κίτρινο), στο κυτταρόπλασμα και κυρίως γύρω από τον πυρήνα.

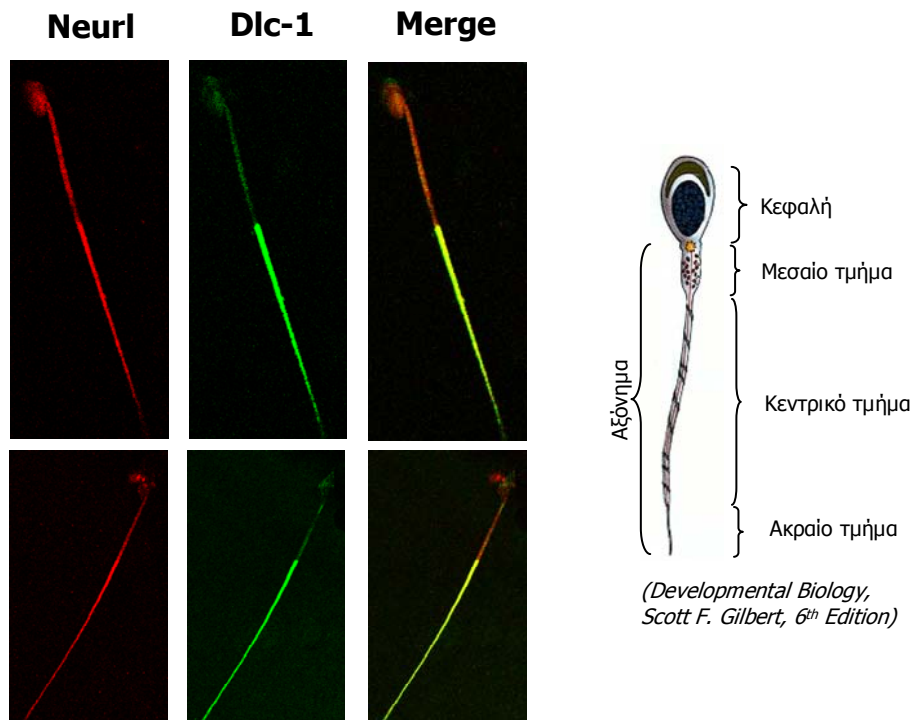


Εικ.62 Συνεντοπισμός Neu1, Dlc-1, β-tubulin σε κύτταρα θηλαστικών. Οπτική τομή κυττάρων COS7 μετά από ταυτόχρονη ανοσοανίχνευση των ενδογενών πρωτεϊνών Neu1 (Α), β-tubulin (Β) και Dlc-1 (Γ). ΑΒ, ΑΓ και ΒΓ είναι επικαλύψεις ανά δύο των αντίστοιχων χρώσεων και ΑΒΓ είναι επικάλυψη και των τριών χρώσεων. Το Neu1 συνεντοπίζεται με την Dlc-1 στο μεγαλύτερο μέρος του κυτταροπλάσματος (ιώδες, ΑΓ) και με την β-tubulin (γαλάζιο, ΑΒ). Και οι τρεις πρωτεΐνες εμφανίζουν έντονη χρώση στην περιφέρεια των πυρήνων (βέλη) και συνεντοπίζονται στο κεντρόσωμα (ΚΣ) των κυττάρων (λευκό, ΑΒΓ).

ενώ στην τελόφαση το Neu1 εντοπίζεται στην περιφέρεια των νεοσχηματιζόμενων πυρήνων, παρουσιάζοντας επικάλυψη με την Dlc-1 (Εικ.58). Τέλος, παρατηρήσαμε σημαντική επικάλυψη των πρωτεϊνών Neu1 και Dlc-1 στο αξόνημα σπερματοζωαρίων ποντικού (Εικ.63), γεγονός που υποδηλώνει αλληλεπίδραση του Neu1, εκτός της κυτταροπλασματικής, και με τη αξονημική δυνείνη και ενδεχομένως συσχετίζεται με την ανώμαλη δομή αξονήματος των σπερματοζωαρίων των Neu1 $-/-$ ποντικών [Vollrath *et al*, 2001, βλ. Εισαγωγή].

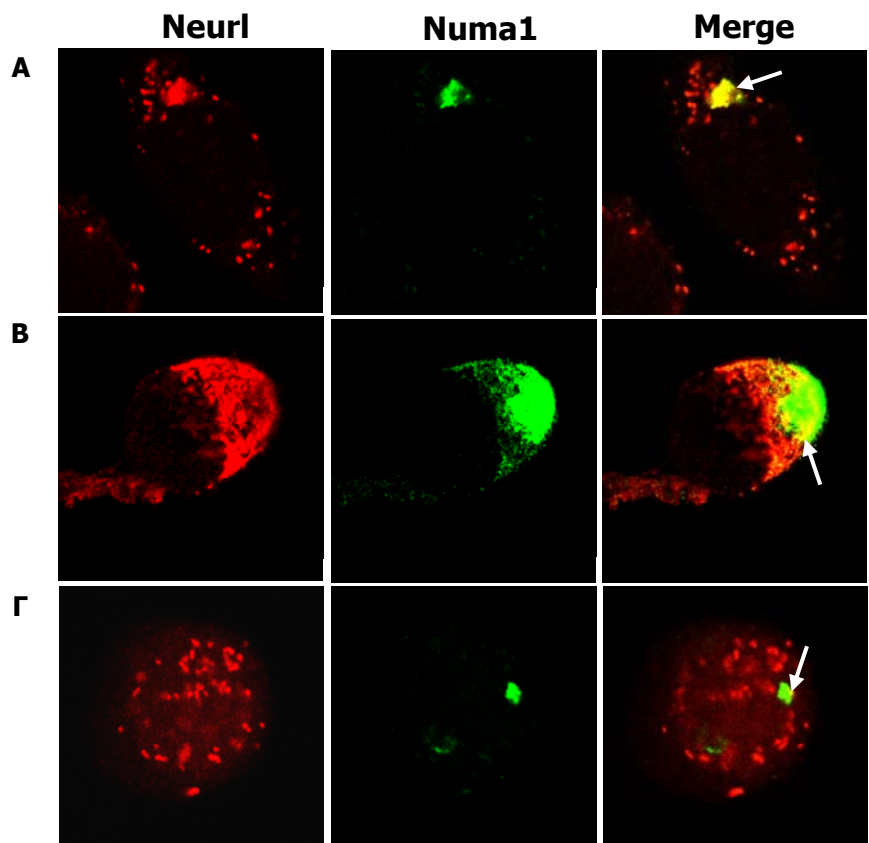
ε. Συνεντοπισμός των πρωτεϊνών Neu1 και Numa1

Για τον υποκυτταρικό εντοπισμό της πρωτεΐνης Numa1 (*Nuclear mitotic protein*) των θηλαστικών είναι γνωστό ότι εξαρτάται από τη φάση της κυτταρικής διαίρεσης [Compton & Cleveland, 1993/ Merdes *et al*, 1996,2000/ Blagden & Glover, 2003]. Επιπλέον, παρά τον σημαντικό της ρόλο στην κυτταρική διαίρεση, η Numa1 απουσιάζει από αρκετούς κυτταρικούς τύπους των θηλαστικών και από κυτταρικές σειρές. Ειδικά στις κυτταρικές σειρές, έχει αναφερθεί απώλεια της Numa1 σε μεγάλο ποσοστό των κυττάρων μετά από συνεχιζόμενη καλλιέργεια [Taimen *et al*, 2000]. Για αυτό το λόγο ήταν δυνατή η ανίχνευση της ενδογενούς Numa1 σε μικρό ποσοστό κυττάρων HeLa, ενώ η πρωτεΐνη δεν ήταν δυνατόν να ανιχνευθεί σε κύτταρα COS7. Συγκεκριμένα, ισχυρό σήμα του μονοκλωνικού αντισώματος α -Numa1 παρατηρήθηκε στο 2% των κυττάρων που παρατηρήθηκαν. Παρατηρώντας ταυτόχρονα τη χρώση του DNA (DAPI) διαπιστώθηκε ότι τα κύτταρα που έφεραν ανιχνεύσιμη ποσότητα Numa1 ήταν κατά κανόνα σε μίτωση, αντιπροσωπεύοντας το 70% των συνολικών μιτώσεων. Στο 50% των θετικών για Numa1 μιτωτικών κυττάρων αυτή εντοπιζόταν σε δύο διακριτά σημεία, συμμετρικά και κάθετα ως προς το ισημερινό επίπεδο, σε συμφωνία με τον γνωστό εντοπισμό της στους πόλους της μιτωτικής ατράκτου. Το 43% των θετικών για Numa1 μιτωτικών κυττάρων βρίσκονταν σε πρόφαση, με εμφανώς συμπυκνωμένη χρωματίνη. Αυτά έφεραν ισχυρό σήμα της Numa1 σε ένα σημείο, πλευρικό του πυρήνα, σε συμφωνία με προηγούμενες παρατηρήσεις εντοπισμού της Numa1 στο κεντρόσωμα. Τέλος, μικρότερο ποσοστό μιτωτικών κυττάρων (7% των θετικών για Numa1) παρατηρήθηκαν κατά την τελόφαση. Σε αυτά ανιχνεύθηκε η Numa1 στην περιφέρεια των σχηματιζόμενων πυρήνων. Σε μικρό ποσοστό μεσοφασικών κυττάρων παρατηρήθηκε πολύ ασθενής και διάχυτη χρώση στον πυρήνα, υποδηλώνοντας την παρουσία μικρής ποσότητας Numa1 και σε αυτή τη φάση. Η ενίσχυση του σήματος κατά τη μίτωση πιθανόν οφείλεται στη συγκέντρωση της πρωτεΐνης στο κεντρόσωμα ή στους πόλους της μιτωτικής ατράκτου ή σχετίζεται με την προσβασιμότητα του συγκεκριμένου επιτόπου σε διαφορετικές φάσεις του κυτταρικού κύκλου. Ανιχνεύοντας ταυτόχρονα και την πρωτεΐνη Neu1, διαπιστώσαμε ότι ένα σημαντικό μέρος της συνεντοπίζεται με τη Numa1 στο κεντρόσωμα κατά την πρόφαση (Εικ.64A), επιβεβαιώνοντας και τον συνεντοπισμό του με τις Dlc-1 και β -tubulin στην ίδια φάση. Επιπλέον, συνεντοπισμός Neu1 και Numa1 παρατηρήθηκε και στην περιφέρεια των νέων πυρήνων κατά την τελόφαση (Εικ.64B). Αντίθετα, κατά τη μετάφαση ο συνεντοπισμός Neu1 και Numa1 ήταν πιο περιορισμένος και εντοπίστηκε στους πόλους του 60% των μεταφάσεων στις οποίες ανιχνεύθηκε η Numa1 (Εικ.64Γ).



Εικ.63 Συνεντοπισμός των πρωτεϊνών Neur1 και Dlc-1 στο αξόνημα. Οι πρωτεΐνες Neur1 (κόκκινο) και Dlc-1 (πράσινο) ανοσοανιχνεύθηκαν σε σπερματοζώαρια ποντικού. Το Neur1 εντοπίζεται κατά μήκος του σπερματοζωαρίου, με εντονότερη χρώση στο κεντρικό τμήμα του μαστιγίου, ενώ η Dlc-1 εμφανίζει πολύ πιο έντονη χρώση στο κεντρικό τμήμα του μαστιγίου σε σχέση με το υπόλοιπο σπερματοζώαριο. Τα σπερματοζώαρια παρατηρήθηκαν σε οπτικό μικροσκόπιο φθορισμού. Ο συνεντοπισμός των δύο πρωτεϊνών στην περιοχή του αξονήματος φαίνεται με κίτρινο (merge).

Εικ.64 Συνεντοπισμός Neur1 και Numa1 σε κύτταρα θηλαστικών. (Α-Γ): Οπτικές τομές κυττάρων COS7, μετά από ταυτόχρονη ανοσοανίχνευση των ενδογενών πρωτεϊνών Neur1 (κόκκινο) και Numa1 (πράσινο). Στην επικάλυψη των δύο τομών (merge) φαίνεται σημαντικός συνεντοπισμός (κίτρινο), στο κεντρώσωμα (Α) και κατά τον σχηματισμό πυρήνων στην τελόφαση (Β). Πιο περιορισμένος είναι ο συνεντοπισμός στους πόλους της μιτωτικής ατράκτου (Γ).



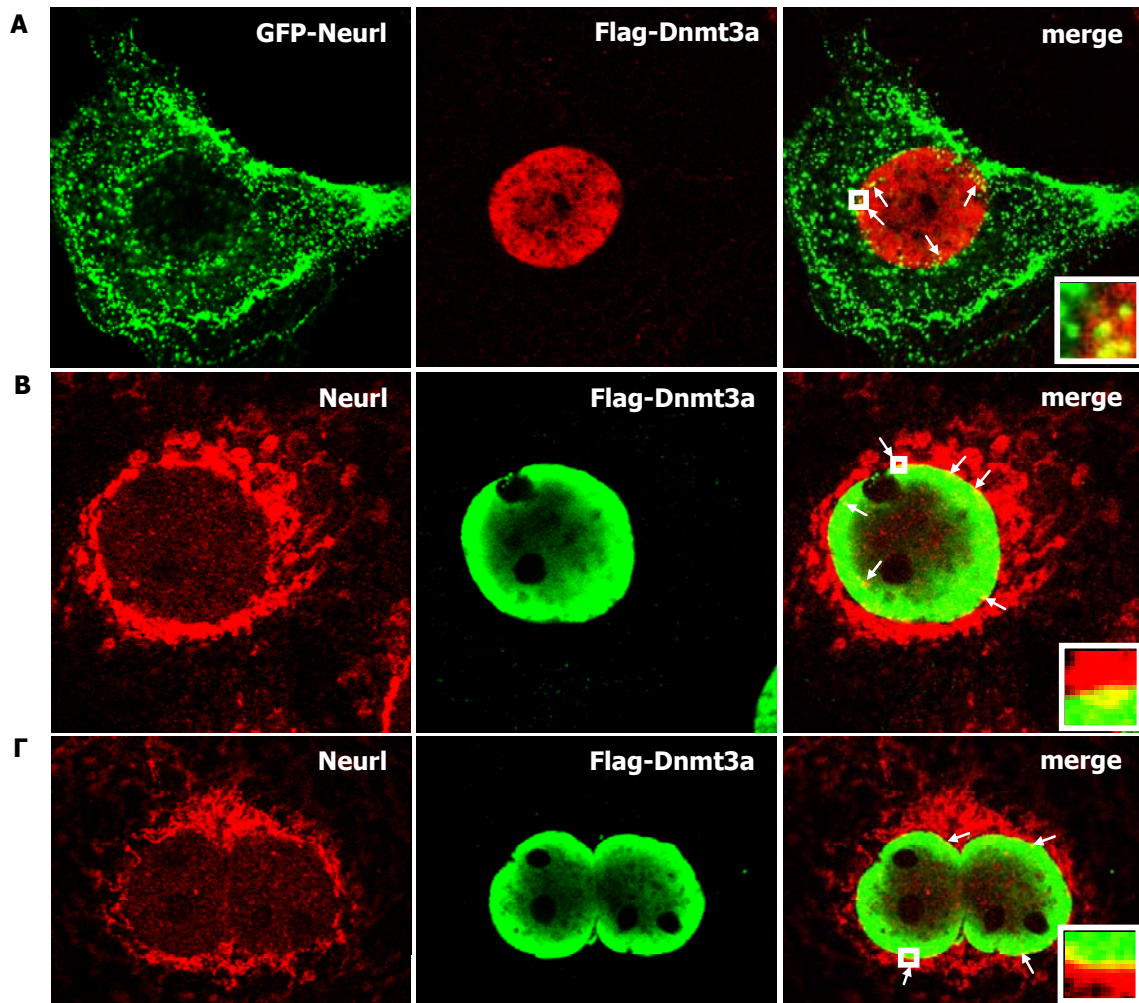
στ. Συνεντοπισμός των πρωτεϊνών *Neur1* και *Dnmt3a*

Η πρωτεΐνη *Dnmt3a* πρόσφατα βρέθηκε να αλληλεπιδρά με την *Ubc9* και να τροποποιείται από SUMO-1 [Ling *et al*, 2004], γεγονός που την καθιστά ενδιαφέρον 'θήραμα' του *Neur1*, δικαιολογώντας το μοντέλο σύμφωνα με το οποίο το *Neur1* λειτουργεί ως 'adaptor' διευκολύνοντας τροποποιήσεις πρωτεϊνικών υποστρωμάτων, από SUMO ή ουβικουϊτίνη (βλ. Συζήτηση). Ο υποκυτταρικός εντοπισμός της πρωτεΐνης *Dnmt3a* μελετήθηκε ανιχνεύοντας την πλήρους μήκους ανασυνδυασμένη πρωτεΐνη Flag-*Dnmt3a* σε κύτταρα που είχαν διαμολυνθεί παροδικά με την πλασμιδιακή κατασκευή *Flag-Dnmt3a(1-908)* (προσφορά από τον K. D. Robertson, Univ. of Florida). Η Flag-*Dnmt3a* βρέθηκε κυρίως στο εσωτερικό του πυρήνα και σε μικρότερο ποσοστό στην περιπυρηνική περιοχή, σύμφωνα και με τις παρατηρήσεις των Ling *et al* (2004). Ο συνεντοπισμός της Flag-*Dnmt3a* με την ενδογενή πρωτεΐνη *Neur1* ή την ανασυνδυασμένη πρωτεΐνη GFP-*Neur1(5-574)* διαπιστώθηκε κυρίως στην περιφέρεια του πυρήνα (Εικ.65). Επιπλέον, διαπιστώθηκε συνεντοπισμός της Flag-*Dnmt3a* με την ανασυνδυασμένη πρωτεΐνη GFP-SUMO-1 στο εσωτερικό του πυρήνα, γεγονός που πρέπει να σχετίζεται με την τροποποίηση της *Dnmt3a* από SUMO-1 (Εικ.66).

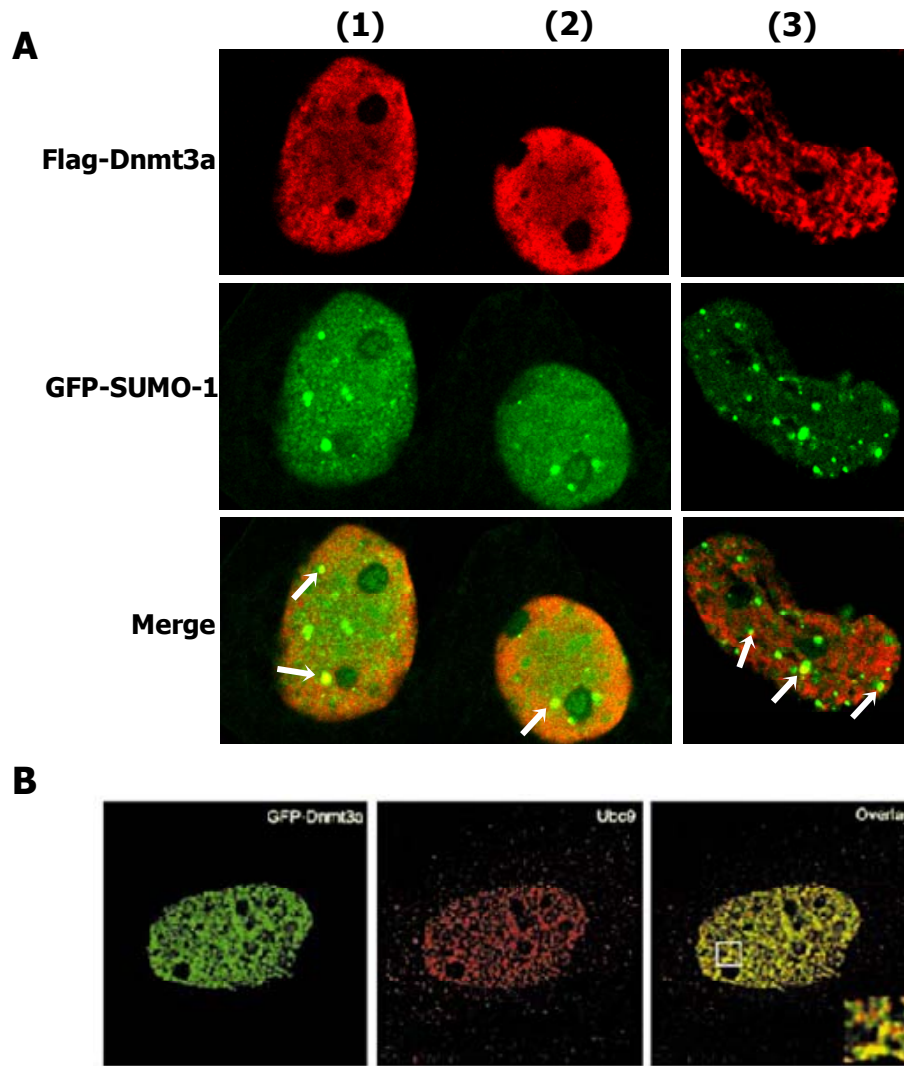
ζ. Συνεντοπισμός *Neur1* και SUMO-1

Η πρωτεΐνη *Ubc9*, η οποία συγκαταλέγεται μεταξύ των 'θήραμάτων' του *Neur1* που ταυτοποιήθηκαν σε αυτή την εργασία, είναι το μοναδικό γνωστό συζευκτικό ένζυμο (E2) της διαδικασίας ομοιοπολικής πρόσδεσης μορίων SUMO (*Small Ubiquitin-like Modifier*) σε πρωτεϊνικά υποστρώματα (βλ. Συζήτηση). Κατά κανόνα, οι πρωτεΐνες που αλληλεπιδρούν με την *Ubc9* είναι είτε ένζυμα που καταλύουν ή διευκολύνουν τη διαδικασία της SUMOλίωσης, είτε υποστρώματα που SUMOλιώνονται [Johnson, 2004]. Επιπλέον, από τα τρία μόρια SUMO (SUMO-1, SUMO-2 και SUMO-3) που συναντώνται στα κύτταρα των σπονδυλωτών, το SUMO-1 (11KD) είναι αυτό που κυρίως εμφανίζεται συζευγμένο με υποστρώματα (βλ. Συζήτηση). Για τους παραπάνω λόγους, διερευνήσαμε τον συνεντοπισμό της πρωτεΐνης *Neur1* με το SUMO-1 των κυττάρων COS7. Από την ανοσοανίχνευση των ενδογενών μορίων *Neur1* και SUMO-1, εντοπίστηκαν διακριτά σημεία συνεντοπισμού στο εσωτερικό και στην περιφέρεια του πυρήνα, τα οποία ήταν περισσότερα στα κύτταρα με μεγαλύτερο ποσοστό πυρηνικού *Neur1* (Εικ.67). Ο συνεντοπισμός των πρωτεϊνών SUMO-1 και *Neur1* επιβεβαιώθηκε και με εγκάρσιες οπτικές τομές των κυττάρων (Εικ. 67B'). Τέλος, σημειώνεται ότι επειδή το ποσοστό του ελεύθερου SUMO-1 είναι εξαιρετικά χαμηλό [Su & Li, 2002], η χρώση του αντισώματος α-SUMO-1 που παρατηρούμε, προέρχεται κατά κύριο λόγο από πρωτεϊνικά υποστρώματα που έχουν SUMOλιωθεί. Κατά συνέπεια, το *Neur1* συνεντοπίζεται με ορισμένα από αυτά τα υποστρώματα.

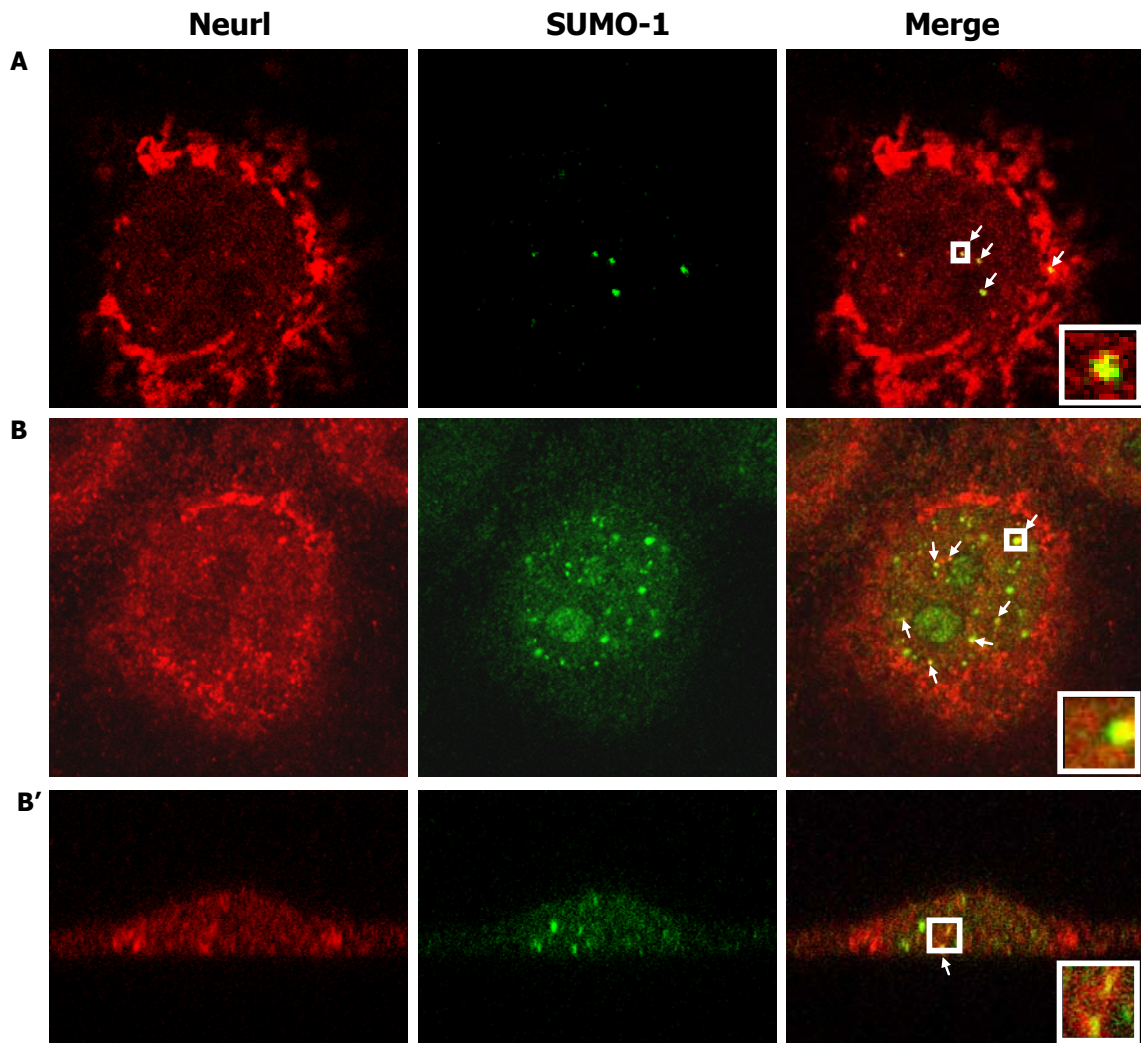
Τα πειράματα ανοσοανίχνευσης που περιγράψαμε, αποτελούν σημαντικό προκαταρκτικό βήμα στην πορεία κατανόησης της βιολογικής λειτουργίας της πρωτεΐνης *Neur1*. Χρησιμοποιώντας ως μοριακούς δείκτες τις πρωτεΐνες Lamin A/C και β-tubulin, οριοθετήσαμε υποκυτταρικά τις θέσεις εντοπισμού του *Neur1* στα κύτταρα θηλαστικών (διαιρούμενα και μη). Σημαντικό εύρημα είναι ο πυρηνικός εντοπισμός της πρωτεΐνης *Neur1*, που παρατηρήθηκε σε μικρό ποσοστό των κυττάρων και, σε συνδυασμό με επιπρόσθετες ενδείξεις άλλων πειραμάτων [Timmusk *et al*, 2002 και E. Κουτέλου, αδημοσίευτα αποτελέσματα], ενδεχομένως υποδηλώνει πυρηνο-κυτταροπλασματική διακίνηση του *Neur1*. Ο εντοπισμός ενός ποσοστού των μο-



Εικ.65 Συνεντοπισμός των πρωτεϊνών Neur1 και Dnmt3a στα κύτταρα θηλαστικών. (Α): Οπτική τομή κυττάρου COS7, ύστερα από ταυτόχρονη παροδική διαμόλυνση με τις πλασμιδιακές κατασκευές που κωδικοποιούν τις ανασυνδυασμένες πρωτεΐνες GFP-Neur1 και Flag-Dnmt3a. Παρατηρείται συνεντοπισμός σε διακριτά σημεία της περιπυρηνικής περιοχής και στο εσωτερικό του πυρήνα (merge, βέλη). (Β,Γ): Οπτικές τομές κυττάρων COS7, στα οποία μετά την παροδική διαμόλυνση με το πλασμίδιο που κωδικοποιεί την Flag-Dnmt3a, πραγματοποιήθηκε ταυτόχρονη ανοσοανίχνευση της Flag-Dnmt3a (πράσινο) και της ενδογενούς πρωτεΐνης Neur1 (κόκκινο). Και το ενδογενές Neur1 εμφανίζει σημεία επικάλυψης (κίτρινο) με την Flag-Dnmt3a στην περιφέρεια του πυρήνα. Σε πλαίσια φαίνονται ενδεικτικά οι αντίστοιχες περιοχές σε μεγέθυνση.



Εικ.66 Συνεντοπισμός Flag-Dnmt3a και GFP-SUMO-1 (A): Κύτταρα COS7 διαμολύνθηκαν παροδικά με τις πλασμιδιακές κατασκευές που κωδικοποιούν τις πρωτεΐνες Flag-Dnmt3a και GFP-SUMO-1 (2μg και 1μg, αντίστοιχα). Ακολούθησε ταυτόχρονη ανίχνευση της χμιαϊρικής πρωτεΐνης GFP-SUMO-1 (πράσινο) και ανοσοανίχνευση της Flag-Dnmt3a (κόκκινο). Το GFP-SUMO-1 εντοπίζεται στον πυρήνα σε σπικτικές δομές, εμφανίζοντας περισσότερο (κύτταρα 1 και 2) ή λιγότερο (κύτταρο 3) διάχυτο πρότυπο. Ο συνεντοπισμός του με την Flag-Dnmt3a (κίτρινο) φαίνεται με βέλη. **(B):** Συνεντοπισμός GFP-Dnmt3a και Ubc9 στο εσωτερικό του πυρήνα από τους Ling *et al* (2004).



Εικ.67 Συνεντοπισμός του Neur1 με SUMO-1 σε κύτταρα θηλαστικών. Οπτικές τομές κυττάρων COS7, ύστερα από ταυτόχρονη ανοσοανίχνευση των ενδογενών πρωτεϊνών Neur1 (κόκκινο) και SUMO-1 (πράσινο). Η χρώση του SUMO-1 εντοπίζεται κυρίως σε διακριτές στικτές δομές του πυρήνα. Β' είναι εγκάρσια τομή του κυττάρου Β. Σημεία συνεντοπισμού Neur1 και SUMO-1 (κίτρινο) που υποδεικνύονται με βέλη παρατηρούνται στα δύο κάθετα μεταξύ τους οπτικά επίπεδα. Σε πλαίσιο φαίνονται ενδεικτικές περιοχές σε μεγέθυνση.

ρίων Neu1 στο εσωτερικό ή στην περιφέρεια του πυρήνα, είναι δυνατόν να δικαιολογήσει τις αλληλεπιδράσεις με τα μόρια Ubc9, Numa1, Lek1 και Dnmt3a, ενώ είναι πιθανές και οι αλληλεπιδράσεις του Neu1 με τις πρωτεΐνες Numa1 και Dlc-1 κατά την κυτταρική διαίρεση. Επιπλέον, συνεντοπισμός του Neu1 με την υπομονάδα Dlc-1 της δυνεΐνης παρατηρήθηκε σε όλες τις φάσεις του κυτταρικού κύκλου που εξετάστηκαν και ενδεχομένως να είναι σημαντικός για τη μετακίνηση της πρωτεΐνης Neu1 πάνω στο δίκτυο των μικροσωληνίσκων. Τέλος, ο συνεντοπισμός με SUMO-1, ενισχύει την υπόθεση ότι η πρωτεΐνη Neu1 εμπλέκεται σε διαδικασίες SUMOλίωσης. Τα παραπάνω αποτελούν σημαντικές ενδείξεις αλληλεπιδράσης του Neu1 με τα συγκεκριμένα 'θηράματα', οι οποίες μας οδήγησαν στον σχεδιασμό των βιοχημικών πειραμάτων που ακολούθησαν.

6. Έλεγχος και χαρτογράφηση αλληλεπιδράσεων του Neu1 σε κύτταρα θηλαστικών

Μετά την επιβεβαίωση της άμεσης αλληλεπιδράσης του Neu1 με τις πρωτεΐνες-'θηράματα' Ubc9, Dlc-1, Numa1 και Lek1 σε *in vitro* συνθήκες (πειράματα 'GST-pulldown') και τις παρατηρήσεις συνεντοπισμού του Neu1 με τις Ubc9, Dlc-1, Dnmt3a και Numa1 δοκιμάστηκε η βιοχημική απομόνωση των συμπλόκων του πλήρους μήκους Neu1 ή περιοχών του με κάθε μια από αυτές τις πρωτεΐνες, από πρωτεϊνικά εκχυλίσματα κυττάρων θηλαστικών. Συγκεκριμένα, σε παροδικά διαμολυσμένα κύτταρα COS7 εκφράστηκαν οι πρωτεΐνες-'θηράματα' (πλήρες μήκος ή περιοχές αυτών, Εικ.68) σημασμένες στο αμινοτελικό τους άκρο με τον επίτοπο Xpress, μαζί με χιμαιρικά (πλήρη ή ελλειμματικά) μόρια Neu1 (Εικ.72a), σε σύντηξη με την GFP (*Green Fluorescence Protein*). Ακολούθησε ανοσοκατακρήμνιση με αντίσωμα α-Xpress και ανοσοανίχνευση των συμπλόκων με αντίσωμα α-GFP. Και για τις τέσσερις ανασυνδυασμένες πρωτεΐνες (Xpress-Ubc9, -Dlc-1, -Lek1, -Numa1) κατακρημνίστηκαν πρωτεϊνικά σύμπλοκα με την σχεδόν πλήρους μήκους πρωτεΐνη GFP-Neu1(5-574) και την GFP-Neu1(5-537) (Εικ.69,73). Η χιμαιρική πρωτεΐνη GFP-Neu1(5-269), που περιέχει το *domain* NEUZ1, ανοσοκατακρημνίστηκε με τις πρωτεΐνες Xpress-Numa1, Xpress-Lek1 και Xpress-Dlc-1, αλλά όχι με την Xpress-Ubc9 (Εικ.75). Τέλος, η GFP-Neu1(227-574), που περιέχει το καρβοξυτελικό τμήμα (*domains* NEUZ2 και RING) του Neu1, αλληλεπιδρά με τις Xpress-Ubc9, Xpress-Numa1 και Xpress-Lek1, αλλά όχι με την Xpress-Dlc-1 (Εικ.74). Επιπλέον, για την πρωτεΐνη Ubc9, πραγματοποιήθηκε ανοσοκατακρήμνιση και επιβεβαιώθηκε η παρουσία του ενδογενούς πρωτεϊνικού συμπλόκου Neu1-Ubc9, σε εκχυλίσματα κυττάρων COS-7 (Εικ.71). Ειδικά για την πρωτεΐνη Dnmt3a, για την οποία πρόσφατα βρέθηκε από τους Ling *et al* (2004) ότι αλληλεπιδρά με την Ubc9 και τροποποιείται με SUMO-1, ανοσοκατακρημνίστηκαν πρωτεϊνικά σύμπλοκα που περιείχαν την ανασυνδυασμένη πρωτεΐνη πλήρους μήκους Flag-Dnmt3a και την GFP-Neu1(5-574), ύστερα από παροδική διαμόλυνση κυττάρων COS7 με τις αντίστοιχες πλασμιδιακές κατασκευές (Εικ.70).

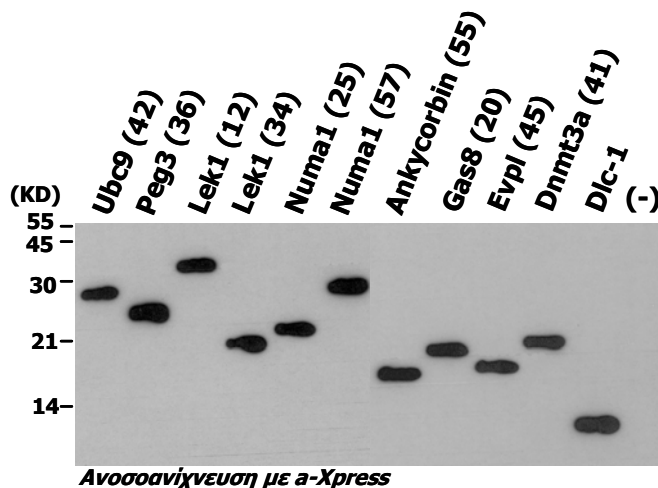
Για τα θηράματα Peg3(#36), Ankycorbin(#55), Gas8(#20) και Evp1(#45), που αρχικά είχαν εξαιρεθεί από τη μελέτη λόγω ελλειπών στοιχείων για τη λειτουργία τους, αλλά στη συνέχεια προέκυψαν στοιχεία που τα συσχετίζουν με τα παραπάνω (βλ. Συζήτηση), πραγματοποιήθηκαν ανοσοκατακρημνίσεις με

a-Chpress, μετά από συνδιαμόλυνση κυττάρων COS7, όπως περιγράφεται παραπάνω. Και τα τέσσερα αυτά θηράματα ανοσοκατακρημνίστηκαν με τις χιμαιρικές πρωτεΐνες GFP-Neur1(5-574), GFP-Neur1(5-537) και GFP-Neur1(5-269), ενώ η πρωτεΐνη Chpress-Gas8 (‘θήραμα’ #20) ήταν η μόνη από τις πέντε που δεν κατακρημνίστηκε με την GFP-Neur1(227-574) (Εικ.74).

α. Συνέκφραση της πρωτεΐνης Neur1 και των πρωτεϊνών–‘θήραμάτων’ σε κύτταρα COS7

Για τη μελέτη των αλληλεπιδράσεων του Neur1 *in vivo*, ήταν αναγκαία η συνέκφραση της ανασυνδυασμένης πρωτεΐνης Neur1 με τις πρωτεΐνες–‘θήραματα’ σε κύτταρα θηλαστικών. Για την έκφραση του Neur1 αρχικά δοκιμάστηκε η παροδική διαμόλυνση κυττάρων COS7 με την πλασμιδιακή κατασκευή *Neur1/pCDNA3.1mycHisA*, η οποία κωδικοποιεί την σχεδόν πλήρους μήκους πρωτεΐνη Neur1 (αμινοξέα 5-574), σημασμένη με τον επίτοπο myc στο καρβοξυτελικό της άκρο (βλ. ‘Πλασμιδιακές Κατασκευές’). Η ανασυνδυασμένη όμως πρωτεΐνη Neur1(5-574)-myc ανιχνεύονταν σε εξαιρετικά χαμηλά επίπεδα στα κυτταρικά εκχυλίσματα και δεν ήταν αρκετή για την πραγματοποίηση πειραμάτων ανοσοκατακρήμνισης. Όπως φάνηκε στη συνέχεια με πειράματα διαμόλυνσης παρουσία του αναστολέα πρωτεασώματος MG132, τα χαμηλά επίπεδα της πρωτεΐνης Neur1(5-574)-myc οφείλονται σε αστάθεια εξαρτώμενη από το πρωτεάσωμα (Ε. Κουτέλου, αδημοσίευτες παρατηρήσεις). Αντίθετα, οι συντηξείς της πρωτεΐνης Neur1 με την πράσινη φθορίζουσα πρωτεΐνη (GFP) δεν εμφάνιζαν παρόμοια αστάθεια, πιθανόν λόγω του μεγαλύτερου μεγέθους (29KD) αυτού του επιτόπου που ενδεχομένως εμποδίζει συγκεκριμένες επαφές του ανασυνδυασμένου μορίου με τον μηχανισμό αποικοδόμησης από το πρωτεάσωμα. Για τον λόγο αυτό, στα πειράματα έκφρασης ανασυνδυασμένου Neur1 που περιγράφονται παρακάτω, έχει πραγματοποιηθεί παροδική διαμόλυνση κυττάρων COS7 με τις πλασμιδιακές κατασκευές που κωδικοποιούν την πλήρη πρωτεΐνη Neur1 ή τμήματα αυτής, σε σύντηξη με την GFP (Εικ.72α,76).

Για να εκφραστούν οι πρωτεΐνες–‘θήραματα’ του Neur1 σε κύτταρα θηλαστικών, τα cDNAs που τις κωδικοποιούν αποκόπηκαν από τον πλασμιδιακό φορέα pVP16 με BamHI-EcoRI ή NotI και κλωνοποιήθηκαν στις ίδιες θέσεις των φορέων pCDNA3.1HisA ή pCDNA3.1HisC, αντίστοιχα (βλ. ‘Πλασμιδιακές Κατασκευές’). Οι ανασυνδυασμένες πρωτεΐνες που εκφράζονται από αυτά τα πλασμίδια φέρουν τον επίτοπο Chpress (34 ή 48 αμινοξέα) στο αμινοτελικό τους άκρο.

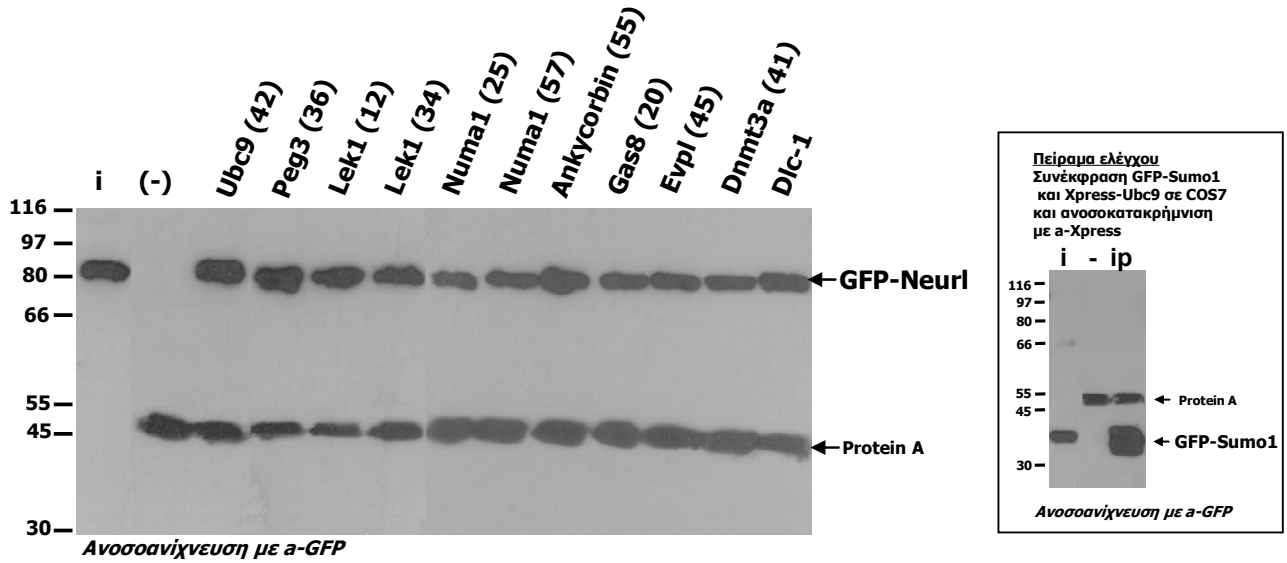


Εικ.68 Έκφραση των πρωτεϊνών-‘θηραμάτων’ του Neur1 σε κύτταρα COS7, μετά από παροδική διαμόλυνση με τις αντίστοιχες πλασμιδιακές κατασκευές. Η διαμόλυνση έγινε σε τρυβλία 10cm, που καλύπτονταν από κύτταρα κατά ~80%, με 6μg DNA από κάθε πλασμίδιο, ακολουθώντας τη μέθοδο του ‘φωσφορικού ασβεστίου’. Τα κύτταρα λύθηκαν 48hrs μετά τη διαμόλυνση, για 1hr στους 4°C, σε 200μl διαλύματος λύσης (50mM Tris-Cl pH 7.5, 250mM NaCl, 1% Triton-X, 1mM EDTA, 1mM PMSF). Το 10% του πρωτεϊνικού εκχυλίσματος ηλεκτροφορήθηκε σε πήκτωμα πολυακρυλαμιδης-SDS (15%) και ακολούθησε ανάλυση *Western* με αντίσωμα α-Xpress (βλ. ‘Υλικά και Μέθοδοι’). Το δείγμα αρνητικού ελέγχου του αντισώματος (-) είναι 50μg πρωτεϊνικού εκχυλίσματος κυττάρων COS7 αγρίου τύπου. Στις παρενθέσεις φαίνονται οι αριθμοί των αντίστοιχων κλώνων cDNA, με βάση την αριθμηση του Πίνακα 21. Η πρωτεΐνη Dlc-1 εκφράστηκε από τον πλήρους μήκους κλώνο cDNA *Dlc-1/pCDNA3.1HisA*.

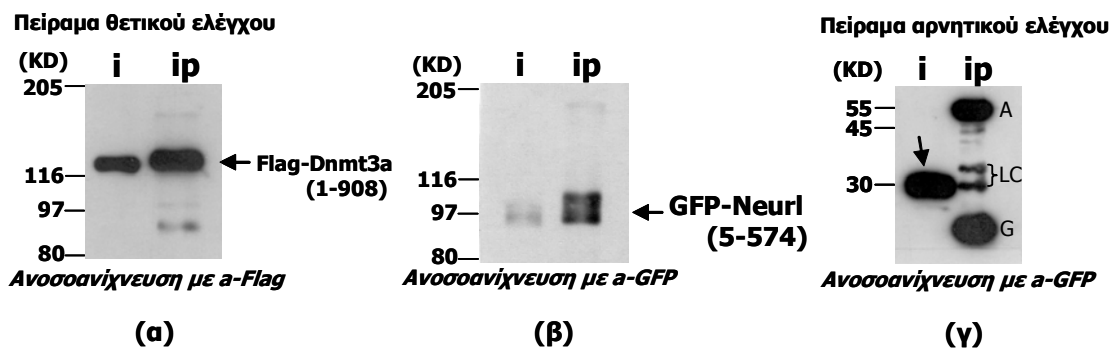
β. Οι πρωτεΐνες-‘θηράματα’ αλληλεπιδρούν με το πλήρους μήκους Neur1

Μετά τη συνέκφραση των ανασυνδυασμένων πρωτεϊνών-‘θηραμάτων’ και της πρωτεΐνης GFP-Neur1 στα κύτταρα COS7, πραγματοποιήθηκε ανοσοκατακρήμνιση με το αντίσωμα α-Xpress, απομόνωση των πρωτεϊνικών συμπλόκων και ανοσοανίχνευση με α-GFP. Διαπιστώθηκε ότι όλες οι πρωτεΐνες-‘θηράματα’ είναι ικανές να αλληλεπιδρούν με το Neur1 στα κύτταρα θηλαστικών (Εικ.69). Αξίζει να σημειωθεί το γεγονός ότι στις συνθήκες του παραπάνω πειράματος, οι πρωτεΐνες-‘θηράματα’ εμφανίζουν υψηλή συγγένεια για την πρωτεΐνη Neur1. Αυτό φαίνεται περισσότερο στην περίπτωση της Dlc-1, της οποίας η αλληλεπίδραση με την Neur1 ήταν ασθενής *in vitro* (Εικ.53, πειράματα *GST-pulldown*). Η αυξημένη συγγένεια *in vivo* της Dlc-1 με την πρωτεΐνη Neur1 πιθανόν οφείλεται στο ότι για το συγκεκριμένο πείραμα χρησιμοποιήσαμε το πλήρους μήκους μόριο Dlc-1, το οποίο είναι ικανό να σχηματίζει ομοδιμερή, σταθεροποιώντας ενδεχομένως τις αλληλεπιδράσεις του με άλλα μόρια, όπως είναι η πρωτεΐνη Neur1 (βλ. Συζήτηση).

Η πρόσφατη διαπίστωση των Ling *et al* (2004) ότι η Dnmt3a αλληλεπιδρά με την πρωτεΐνη Ubc9 και τροποποιείται μεταμεταφραστικά με SUMO-1 καθιστά αυτή την πρωτεΐνη πολύ ενδιαφέρον ‘θήραμα’ του Neur1, εφόσον είναι πιθανόν το Neur1 να συμμετέχει σε αυτή τη διαδικασία, αλληλεπιδρώντας και με τις δύο πρωτεΐνες. Έτσι, δοκιμάσαμε επιπρόσθετα την αλληλεπίδραση του Neur1 με την πλήρους μήκους Dnmt3a (Εικ.70), συνδιαμολύνοντας κύτταρα COS7 με τα πλασμίδια *pEGFPC1-Neur1* και *Flag-Dnmt3a(1-908)* (προσφορά από τον K. D. Robertson, Univ. of Florida). Η εξακρίβωση της *in vivo* αλληλεπίδρασης των δύο πλήρους μήκους πρωτεϊνών (Neur1 και Dnmt3a) είναι σημαντική συμβολή στην πορεία διερεύνησης της λειτουργίας του Neur1.



Εικ.69. Ανοσοκατακρήμνιση του GFP-Neur1 από τις ανασυνδυασμένες πρωτεΐνες-‘θηράματα’ σε κύτταρα COS7, μετά από παροδική συνδιαμόλυνση με την πλασμιδιακή κατασκευή *pEGFPC1/Neur1(5-574)* και κάθε μία από τις πλασμιδιακές κατασκευές των ‘θηραμάτων’ (βλ. ‘Πλασμιδιακές Κατασκευές’). Η διαμόλυνση έγινε σε τρυβλία 10cm, που καλύπτονταν από κύτταρα κατά ~80%, με 18μg DNA *pEGFPC1/Neur1(5-574)* και 6μg από κάθε πλασμιδίο ‘θηράματος’, ακολουθώντας τη μέθοδο του ‘φωσφορικού αβσεπτιού’. Τα κύτταρα λύθηκαν 48hrs μετά τη διαμόλυνση (Εικ.68) και το 40% του δείγματος (~500μg ολικής πρωτεΐνης) επώαστηκε με 1μg αντισώματος α-GFP για 3hrs στους 4°C. Ίση ποσότητα δείγματος επώαστηκε στους 4°C, χωρίς αντίσωμα (πείραμα αρνητικού ελέγχου). Τα πρωτεϊνικά σύμπλοκα δεσμεύτηκαν σε 20μl σφαιρίδια αραρόζης Protein A/G (Santa Cruz, sc-2003) με επώαση στους 4°C για 1 ώρα. Ακολούθησε έκλυση από τα σφαιρίδια με βρασμό για 5min σε 1x διαλύματος *Laemli*, ηλεκτροφόρηση σε πήκτωμα πολυακρυλαμιδης-SDS (10%) και ανάλυση *Western* με αντίσωμα α-GFP. Το δείγμα ‘i’ (input) είναι το 10% του πρωτεϊνικού εκχυλίσματος πριν την ανοσοκατακρήμνιση. Το δείγμα ‘(-)’ είναι το πείραμα αρνητικού ελέγχου (κατακρήμνιση απουσία αντισώματος). Στις παρενθέσεις φαίνονται οι αριθμοί των αντίστοιχων κλώνων cDNA, με βάση την αρίθμηση του Πίνακα 21. Η πρωτεΐνη Dlc-1 εκφράστηκε από τον πλήρους μήκους κλώνο cDNA *Dlc-1/pCDNA3.1HisA*. Σε πλαίσιο φαίνεται το πείραμα θετικού ελέγχου, στο οποίο η πρωτεΐνη Xpress-Ubc9 κατακρημνίζει την GFP-Sumo1, στις ίδιες συνθήκες.

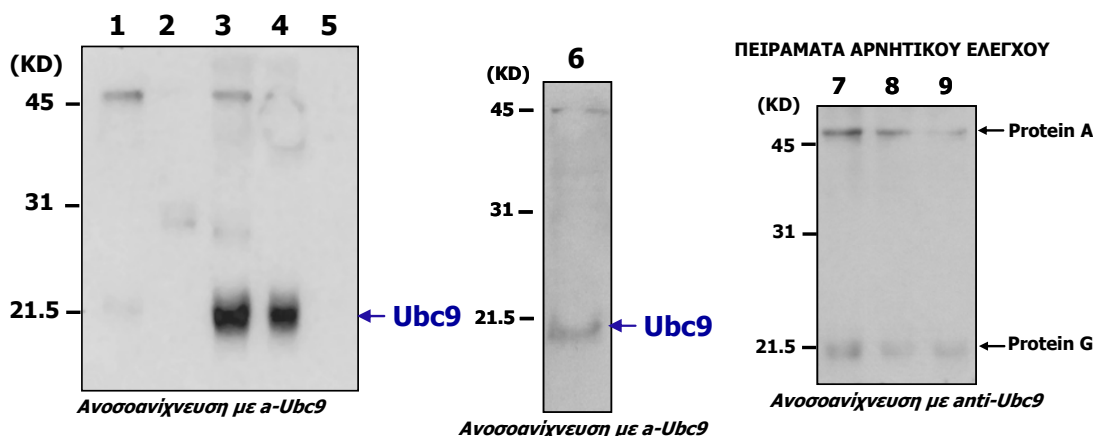


Εικ.70 Οι πρωτεΐνες *Neur1* και *Dnmt3a* αλληλεπιδρούν στα κύτταρα των θηλαστικών. Κύτταρα COS7 διαμολύνθηκαν παροδικά με τις πλασμιδιακές κατασκευές *Flag-Dnmt3a(1-908)* και *pEGFPC1-Neur1(5-574)* (12μg από κάθε πλασμιδίο σε ένα τρυβλίο 10cm). Ακολούθησε λύση των κυττάρων σε 400μl διαλύματος RIPA-1xPBS (1%NP40, 0.5% deoxycholate, 0.1% SDS, 10% γλυκερόλη, 1x PBS, pH 7.5). Για την ανοσοκατακρήμνιση χρησιμοποιήθηκαν περίπου 800μg πρωτεϊνικού εκχυλίσματος αραιωμένου σε 1ml, μετά την προσθήκη κατάλληλου όγκου διαλύματος 0.5% NP40-1xPBS. Τα κυτταρικά εκχυλίσματα επώαστηκαν με 1μg α-Flag και 20μl σφαιρίδια Protein A/G. Ακολούθησε ηλεκτροφόρηση (SDS-PAGE) σε πήκτωμα 8% και ανοσοανίχνευση με α-Flag (α) και α-GFP (β). Το δείγμα ‘i’ είναι το 10% του κυτταρικού εκχυλίσματος που χρησιμοποιήθηκε στην ανοσοκατακρήμνιση και το δείγμα ‘ip’ είναι το 50% του δείγματος της ανοσοκατακρήμνισης. Στο πείραμα αρνητικού ελέγχου (γ) αποκλείστηκε η πιθανότητα αλληλεπίδρασης της *Flag-Dnmt3a* με την *GFP*. Πραγματοποιήθηκε ανοσοκατακρήμνιση με α-Flag, όπως προηγουμένως, σε εκχύλισμα κυττάρων που είχαν διαμολυνθεί με τα πλασμιδία *Flag-Dnmt3a(1-908)* και *pEGFPC1*. Όπως φαίνεται στην εικόνα (γ), η *GFP* (δείγμα ‘i’, βέλος) δεν ανοσοκατακρημνίζεται από το α-Flag. Με ‘Α’ και ‘G’ συμβολίζονται οι ζώνες των Protein A και G, αντίστοιχα, ενώ οι ζώνες ‘LC’ είναι οι μικρού μοριακού βάρους αλυσίδες του α-Flag. Επομένως, η πρωτεΐνη *GFP-Neur1* ανιχνεύεται στο ίδιο πρωτεϊνικό σύμπλοκο με την *Flag-Dnmt3a*, εφόσον και οι δύο κατακρημνίζονται με το αντίσωμα α-Flag.

γ. Οι ενδογενείς πρωτεΐνες Neur1 και Ubc9 αλληλεπιδρούν στα κύτταρα θηλαστικών

Με σκοπό την ανίχνευση του πρωτεϊνικού συμπλόκου των ενδογενών πρωτεϊνών Neur1 και Ubc9, εφαρμόστηκε η μέθοδος της ανοσοκατακρήμνισης σε εκχυλίσματα κυττάρων COS7, χρησιμοποιώντας τα αντισώματα a-Neur1 και a-Ubc9 (Εικ.71). Στο ανοσοκατακρήμνισμα και των δύο αντισωμάτων ανιχνεύθηκε η ίδια ζώνη με a-Ubc9, η οποία αντιστοιχεί στην ενδογενή πρωτεΐνη Ubc9 (~18KD). Για να αποκλειστεί το ενδεχόμενο το σήμα που ανιχνεύθηκε στις ανοσοκατακρήμνισεις με a-Neur1 και a-Ubc9 να προκύπτει από την protein G που έχει παρόμοιο μοριακό βάρος (22KD), πραγματοποιήθηκαν στα ίδια κυτταρικά εκχυλίσματα παράλληλες ανοσοκατακρήμνισεις με τα αντισώματα a-Numa1, a-Delta1 και a-GST (πειράματα αρνητικού ελέγχου, Εικ.71). Από τα πειράματα αυτά διαπιστώνεται ότι στο πρωτεϊνικό σύμπλοκο που κατακρήμνίζεται από το αντίσωμα a-Neur1 περιέχεται η πρωτεΐνη Ubc9, γεγονός που αποτελεί σημαντική ένδειξη για την *in vivo* αλληλεπίδραση των δύο πρωτεϊνών.

ΑΛΛΗΛΕΠΙΔΡΑΣΗ Neur1 ΚΑΙ Ubc9 *IN VIVO*



Διαδρομές

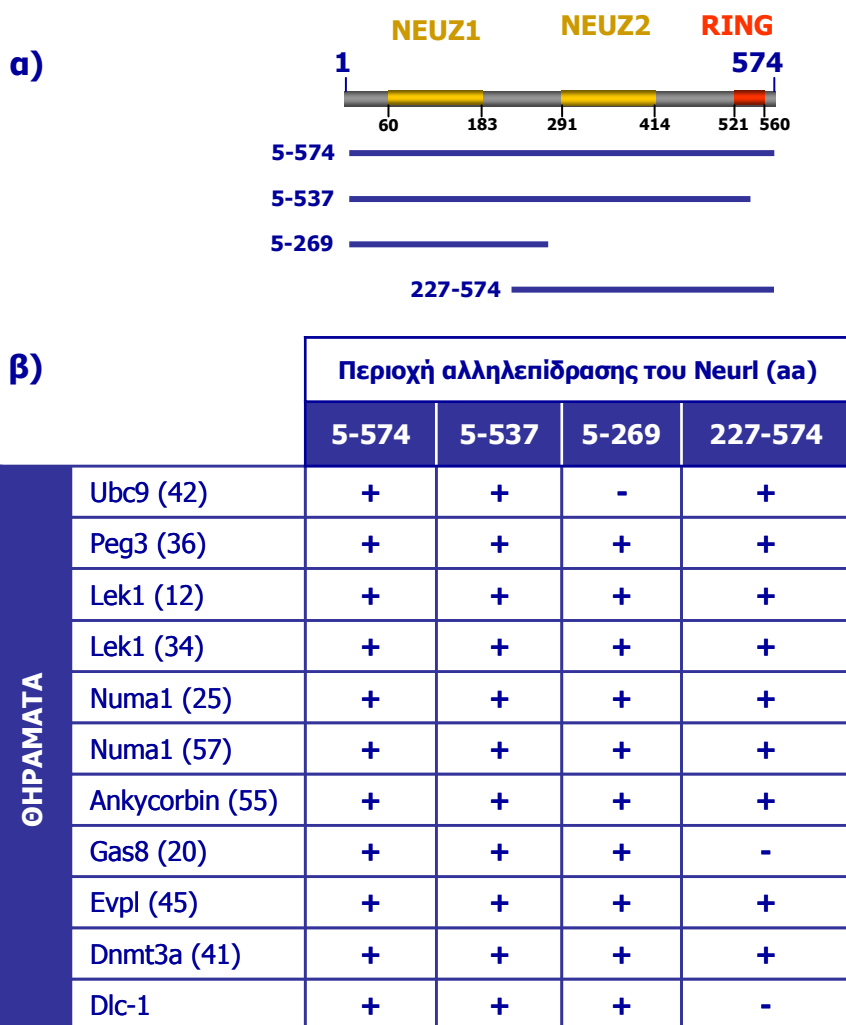
- 1) 8µl σφαιρίδια αραρόζης-Protein A/G agarose beads,
- 2) 250ng a-Ubc9,
- 3) δείγμα ανοσοκατακρήμνισης με a-Ubc9,
- 4) δείγμα ανοσοκατακρήμνισης με a-Neur1,
- 5) 2µl a-Neur1
- 6) 40 µg πρωτεϊνικού εκχυλίσματος κυττάρων COS7

- 7) δείγμα ανοσοκατακρήμνισης με a-Numa1
- 8) δείγμα ανοσοκατακρήμνισης με a-Delta1
- 9) δείγμα ανοσοκατακρήμνισης με a-GST

Εικ.71 Αλληλεπίδραση των ενδογενών μορίων Neur1 και Ubc9 στα κύτταρα θηλαστικών. Για κάθε ανοσοκατακρήμνιση χρησιμοποιήθηκαν 750µg πρωτεϊνικού εκχυλίσματος κυττάρων COS7 (διάλυμα λύσης: 150mM NaCl, 0.1% NP40, 50mM HEPES pH 7.0, 10% γλυκερόλη, 1mM PMSF), 1µg αντισώματος και 20µl σφαιρίδια αραρόζης συζευγμένα με Protein A (50KD) και Protein G (22KD). Η ανοσοκατακρήμνιση πραγματοποιήθηκε στο διάλυμα λύσης, σε όγκο 500µl, για 16 ώρες στους 4°C σε περιστρεφόμενο αναδευτήρα. Το 25% κάθε ανοσοκατακρήμνισης ηλεκτροφορήθηκε σε πήκτωμα πολυακρυλαμιδης-SDS (15%). Ακολούθησε ανάλυση *Western* με a-Ubc9. Η ενδογενής πρωτεΐνη Ubc9 (18KD) ανιχνεύθηκε μόνο στις ανοσοκατακρήμνισεις με τα αντισώματα a-Neur1 και a-Ubc9.

δ. Χαρτογράφηση των αλληλεπιδράσεων σε περιοχές της πρωτεΐνης Neurl

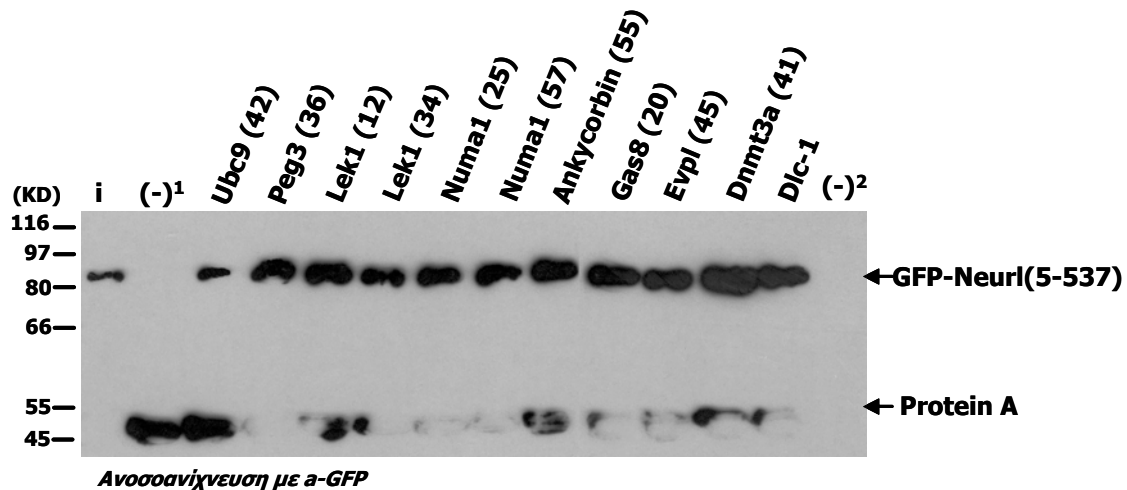
Για τη διερεύνηση των περιοχών της πρωτεΐνης Neurl που ευθύνονται για τις αλληλεπιδράσεις της με τα 'θηράματα', πραγματοποιήθηκαν πειράματα ανοσοκατακρήμνισης μετά από από παροδική συνέκφρασή τους σε κύτταρα COS7, με τις χιμαιρικές πρωτεΐνες GFP-Neurl(5-537), GFP-Neurl(227-574) και GFP-Neurl(5-269), (Εικ.72α), στις ίδιες συνθήκες με αυτές της Εικ.69. Τα αποτελέσματα παρουσιάζονται συνολικά στην Εικ.72.



Εικ.72 Χαρτογράφηση των περιοχών αλληλεπιδράσεως του Neurl με τις πρωτεΐνες-‘θηράματα’. **α)** Οι περιοχές της πρωτεΐνης Neurl που εκφράστηκαν σε σύντηξη με την GFP για τις δοκιμές της *in vivo* αλληλεπιδράσεως με τα ‘θηράματα’. **β)** Το αποτέλεσμα της ανοσοκατακρήμνισης κάθε περιοχής του Neurl με κάθε πρωτεΐνη-‘θήραμα’. Με ‘+’ συμβολίζεται η αλληλεπιδράση των δύο ανασυνδυασμένων μορίων στα κύτταρα COS7, ενώ με ‘-’ η απουσία αλληλεπιδράσεως.

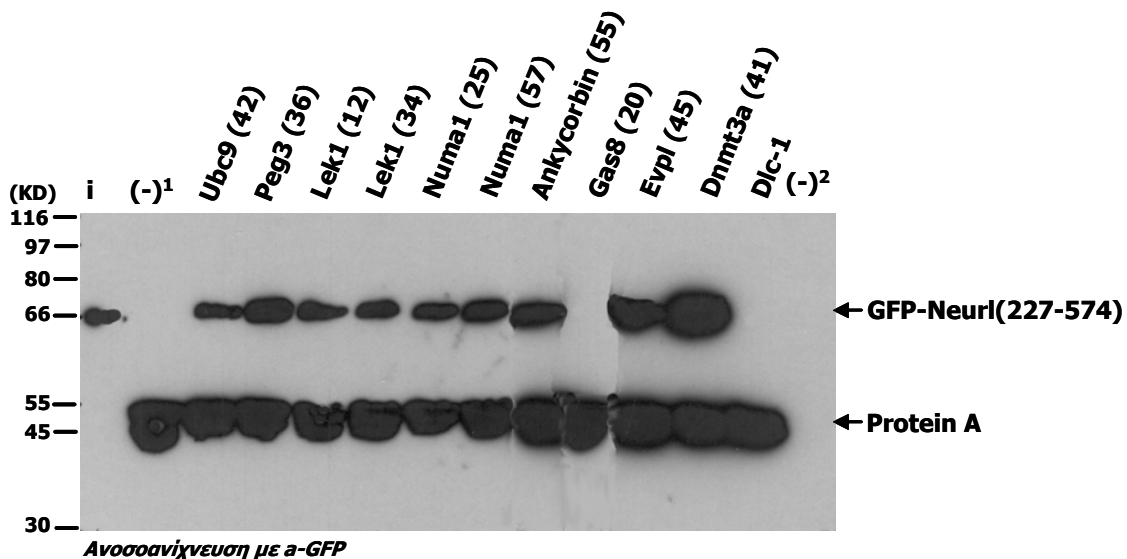
Συμπερασματικά, διαπιστώνουμε τα εξής:

- 1) Το C-τελικό μισό του RING δεν είναι απαραίτητο για τις αλληλεπιδράσεις του Neurl με τις πρωτεΐνες-‘θηράματα’. Αυτό προκύπτει από το γεγονός ότι το μόριο GFP-Neurl(5-537), στο οποίο έχει αποκοπεί το C-τελικό μισό του RING, αλληλεπιδρά με όλες τις πρωτεΐνες-‘θηράματα’, όπως και το πλήρους μήκους Neurl (Εικ.73).



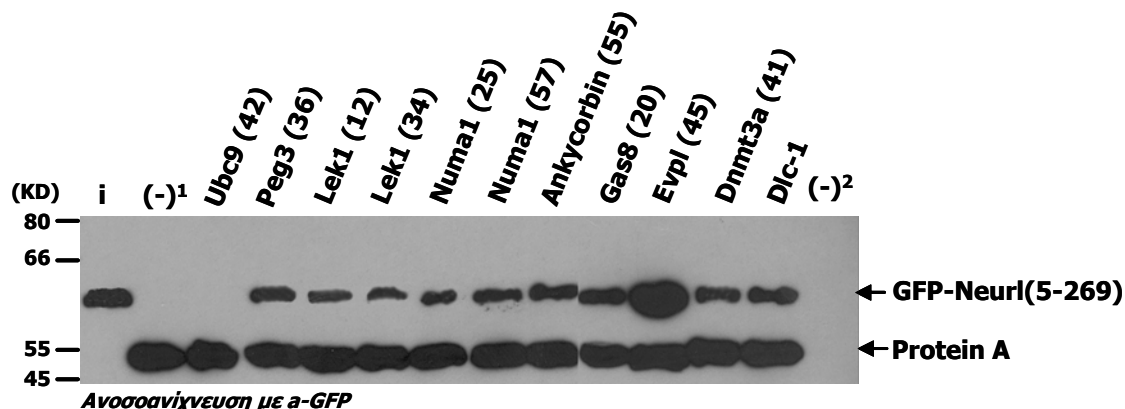
Εικ.73 Ανοσοκατακρήμνιση του GFP-Neur1 (5-537) από τις ανασυνδυασμένες πρωτεΐνες-ΰθηράματα΄ σε κύτταρα COS7, μετά από παροδική συνδιαμόλυνση με την πλασμιδιακή κατασκευή *pEGFPc1/Neur1(5-537)* και κάθε μία από τις πλασμιδιακές κατασκευές των ΰθηραμάτων΄. Μετά την ανοσοκατακρήμνιση με α-Xpress, ανιχνεύθηκε η πρωτεΐνη GFP-Neur1(5-537) στο ανοσοκατακρήμνισμα, με α-GFP.΄ι΄ (*input*): 10% του πρωτεϊνικού εκχυλίσματος πριν την ανοσοκατακρήμνιση, (-)¹: πείραμα αρνητικού ελέγχου (κατακρήμνιση απουσία αντισώματος), (-)²: εκχύλισμα μη διαμολυσμένων κυττάρων COS7.

2) Τα 226 N-τελικά αμινοξέα του Neur1 είναι απαραίτητα για την αλληλεπίδρασή του με τις πρωτεΐνες Dlc-1 και Gas-8. Το μόριο GFP-Neur1(227-574), στο οποίο λείπουν τα 226 N-τελικά αμινοξέα αλληλεπιδρά με όλες τις πρωτεΐνες-ΰθηράματα΄, εκτός των Dlc-1 και Gas8 (Εικ.74).



Εικ.74 Ανοσοκατακρήμνιση του GFP-Neur1 (227-574) από τις ανασυνδυασμένες πρωτεΐνες-ΰθηράματα΄ σε κύτταρα COS7, μετά από παροδική συνδιαμόλυνση με την πλασμιδιακή κατασκευή *pEGFPc1/Neur1(227-574)* και κάθε μία από τις πλασμιδιακές κατασκευές των ΰθηραμάτων΄. Μετά την ανοσοκατακρήμνιση με α-Xpress, ανιχνεύθηκε η πρωτεΐνη GFP-Neur1(227-574) στο ανοσοκατακρήμνισμα, με α-GFP. Δείγματα ΄ι΄, (-)¹, (-)², όπως στην Εικ.73.

3) Το domain NEUZ1 του Neur1 δεν είναι απαραίτητο για την αλληλεπίδρασή του με την πρωτεΐνη Ubc9. Το GFP-Neur1(5-269), το οποίο περιέχει το *domain* NEUZ1 και τις πλευρικές του περιοχές, αλληλεπιδρά με όλα τα ΰθηράματα΄, εκτός του Ubc9 (Εικ.75).



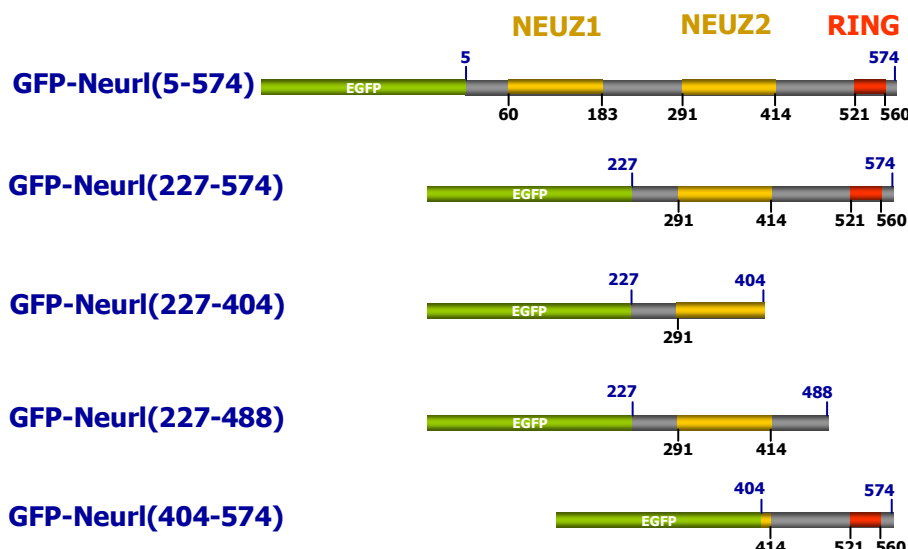
Εικ.75 Ανοσοκατακρήμνιση του GFP-Neur1 (5-269) από τις ανασυνδυασμένες πρωτεΐνες-‘θηράματα’ σε κύτταρα COS7, μετά από παροδική συνδιαμόλυνση με την πλασμιδιακή κατασκευή *pEGFPC1/Neur1(5-269)* και κάθε μία από τις πλασμιδιακές κατασκευές των ‘θηραμάτων’. Μετά την ανοσοκατακρήμνιση με α-Xpress, ανιχνεύθηκε η πρωτεΐνη GFP-Neur1(5-269) στο ανοσοκατακρήμνισμα, με α-GFP. Δείγματα ‘i’, (-)¹, (-)², όπως στην Εικ.73.

Σημειώνεται ότι σε όλα τα παραπάνω πειράματα ελέγχθηκε η παρουσία και των δύο ανασυνδυασμένων πρωτεϊνών στο εκχύλισμα των κυττάρων που χρησιμοποιήθηκε για την ανοσοκατακρήμνιση.

Συνοψίζοντας, η δομή του RING φαίνεται να μην επηρεάζει σημαντικά τις αλληλεπιδράσεις του Neur1 με τα συγκεκριμένα ‘θηράματα’, εφόσον η συγγένεια του Neur1(5-537) με αυτά είναι συγκρίσιμη με αυτήν του πλήρους μήκους Neur1. Τα περισσότερα ‘θηράματα’ (όλα εκτός των Ubc9, Gas8 και Dlc-1) αλληλεπιδρούν ισχυρά και με τις δύο επαναλήψεις NEUZ του Neur1. Ειδικά για τα μόρια Dlc-1 και Gas8, η ακολουθία των αμινοξέων 1-226 που περιέχει και το *domain* NEUZ1, είναι απαραίτητη για την αλληλεπίδρασή τους με την πρωτεΐνη Neur1. Αξίζει να σημειωθεί ότι αυτή η περιοχή του Neur1 καθορίζει και τον υποκυτταρικό του εντοπισμό [Η. Παυλόπουλος, Διδακτορική διατριβή, Παν/μιο Κρήτης, 2002]. Ενδεχομένως, η αλληλεπίδρασή του με τα συστατικά του κυτταροσκελετού Dlc-1 και Gas8 (βλ. Συζήτηση) να επιτρέπει τη συγκεκριμένη τοπολογία του Neur1 στα κύτταρα (βλ. Εισαγωγή). Αντίθετα, για την *in vivo* αλληλεπίδραση του Neur1 με την πρωτεΐνη Ubc9 δεν είναι απαραίτητο το NEUZ1, αλλά το τμήμα του μορίου που περιέχει τις περιοχές NEUZ2 και το N-τελικό μισό του RING (Εικ.75,74,73). Το αποτέλεσμα αυτό, σε συνδυασμό με τις *in vitro* δοκιμές αλληλεπίδρασης (Εικ.53-55) και το πείραμα των δύο υβριδίων (Πίνακας 24) υποδηλώνει ότι τα αμινοξικά κατάλοιπα του Neur1 που είναι απαραίτητα για την αλληλεπίδραση με το Ubc9 εντοπίζονται στο C-τελικό τμήμα του μορίου, που περιέχει τις περιοχές NEUZ2 και RING.

ε. Περαιτέρω ανάλυση των περιοχών του Neur1 που αλληλεπιδρούν με Ubc9

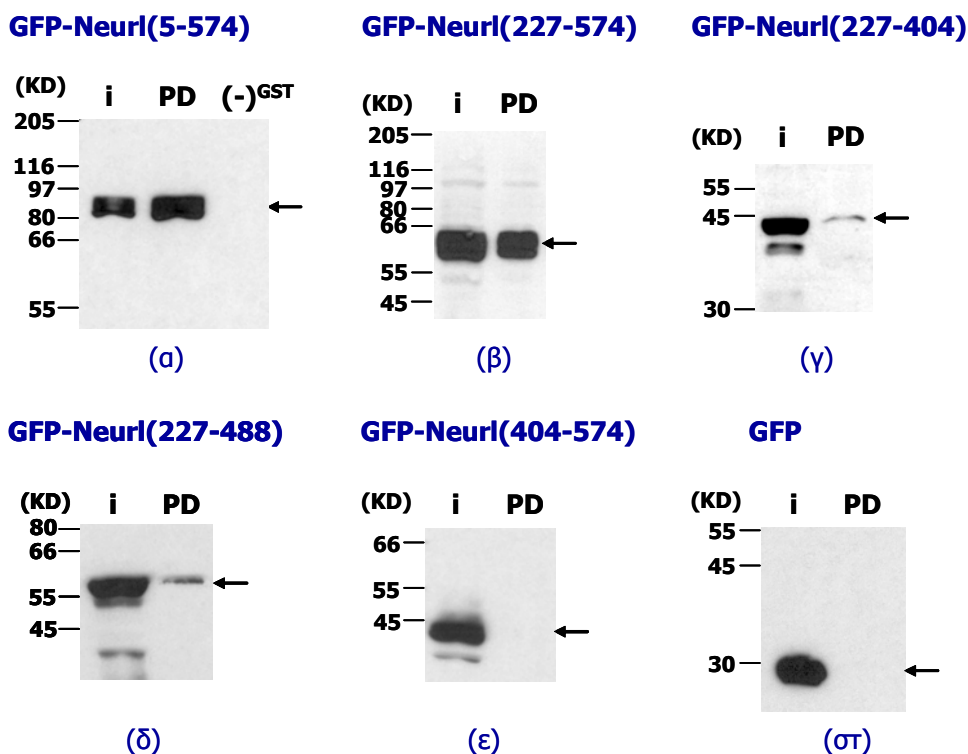
Από τη χαρτογράφηση των αλληλεπιδράσεων του Neur1 με συν-ανοσοκατακρήμνιση των συντηγμένων με GFP τμημάτων του και των πρωτεϊνών-‘θηραμάτων’, προκύπτει ότι η αλληλεπίδραση του Neur1 με την πρωτεΐνη Ubc9 εντοπίζεται στην περιοχή του Neur1 μετά το αμινοξύ 227, επομένως οφείλεται στα *domains* NEUZ2 και RING. Επειδή η Ubc9 είναι το μοναδικό γνωστό συζευκτικό ένζυμο (E2) της τροποποίησης πρωτεϊνών με SUMO (βλ. Συζήτηση), η μελέτη της αλληλεπίδρασης του Neur1 με αυτή είναι δυνατόν να αποκαλύψει σημαντικά στοιχεία για τη λειτουργία ή και τη ρύθμιση του μορίου.



Εικ.76 Το σχεδόν πλήρες και τα ελλιμμηματικά μέρη Neu1, που εκφράστηκαν συντηγμένα με την πράσινη φθορίζουσα πρωτεΐνη (GFP), για να χρησιμοποιηθούν στις δοκιμές *GST-pulldown*, με σκοπό τη χαρτογράφηση των αλληλεπιδράσεων του Neu1 με την πρωτεΐνη Ubc9. Οι αριθμοί αντιστοιχούν σε αμινοξέα της πρωτεΐνης Neu1. Η κλωνοποίηση των αντίστοιχων cDNAs περιγράφεται στα Κεφάλαια 'Υλικά και Μέθοδοι' [για το χιμαιρικό μόριο GFP-Neu1(5-574)] και 'Πλασμιδιακές Κατασκευές' [για τα μέρη GFP-Neu1(227-404, 227-488 και 404-574)].

Για να διερευνηθεί η συμμετοχή του κάθε ενός από τα δύο *domains* (NEUZ2 και RING) της πρωτεΐνης Neu1 στην αλληλεπίδρασή της με το Ubc9, δοκιμάστηκαν οι αλληλεπιδράσεις του Ubc9 με τρία επιπλέον τμήματα του Neu1, τα GFP-Neu1(227-404) και GFP-Neu1(227-488) που περιέχουν το NEUZ2, και το GFP-Neu1(404-574) που περιέχει το RING (Εικ.76). Οι δοκιμές αυτές πραγματοποιήθηκαν σε *in vitro* συνθήκες, όπου ευνοείται η ανίχνευση ασθενών αλληλεπιδράσεων λόγω της μεγαλύτερης αντιπροσώπευσης των βακτηριακά εκφραζόμενων πρωτεϊνών. Συγκεκριμένα, εκχυλίσματα κυττάρων COS7 που είχαν διαμολυνθεί παροδικά με τις πλασμιδιακές κατασκευές που κωδικοποιούν τα συντηγμένα με GFP τμήματα του Neu1 αναμίχθηκαν με 50μg πρωτεΐνης GST-Ubc9, η οποία είχε εκφραστεί σε βακτήρια και δεσμευθεί σε σφαιρίδια σεφαρόζης με επικάλυψη γλουταθειόνης. Μετά τη δοκιμή αλληλεπίδρασης (Υλικά και Μέθοδοι), τα τμήματα του Neu1 που δεσμεύθηκαν στη Ubc9 ανιχνεύθηκαν σε ανάλυση *Western* με α-GFP (Εικ.77).

Από την παραπάνω ανάλυση φαίνεται ότι η αλληλεπίδραση του Neu1 με την πρωτεΐνη Ubc9 επιτυγχάνεται πιο αποτελεσματικά με την ταυτόχρονη παρουσία των *domains* NEUZ2 και RING στο μόριο του Neu1. Πιθανόν οι επαφές των δύο μορίων να οφείλονται σε αμινοξικά κατάλοιπα κατανεμημένα σε αυτά τα δύο *domains* ή και σε ολόκληρο το Neu1. Αντίστοιχα, σε μελέτες αλληλεπιδράσεων μεταξύ πρωτεϊνών που φέρουν RING και μορίων E2 έχει διαπιστωθεί ότι για την αλληλεπίδραση είναι απαραίτητες και περιοχές εκτός του RING [Zheng *et al*, 2000]. Επιπλέον, πολλές από τις γνωστές πρωτεΐνες που αλληλεπιδρούν με την Ubc9, δεν φέρουν RING (βλ. Συζήτηση).



Εικ.77 Η αλληλεπίδραση του *Neur1* με την πρωτεΐνη *Ubc9* οφείλεται κυρίως στα *domains* *NEUZ2* και *RING* του *Neur1*. *In vitro* δοκιμές αλληλεπίδρασης (*GST-pulldown*) της πρωτεΐνης *GST-Ubc9* με το πλήρες (α) ή ελλειμματικά (β-ε) μόρια *Neur1*, συντηγμένα με *GFP*. Οι πρωτεΐνες *GFP-Neur1*(5-574), (227-574), (227-404), (227-488), (404-574) και η *GFP* εκφράστηκαν σε κύτταρα *COS7*, ύστερα από παροδική διαμόλυση με τις αντίστοιχες πλασμιδιακές κατασκευές ή με τον πλασμιδιακό φορέα *pEGFP1* (στ). Τα κύτταρα λύθηκαν σε διάλυμα 50mM *Tris-Cl* pH 8.0, 1mM *EDTA* pH 8.0, 100mM *NaCl*, 0.5% *NP40*, 1mM *PMSF* και η δοκιμή αλληλεπίδρασης πραγματοποιήθηκε όπως περιγράφεται στο Κεφ. 'Υλικά και Μέθοδοι'. Οι συντηγμένες σε *GFP* πρωτεΐνες ανιχνεύθηκαν σε ανάλυση *Western* με *α-GFP*. Στις εικόνες (α) - (στ) οι συντηγμένες με *GFP* (α-ε) πρωτεΐνες ή η *GFP* (στ) υποδεικνύονται με βέλη. Τα δείγματα 'i' (*input*) αντιστοιχούν στο 10% του κυτταρικού εκχυλίσματος που χρησιμοποιήθηκε σε κάθε πείραμα. Τα δείγματα 'PD' (*pull down*) αντιστοιχούν στο 50% των μορίων που δεσμεύθηκαν στην πρωτεΐνη *Ubc9*. Το δείγμα (-)^{GST} (α) αντιπροσωπεύει το πείραμα αρνητικού ελέγχου που δείχνει ότι δεν υπάρχει αλληλεπίδραση μεταξύ της πρωτεΐνης *GFP-Neur1*(5-574) και της *GST* στις ίδιες συνθήκες. Τέλος, στο πείραμα αρνητικού ελέγχου της εικόνας (στ) φαίνεται ότι η *GFP* δεν αλληλεπιδρά με την *GST-Ubc9*.

ΣΥΖΗΤΗΣΗ

Η εργασία αυτή εντάσσεται στα πλαίσια της συμμετοχής του εργαστηρίου μας στο Πρόγραμμα Ανάλυσης του Ανθρώπινου Γονιδιώματος, μεταξύ των βασικών στόχων του οποίου ήταν α) η κατασκευή χρωμοσωμικών χαρτών μεγάλης ευκρίνειας και β) η διερεύνηση της λειτουργίας νέων γονιδίων. Η ομάδα μας έχει εστιάσει στην ανάλυση μιας σειράς γονιδίων της περιοχής 10q23.31-qter με στόχο τη διερεύνηση της μοριακής βάσης μενδελικών νοσημάτων (γονίδια *LGI1* και *FRA10AC1*) [Θ. Σαραφίδου, Διδακτορική Διατριβή, υπό συγγραφή] και τη μελέτη της λειτουργίας γονιδίων με ιδιαίτερο βιολογικό ενδιαφέρον. Για τη λειτουργική ανάλυση του γονιδίου *NeurL* του ανθρώπου, έχουν αξιοποιηθεί μέχρι τώρα τα ετερόλογα συστήματα της δροσόφιλας, του ποντικού και κυτταρικές σειρές θηλαστικών (Η. Παυλόπουλος, Διδακτορική Διατριβή, Παν/μιο Κρήτης, 2002/ Ε. Κουτέλου, Διδακτορική Διατριβή, Παν/μιο Κρήτης, υπό εξέλιξη και παρούσα Διδακτορική Διατριβή). Βασικοί στόχοι αυτής της εργασίας ήταν η συμβολή στην κατασκευή του φυσικού και γονιδιακού χάρτη του χρωμοσώματος 10, η ανάλυση της γονιδιωματικής οργάνωσης και της δομής του γονιδίου *NeurL* του ποντικού και ο προσδιορισμός των πρωτεϊνών που αλληλεπιδρούν άμεσα με την πρωτεΐνη *NeurL*.

1. Το χρωμόσωμα 10 του ανθρώπου

Το πρώτο μέρος της διατριβής συνέβαλλε στην αλληλούχηση του χρωμοσώματος 10 του ανθρώπου με 338 γονιδιωματικούς κλώνους PAC και BAC, οι οποίοι χαρτογραφήθηκαν στην περιοχή 10q23.31-q23.33, μεταξύ των γενετικών δεικτών D10S541 και D10S583. Η πλήρης νουκλεοτιδική αλληλουχία της περιοχής αυτής (4.38 Mb) αλλά και ολόκληρου του χρωμοσώματος (~135 Mb) που είναι σήμερα διαθέσιμη, αποτελεί σημαντικό εργαλείο για την ταυτοποίηση και κλωνοποίηση νέων γονιδίων [Deloukas *et al*, 2004/ www.ensembl.org/Homo_sapiens/mapview?chr=10] και έχει ήδη συνεισφέρει στην ταυτοποίηση και τον χαρακτηρισμό γονιδίων που σχετίζονται με μενδελικά νοσήματα (πχ. το γονίδιο της μερικής επιληψίας κροταφικού λοβού *LGI1*, Θ. Σαραφίδου, Διδακτορική διατριβή, υπό συγγραφή).

Η συστηματική προσπάθεια χαρακτηρισμού των γονιδίων (*gene annotation*) του χρωμοσώματος 10 είχε στόχο την ταυτοποίηση των μονάδων γενετικής πληροφορίας που περιέχονται στη νουκλεοτιδική αλληλουχία, διαδικασία απαραίτητη για τη διερεύνηση της λειτουργίας τους σε μοριακό και κυτταρικό επίπεδο. Βασίστηκε κυρίως σε *de novo* προβλέψεις γονιδιακών δομών από βιοϋπολογιστικά προγράμματα της βάσης δεδομένων *Ensembl*, που είχαν δοκιμαστεί και σε άλλα χρωμοσώματα (GENESCAN, FGENESH, Mungall *et al*, 2003). Ο διαφορετικός βαθμός στον οποίο οι προβλεπόμενες δομές γονιδίων μπορούν να επιβεβαιωθούν από πειραματικά δεδομένα αποτελεί το κριτήριο της ταξινόμησής τους σε διάφορες κατηγορίες (Πίνακας 25).

Στον Πίνακα 25 παραβάλλονται οι γονιδιακές πυκνότητες κάθε κατηγορίας στην περιοχή 10q23.31-q23.33 με αυτές του συνολικού χρωμοσώματος 10.

Πίνακας 25. Ταξινόμηση γονιδίων του χρωμοσώματος 10

	chr 10	10q23.31-q23.33
Μέγεθος (Mb)	~130 (ευχρωματίνη) ~135 (σύνολο)	4.38
Γνωστά γονίδια που κωδικοποιούν πρωτεΐνη	654 (5/Mb)	25 (5.7/Mb)
Νέα γονίδια που κωδικοποιούν πρωτεΐνη	162 (1.2/Mb)	9 (2.1/Mb)
Σύνολο γονιδίων που κωδικοποιούν πρωτεΐνη	816 (6.2/Mb)	34 (7.8/Mb)
Νέα γονίδια που κωδικοποιούν RNA (ncRNAs)	219 (1.7/Mb)	10 (2.3/Mb)
Θεωρητικά γονίδια	322 (2.5/Mb)	7 (1.6/Mb)
Σύνολο λειτουργικών γονιδίων	1357 (10.4/Mb)	51 (11.7/Mb)
Επεξεργασμένα ψευδογονίδια	371 (2.8/Mb)	16 (3.6/Mb)
Μη επεξεργασμένα ψευδογονίδια	59 (0.4/Mb)	1 (0.2/Mb)

Σημ: Σύγκριση της πυκνότητας κάθε κατηγορίας γονιδίων στην περιοχή 10q23.31-q23.33 με αυτές του συνολικού χρωμοσώματος 10. Στις παρενθέσεις φαίνονται οι γονιδιακές πυκνότητες (γονίδια/Mb). Για το σύνολο του χρωμοσώματος αυτές υπολογίστηκαν με βάση το μήκος της ευχρωματίνης.

Τα γονίδια με ανοικτό πλαίσιο ανάγνωσης (ORF) ταξινομούνται σε γνωστά (*known*) ή νέα (*novel*), με βάση το αν αντιπροσωπεύουν ή όχι ήδη γνωστές αλληλουχίες πρωτεϊνών ή κλώνων cDNA. Συγκεκριμένα, τα γονίδια που χαρακτηρίζονται ως 'γνωστά' αντιστοιχούν σε κλώνους cDNA ή πρωτεΐνες του ανθρώπου που έχουν καταχωρηθεί στη βάση δεδομένων *LocusLink*, ενώ τα προβλεπόμενα πεπτιδία που προκύπτουν από την *in silico* μετάφραση των 'νέων' μπορεί να είναι ταυτόσημα ή ομόλογα με πρωτεΐνες άλλων οργανισμών. Στη νουκλεοτιδική αλληλουχία μήκους ~130Mb, που αντιπροσωπεύει το 99.4% της ευχρωματίνης του χρωμοσώματος 10, προβλέφθηκαν 1357 λειτουργικά γονίδια. Συνεπώς, κατά μέσο όρο περιέχονται 10.4 λειτουργικά γονίδια ανά Mb του χρωμοσώματος 10. Τα 816 από αυτά είναι πλήρως χαρακτηρισμένα γονίδια που κωδικοποιούν πρωτεΐνες και διακρίνονται σε 654 γνωστά και 162 νέα. Επιπλέον, διακρίνονται και τα θεωρητικά (*putative*) γονίδια, τα οποία εμφανίζουν ομολογία με ανθρώπινα ESTs και οργάνωση εξονίων/ιντρονίων, όμως δεν έχουν ORF και συνήθως δεν είναι πλήρους μήκους (χωρίς χαρακτηρισμένες περιοχές 5' και 3' UTR). Αξίζει

να σημειωθεί ότι το 18% (322/1787) των προβλεπόμενων γονιδίων του χρωμοσώματος ανήκουν σε αυτή την κατηγορία [Deloukas *et al*, 2004]. Αντίστοιχα, η περιοχή q23.31-q23.33 είναι από τις πιο πλούσιες σε γονίδια, εμφανίζοντας μεγαλύτερη πυκνότητα λειτουργικών γονιδίων σε σχέση με το υπόλοιπο χρωμόσωμα.

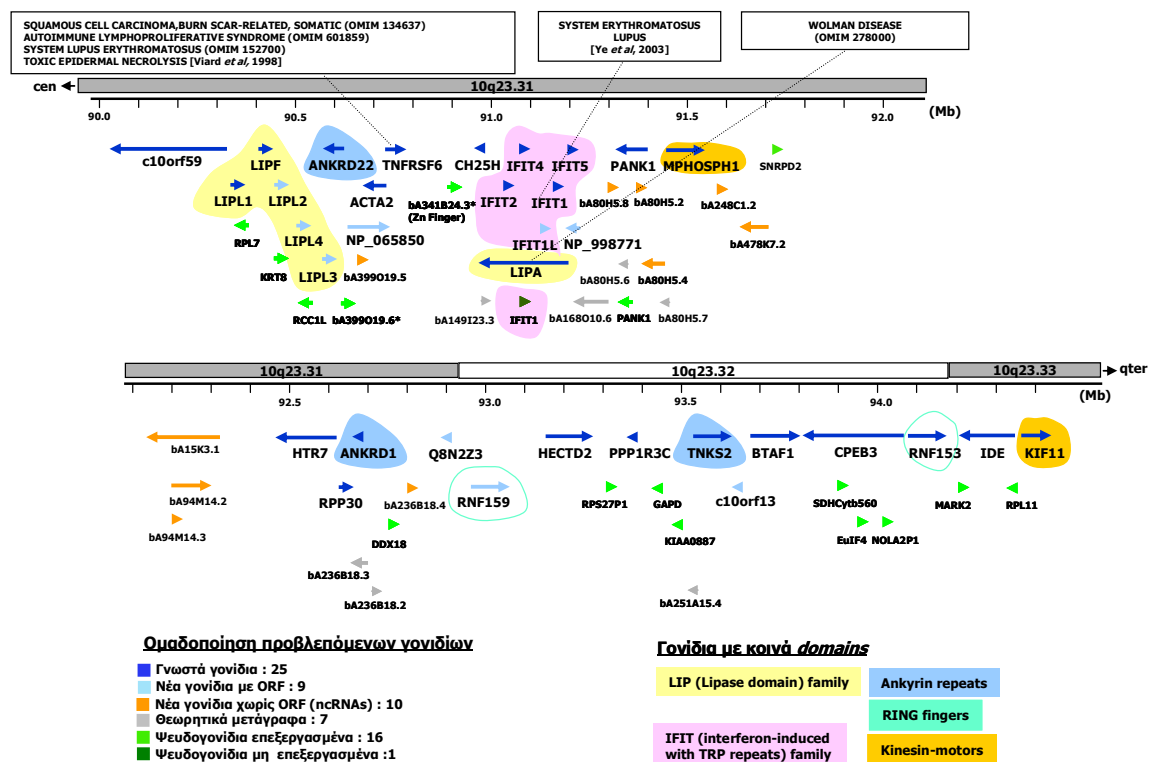
α. Τα γονίδια της περιοχής D10S541-D10S583 (10q23.31-q23.33)

Το μήκος αυτής της χρωμοσωμικής περιοχής, με βάση την πλήρη αλληλούχησή της, είναι 4.38Mb και αντιστοιχεί στο 3% του συνολικού μήκους του χρωμοσώματος 10. Στα πλαίσια ταυτοποίησης των γονιδίων του χρωμοσώματος 10 [Deloukas *et al*, 2004], στη συγκεκριμένη περιοχή εντοπίστηκαν 51 λειτουργικά γονίδια, από τα οποία τα 34 είναι πλήρως χαρακτηρισμένα και κωδικοποιούν πρωτεΐνες (Εικ.78). Από τα τελευταία, τα 25 είναι γνωστά, ενώ 9 είναι επιβεβαιωμένες προβλέψεις νέων γονιδίων. Επιπλέον, ανιχνεύθηκαν και 10 νέα γονίδια που κωδικοποιούν ncRNAs, καθώς και 7 θεωρητικά (*putative*) γονίδια. Τα αποτελέσματα αυτά δικαιολογούν τον χαρακτηρισμό της περιοχής ως 'πλούσιας σε γονίδια', όπως φαίνεται από τη σύγκριση της συχνότητας λειτουργικών γονιδίων της συγκεκριμένης περιοχής (11.4/Mb) και ολόκληρου του χρωμοσώματος (10.4/Mb). Η διαφορά αυτή οφείλεται κυρίως στα γονίδια που κωδικοποιούν πρωτεΐνες, τα οποία εμφανίζονται με σαφώς μεγαλύτερη συχνότητα στη συγκεκριμένη περιοχή (7.8/Mb) σε σχέση με το υπόλοιπο χρωμόσωμα (6.2/Mb), όπως φαίνεται στον Πίνακα 25. Αντίθετα, τα ψευδογονίδια συναντώνται με την ίδια περίπου συχνότητα (~3/Mb).

Στην περιοχή χαρτογραφούνται 17 ψευδογονίδια (Εικ.78), από τα οποία δύο αποτελούν νέες προβλέψεις, ενώ τα υπόλοιπα αντιπροσωπεύουν στην πλειονότητά τους (14/15) επεξεργασμένα μετάγραφα λειτουργικών γονιδίων που εντοπίζονται είτε σε άλλο χρωμόσωμα, είτε και στην ίδια χρωμοσωμική περιοχή (πχ το ψευδογονίδιο *PANK1*). Τα περισσότερα από τα ψευδογονίδια της περιοχής αντιπροσωπεύουν λειτουργικά γονίδια ριβονουκλεοπρωτεϊνών (RPL7, RPS27P11, SNRPD2 και RPL7), δομικών πρωτεϊνών (κερατίνη KRT8), και ενζύμων του μεταβολισμού (δεϋδρογενάσες GAPD και SDHCytb560), οι οποίες αποτελούν τις πιο χαρακτηριστικές κατηγορίες 'housekeeping' γονιδίων, με υψηλά επίπεδα έκφρασης και μεγάλο αριθμό ψευδογονιδίων [Zhang & Gerstein, 2004].

Ο αυξημένος αριθμός γονιδίων που παρατηρείται στην περιοχή αυτή πιθανόν σχετίζεται με τους διπλασιασμούς τμημάτων της 10q23 που συνέβησαν κατά την εξέλιξη των πρωτευόντων. Αυτοί οδήγησαν στη δημιουργία των δύο γονιδιακών αθροισμάτων της περιοχής (LIP και IFIT). Το άθροισμα γονιδίων LIP περιλαμβάνει 6 λειτουργικά γονίδια με *domain* λιπάσης, από τα οποία μόνο τα 4 ήταν γνωστά πριν την πλήρη αλληλούχηση του χρωμοσώματος. Το άθροισμα αυτό έχει ιδιαίτερη σημασία λόγω του συσχετισμού της συγκεκριμένης περιοχής με τη νόσο Wolman ή CESD (*Cholesteryl Ester Storage Disease*, OMIM ID:278000), μια θανατηφόρα ασθένεια της βρεφικής ηλικίας που εκδηλώνεται με ιστολογικές αλλοιώσεις των περισσότερων εσωτερικών οργάνων και σχετίζεται με διαταραχές στον μεταβολισμό λιπιδίων. Αν και διάφορες μεταλλάξεις του γονιδίου *LIPA* (*Lysosomal acid lipase*) έχουν βρεθεί σε ασθενείς της νόσου, δεν αρκούν για τον πλήρη συσχετισμό με τα ετερογενή κλινικά και βιοχημικά ευρήματα [Redonnet-Vernhet *et al*, 1997], γεγονός που δεν αποκλείει τη συνεισφορά και άλλων γονιδίων του αθροίσματος στην εκδήλωση του φαινοτύπου. Από τα

υπόλοιπα γονίδια του αθροίσματος, το *LIPF* (γαστρική λιπάση) είναι αυτό με αποδεδειγμένη λειτουργία λιπάσης, σημαντικής για τον καταβολισμό τριγλυκεριδίων στον γαστρεντερικό σωλήνα [Lohse *et al*, 1997/ Lund *et al*, 1998], ιδιαίτερα σε άτομα με ανεπάρκεια της παγκρεατικής λιπάσης, η οποία κωδικοποιείται από το γονίδιο *PNLIP*, ένα από τα τέσσερα μέλη του αντίστοιχου αθροίσματος στην περιοχή 10q25.3. Τέλος, στην περιοχή του αθροίσματος LIP βρίσκεται και το γονίδιο *CH25H* (*Cholesterol-25 Hydroxylase*), του οποίου το προϊόν, αν και δεν εμφανίζει δομική ομοιότητα με τις λιπάσες του αθροίσματος LIP, λειτουργεί στον καταβολισμό της χοληστερόλης αλλά και στη μεταγραφική ρύθμιση γονιδίων που σχετίζονται με τον μεταβολισμό λιπιδίων, ενισχύοντας το ρόλο της χρωμοσωμικής περιοχής 10q23.31 στη συγκεκριμένη μεταβολική διαδικασία.



Εικ.78 Τα γονίδια της περιοχής 10q23.31-q23.33 (ανάμεσα στους γενετικούς δείκτες D10S541 και D10S583). Οι διαφορετικές κατηγορίες γονιδίων συμβολίζονται με βέλη διαφορετικών χρωμάτων. Φαίνεται η θέση, το μέγεθος (Mb) και η κατεύθυνση της μεταγραφής κάθε γονιδίου. Οι γενετικές ασθένειες που σχετίζονται με γονίδια της περιοχής φαίνονται σε πλαίσιο.

Το άθροισμα IFIT (*Interferon inducible with Tetratricopeptide repeats*) αποτελείται από 5 γονίδια τοποθετημένα πολύ κοντά ή και μέσα στο γονίδιο *LIPA*, τα οποία μεταγράφονται προς την αντίθετη κατεύθυνση. Αξίζει να σημειωθεί ότι έξι από τα οκτώ εναλλακτικά μετάγραφα του *LIPA* επικαλύπτονται με μέλη του αθροίσματος IFIT. Συγκεκριμένα, στα δύο πρώτα ιντρόνια του *LIPA* περιέχονται 4 IFIT γονίδια, από τα οποία το *IFIT5* φαίνεται να μεταγράφεται από κοινό υποκινητή διπλής κατεύθυνσης με το εναλλακτικό μετάγραφο LIPA-004. Αυτό αποτελεί χαρακτηριστικό παράδειγμα επικαλυπτόμενων γονιδίων για το χρωμόσωμα 10, στο οποίο έχουν εντοπιστεί 101 τέτοια ζεύγη, στα 38 από τα οποία το ένα γονίδιο περιέχεται σε ιντρόνιο του άλλου και μεταγράφεται προς αντίθετη κατεύθυνση. Στο 34% των μερικώς επικαλυπτόμενων γονιδίων, η επικάλυψη αφορά τα 5' εξόνια, και σχετίζεται με κοινή ρύθμιση από υποκινητές διπλής κατεύθυνσης [Delou-

kas *et al*, 2004]. Από τα γονίδια του αθροίσματος IFIT, περισσότερα πειραματικά δεδομένα υπάρχουν για τη λειτουργία του *IFIT1*. Η πρωτεΐνη IFIT1 (56KD) αρχικά απομονώθηκε λόγω της υπερέκφρασής της σε ανθρώπινα κύτταρα μετά την επίδραση ιντερφερόνης [Wathelet *et al*, 1986]. Η ανάλυση των υποκινητών των γονιδίων *IFIT1* και *IFIT2* έδειξε ότι περιέχουν ρυθμιστικά στοιχεία απόκρισης σε ιντερφερόνη, αλλά μπορούν να επάγονται και από άλλα σήματα (ιντερλευκίνη-1, δίκλωνο RNA, παρουσία ιών) [Wathelet *et al*, 1987]. Ο ακριβής ρόλος των επαναλήψεων TPR (*tetratricopeptide repeats*) σε αυτές τις πρωτεΐνες δεν έχει διερευνηθεί, όμως το μοτίβο αυτό συναντάται σε μεγάλη ποικιλία πρωτεϊνών, των οποίων διευκολύνει τις αλληλεπιδράσεις με άλλες πρωτεΐνες ή πολυπρωτεϊνικά σύμπλοκα [Das *et al*, 1998]. Το γονίδιο *IFIT1* θεωρήθηκε πρόσφατα υποψήφιο για την εκδήλωση του συστηματικού ερυθροματώδους λύκου (*System Lupus Erythematosus*, *SLE*), λόγω της υπερέκφρασής του στους ασθενείς με SLE [Ye *et al*, 2003]. Στην ίδια μελέτη διαπιστώθηκε η αλληλεπιδράση της πρωτεΐνης IFIT1 με την Rho/RacGEF (*Rho/Rac Guanine Exchange Factor*), διευρύνοντας το πεδίο έρευνας της μοριακής βάσης της νόσου.

Εκτός από τα παραπάνω γονιδιακά αθροίσματα, στην περιοχή συναντώνται και τρεις ομάδες γονιδίων που κωδικοποιούν πρωτεΐνες με κοινά *domains*: 1) τα γονίδια *ANKRD22*, *ANKRD1* και *TNKS2*, των οποίων τα πρωτεϊνικά προϊόντα φέρουν επαναλήψεις αγκυρίνης, 2) τα *RNF159* και *RNF153*, που κωδικοποιούν πρωτεΐνες με RING *domain* και 3) τα γονίδια *MPHOSPH1* και *KIF11*, που κωδικοποιούν πολυπεπτιδία με *domain* της πρωτεΐνης-κινητήρα κινεσίνης. Η λειτουργική σχέση των γονιδίων αυτών των ομάδων δεν έχει διερευνηθεί. Για την πρωτεΐνη TNKS2 (*Tankyrase 2*) είναι γνωστό ότι αλληλεπιδρά με τον παράγοντα TRF-1 (*Telomere Repeat Binding Factor 1*), ρυθμίζοντας το μήκος των τελομερών, με την πρωτεΐνη IRAP (*Insulin Responsive Aminopeptidase*) της εξωτερικής επιφάνειας μεμβρανικών κυστιδίων του *Golgi*, ρυθμίζοντας τον υποκυτταρικό τους προορισμό, αλλά σχηματίζει και ολιγομερή με την *Tankyrase-1*, με την οποία συνεργάζεται στις παραπάνω λειτουργίες [Sbodio *et al*, 2002]. Η MPHOSPH1 (*M phase phosphoprotein 1* ή MPP1) προσδένεται σε μικροσωληνίσκους και κινείται προς τα (+) άκρα τους, ρυθμίζει τον κυτταρικό κύκλο κατά τη μετάβαση από τη φάση G2 στην M και είναι απαραίτητη στα τελικά στάδια της κυτταρικής κίνησης [Abaza *et al*, 2003]. Παρόμοια λειτουργία κινεσίνης έχει διαπιστωθεί και για την KIF11 (ή EG5), η οποία επιπλέον ρυθμίζει τον διαχωρισμό των κεντροσωμάτων και τη δημιουργία των μιτωτικών αστέρων στα αρχικά στάδια της μίτωσης, αλλά είναι απαραίτητη και για τον διαχωρισμό των θυγατρικών κυττάρων μετά τη διαίρεση [Whitehead & Rattner, 1998].

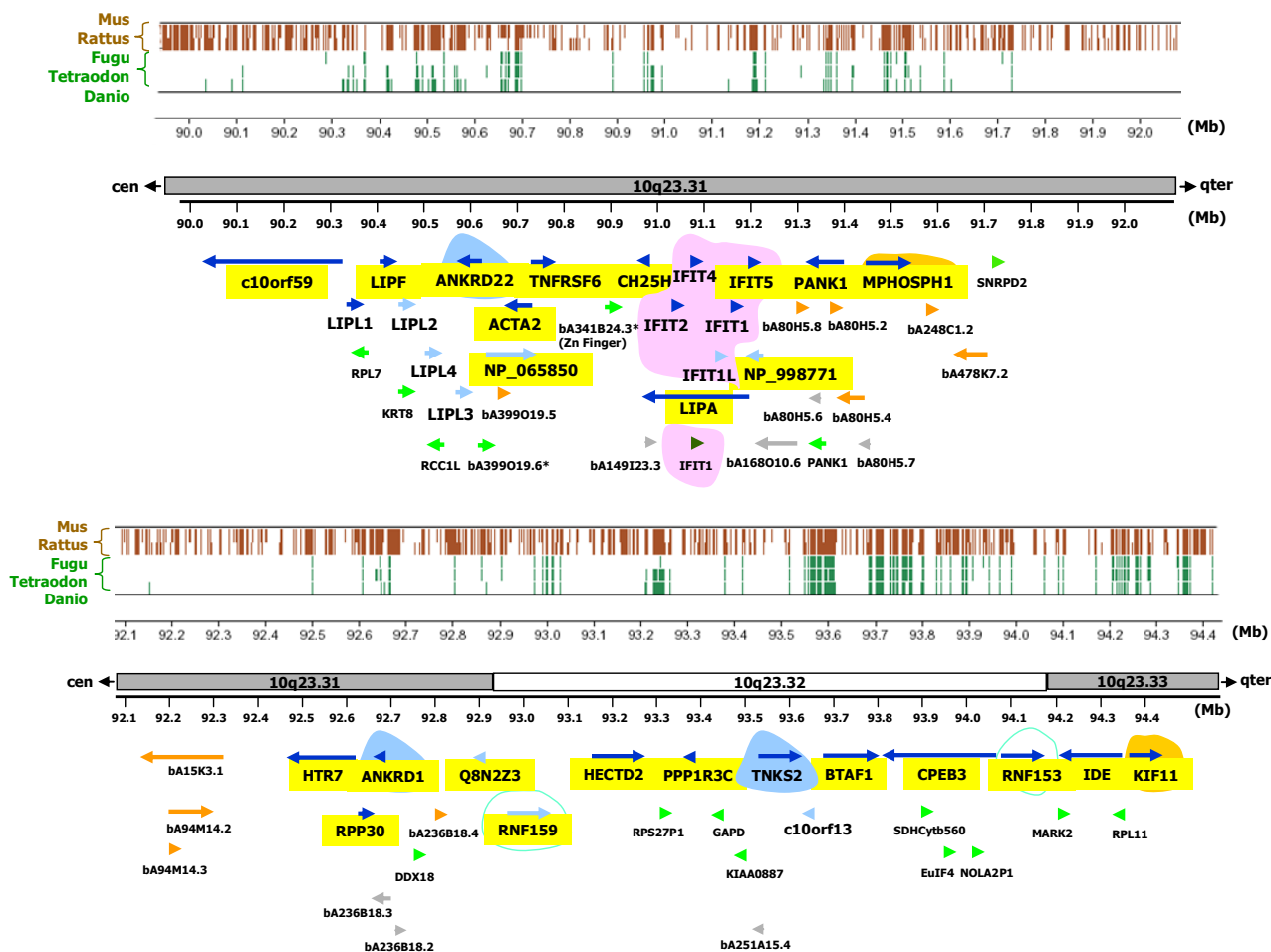
Για το γονίδιο *TNFRSF6* (*Tumour Necrosis Factor Receptor Superfamily 6*), γνωστό και ως *FAS* ή *CD95*, υπάρχουν πολλαπλές αναφορές που το συσχετίζουν με αυτοάνοσες ασθένειες και με ογκογένεση. Συγκεκριμένα, μεταλλάξεις του έχουν ανιχνευθεί σε νεοπλασίες του δέρματος που εκδηλώνονται μετά από εγκαύματα (*Burn-scar related squamous cell carcinoma*, OMIM 134637), καθώς και σε ασθενείς συστηματικού ερυθροματώδη λύκου (OMIM 152700) και του αυτοάνοσου συνδρόμου παιδικής ηλικίας ALPS (*Autoimmune lymphoproliferative syndrome*, OMIM 601859). Επιπλέον, έχει διαπιστωθεί ότι ευθύνεται για την απόπτωση των κερατινοκυττάρων του δέρματος ασθενών της αυτοάνοσου νόσου TEN (*Toxic Epidermal Necrolysis*) [Viard *et al*, 1998].

Μεταξύ των γονιδίων της συγκεκριμένης χρωμοσωμικής περιοχής περιλαμβάνονται αρκετά επιπλέον που κωδικοποιούν πρωτεΐνες γνωστής λειτουργίας, όπως αυτό της υπομονάδας α2 της ακτίνης (*ACTA2*), της παντοθενικής κινάσης (*PANK1*) που καταλύει τη σύνθεση του συνενζύμου CoA, του υποδοχέα σεροτονίνης (*HTR7, hydroxytryptamine receptor 7*), της υπομονάδας 30KD της ριβονουκλεάσης P (*RPP30*) η οποία καταλύει την ωρίμανση των tRNAs στον πυρηνίσκο των ευκαρυωτικών κυττάρων, της ρυθμιστικής υπομονάδας 3C της πρωτεϊνικής φωσφατάσης 1 (*PPP1R3C*) που ρυθμίζει διάφορες κυτταρικές λειτουργίες αποφωσφορυλιώνοντας μεγάλη ποικιλία πρωτεϊνικών υποστρωμάτων, του μεταγραφικού παράγοντα BTAF1 (συστατικού του συμπλόκου TFIID, απαραίτητου για την έναρξη της μεταγραφής από την RNA πολυμεράση II), της πρωτεΐνης CPEB3 (*Cytoplasmic polyadenylation element binding protein 3*) για την οποία έχει διαπιστωθεί η συμμετοχή στη ρύθμιση της μετάφρασης σε αρχικά στάδια της εμβρυϊκής ανάπτυξης και σε συγκεκριμένες περιοχές του ενήλικου εγκεφάλου και του ενζύμου αποικοδόμησης της ινσουλίνης IDE (*insulin degrading enzyme*), το οποίο θεωρείται υπεύθυνο για τον τερματισμό της απόκρισης των κυττάρων σε ινσουλίνη. Τέλος, τα γονίδια *LIPL(1-4)*, *IFIT(2, 4, 5, 1L)*, *ANKRD(22, 1)*, *RNF(159, 153)* και *HECTD2* είναι άγνωστης λειτουργίας, όμως φέρουν χαρακτηριστικά *domains* (λιπάσης, TPR, Ankyrin repeats, RING και HECT, αντίστοιχα) που υποδηλώνουν την πιθανή λειτουργία τους. Επιπλέον, η πρωτεΐνη άγνωστης λειτουργίας c10orf59 φέρει χαρακτηριστικά μοτίβα που συναντώνται σε ένζυμα μεταβολισμού (υδροξυλάσες και αναγωγάσες) και σε αντλίες καλίου, ενώ η NP_065850 φέρει το *domain* Mon34 που συναντάται συχνά σε ρυθμιστικές υπομονάδες του πρωτεασώματος και σε ρυθμιστές μεταγραφικών παραγόντων. Η ίδια πρωτεΐνη, όπως και οι Q8N2Z3 και c10orf13, πιθανόν εντοπίζονται στον πυρήνα λόγω της παρουσίας σε αυτές σημάτων πυρηνικού εντοπισμού που ανιχνεύονται *in silico*, ενώ η NP_998771 φέρει 12 διαμεμβρανικές περιοχές. Η c10orf13 φέρει επιπλέον και μια περιοχή που συναντάται συχνά σε πρωτεΐνες που προσδένονται στον FGF (*Fibroblast Growth Factor*).

β. Εξελικτική συντήρηση των γονιδίων της περιοχής

Στα πλαίσια της συγκριτικής ανάλυσης του χρωμοσώματος 10 με τα γονιδιώματα τρωκτικών και ιχθύων, για την περιοχή 10q23.31-q23.33 εντοπίστηκαν οι συντηρημένες περιοχές, οι οποίες αντιστοιχούν σε γονίδια ή σε διαγονιδιακές περιοχές. Όπως αναμένεται, οι συντηρημένες περιοχές μεταξύ ανθρώπου και τρωκτικών είναι πολύ περισσότερες σε σχέση με αυτές που εντοπίζονται μεταξύ ανθρώπου και ιχθύων (Εικ.79). Επιπλέον, ενώ στο γονιδίωμα των ιχθύων αυτές κυρίως εντοπίζονται σε γονίδια, στα τρωκτικά παρατηρείται συντήρηση και στις διαγονιδιακές περιοχές.

Μεγαλύτερη συντήρηση εμφανίζουν τα γονίδια που κωδικοποιούν πρωτεΐνες. Από τα 34 γονίδια της περιοχής που κωδικοποιούν πρωτεΐνη, τα 22 έχουν ορθόλογα στους ιχθύες (Πίνακας 26), σύμφωνα με στοιχεία της βάσης δεδομένων *Ensembl*. Από αυτά, τα 11 υπάρχουν και στα έντομα (*Drosophila melanogaster* και *Anopheles gambiae*), ενώ 4 υπάρχουν και στον *C. elegans*. Όπως είναι φανερό από το πρότυπο εξελικτικής συντήρησης των γονιδίων γνωστής λειτουργίας της περιοχής, οι πιο συντηρημένες πρωτεΐνες είναι αυτές που σχετίζονται με τη διαδικασία της μετάφρασης, όπως οι CPEB3 και RPP30 ή αποτελούν συστατικά του βασικού μηχανισμού μεταγραφής (BTAF1). Ο μεταβολισμός των λιπιδίων που αντιπροσωπεύεται από τα γονι-



Εικ.79 Εξελικτική συντήρηση της περιοχής 10q23.31-q23.33 στα θηλαστικά και στους ιχθύες. Στο διάγραμμα πάνω από το ιδεόγραμμα της χρωμοσωμικής περιοχής, παριστάνεται με κατακόρυφες γραμμές η συντήρηση της γονιδιωματικής αλληλουχίας (ταυτότητα >90%) στα τρωκτικά (καφέ) και στους ιχθύες (πράσινο). Οι συμβολισμοί των γονιδίων είναι σύμφωνα με την Εικ.78. Σε κίτρινο υπόβαθρο παρουσιάζονται τα γονίδια που κωδικοποιούν πρωτεΐνες και είναι συντηρημένα και στους πέντε οργανισμούς που συγκρίνονται.

για *LIPF* και *LIPA*, φαίνεται επίσης συντηρημένη λειτουργία από τα θηλαστικά μέχρι και τα έντομα, ενώ ορθόλογο του *LIPF* συναντάται και στον νηματώδη. Από τα γονίδια της περιοχής που λειτουργούν σε μονοπάτια κυτταρικής σηματοδότησης, πιο συντηρημένο είναι αυτό του υποδοχέα του νευροδιαβιβαστή σεροτονίνη (*HTR7*), που συναντάται ακόμα και στο απλούστερο νευρικό σύστημα του *C. elegans*. Αντίθετα, το γονίδιο *TNFRSF6* του υποδοχέα TNF (*Tumour Necrosis Factor*) σχετίζεται με την ανοσολογική απόκριση, εξειδικευμένη λειτουργία των θηλαστικών. Πιθανόν η αυξημένη πολυπλοκότητα αλλά και οι φυσιολογικές διαφορές που χαρακτηρίζουν το σύστημα των θηλαστικών να οδήγησαν στην επέκταση των γονιδιακών αθροισμάτων LIP και IFIT ειδικά σε αυτά (Πίνακας 26), δημιουργώντας ενδεχομένως νέες πρωτεΐνες με πιο εξειδικευμένες λειτουργίες στον μεταβολισμό λιπιδίων (LIPL1-4) ή στην απόκριση σε ιντερφερόνη (IFIT2, -4, -1L). Ενδιαφέρον παρουσιάζει και το γεγονός της παρουσίας ορθόλογων γονιδίων *Tankyrase 2* (*TNKS2*) στα θηλαστικά και στα έντομα, αλλά όχι στους ιχθύες. Η λειτουργία της Tankyrase 2 στον καθορισμό του μήκους των τελομερών στους ιχθύες πιθανόν υποκαθίσταται από την Tankyrase 1. Σύμφωνα με τα δεδομένα της βάσης δεδομένων *Ensembl*, το γονίδιο *TNKS* που χαρτογραφείται στο ανθρώπινο χρωμόσωμα 8 και κωδικοποιεί την Tankyrase 1, έχει ορθόλογα στους ιχθύες αλλά όχι στα έντομα. Το ίδιο ισχύει και για το γονίδιο της ακτίνης

ACTA2, που ανήκει σε μια μεγάλη οικογένεια γονιδίων (>100 μέλη στο ανθρώπινο γονιδίωμα), ορισμένα από τα οποία έχουν ορθόλογα και στα ασπόνδυλα.

Πίνακας 26. Εξελικτική συντήρηση γονιδίων της περιοχής 10q23.31-q23.33 που κωδικοποιούν πρωτεΐνες

	Ευκαρυωτικοί	Πολυκύτταροι	Σπονδυλωτά	Θηλαστικά	Θηλαστικά & Έντομα	Πρωτεύοντα
Γονίδια γνωστής λειτουργίας	<i>LIPF</i> <i>HTR7</i> <i>RPP30</i> <i>CPEB3</i>	<i>LIPA</i> <i>PANK1</i> <i>PPP1R3C</i> <i>BTAF1</i> <i>IDE</i> <i>KIF11</i>	<i>ACTA2</i> <i>CH25H</i> <i>MPHOSPH1</i>	<i>IFIT1</i> <i>TNFRSF6</i>	<i>TNKS2</i>	
Γονίδια άγνωστης λειτουργίας		<i>RNF153</i>	<i>C10orf59</i> <i>ARKRD22</i> <i>NP_065850</i> <i>IFIT5</i> <i>NP_998771</i> <i>ANKRD1</i> <i>RNF159</i> <i>HECTD2</i>	<i>c10orf13</i> <i>LIPL1</i> <i>LIPL2</i> <i>LIPL3</i> <i>LIPL4</i> <i>IFIT2</i> <i>IFIT4</i> <i>IFIT1L</i>		<i>Q8N2Z3</i>

ΟΡΓΑΝΙΣΜΟΙ						
<i>Pan</i>	+	+	+	+	+	+
<i>Mus</i>	+	+	+	+	+	-
<i>Rattus</i>	+	+	+	+	+	-
<i>Tetraodon</i>	+	+	+	-	-	-
<i>Fugu</i>	+	+	+	-	-	-
<i>Danio</i>	+	+	+	-	-	-
<i>Drosophila</i>	+	+	-	-	+	-
<i>Anopheles</i>	+	+	-	-	+	-
<i>Caenorhabditis</i>	+	-	-	-	-	-

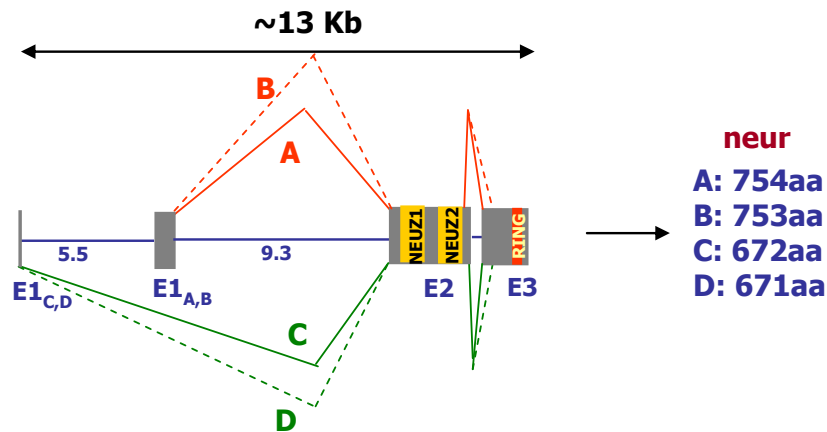
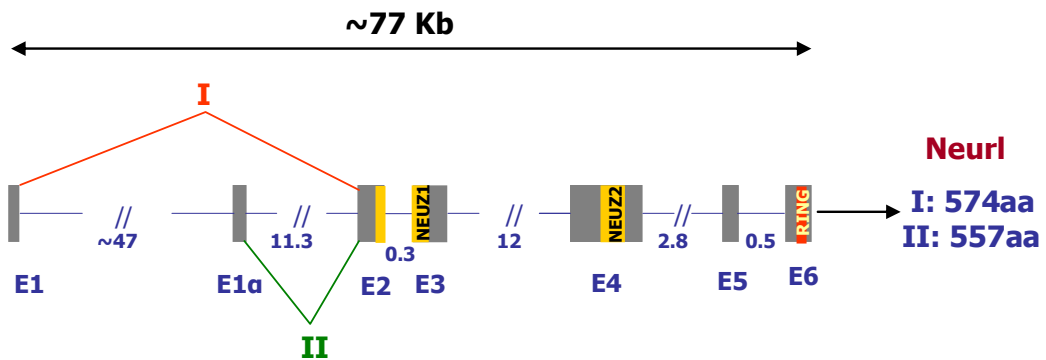
Σημ. 1) Οι πέντε στήλες του πίνακα ανηπρωσωπεύουν ομάδες γονιδίων που διακρίνονται με βάση τη φυλογενετική ομάδα των οργανισμών στους οποίους συναντώνται. Με (+) συμβολίζεται η παρουσία ορθόλογου γονιδίου σε κάθε οργανισμό.

Σημ. 2) Τα γονίδια γνωστής λειτουργίας ομαδοποιούνται σε τέσσερις λειτουργικές ομάδες:

- Μεταβολισμός
- Σηματοδότηση
- Μεταγραφή – Μετάφραση
- Μίτωση

2. Συντήρηση της γονιδιωματικής οργάνωσης του *Neurl*

Όπως διαπιστώσαμε μελετώντας την οργάνωση εξονίων-ιντρονίων στα ορθόλογα γονίδια *Neurl* (*Mus musculus*) και *NEURL* (*Homo sapiens*), και στα δύο είδη διατηρείται ο αριθμός και το μέγεθος των εξονίων, όπως και τα σχετικά μεγέθη των ιντρονίων. Ειδικότερα, διαπιστώσαμε συντήρηση της νουκλεοτιδικής αλληλουχίας, τόσο στις κωδικές όσο και σε μη κωδικές περιοχές, η οποία υποδηλώνει την παρουσία συντηρημένων λειτουργικών στοιχείων. Χαρακτηριστική είναι η συντήρηση της εναλλακτικής συναρμογής εξονίων και στα δύο γονίδια, που οδηγεί στη σύνθεση των δύο εναλλακτικών μεταγράφων, 4.4Kb και 3.7Kb. Αυτά κωδικοποιούν τις δύο ισομορφές του *Neurl* των θηλαστικών (574 και 557 αμινοξέα, αντίστοιχα), οι οποίες διαφέρουν στο αμινοτελικό τους άκρο. Ο ίδιος μηχανισμός διατηρείται και στο ορθόλογο γονίδιο της δροσόφιλας, προκαλώντας και σε αυτό ετερογένεια του αμινοτελικού άκρου της πρωτεΐνης *neur* (Εικ.80) Αν και δεν έχουν πραγματοποιηθεί αναλυτικές μελέτες σύγκρισης του προτύπου έκφρασης των εναλλακτικών ισομορφών κατά την ανάπτυξη των θηλαστικών, οι Vollrath *et al* (2001) αναφέρουν ότι το μικρότερο μετάγραφο εκφράζεται στους σκελετικούς μύες των ενήλικων ποντικών, ενώ το μεγαλύτερο στον εγκέφαλο. Άγνωστο παραμένει και το ενδεχόμενο λειτουργικής διάκρισης των δύο πρωτεϊνικών ισομορφών του *Neurl*. Με βάση τα αποτελέσματά μας, οι πρωτεϊνικές αλληλεπιδράσεις που ανιχνεύσαμε για την ισομορφή των 574 αμινοξέων πρέπει να ισχύουν και για αυτήν των 557 αμινοξέων, εφόσον στο αμινοτελικό τμήμα του *Neurl* δεν περιλαμβάνονται τα *domains* NEUZ1 και NEUZ2 που καθορίζουν τις αλληλεπιδράσεις του. Όμως, δεν αποκλείεται η ιστοειδική έκφραση ή ο διαφορετικός υποκυτταρικός εντοπισμός των δύο ισομορφών. Μια ένδειξη που υποστηρίζει το τελευταίο ενδεχόμενο προκύπτει από την *in silico* πρόβλεψη θέσεων μυριστιλίωσης στις δύο πρωτεΐνες *Neurl*. Η N-μυριστιλίωση είναι πολύ συχνή τροποποίηση σε πρωτεΐνες που λειτουργούν σε σηματοδοτικά μονοπάτια, η οποία ενισχύει την αλληλεπίδρασή τους με μεμβράνες, επηρεάζοντας τον υποκυτταρικό τους εντοπισμό [McCabe & Berthiaume, 2001]. Το πρόγραμμα '*Myristoylator*' (au.expasy.org/tools/myristoylator) ανιχνεύει πιθανές θέσεις μυριστιλίωσης για τις γλυκίνες του αμινοτελικού άκρου πρωτεϊνών. Σύμφωνα με αυτό, η μεγαλύτερη ισομορφή του *Neurl* (574 αμινοξέα) περιέχει μια πιθανή θέση μυριστιλίωσης στην πρώτη γλυκίνη του N-τελικού της άκρου, ενώ η μικρότερη ισομορφή (557 αμινοξέα) δεν μυριστιλιώνεται. Παρόμοια διάκριση ισχύει και για τις εναλλακτικές μορφές *neur* της δροσόφιλας (Εικ.80). Με δεδομένη τη συντήρηση αυτής της διάκρισης των ισομορφών του *Neurl* από τη δροσόφιλα μέχρι τον άνθρωπο και τον μεμβρανικό συσχετισμό των δύο πρωτεϊνών [Pavlopoulos *et al*, 2001], θα ήταν ενδιαφέρον να διερευνηθεί πειραματικά η λειτουργική της σημασία. Αξίζει, όμως, να σημειωθεί ότι προηγούμενα βιοχημικά πειράματα στα οποία είχε χρησιμοποιηθεί η ανασυνδυασμένη πρωτεΐνη GFP-*Neurl*(5-574), έδειξαν ότι αυτή αλληλεπιδρά με την εξωτερική επιφάνεια μεμβρανικών κυστιδίων, χωρίς να περιέχει την προβλεπόμενη θέση N-μυριστιλίωσης [Η. Παυλόπουλος, Διδακτορική Διατριβή, Παν/μιο Κρήτης, 2001].

(α) *Drosophila melanogaster neur***(A,B) N-ter (92aa): MGLSD.....FGLVR****(C,D) N-ter (10aa): M****(β) *Mus musculus Neurl*****(I) N-ter (28aa): M****(II) N-ter (10aa): M**

Εικ.80 Σύγκριση γονιδιωματικής οργάνωσης των γονιδίων *neur* (*D.m*) και *Neurl* (*M.m*). Το γονίδιο των θηλαστικών (β) έχει περισσότερα και μεγαλύτερα ιντρόνια. Και στα δύο γονίδια συμβαίνει εναλλακτική συναρμογή εξονίων, που προκαλεί ετερογένεια στο αμινοτελικό άκρο των πρωτεϊνών *neur* και *Neurl*. Στα σχήματα των δύο γονιδίων σημειώνονται οι περιοχές των εξονίων που κωδικοποιούν τα *domains* NEUZ1, NEUZ2 και RING. Για κάθε εναλλακτική ισομορφή των πρωτεϊνών *neur* και *Neurl*, φαίνεται το διαφορετικό N-τελικό άκρο και σημειώνονται οι προβλεπόμενες θέσεις μυριστιλίωσης, με βάση το πρόγραμμα 'Myristilator'.

3. Εξελικτική ιστορία της πρωτεΐνης Neuralized

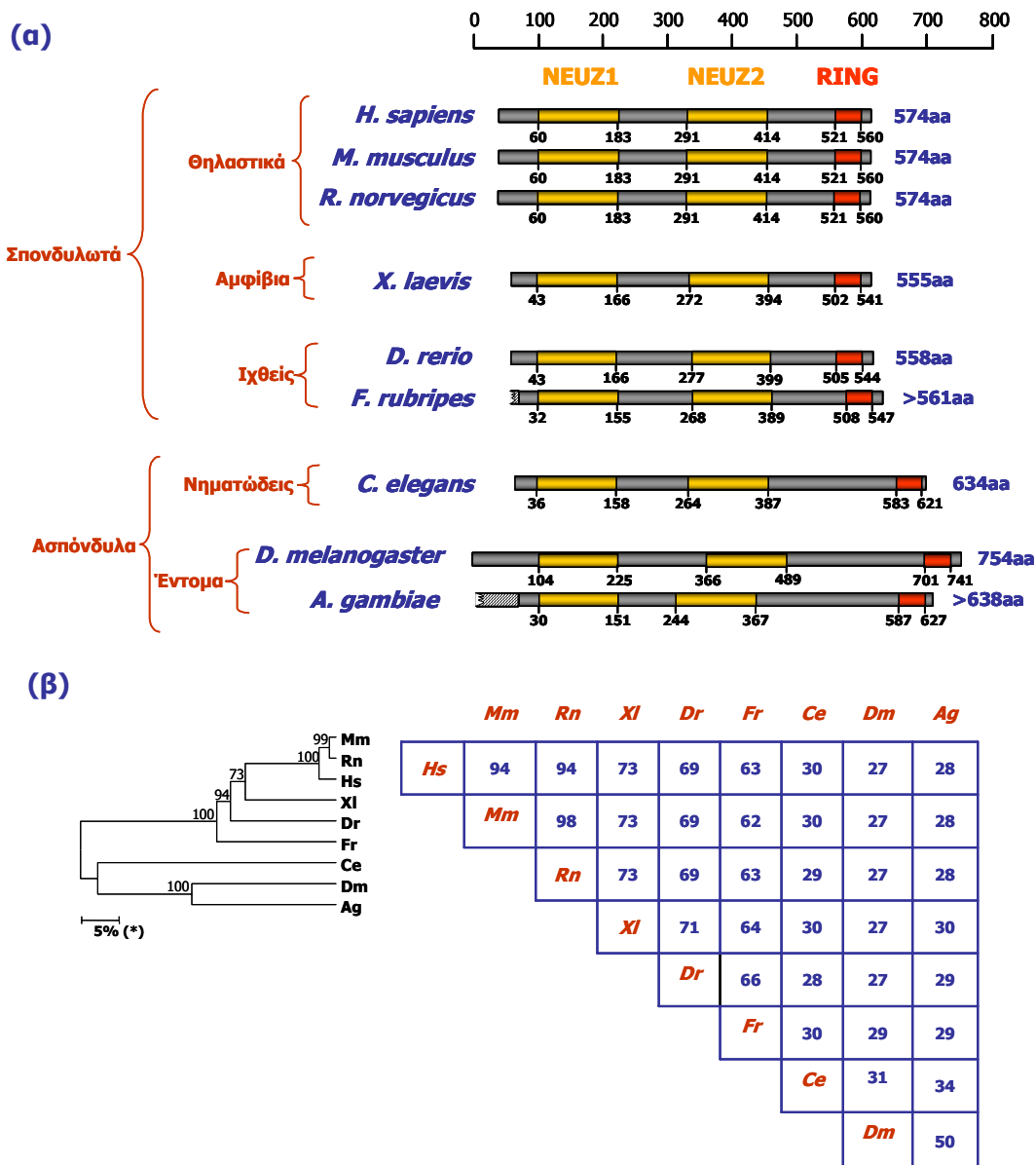
Η πρωτεΐνη Neuralized είναι συντηρημένη στους πολυκύτταρους ζωικούς οργανισμούς. Προηγούμενη σύγκριση μεταξύ των ορθόλογων πρωτεϊνών των θηλαστικών *Mus musculus* και *Homo sapiens* με αυτές της *Drosophila melanogaster* και του *Caenorhabditis elegans* [Pavlopoulos *et al*, 2002] υπέδειξε ταυτότητα αμινοξικών καταλοίπων μεταξύ των ορθόλογων μορίων σε ποσοστά 27% και 30%, αντίστοιχα. Τα πρόσφατα στοιχεία γονιδιωματικής αλληλούχησης και άλλων οργανισμών μας επιτρέπουν επιπρόσθετες δομικές συγκρίσεις των πρωτεϊνών Neuralized στα είδη *Anopheles gambiae*, *Fugu rubripes*, *Danio rerio*, *Xenopus laevis* και *Rattus norvegicus*. Επιπλέον, συγκρίνοντας ξεχωριστά τις περιοχές των επαναλήψεων NEUZ1, NEUZ2 και το RING των ορθόλογων πρωτεϊνών Neuralized, είναι δυνατόν να ελεγχθεί η αμινοξική συντήρηση σε αυτές, γεγονός που υποδηλώνει τη σημασία της κάθε περιοχής στη λειτουργία του μορίου. Συγκρίνοντας τις ορθόλογες πρωτεΐνες Neuralized των θηλαστικών με αυτές άλλων σπονδυλωτών (*Fugu rubripes*, *Danio rerio*, *Xenopus laevis*), εντόμων και του νηματώδη προκύπτουν τα εξής:

α. Διαφορά στο μήκος των πρωτεϊνών Neuralized μεταξύ σπονδυλωτών και ασπονδύλων

Οι πρωτεΐνες των ασπονδύλων είναι κατά πολύ (100-200 αμινοξέα) μεγαλύτερες αυτών των σπονδυλωτών, με χαρακτηριστικά μεγαλύτερη την περιοχή ανάμεσα στο NEUZ2 και το RING domain (Εικ.81). Οι επιπρόσθετες περιοχές που υπάρχουν στα μόρια των ασπονδύλων, είτε ευθύνονται για λειτουργίες ή αλληλεπιδράσεις του Neu1 ειδικές για το σύστημα των ασπονδύλων, είτε αποτελούν απλώς 'συνδετικές' περιοχές μεταξύ των λειτουργικών περιοχών του Neu1 που έχουν διατηρηθεί κατά την εξέλιξη στα σπονδυλωτά.

β. Αμινοξική συντήρηση κατά μήκος της πρωτεΐνης Neuralized

Η πρωτεΐνη NEURL του ανθρώπου εμφανίζει ταυτότητα αμινοξέων σε ποσοστά 60-70% με τις ορθόλογες πρωτεΐνες άλλων σπονδυλωτών (αμφιβίων και ιχθύων), 30% με την πρωτεΐνη του νηματώδη και ελάχιστα μικρότερο ποσοστό (27-28%) με τις πρωτεΐνες των εντόμων (Εικ.81). Με βάση αυτά τα ποσοστά είναι αναμενόμενη και η συντήρηση της λειτουργίας του Neu1, ιδιαίτερα μεταξύ των σπονδυλωτών. Μεταξύ σπονδυλωτών και ασπονδύλων, αν και τα βασικά δομικά χαρακτηριστικά του μορίου διατηρούνται, ενδεχομένως υπάρχουν διαφοροποιήσεις στη λειτουργία του, αντανακλώντας τη διαφορά στην πολυπλοκότητα των δύο συστημάτων.



Εικ.81 Τα ορθόλογα μόρια Neuralized. (α) Διαγραμματική απεικόνιση των πρωτεϊνών, στην οποία φαίνονται οι συντηρημένες περιοχές NEUZ1, NEUZ2 και RING. **(β)** Το φυλογενετικό δέντρο και τα ποσοστά (%) ταυτότητας αμινοξέων που προκύπτουν από τη σύγκριση των ορθόλογων μορίων Neuralized. (*): ποσοτικοποίηση των εξελικτικών αποστάσεων ως διαφορά της αμινοξικής ακολουθίας κατά 5%. Οι αριθμοί στα κομβικά σημεία του δέντρου αντιστοιχούν σε βαθμούς του 'Bootstrap test', για τη στατιστική αξιολόγηση κάθε διακλάδωσης (βλ. Ύλικά και Μέθοδοι). Για τις συγκρίσεις χρησιμοποιήθηκαν οι αλληλουχίες των βάσεων δεδομένων GenBank και Ensembl με κωδικούς πρόσβασης: S35503 (*D. melanogaster*), XP_321060 (*A. gambiae*), U40945 (*C. elegans*), SINFRUP00000154136 (*F. rubripes*), ENSDARG00000023281 (*D.rerio*), Q8UVQ6 (*X. laevis*), ENSRNOP00000027567 (*R. norvegicus*), Y15160 (*Mus musculus*) και AAC17474 (*H. sapiens*).

γ. Οι περιοχές NEUZ1 εμφανίζονται ιδιαίτερα συντηρημένες μεταξύ των ορθόλογων πρωτεϊνών Neuralized

Συγκεκριμένα, η περιοχή NEUZ1 της πρωτεΐνης του *Homo sapiens*, εμφανίζει υψηλά ποσοστά ταυτότητας αμινοξέων με τις αντίστοιχες πρωτεϊνικές περιοχές των πιο απομακρυσμένων εξελικτικά ειδών *F.rubripes* (84%), *D.melanogaster* και *C.elegans* (44%). Αξίζει να σημειωθεί ότι η NEUZ1 είναι η πιο συντηρημένη περιοχή του μορίου στα σπονδυλωτά (Εικ.82). Αντίθετα, η περιοχή NEUZ2 εμφανίζει μικρότερη συντήρηση τόσο μεταξύ των σπονδυλωτών όσο και συγκρίνοντας σπονδυλωτά με ασπόνδυλα, με ποσοστά ταυτότητας αμινοξέων 62% μεταξύ *H. sapiens* και *F.rubripes* και 32% ή 34% μεταξύ *H. sapiens* και

D.melanogaster ή *C.elegans* (Εικ.82). Από τα 124 αμινοξέα της NEUZ1 που στοιχήθηκαν, τα 82 είναι συντηρημένα μεταξύ των ειδών (66%, Εικ.83), ενώ το αντίστοιχο ποσοστό για την NEUZ2 είναι 69 από τα 123 αμινοξέα (56%, Εικ.84). Τέλος, συγκρίνοντας την ακολουθία των συντηρημένων αμινοξέων (*consensus*) των περιοχών NEUZ1 και NEUZ2 προκύπτει το μοτίβο της Εικ.85, στο οποίο 62 αμινοξέα κατά μήκος του NEUZ *domain* διατηρούνται και στις δύο επαναλήψεις του μορίου Neutralized. Η υψηλή συντήρηση αυτών των αμινοξέων μπορεί να οφείλεται στον καθοριστικό ρόλο της περιοχής των επαναλήψεων NEUZ για την υποκυτταρική τοπογραφία του NeurI [Pavliopoulos *et al*, 2002] αλλά και για τις πρωτεϊνικές του αλληλεπιδράσεις.

		<i>Mm</i>	<i>Rn</i>	<i>Xl</i>	<i>Dr</i>	<i>Fr</i>	<i>Ce</i>	<i>Dm</i>	<i>Ag</i>
<i>Hs</i>	N1	99%	99%	91%	87%	84%	44%	44%	44%
	N2	96%	98%	69%	68%	62%	32%	34%	32%
	R	97%	97%	80%	79%	77%	55%	48%	46%
<i>Mm</i>	N1	100%	92%	88%	85%	45%	45%	44%	
	N2	96%	68%	68%	59%	32%	33%	32%	
	R	100%	80%	79%	77%	55%	48%	46%	
<i>Rn</i>	N1		92%	88%	85%	45%	45%	44%	
	N2		69%	68%	61%	34%	33%	33%	
	R		80%	79%	77%	55%	48%	46%	
<i>Xl</i>	N1			87%	86%	44%	44%	42%	
	N2			75%	60%	33%	36%	39%	
	R			84%	90%	50%	51%	48%	
<i>Dr</i>	N1				89%	45%	41%	42%	
	N2				59%	31%	37%	41%	
	R				84%	51%	47%	42%	
<i>Fr</i>	N1					45%	42%	45%	
	N2					34%	41%	40%	
	R					52%	53%	51%	
<i>Ce</i>	N1						49%	52%	
	N2						41%	44%	
	R						51%	51%	
<i>Dm</i>	N1							74%	
	N2							65%	
	R							87%	

Εικ.82 Σύγκριση των επιμέρους περιοχών των ορθόλογων πρωτεϊνών Neutralized NEUZ1 (N1), NEUZ2 (N2) και RING (R). Στον πίνακα αναφέρονται τα ποσοστά ταυτότητας αμινοξέων για κάθε περιοχή των συγκρινόμενων μορίων. Με τα διαφορετικά χρώματα παρουσιάζονται για κάθε ορθόλογη πρωτεΐνη τα ποσοστά συντήρησης των περιοχών με την υψηλότερη (κόκκινο) και την δεύτερη υψηλότερη (πράσινο) συντήρηση, στις περιπτώσεις που αυτές είναι ευδιάκριτες (με διαφορά >3%). Η περιοχή NEUZ1 είναι η πιο συντηρημένη περιοχή του μορίου μεταξύ των σπονδυλωτών, ενώ το RING εμφανίζεται αρκετά συντηρημένο μέχρι και τα ασπόνδυλα.

Συγκεκριμένα, η περιοχή NEUZ1 φαίνεται απαραίτητη για την αλληλεπίδραση του NeurI με πρωτεΐνες που σχετίζονται με τον κυτταροσκελετό (βλ. Αποτελέσματα), γεγονός που ενδεχομένως ρυθμίζει και τον υποκυτταρικό του εντοπισμό. Επιπλέον, όλα τα 'θηράματα' του NeurI που χαρακτηρίσαμε αλληλεπιδρούν με μία από τις δύο ή και με τις δύο επαναλήψεις NEUZ. Έτσι, διαπιστώσαμε ότι η αλληλεπίδραση του Ubc9 απαιτεί την παρουσία του NEUZ2, το οποίο συνεισφέρει μαζί με το RING στη συγκεκριμένη αλληλεπίδραση.

NEUZ1

(a)

```

1      10      20      30      40      50      60      70
Hs ATPLLFHPHTKGSQILMDLSHKAVKRQASFCNAITFSNRPVLIYEQVRLKITKKQCCWSGALRLGFTSKDP
Mm ATPLLFHPHTKGSQILMDLSHKAVKRQASFCNAITFSNRPVLIYEQVRLKITKKQCCWSGALRLGFTSKDP
Rn ATPLLFHPHTKGSQILMDLSHKAVKRQASFCNAITFSNRPVLIYEQVRLKITKKQCCWSGALRLGFTSKDP
Xl ITPLLFHPHAKGSQILMDTSQKAVKRQASFCNAITFSNRPVLIYEQVRLKITKKQCCWSGALRLGFTSKDP
Dr MSPLLFHPNAKGSQIVMDLAQRSVKRQASFCNAITFSNRPVLIYEQVRLKITKKQCCWSGALRLGFTSKDP
Fr SVPLLFHPSTKGSQIVMDMSQRTVKRQASFCNAITFSNRPIAVYEQVRLKITKKQCCWSGALRLGFTCKDP
Ce RGPLQFH-CIHGSNVVILKNGRLAKRRESFCKLAFSNRPIEIDENVCLRLCEVGTNWSGVLRFGVTNDDP
Dm LPPLQFH-SVHGDNIRISRDGTLARRFESFCRAITFSARPVRINERICVKFAEISNNWNGGIRFGFTSNDP
Ag LPPLQFH-TVHGDNIRISREGTVAKRYESFCKGITFSARPVRVNERVCVKFLDISNNWSGVIRFGFTCNDP

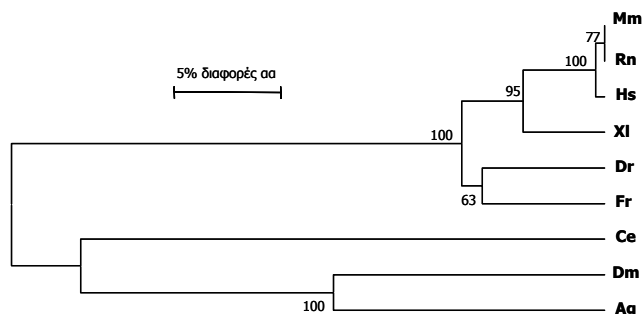
Co *$PLlFH.%.%G%!!%.%.+!+R%.SFC%$!tFSnR!!.E%!r!+*$%.%$SWSG!!R*GfT$%DP

80      90      100      110      120
Hs SRIHPDSLPKYACPDLVSQSGFWAKALPEEFANEGNIIAFWVDKKGRVFHRIN
Mm SRIHPDSLPKYACPDLVSQSGFWAKALPEEFANEGNIIAFWVDKKGRVFYRIN
Rn SRIHPDSLPKYACPDLVSQSGFWAKALPEEFANEGNIIAFWVDKKGRVFYRIN
Xl SRINPDTLPKYACPDLVSQSGFWAKALPEEFANEGNIIAFWVDKKGRVFYRVN
Dr SRINPDSLPKYACPDLVSQSGFWAKALPEEFANEGLISFWVDKKGRVFYRIN
Fr SRINPDTLPKYACPDLVSQSGFWAKALPEELSNEGTVIAFWVDKKGRVFYRIN
Ce EMYRDIPVPTFACPDLTKDGYWAKALPERYSNEGNILHFYVNAHGELFYGIN
Dm VTLEGT-LPKYACPDLTNRPGFWAKALHEQYCEKDNILYYVNGAGDVIYGIN
Ag ASLRGN-LPKYACPDLTNKPGFWAKALNERYCYRGNVLFYYVTPSGDVHFGIN

Co .%*$%.!P*ACPDL.%*$G*WAKALPE*%$%gn!!.**V%.%G%!**r!N

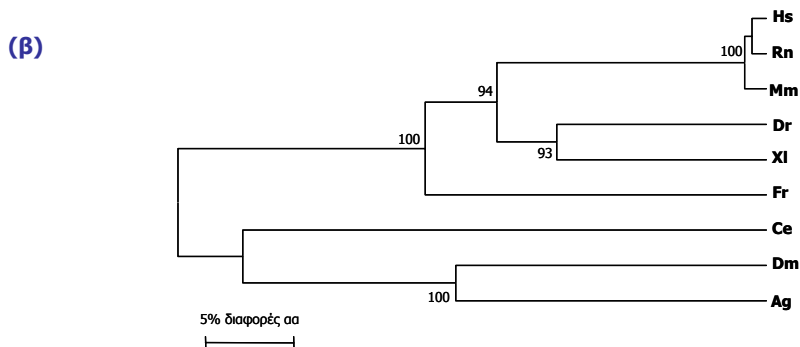
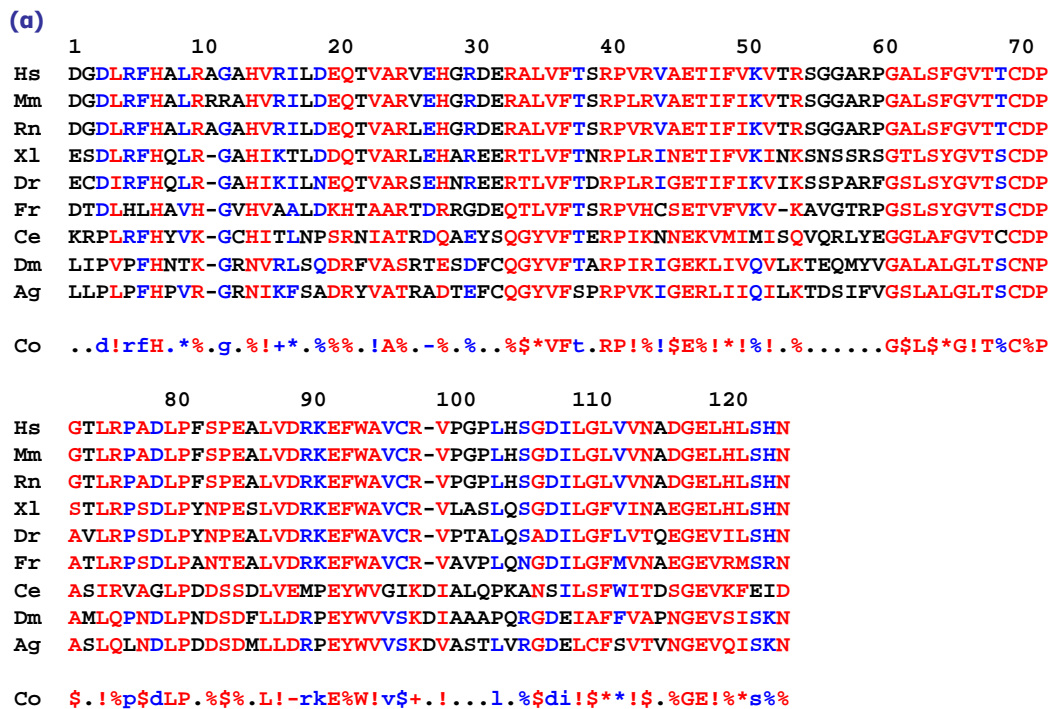
```

(β)

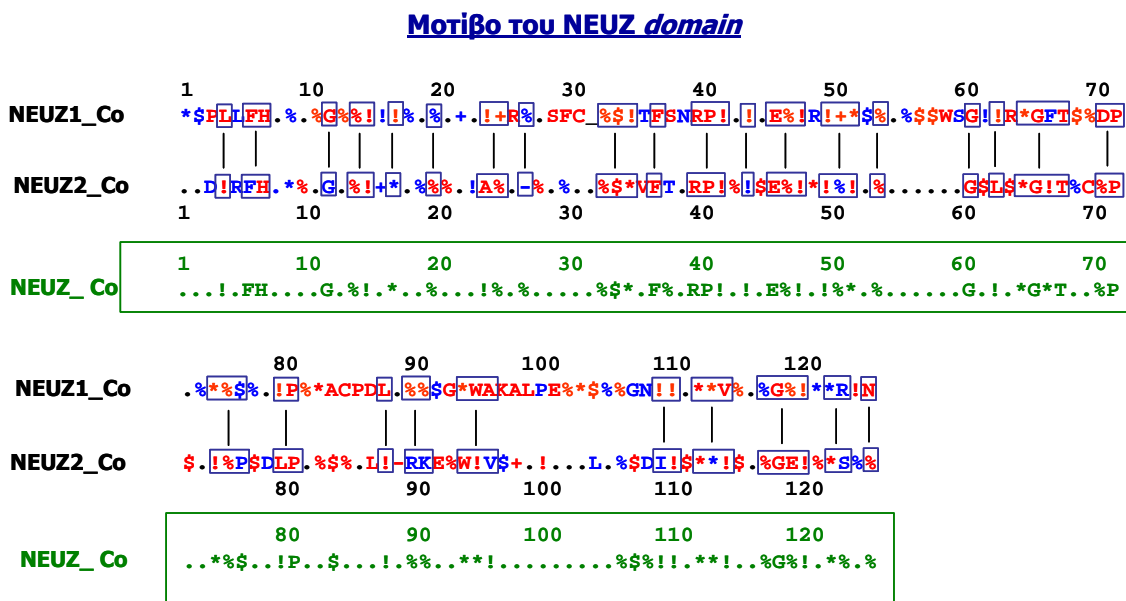


Εικ.83 Σύγκριση των περιοχών NEUZ1 των ορθόλογων πρωτεϊνών Neuralized. (α) Στοιχισμός αμινοξικών αλληλουχιών (Clustal W). Τα αμινοξέα που φαίνονται με κόκκινο είναι συντηρημένα σε όλους τους οργανισμούς που συγκρίθηκαν. Τα μπλε είναι συντηρημένα τουλάχιστον στο 70% των οργανισμών. (β) φυλογενετικό δέντρο (MEGA2.1). Στα κομβικά σημεία του δέντρου φαίνονται οι τιμές του 'Bootstrap test' (βλ. 'Υλικά και Μέθοδοι'). Στην αλληλουχία Co (*consensus*) ισχύουν οι συμβολισμοί: '*'= υδρόφοβα (L,I,Y,F,W,V,M,A), '!' = αλειφατικά (L,I,V,A,M), '\$' = μικρά (G,A,S,N,T,C,P), '%' = φορτισμένα (S,T,N,R,E,Q,H,D,K), '+' = θετικά φορτισμένα (K,R), '-' = αρνητικά φορτισμένα αμινοξέα (D,E).

NEUZ2



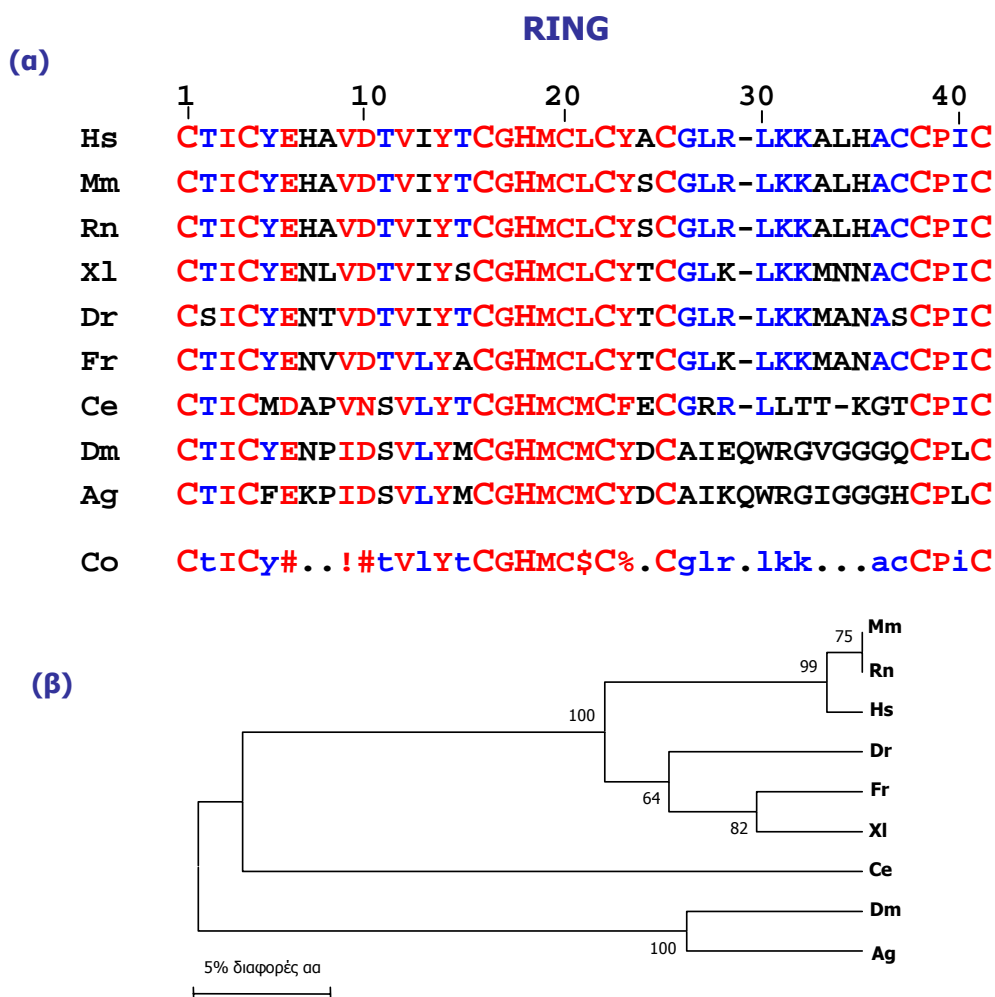
Εικ.84 Σύγκριση των περιοχών NEUZ2 των ορθόλογων πρωτεϊνών Neuralized. (α) Στοίχιση αμινοξικών αλληλουχιών (Clustal W). Τα αμινοξέα που φαίνονται με κόκκινο είναι συντηρημένα σε όλους τους οργανισμούς που συγκρίθηκαν. Τα μπλε είναι συντηρημένα τουλάχιστον στο 70% των οργανισμών. (β) φυλογενετικό δέντρο (MEGA2.1). Στα κομβικά σημεία του δέντρου φαίνονται οι τιμές του 'Bootstrap test' (βλ. 'Υλικά και Μέθοδοι'). Στην αλληλουχία Co (*consensus*) ισχύουν οι συμβολισμοί: '*'= υδρόφοβα (L,I,Y,F,W,V,M,A), '!' = αλειφατικά (L,I,V,A,M), '\$'= μικρά (G,A,S,N,T,C,P), '%'= φορτισμένα (S,T,N,R,E,Q,H,D,K), '+ '= θετικά φορτισμένα (K,R), '- '= αρνητικά φορτισμένα αμινοξέα (D,E).



Εικ.85. Η συντηρημένη αλληλουχία αμινοξέων (*consensus*) του NEUZ domain (NEUZ_Co, σε πλαίσιο), που προκύπτει από τη σύγκριση των συντηρημένων αμινοξέων των περιοχών NEUZ1 και NEUZ2 του Neutralized (αλληλουχίες NEUZ1_Co και NEUZ2_Co, αντίστοιχα). Στις αλληλουχίες Co (*consensus*) ισχύουν οι συμβολισμοί: '*'= υδρόφοβα (L,I,Y,F,W,V,M,A), '!'=αλειφατικά (L,I,V,A,M), '\$'=μικρά (G,A,S,N,T,C,P), '%'=φορτισμένα (S,T,N,R,E,Q,H,D,K), '+'=θετικά φορτισμένα (K,R), '-'=αρνητικά φορτισμένα αμινοξέα (D,E).

δ. Το RING είναι η δεύτερη πιο συντηρημένη περιοχή στα σπονδυλωτά

Το καρβοξυτελικό άκρο του Neutralized που περιέχει το μοτίβο RING εμφανίζει τα μεγαλύτερα ποσοστά συντήρησης μετά από αυτά της περιοχής NEUZ1 στις συγκρίσεις μεταξύ σπονδυλωτών (Εικ.82). Επιπλέον, το RING είναι η πιο συντηρημένη περιοχή της πρωτεΐνης Neutralized, στις συγκρίσεις μεταξύ σπονδυλωτών και ασπονδύλων. Εξελικτικά απομακρυσμένα είδη όπως ο νηματώδης και ο ποντικός απέχουν μόνο κατά 44% στην περιοχή του RING. Από τη σύγκριση των ορθόλογων πρωτεϊνών Neutralized προκύπτει για την περιοχή του RING το μοτίβο: **C₁ T I C₂ X₇ V (Y/L) Y X C₃ G H₁ M C (L/M) C₄ (Y/F) X C₅ X₉₋₁₁ C₆ P (I/L) C₇**, με βάση το οποίο επτά κατάλοιπα κυστεΐνης (C₁-C₇) και ένα κατάλοιπο ιστιδίνης (H₁) προσδένουν δύο ιόντα ψευδαργύρου (Εικ.86). Στο RING των ορθόλογων μορίων Neutralized, εμφανίζεται συντηρημένος ο αριθμός και το είδος των αμινοξέων μεταξύ των κυστεϊνών C₁-C₂ και C₆-C₇. Συνολικά, τα 20 από τα 40 αμινοξέα του RING εμφανίζονται συντηρημένα μεταξύ των ορθόλογων μορίων που συγκρίθηκαν. Σε αυτά περιλαμβάνονται τα επτά συντηρημένα κατάλοιπα κυστεΐνης και ιστιδίνης που συνιστούν το μοτίβο C3HC4, τα αμινοξέα Ile3, Val12, Tyr14, Gly17 και Pro38, καθώς και το τριπεπτίδιο Met-Cys-Leu στις θέσεις 19-21 (Εικ.86). Επιπλέον, υπάρχουν πέντε θέσεις στις οποίες παρατηρούνται δύο μόνο είδη αμινοξέων στα ορθόλογα μόρια που συγκρίνονται: η 6^η (Asp στον *C.elegans* και Glu στα υπόλοιπα είδη), η 9^η (Ile στα έντομα και Val στα υπόλοιπα είδη), η 10^η (Asn στον *C.elegans* και Asp στα υπόλοιπα είδη), η 21^η (Leu στα σπονδυλωτά και Met στα ασπόνδυλα) και η 23^η (Phe στον *C.elegans* και Tyr στα υπόλοιπα είδη). Μεγαλύτερη ποικιλότητα στην περιοχή RING μεταξύ των ορθόλογων μορίων Neu1 παρατηρείται ανάμεσα στην πέμπτη και έκτη συντηρημένη κυστεΐνη, και ιδιαίτερα μεταξύ σπονδυλωτών και ασπονδύλων. Οι θέσεις 35-37 του RING εμφανίζουν ποικιλότητα ακόμα και μεταξύ των σπονδυλωτών (Ala-Leu-His στα θηλαστικά, Met-Ala-Asn στα μη θηλαστικά).



Εικ.86 Σύγκριση των περιοχών RING των ορθόλογων πρωτεϊνών Neuralized. (α) Στοίχιση αμινοξικών αλληλουχιών (Clustal W). Τα αμινοξέα που φαίνονται με κόκκινο είναι συντηρημένα σε όλους τους οργανισμούς που συγκρίθηκαν. Τα μπλε είναι συντηρημένα τουλάχιστον στο 50% των οργανισμών. Τα αμινοξέα Cys/His που συνιστούν το συντηρημένο μοτίβο C3HC4 του RING φαίνονται με μεγαλύτερους χαρακτήρες. (β) Εξελικτικό δέντρο (MEGA2.1). Στα κομβικά σημεία του δέντρου φαίνονται οι τιμές του 'Bootstrap test' (βλ. 'Υλικά και Μέθοδοι'). Στην αλληλουχία Co (*consensus*) ισχύουν οι συμβολισμοί: '!' = αμινοξέα I ή V, '\$' = αμινοξέα L ή M, '%' = αμινοξέα F ή Y και '#' = αμινοξέα N, D, Q ή E.

Τα παραπάνω συνηγορούν στο ότι η λειτουργία του RING στο μόριο Neu1 είναι συντηρημένη και πιθανόν κοινή μεταξύ σπονδυλωτών και ασπονδύλων.

4. Πρωτεΐνες που αλληλεπιδρούν με το Neu1

Πρωτεΐνες που φέρουν RING συχνά συμμετέχουν σε ετερογενείς βιολογικές λειτουργίες αλληλεπιδρώντας με μεγάλη ποικιλία πρωτεϊνών, τις οποίες συνδέουν ως 'adaptors' με πολυπρωτεϊνικά σύμπλοκα [Borden, 2000/ Jackson *et al*, 2000/ Joazeiro & Weissman, 2000]. Η ταυτοποίηση των πρωτεϊνών που αλληλεπιδρούν με μια RING *finger*-πρωτεΐνη προσφέρει σημαντικές ενδείξεις για το βιολογικό της ρόλο. Στην εργασία αυτή εντοπίσαμε συνολικά 9 πρωτεΐνες, οι οποίες αλληλεπιδρούν με το Neu1 σε κύτταρα σακχαρομύκητα, σε *in vitro* συνθήκες και σε κύτταρα θηλαστικών και ομαδοποιούνται σε τρεις βασικές καταγορίες.

➤ **Πρωτεΐνες που αλληλεπιδρούν με τον κυτταροσκελετό ή αποτελούν συστατικά ινιδίων**

Ένα σημαντικό ποσοστό των πρωτεϊνών που αλληλεπιδρούν με το Neu1 ανήκουν σε αυτή την κατηγορία. Είναι αξιοσημείωτο το ότι η πρωτεΐνη Numa1 αντιστοιχεί στο 40% των επιλεγμένων κλώνων cDNA του πειράματος δύο υβριδίων. Τρεις από τις πρωτεΐνες αυτής της κατηγορίας (Numa1, Enp1, Ankycorbin) έχουν μεγάλο μοριακό βάρος (>100KD) και αλληλεπιδρούν με μικροσωληνίσκους ή ενδιάμεσα ινίδια, ενώ οι πρωτεΐνες μικρότερου μοριακού βάρους Dlc-1 (8KD) και Gas8 (50KD) αλληλεπιδρούν με μεγαλομοριακά σύμπλοκα πρωτεϊνών-κινητήρων. Χαρακτηριστική είναι η παρουσία της ελικοειδούς (*coiled coil*) διαμόρφωσης, η οποία συναντάται στις πρωτεΐνες Numa1, Enp1 και Ankycorbin, εκτείνεται στο μεγαλύτερο μέρος αυτών των μορίων και περιέχει τη θέση αλληλεπίδρασης με το Neu1 (Εικ.87). Πρόκειται για μια υπερελικωμένη δομή α-ελίκων με τη χαρακτηριστική επανάληψη 7 αμινοξέων, από τα οποία το 1^ο και το 4^ο είναι υδρόφοβα. Η δομή αυτή συναντάται συχνά σε πρωτεΐνες του κυτταροσκελετού που σχηματίζουν ινίδια και συμμετέχουν σε υπερμοριακές δομές, ευνοώντας τις αλληλεπιδράσεις τους με άλλες πρωτεΐνες ή τον ολιγομερισμό τους. Οι αλληλεπιδράσεις της στηρίζονται κυρίως στην παρουσία φορτισμένων αμινοξέων στην εξωτερική επιφάνεια της έλικας [Burkhard *et al*, 2001].

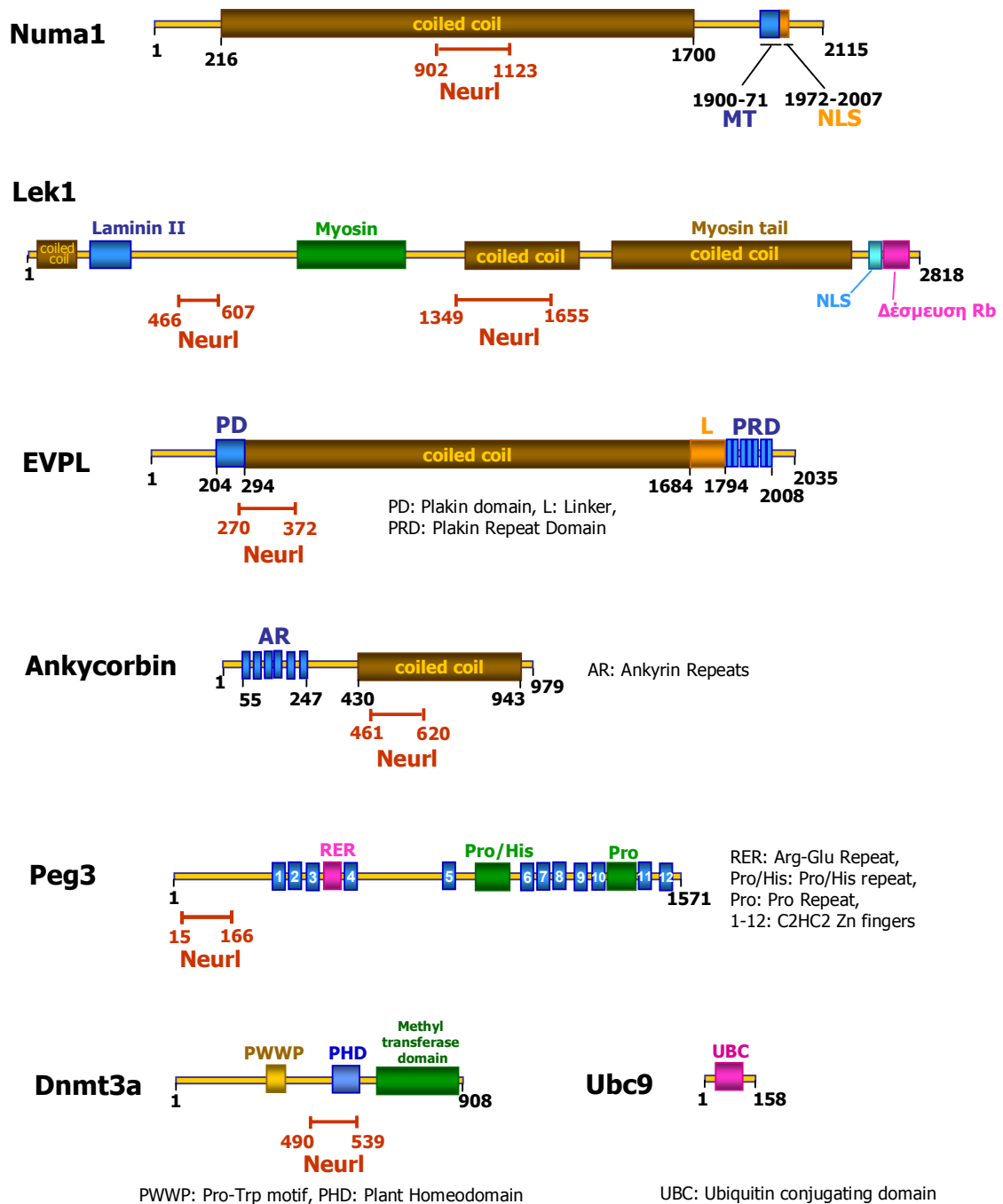
Η σχέση της πρωτεΐνης Neu1 με τον κυτταροσκελετό δεν είχε έως τώρα διερευνηθεί. Όμως, τα δεδομένα που έχουμε για τη λειτουργία του στη δροσόφιλα, όπου επάγει την ενδοκύττωση του Delta, και τα βιοχημικά πειράματα σε κύτταρα θηλαστικών, όπου ανιχνεύεται στην εξωτερική επιφάνεια μεμβρανικών κυστιδίων, δεν αποκλείουν την αλληλεπίδρασή του με πρωτεΐνες του κυτταροσκελετού. Μέσω των πρωτεϊνών-κινητήρων και των ινιδίων του κυτταροσκελετού είναι πιθανόν να διευκολύνεται η μετακίνηση του Neu1 από την περιπυρηνική περιοχή μέχρι και την πλασματική μεμβράνη και αντίστροφα, γεγονός που δικαιολογεί και τον εντοπισμό του σε διάφορες περιοχές του κυττάρου. Η πληροφορία που έχει συσσωρευτεί τα τελευταία χρόνια για τη λειτουργία κυρίως των πρωτεϊνών Numa1 και Dlc-1 συσχετίζει έμμεσα το Neu1 και με άλλες λειτουργίες εκτός της ενδοκύττωσης, διευρύνοντας το πεδίο δράσης αυτού του μορίου και παρέχοντας σημαντικές ενδείξεις για τη μοριακή βάση του φαινοτύπου απενεργοποίησης της λειτουργίας του και τον βιολογικό του ρόλο στα θηλαστικά. Παρακάτω αναφέρονται τα στοιχεία που υπάρχουν για κάθε μία από τις πρωτεΐνες αυτής της ομάδας και ο συσχετισμός τους με το Neu1.

α. Numa1

Η Numa1 (*Nuclear mitotic apparatus protein 1*) είναι μεγάλου μοριακού βάρους πρωτεΐνη (~230KD) με μια κεντρική περιοχή ελικοειδούς (*coiled-coil*) διαμόρφωσης (Εικ.87), στην οποία οφείλει την ικανότητα ολιγομερισμού και την ενσωμάτωσή της στα ινίδια του πυρηνοσκελετού [Yang *et al*, 1992/ Tang *et al*, 1994/ Harborth *et al*, 1999]. Όπως φαίνεται από τη χαρτογράφηση των επιλεγμένων κλώνων του πειράματος δύο υβριδίων, μέρος αυτής της περιοχής ευθύνεται και για την αλληλεπίδρασή της με το Neu1.

Η αλληλεπίδραση των μορίων Numa1 και Neu1 υποστηρίζεται εκτός από τις *in vitro* δοκιμές και τα βιοχημικά πειράματα κυτταρικών εκχυλισμάτων (βλ. Αποτελέσματα) και από την επικάλυψη στο πρότυπο

έκφρασης των αντίστοιχων γονιδίων. Το αναπτυξιακό πρότυπο έκφρασης των γονιδίων *Numa1* και *Neur1* του ποικίλου επικαλύπτεται σε αρκετά σημεία, στο νευρικό σύστημα, σε αναπτυσσόμενα αισθητήρια όργανα και σε επιθήλια εσωτερικών οργάνων, όπως έχει δείξει συγκριτική τους μελέτη [Μπαζοπούλου ΜΔ, Διπλωματική Εργασία, Παν/μιο Κρήτης, 2002], υποστηρίζοντας τη σημασία αυτής της αλληλεπίδρασης στην ανάπτυξη των



Εικ.87 Πρωτεΐνες που αλληλεπιδρούν με το Neur1. Απεικονίζονται οι λειτουργικές περιοχές (*domains*) των πρωτεϊνών και η περιοχή αλληλεπίδρασης με το Neur1 (κόκκινο). Οι αριθμοί αντιστοιχούν σε αμινοξέα με βάση τις καταχωρημένες αμινοξικές αλληλουχίες. NLS: σήμα πυρηνικού εντοπισμού, MT: θέση πρόσδεσης σε μικροσωληνίσκου, Rb: πρωτεΐνη ρετινοβλαστώματος.

συγκεκριμένων ιστών. Είναι αξιοσημείωτο το ότι μετάγραφα και των δύο γονιδίων ανιχνεύονται κυρίως σε επιθηλιακά κύτταρα, σε συμφωνία με προηγούμενες αναφορές επιλεκτικής έκφρασης της *Numa1* σε διαιρού-

μενα κύτταρα μεγάλης ποικιλίας ιστών του ανθρώπου και του ποντικού [Taimen *et al*, 2000/ Tang *et al*, 2004].

Η λειτουργία που προτείνεται για τη Numa1 προκύπτει από πειράματα απενεργοποίησής της με μικροενέσεις αντισωμάτων α-Numa και υπερέκφρασης ελλειμματικών μορίων της σε κυτταρικές σειρές. Σύμφωνα με αυτά, η Numa1 λειτουργεί στα αρχικά στάδια της πρόφασης, πριν τη διάλυση του πυρηνικού φακέλου, όταν μεταφέρεται μέσω του δικτύου των μικροσωληνίσκων και του συμπλόκου δυνείνης-δυνακτίνης στους μελλοντικούς πόλους της μιτωτικής ατράκτου, όπου προσδένεται στα (-) άκρα των μικροσωληνίσκων του κεντροσώματος [Merdes *et al*, 1996,2000/ Blagden & Glover, 2003]. Όπως παρατηρήσαμε, το Neur1 συνεντοπίζεται στην περιοχή του κεντροσώματος με τη Numa1, καθώς και με τις πρωτεΐνες Dlc-1 και β-tubulin. Ενδεχομένως, ο βιολογικός ρόλος της αλληλεπίδρασης Neur1 και Numa1 να σχετίζεται με τη λειτουργία της Numa1 στη φάση αυτή του κυτταρικού κύκλου. Επιπλέον, θα ήταν ενδιαφέρον να διερευνηθεί και ο πιθανός ρυθμιστικός ρόλος του Neur1 κατά τη μετακίνηση της Numa1 προς το κεντρόσωμα ή γενικά για την οργάνωση των μικροσωληνίσκων γύρω από το κεντρόσωμα, με δεδομένη την ικανότητά του να αλληλεπιδρά άμεσα τόσο με τη Numa1 όσο και με την υπομονάδα Dlc-1 της κυτταροπλασματικής δυνείνης, σύμπλοκο που λειτουργεί ως 'κινητήρας' για τη μεταφορά της Numa1 προς τα (-) άκρα των μικροσωληνίσκων [Merdes *et al*, 2000].

Επίσης, η παρουσία της Numa1 έχει συσχετιστεί σε ορισμένες κυτταρικές σειρές θηλαστικών με τη σταθεροποίηση της μιτωτικής ατράκτου [Compton & Cleveland, 1993/ Gaglio *et al*, 1995/ Zeng, 2000]. Συγκεκριμένα, οι Du *et al* (2002) παρατηρούν ότι ένα μικρό τμήμα του καρβοξυτελικού άκρου της Numa1 (αμινοξέα 1900-1971) είναι αρκετό για την πρόσδεσή της στους μικροσωληνίσκους και τη σταθεροποίηση της μιτωτικής ατράκτου. Επιπλέον, διαπιστώνουν το ρυθμιστικό ρόλο της πρωτεΐνης LGN (ομόλογης της δροσοφιλικής Pins, *Partner of inscutable*), στη λειτουργία της Numa1. Η LGN, αλληλεπιδρώντας με τη Numa1, ρυθμίζει τη συνεχή σύνδεση και αποσύνδεσή της από τα δυναμικά αποπολυμεριζόμενα (-) άκρα των μικροσωληνίσκων, διαδικασία απαραίτητη για την οργάνωση της μιτωτικής ατράκτου. Επειδή τα σχετικά επίπεδα ελεύθερων μορίων Numa1 και LGN είναι κρίσιμα για τη λειτουργία τους στους πόλους της ατράκτου και με δεδομένη την αλληλεπίδραση της Numa1 με το Neur1, θα ήταν ενδιαφέρον να διερευνηθεί αν το Neur1 επηρεάζει αυτή την τοπική στοιχειομετρία μορίων Numa1, LGN και μικροσωληνίσκων.

Η αλληλεπίδραση των LGN και Numa1 με την προσδεδεμένη σε GDP μορφή της πρωτεΐνης Gai και με άλλους παράγοντες του κυτταρικού φλοιού και πρωτεΐνες-κινητήρες (πιθανόν δυνείνη), φαίνεται να είναι σημαντική και για τον καθορισμό του προσανατολισμού της μιτωτικής ατράκτου και τις ασύμμετρες κυτταρικές διαιρέσεις στα κύτταρα των σπονδυλωτών, σύμφωνα με τα πειράματα των Du & Macara (2004). Η ασύμμετρη κυτταρική διαίρεση είναι σημαντικός μηχανισμός καθορισμού της κυτταρικής τύχης, συντηρημένος από τα ασπόνδυλα έως και τα θηλαστικά [Knoblich, 2001]. Ήδη για τη δροσοφιλική πρωτεΐνη neur είναι γνωστό ότι η επιλεκτική παρουσία της στο μελλοντικό νευρικό κύτταρο είναι αποτέλεσμα της ασύμμετρης κατανομής της στον διαιρούμενο πρόγονο αυτού του κυττάρου, όπου συνεντοπίζεται με την επίσης ασύμμε-

τρα κατανεμημένη πρωτεΐνη numb [Le Borgne & Schweisguth, 2003a] (βλ. Εισαγωγή). Ανιχνεύοντας το *Neur1* σε κυτταρικές σειρές θηλαστικών δεν παρατηρήσαμε ασύμμετρη κατανομή του κατά την κυτταρική διαίρεση. Αυτό όμως πιθανόν να μην αντιπροσωπεύει την *in vivo* κατάσταση, λόγω της πιθανής έλλειψης κυτταροειδικών παραγόντων στις κυτταρικές σειρές. Το ίδιο ισχύει και στην περίπτωση του LGN, του οποίου η ασύμμετρη κατανομή επηρεάζεται από τη φάση του κυτταρικού κύκλου και τις συνθήκες καλλιέργειας και παρατηρείται μόνο σε ένα ποσοστό πρόδρομων νευρικών κυττάρων [Du *et al*, 2002]. Η περαιτέρω διερεύνηση της συμμετοχής του *Neur1* σε ασύμμετρες κυτταρικές διαιρέσεις κρίνεται απαραίτητη με βάση τα παραπάνω αλλά και λόγω της αποδεδειγμένης ασύμμετρης κατανομής της πρωτεΐνης Numb των θηλαστικών και των πτηνών κατά τη διαφοροποίηση νευρώνων του φλοιού του πρόσθιου εγκεφάλου [Zhong *et al*, 1996, 2000/ Wakamatsu *et al*, 1999], καθώς και άλλων κυτταρικών τύπων που διαφοροποιούνται μέσω σηματοδότησης από τον υποδοχέα Notch [Cayouette *et al*, 2001/ Cayouette M & Raff, 2002] και στους οποίους παρατηρείται έκφραση του *Neur1* σε πρώιμα αναπτυξιακά στάδια.

Στο τέλος της ανάφασης και κατά την τελόφαση, η *Numa1* παύει να είναι προσδεμένη στους πόλους της ατράκτου και φαίνεται ότι ρυθμίζει τον επανασχηματισμό του πυρηνικού φακέλου [Price & Pettijohn, 1986]. Αυτό μπορεί να είναι έμμεσο αποτέλεσμα της οργάνωσης των χρωμοσωμάτων μέσω των πόλων της ατράκτου για την ένθεση συστατικών του πυρηνικού φακέλου γύρω από αυτά ή πιθανόν η *Numa1*, ως συστατικό του πυρηνοσκελετού, να οργανώνει τον επανασχηματισμό του στρατολογώντας και άλλες πρωτεΐνες γύρω από αυτήν [Cleveland, 1995]. Στη φάση αυτή, και το *Neur1* ανιχνεύεται μακριά από τους πόλους της ατράκτου και συνεντοπίζεται μερικώς με τη *Numa1*. Η αλληλεπίδραση του *Neur1* με τη *Numa1* και πιθανόν και με άλλες πυρηνικές πρωτεΐνες σε αυτή τη φάση, πιθανόν να είναι σημαντική για την είσοδο μέρους των μορίων του στον πυρήνα.

Η παρουσία της *Numa1* στους πόλους της ατράκτου των διαιρούμενων πρόδρομων σπερματοκυττάρων και η απουσία της από τα διαφοροποιημένα σπερματοζωάρια ενδεχομένως σημαίνει ότι έχει παρόμοιο ρόλο και κατά τη μείωση [Taimen *et al*, 2004]. Όμως, η παρουσία της πρωτεΐνης *Neur1* τόσο στα πρόδρομα όσο και στα διαφοροποιημένα σπερματοζωάρια, δείχνει ότι ο ρόλος της στα συγκεκριμένα κύτταρα συνεχίζεται και μετά τη διαφοροποίηση και επομένως δεν εξαρτάται αποκλειστικά από την παρουσία της *Numa1*. Παρά τον σημαντικό της ρόλο στην κυτταρική διαίρεση, η *Numa1* απουσιάζει από αρκετούς κυτταρικούς τύπους των θηλαστικών, ιδιαίτερα από τελικά διαφοροποιημένα κύτταρα, ενώ στους ιστούς των ενηλίκων ανιχνεύεται κυρίως σε επιθήλια [Taimen *et al*, 2000]. Αυτό συμφωνεί και με την πιο περιορισμένη έκφραση του *Neur1* στο ενήλικο ποντίκι σε σχέση με το έμβρυο [Pavliopoulos *et al*, 2002]. Επιπλέον, έχει παρατηρηθεί απώλεια της *Numa1* λόγω πρωτεόλυσης μετά από συνεχιζόμενη καλλιέργεια ή επαγωγή της διαφοροποίησης, χωρίς να έχει εξακριβωθεί πλήρως η μοριακή βάση αυτού του φαινομένου [Gueth-Hallonet *et al*, 1997/ Taimen *et al*, 2000/ Taimen & Kallajoki, 2003]. Η αλληλεπίδραση του *Neur1* με μόρια όπως η *Peg3*, που σχετίζονται με τον αποπτωτικό μηχανισμό, αλλά και οι ενδείξεις που υπάρχουν για τη σχέση του με τον μηχανισμό ουβικουιτίνιλίωσης και πρωτεόλυσης μέσω πρωτεασώματος [Εισαγωγή και Ε. Κουτέλου, αδημοσί-

ευτα αποτελέσματα] το καθιστούν ένα ενδιαφέρον υποψήφιο για τον έλεγχο της πιθανής συμμετοχής του στη μετα-μεταφραστική ρύθμιση της Numa1.

Τέλος, σε μελέτες των σταδίων διαφοροποίησης νευρικών κυττάρων θηλαστικών, η πρωτεΐνη Numa1 έχει ανιχνευθεί στο κυτταρόπλασμα πριν τα τελικά στάδια διαφοροποίησης, σε κύτταρα που έχουν σταματήσει να διαιρούνται, όπου φαίνεται να συμμετέχει στην οργάνωση των μικροσωληνίσκων για τον σχηματισμό των δενδριτών του νευρικού κυττάρου [Ferhat *et al*, 1998]. Ο ακριβής της ρόλος στη ρύθμιση των αρχικών σταδίων σχηματισμού κυτταρικών προεκβολών από μικροσωληνίσκους, όπως είναι ο άξονας των νευρώνων και το αξόνημα των σπερματοζωαρίων, κυτταρικοί τύποι που επηρεάζονται στην έλλειψη λειτουργικού Neur1 [Vollrath *et al*, 2001/ Ruan *et al*, 2001], δεν έχει διερευνηθεί. Ο περαιτέρω έλεγχος του βιολογικού ρόλου αυτής της αλληλεπίδρασης μπορεί να αποκαλύψει ενδιαφέρουσες πτυχές της λειτουργίας του Neur1 στην κυτταρική διαίρεση, στον καθορισμό της αναπτυξιακής τύχης και στη διαφοροποίηση κυττάρων, κυρίως κατά τη νευρογένεση και τη σπερματογένεση των θηλαστικών.

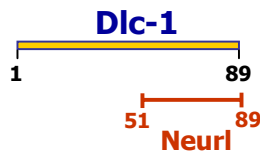
β. Dlc-1

Η Dlc-1 (ή LC8, DLC8, hdlc1, PIN) είναι μία από τις μικρές υπομονάδες (*light chains*) της κυτταροπλασματικής δυνεινης, της αξονημικής δυνεινης και της μυοσίνης V. Είναι μια πρωτεΐνη 89 αμινοξέων, πολύ συντηρημένη από την *Chlamydomonas* μέχρι τον άνθρωπο [Wilson *et al*, 2001]. Οι υπομονάδες μικρού μοριακού βάρους όπως η Dlc-1 (8KD), προσφέρουν στα σύμπλοκα πρωτεϊνών-κινητήρων εξειδίκευση για πρόσδεση σε μεγάλο αριθμό μορίων, τα οποία μεταφέρουν προς συγκεκριμένη κατεύθυνση κατά μήκος ινιδίων του κυτταροσκελετού. Ειδικότερα, η δυνεΐνη κινείται μέσω δύο μεγάλου μοριακού βάρους (~500KD) αλυσίδων (DHC, *Dynein Heavy Chains*), υδρολύοντας ATP, κατά μήκος των μικροσωληνίσκων με κατεύθυνση προς τα (-) άκρα τους. Διαφορετικές ισομορφές των υπομονάδων μικρού (8-22KD, DLCs), μικρού-ενδιάμεσου (53-59KD, DLICs) και ενδιάμεσου (70-74KD, DICs) μοριακού βάρους δημιουργούν την ποικιλομορφία μεταξύ των διαφορετικών ολοενζύμων δυνεινης που συναντώνται στα ευκαρυωτικά κύτταρα (Εικ.88). Με βάση το πείραμα των δύο υβριδίων και τις *in vitro* δοκιμές αλληλεπίδρασης, το C-τελικό τμήμα της Dlc-1 (αμινοξέα 51-89) αρκεί για την αλληλεπίδραση με την πρωτεΐνη Neur1. Η αλληλεπίδραση των δύο πρωτεϊνών φαίνεται να ενισχύεται στα κύτταρα των θηλαστικών, όπου χρησιμοποιήσαμε την ανασυνδυασμένη πλήρους μήκους Dlc-1 στα πειράματα συν-ανοσοκατακρήμνισης με το GFP-Neur1.

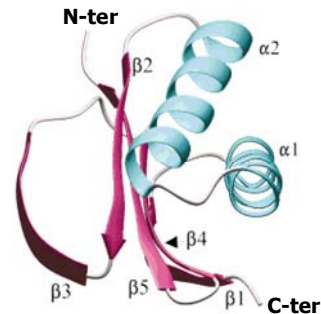
Ο όρος 'κυτταροπλασματική' δεν χαρακτηρίζει όλες τις υπομονάδες της δυνεινης εφόσον αρκετές από αυτές, όπως η Dlc-1, εντοπίζονται τόσο στο κυτταρόπλασμα όσο και στον πυρήνα κατά τη μεσόφαση. Από τα πειράματα ανοσοανίχνευσης των ενδογενών μορίων Neur1 και Dlc-1, διαπιστώνουμε συνεντοπισμό των δύο πρωτεϊνών κυρίως γύρω από τον πυρήνα κατά τη μεσόφαση. Η αλληλεπίδραση του Neur1 με την κυτταροπλασματική δυνεΐνη πιθανόν σχετίζεται με τον ρόλο του στην ενδοκύτωση του Delta [Pavlopoulos *et al*, 2001] ή και άλλων μεμβρανικών πρωτεϊνών. Η κυτταροπλασματική δυνεΐνη, σχηματίζοντας σύμπλοκο με την ενεργοποιητική της πρωτεΐνη δυνακτίνη, αποτελεί τον κυρίαρχο μηχανισμό μετακίνησης μεμβρανικών κυστιδίων και οργανιδίων προς την περιφέρεια του κυττάρου, διαδικασία απαραίτητη για την ανακύκλωση

μεμβρανών και μεμβρανικών υποδοχών και τη βιογένεση πολλών οργανιδίων [Oda *et al*, 1995/ Hirokawa, 1998]. Επιπλέον, συμμετέχει στον καθορισμό της θέσης του πυρήνα στα κύτταρα των μυκήτων, αλλά και σε μεταναστευτικούς νευρώνες του εγκεφάλου των θηλαστικών [Bloom, 2001], λειτουργία που σχετίζεται και με τον καθορισμό του ισημερινού επιπέδου της μιτωτικής ατράκτου.

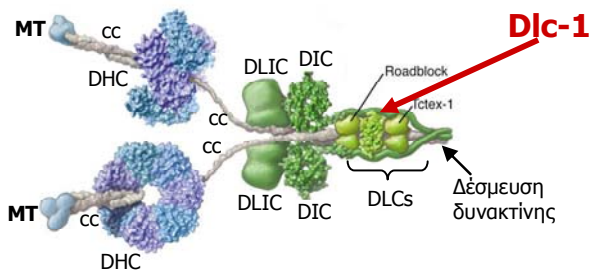
(α) 'Θήραμα' του Neur1



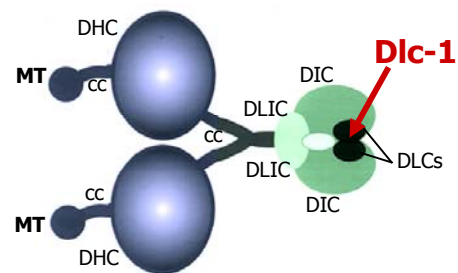
(β) Μονομερές Dlc-1



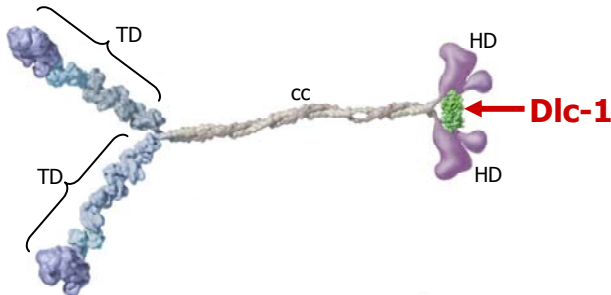
(γ) Κυτταροπλασματική δυνεΐνη



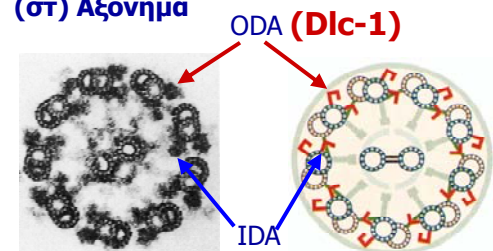
(δ) Αξονημική δυνεΐνη



(ε) Μυοσίνη V



(στ) Αξόνημα



Εικ.88 Η πρωτεΐνη Dlc-1. **α)** Η περιοχή αλληλεπίδρασης με το Neur1, **β)** Δευτεροταγής δομή του μονομερούς Dlc-1, **(γ-ε)** Κυτταροπλασματική δυνεΐνη, αξονημική δυνεΐνη και μυοσίνη V. Η θέση της Dlc-1 φαίνεται με κόκκινο βέλος (DHC: Dynein Heavy Chain, DLIC: Dynein Light Intermediate Chain, DIC: Dynein Intermediate Chain, DLCs: Dynein Light Chains, cc: coiled coil, MT: θέση πρόσδεσης στους μικροσωληνίσκους, TD: Tail Domain, HD: Head Domain) **στ)** Αξόνημα (δομή 9+2), εικόνα ηλεκτρονικής μικροσκοπίας από εγκάρσια τομή μαστιγίου και γραφική αναπαράσταση. ODA: Outer Dynein Arm (περιέχει την υπομονάδα Dlc-1) και IDA: Inner Dynein Arm.

Αλληλεπιδρώντας με την Dlc-1, όπως και με τη Numa1, το Neur1 ενδεχομένως συμμετέχει στον μηχανισμό της κυτταρικής διαίρεσης. Για την κυτταροπλασματική δυνεΐνη του κυτταρικού φλοιού (*cell cortex*) είναι γνωστό ότι συμμετέχει στον καθορισμό της θέσης της μιτωτικής ατράκτου κατά την ασύμμετρη κυτταρική διαίρεση στους ανώτερους ευκαρυωτικούς οργανισμούς [Bloom, 2001/ Faulkner *et al*, 2000], ενώ ο ρόλος της στα μετάζωα περιλαμβάνει επιπλέον τη μετακίνηση ενδιάμεσων ινιδίων και πρωτεϊνών του κε-

ντροσώματος (πχ της πρωτεΐνης Numa1), τη διάλυση του πυρηνικού φακέλου κατά την κυτταρική διαίρεση, τη δημιουργία της μιτωτικής ατράκτου και τη λειτουργία των κινητοχώρων [Karki & Holzbaaur, 1999]. Πιθανόν η πρωτεΐνη *Neur1* να παίζει ρόλο ενδιάμεσου 'adaptor' και να διευκολύνει την αλληλεπίδραση της Numa1 με την κυτταροπλασματική δυνεΐνη, απαραίτητη για τη μεταφορά της στο κεντρόσωμα. Ο συνεντοπισμός και των τριών πρωτεϊνών που παρατηρήσαμε στο κεντρόσωμα ενδεχομένως σχετίζεται με την ίδια διαδικασία.

Επιπλέον, διαπιστώσαμε συνεντοπισμό των πρωτεϊνών *Neur1* και *Dlc-1* στο μαστίγιο σπερματοζωαρίων. Η *Dlc-1* είναι υπομονάδα τόσο της κυτταροπλασματικής όσο και της αξονημικής δυνεΐνης. Η τελευταία λειτουργεί στο αξόνημα των μαστιγίων ή των βλεφαρίδων μονοκύτταρων (*Chlamydomonas*) ή βλεφαριδοφόρων και μαστιγοφόρων κυττάρων πολυκύτταρων οργανισμών (σπερματοζωάρια). Η αξονημική δυνεΐνη αποτελείται από συγκεκριμένο συνδυασμό υπομονάδων DHC-DIC-DLIC-DLC, περιλαμβάνοντας και ορισμένες κοινές υπομονάδες με την κυτταροπλασματική δυνεΐνη [King *et al*, 1996]. Οι μηχανισμοί ρύθμισης της συγκρότησης του ολοενζύμου από τις επιμέρους υπομονάδες δεν έχουν διαλευκανθεί πλήρως, όμως σημαντικός στη διαδικασία είναι ο ρόλος της υπομονάδας *Dlc-1* [King, 2000/2003]. Η αλληλεπίδραση του *Neur1* με την αξονημική *Dlc-1* πιθανόν εξηγεί τη διαταραγμένη δομή του αξονήματος των σπερματοζωαρίων των *Neur1*-/- ζώων που αναφέρονται από τους Vollrath *et al* (2001). Η αξονημική δυνεΐνη προσδένεται στους μικροσωληνίσκους του εξωτερικού δακτυλίου του αξονήματος (Εικ.88). και διευκολύνει της κινήσεις των μικροσωληνίσκων της δομής (9+2). Εκτός από την αξονημική δυνεΐνη που αποτελεί δομικό συστατικό του αξονήματος, τα μαστίγια και οι βλεφαρίδες κυττάρων θηλαστικών έχουν επίσης κυτταροπλασματική δυνεΐνη και κινεσίνη, οι οποίες κινούνται κατά μήκος του αξονήματος μεταφέροντας 'φορτία' (κίνηση IFT, *Intraflagellar Transport*). Συγκεκριμένα, η κυτταροπλασματική δυνεΐνη μεταφέρει 'φορτία' με κατεύθυνση προς το κυτταρικό σώμα (*retrograde*), ενώ η κινεσίνη προς το άκρο του μαστιγίου (*anterograde*). Τα 'φορτία' αυτά μπορεί να είναι μεμβρανικά κυστίδια ή πρωτεΐνες (κυρίως *tubulin*), των οποίων η συνεχής μετακίνηση κατά μήκος του αξονήματος είναι απαραίτητη για την επιμήκυνσή του και την κινητικότητα του μαστιγίου. Με δεδομένο τον εντοπισμό του *Neur1* στην εξωτερική επιφάνεια μεμβρανικών κυστιδίων κυττάρων θηλαστικών και το ρόλο του στην μετακίνησή τους κατά την ενδοκύτωση [Pavlopoulos *et al*, 2001] είναι πιθανόν το *Neur1* να δρα με παρόμοιο τρόπο και στο αξόνημα. Μέσω της αλληλεπίδρασής του με την *Dlc-1* πιθανόν να διευκολύνει την κίνηση ορισμένων μεμβρανικών κυστιδίων. Η απουσία του *Neur1* ενδεχομένως να επηρεάζει τον ρυθμό πολυμερισμού των μικροσωληνίσκων του αξονήματος, με αποτέλεσμα να συντίθενται μικρότεροι μικροσωληνίσκοι και σε ορισμένες εγκάρσιες τομές των μαστιγίων να παρατηρούνται οι αποκλίσεις από τη δομή (9+2), όπως διαπιστώνουν οι Vollrath *et al* (2001). Ο ίδιος μηχανισμός ισχύει και για τη μετακίνηση 'φορτίων' (πχ μεμβρανικών κυστιδίων νευροδιαβιβαστών) κατά μήκος του άξονα των νευρικών κυττάρων. Ο ρόλος του *Neur1* σε αυτό το σύστημα δεν έχει μελετηθεί σε μοριακό επίπεδο, όμως οι Ruan *et al* (2001) διαπιστώνουν ότι η απενεργοποίηση του *Neur1* στα ποντίκια σχετίζεται με την αδυναμία τους να εκτελέσουν συγκεκριμένες λειτουργίες του νευρικού συστήματος, όπως η διάκριση των οσμών και ο συντονισμός των κινήσεων. Επομένως, οι δύο φαινότυποι της έλλειψης λειτουργίας του *Neur1* που έχουν αναφερθεί, δεν αποκλείουν τη δράση του μέσω της *Dlc-1*. Το γεγονός ότι και οι δύο είναι πολύ πιο 'ήπιοι' σε σχέση με αυτόν της έλλειψης της *Dlc-1*

[Dick *et al*, 1996], υποδηλώνει ότι υπάρχουν και άλλα μόρια που επιτελούν τις ίδιες ή ανάλογες λειτουργίες (πχ την κίνηση IFT) αλληλεπιδρώντας με την Dlc-1 ή και με άλλες υπομονάδες της δυνεΐνης.

Η Dlc-1 των σπονδυλωτών έχει ανιχνευθεί και ως υπομονάδα της μυοσίνης V. Και αυτή η πρωτεΐνη-κινητήρας χρησιμοποιείται στα κύτταρα, όπως και η δυνεΐνη, για μετακινήσεις οργανιδίων στο δίκτυο ινιδίων ακτίνης. Παραδείγματα αποδεδειγμένου ρόλου της μυοσίνης V σε ενδοκυττάρια μεταφορά περιλαμβάνουν την κίνηση του ενδοπλασματικού δικτύου στο κυτταρόπλασμα νευρικών αξόνων κεφαλόποδων [Tabb *et al*, 1998] και την κατανομή των μελανοφόρων κοκκίων στα μελανοκύτταρα του ποντικού [Reck-Peterson *et al*, 2000]. Εκτός των ινιδίων ακτίνης, έχουν διαπιστωθεί αλληλεπιδράσεις της μυοσίνης V και με το δίκτυο των μικροσωληνίσκων στο κυτταρόπλασμα νευρικών κυττάρων, στο κεντρόσωμα και στους πόλους της ατράκτου διαιρούμενων κυττάρων [Evans *et al*, 1997/ Espreafico *et al*, 1998]. Μελέτες μετακίνησης κυστιδίων του ενδοπλασματικού δικτύου μέσω της μυοσίνης V προτείνουν ότι η αλληλεπίδρασή της με τους μικροσωληνίσκους εξασφαλίζει τη μεταφορά της μέσω της κυτταροπλασματικής δυνεΐνης από το κέντρο του κυττάρου μέχρι τον πλούσιο σε ινίδια ακτίνης κυτταρικό φλοιό, όπου λειτουργεί η μυοσίνη για μετακίνηση ή αγκυροβόληση των κυστιδίων [Tabb *et al*, 1998]. Η περαιτέρω διερεύνηση της λειτουργικής σχέσης του Neur1 με αυτό το σύμπλοκο ενδεχομένως να αποκαλύψει και νέες λειτουργίες για το Neur1, πέρα από το ρόλο του στην ενδοκύτωση. Θα ήταν ιδιαίτερα ενδιαφέρον να συσχετιστεί η αλληλεπίδραση Neur1 και Dlc-1 με το ρόλο της μυοσίνης και της δυνεΐνης στο κεντρόσωμα και στους πόλους της ατράκτου κατά την κυτταρική διαίρεση.

Εκτός από συνδεδεμένη στα παραπάνω σύμπλοκα, η Dlc-1 βρίσκεται και ελεύθερη σχηματίζοντας ομοδιμερή και ρυθμίζοντας ετερογενείς βιολογικές διεργασίες μέσω των αλληλεπιδράσεών της με μεγάλη ποικιλία πρωτεϊνών. Για παράδειγμα, διμερές της Dlc-1 των σπονδυλωτών αλληλεπιδρά με τον προαποπτωτικό παράγοντα Bim [Fan *et al*, 2001], τον οποίο απελευθερώνει παρουσία σημάτων επαγωγής της απόπτωσης [Puthalakath *et al*, 1999], με τον αναστολέα του μεταγραφικού παράγοντα NFκB, IκB, επηρεάζοντας τον υποκυτταρικό εντοπισμό του συμπλόκου NFκB-IκB [Crepieux *et al*, 1997], με την πρωτεΐνη του νευρικού ιστού GKAP (*Guanylate Kinase domain-Associated Post Synaptic Density-95*), διευκολύνοντας τη μεταφορά της στην μετασυναπτική περιοχή των νευρώνων ασύμμετρων συνάψεων του εγκεφάλου [Naisbitt *et al*, 2000], με τη συνθάση nNOS (*neuronal Nitric Oxide Synthase*), ρυθμίζοντας τη λειτουργία της στα νευρικά κύτταρα [Jaffrey & Snyder, 1996/ Wiggins *et al*, 2003], με το mRNA της παραθυροειδούς ορμόνης [Erstein *et al*, 2000], ρυθμίζοντας τον υποκυτταρικό του εντοπισμό, με μεταγραφικούς παράγοντες στον πυρήνα κυττάρων θηλαστικών και δροσόφιλας, επηρεάζοντας τον εντοπισμό και την ενεργότητά τους [Herzig *et al*, 2000/ Kaiser *et al*, 2003] καθώς και με πρωτεΐνες ιών των θηλαστικών (*Lyssavirus, Rabies virus, African Swine Fever Virus*), διευκολύνοντας ενδεχομένως τη μετακίνησή τους από την περιφέρεια προς τον πυρήνα των κυττάρων στα αρχικά στάδια της μόλυνσης [Jacob *et al*, 2000/ Raux *et al*, 2000/ Alonso *et al*, 2001]. Είναι ενδιαφέρον ότι όλες οι παραπάνω πρωτεΐνες που αλληλεπιδρούν με την Dlc-1 εμφανίζουν μεγάλη ετερογένεια στο μέγεθος, στον υποκυτταρικό εντοπισμό και στη λειτουργία τους, ενώ δεν εμφανίζουν σημαντική ομολογία στην αμινοξική τους σύσταση. Όπως έδειξε η συγκριτική ανάλυση της δομής των συμπλόκων του

διμερούς Dlc-1 με τις πρωτεΐνες Bim και nNOS, το διμερές Dlc-1 σχηματίζει δύο συμμετρικές κοιλότητες, οι οποίες λόγω της αμινοξικής τους σύστασης έχουν τη δυνατότητα να αλληλεπιδρούν ειδικά με μεγάλη ποικιλία πρωτεϊνών [Fan *et al*, 2001]. Τα αποτελέσματά μας δεν αποκλείουν την αλληλεπίδραση του Neu1 με την ελεύθερη Dlc-1. Όμως, σε πρόσφατα βιοχημικά πειράματα διαπιστώσαμε την παρουσία και άλλων υπομονάδων της κυτταροπλασματικής και της αξονικής δυνείνης στα πρωτεϊνικά σύμπλοκα που κατακρημνίζονται με anti-Neu1 [Ε. Κουτέλου, αδημοσίευτα αποτελέσματα].

γ. Growth Arrest Specific protein 8

Η πρωτεΐνη Gas8 του ποντικού κωδικοποιείται από το γονίδιο *Gas8*, το οποίο είχε κλωνοποιηθεί αρχικά από μια βιβλιοθήκη κλώνων cDNA κυττάρων NIH3T3 εμπλουτισμένη σε μετάγραφα που εκφράζονται σε συνθήκες επαγωγής 'G1 arrest' [Schneider *et al*, 1988], αντιπροσωπεύοντας πρωτεΐνες που πιθανόν ρυθμίζουν τη μετάβαση του κυττάρου στη μίτωση. Το ορθόλογο ανθρώπινο γονίδιο (*GAS11*) είχε εξεταστεί ως υποψήφιο ογκοκατασταλτικό σε μελέτη απώλειας ετεροζυγωτίας (*Loss Of Heterozygosity*), λόγω της χαρτογράφησης του στη χρωμοσωμική θέση 16q24.3, ελλείμματα της οποίας παρατηρούνται συχνά σε καρκίνους μαστού και προστάτη. Με βάση αυτή τη μελέτη, αποκλείστηκε ο ρόλος του *GAS11* στον καρκίνο του μαστού, αλλά θεωρήθηκε πιθανός ο ρόλος του στον καρκίνο του προστάτη [Whitmore *et al*, 1998].

Πρόσφατες μελέτες του αναπτυξιακού προτύπου έκφρασης του *Gas8* στον ποντικό και του υποκυτταρικού εντοπισμού της πρωτεΐνης συσχετίζουν το γονίδιο με τα τελικά στάδια της σπερματογένεσης και τον μηχανισμό κίνησης του ώριμου σπερματοζωαρίου, γεγονός που καθιστά ιδιαίτερα ενδιαφέρουσα την αλληλεπίδρασή του με την πρωτεΐνη Neu1. Αναλυτικότερα, το mRNA και η πρωτεΐνη του *Gas8* ανιχνεύονται κυρίως στους όρχεις ενήλικων ποντικών, όπου τα επίπεδά τους είναι υψηλότερα σε σύγκριση με αυτά των νεογέννητων ή νεαρής ηλικίας ποντικών, υποδηλώνοντας τη σχέση του με τα στάδια της διαφοροποίησης των σπερματοκυττάρων που ακολουθούν τη μείωση. Η πρωτεΐνη Gas8 συσσωρεύεται στα πρόδρομα σπερματοζωάρια, όταν αυτά πλησιάζουν στον αυλό των σπερματικών σωληναρίων και εντοπίζεται στο κυτταρόπλασμα των στρογγυλών και στην ουρά των επιμηκυσμένων και ώριμων σπερματιδίων. Στα ώριμα σπερματοζωάρια της επιδιδυμίδας, η πρωτεΐνη Gas8 ανιχνεύεται στα μαστίγια. Το ίδιο διαπιστώνεται και σε ανθρώπινα σπερματοζωάρια, τόσο υγιών όσο και ασθενών με μη αποφρακτική αζωοσπερμία. Όμως, το ότι ο ρόλος της πρωτεΐνης αφορά αποκλειστικά τα γαμετοκύτταρα επιβεβαιώνεται από την απουσία της σε δείγματα βιοψίας ασθενών με αζωοσπερμία συνδρόμου κυττάρων Sertoli, η οποία χαρακτηρίζεται από την απουσία σπερματοζωαρίων στα σπερματικά σωληνάκια [Yeh *et al*, 2002]. Ενδιαφέρουσα είναι και η παρατήρηση της ίδιας ομάδας ότι η Gas8 του ποντικού ανιχνεύεται επίσης και στις βλεφαρίδες του βρογχικού επιθηλίου των πνευμόνων και του επιθηλίου του αυλού των ωαγωγών, γεγονός που συμφωνεί με τον προτεινόμενο ρόλο της στη λειτουργία του αξονήματος, αλλά και ο εντοπισμός του ορθόλογου μορίου στο μαστίγιο μονοκύτταρων ευκαρυωτικών οργανισμών (*Trypanosoma* και *Chlamydomonas*). Το τελευταίο επιβεβαιώνει τη συμμετοχή του στην εξελικτικά συντηρημένη λειτουργία των μαστιγίων. Συγκεκριμένα, η πρωτεΐνη PF2 της *Chlamydomonas* (ορθόλογη της Gas8/GAS11 των θηλαστικών) έχει χαρακτηριστεί ως συστατικό του ρυθμιστικού συμπλόκου δυνείνης DRC (*Dynein Regulatory Complex*), το οποίο αποτελεί σημαντικό ενδιάμεσο στο

μονοπάτι που ρυθμίζει την κινητικότητα του μαστιγίου. Το σύμπλοκο DRC της *Chlamydomonas* αποτελείται από τουλάχιστον 7 πολυπεπίδια, στενά συνδεδεμένα με τους μικροσωληνίσκους του εξωτερικού δακτυλίου του αξονήματος, μεταξύ των οποίων και η PF2. Ρόλος του είναι να μεταφέρει σήματα, η ταυτότητα των οποίων δεν έχει αποσαφηνιστεί, από το κεντρικό ζεύγος των μικροσωληνίσκων του αξονήματος προς τους βραχίονες της δυνείνης, ρυθμίζοντας τη λειτουργία της ως κινητήρα και κατ' επέκταση την κίνηση του μαστιγίου [Rupp & Porter, 2003]. Το Neurl ενδεχομένως παίζει ρόλο 'γέφυρας' αλληλεπιδρώντας ταυτόχρονα με τις πρωτεΐνες Gas8 και Dlc-1, ώστε να συμβάλλει στην κινητικότητα των μαστιγίων.

δ. Envoiplakin και Ankycorbin

Για τη λειτουργία αυτών των δύο 'θηραμάτων' του Neurl υπάρχουν περιορισμένες πληροφορίες. Είναι πρωτεΐνες μεγάλου μοριακού βάρους με ελικοειδείς (*coiled coil*) περιοχές, που περιγράφηκαν πρόσφατα, μετά την ολοκλήρωση της αλληλούχησης του ανθρώπινου γονιδιώματος. Η *Envoiplakin* (EVPL) κατατάχθηκε λόγω ομολογίας μεταξύ των πρωτεϊνών της οικογένειας των πλακινών, που περιλαμβάνει επιπλέον τις πρωτεΐνες *periplakin*, *desmoplakin*, *plectin* και BPAG1 (*Bullous pemphigoid antigen 1*) [Leung *et al*, 2001] (Εικ.87). Το ελικοειδές κεντρικό τμήμα και οι επτά επαναλήψεις πλακίνης (*Plakin Repeat Domain*) στο C-τελικό της άκρο της EVPL είναι απαραίτητα για τον σχηματισμό ετεροτετραμερών με την *periplakin* και για τις αλληλεπιδράσεις με ενδιάμεσα ινίδια, αντίστοιχα [Leung *et al*, 2001/ Määta *et al*, 2000]. Οι πλακίνες συνδέουν ινίδια του κυτταροσκελετού (μικροσωληνίσκους ή ενδιάμεσα ινίδια) με θέσεις επαφής του κυττάρου με γειτονικά κύτταρα ή με την εξωκυττάρια ουσία, λειτουργίες που δικαιολογούν τον υποκυτταρικό εντοπισμό των μελών γνωστής λειτουργίας αυτής της οικογένειας (*desmoplakin*, *plectin* και BPAG1) στα δεσμοσώματα και ημιδεσμοσώματα κυττάρων θηλαστικών. Σύμφωνα με πρόσφατες μελέτες και τα νεότερα μέλη της οικογένειας (*envoiplakin* και *periplakin*) εντοπίζονται στα δεσμοσώματα, εκτελώντας ενδεχομένως παρόμοια λειτουργία. Επιπλέον, έχουν ανιχνευθεί στον κερατινοποιημένο φάκελο (*cornified envelope*) πλήρως διαφοροποιημένων στρωματοποιημένων επιθηλιακών κυττάρων, όπως αυτά των ανώτερων στρωμάτων της επιδερμίδας [DiColandrea *et al*, 2000/ Kalinin *et al*, 2001].

Η Ankycorbin περιγράφεται από τους Peng *et al* (2000) ως μια νέα πρωτεΐνη του κυτταροσκελετού των ινιδίων ακτίνης. Στο N-τελικό της άκρο φέρει έξι επαναλήψεις αγκυρίνης, γνωστές για την ιδιότητά τους να επιτρέπουν πρωτεϊνικές αλληλεπιδράσεις με μεγάλη ποικιλία μορίων, ενώ το μεγαλύτερο ποσοστό του μορίου καταλαμβάνεται από μια ελικοειδή (*coiled coil*) περιοχή, πολύ συνηθισμένη στις πρωτεΐνες του κυτταροσκελετού, η οποία συνήθως ευνοεί τον διμερισμό τους ή την ενσωμάτωσή τους σε υπερμοριακές δομές. Για τη λειτουργία της υπάρχει περιορισμένη πληροφορία. Το γονίδιο που την κωδικοποιεί εκφράζεται καθολικά στα όργανα των θηλαστικών. Η πρωτεΐνη συνεντοπίζεται με τα ινίδια ακτίνης στα σημεία διακυτταρικών συνδέσμων και στην περιφέρεια καλλιεργούμενων κυττάρων θηλαστικών. Επιπλέον, στους κυτταρικούς τύπους που εμφανίζουν μικρολάχνες στην κορυφαία τους επιφάνεια, όπως τα κύτταρα του εντερικού επιθηλίου που ελέγχθηκαν από τους Peng *et al* (2000), η ankycorbin ανιχνεύεται στο τερματικό δίκτυο (*terminal web*) ινιδίων, το οποίο διατρέχει την κορυφαία πλευρά των κυττάρων, κάθετα προς τις μικρολάχνες, καταλήγοντας σε συνδέσμους επαφής (*adherent junctions*) του πλευρικού τμήματος των επιθηλιακών κυττάρων. Όπως

έδειξαν οι *in vitro* δοκιμές αλληλεπίδρασης των Peng *et al* (2000), η ankycorbin δεν αλληλεπιδρά άμεσα με την ακτίνη, υποδηλώνοντας την παρουσία ενδιάμεσων μορίων που την συγκρατούν στα ινίδια της ακτίνης. Αν και ο ακριβής ρόλος αυτής της πρωτεΐνης δεν έχει πλήρως διαλευκανθεί, ο χαρακτηριστικός υποκυτταρικός εντοπισμός της υποστηρίζει το ενδεχόμενο να συμμετέχει στη διατήρηση ή/και την οργάνωση του κυτταροσκελετού της ακτίνης.

Ο βιολογικός ρόλος της αλληλεπίδρασης του *Neur1* με αυτά τα μόρια δεν μπορεί να εξηγηθεί πλήρως με βάση τα υπάρχοντα δεδομένα. Χρειάζεται αρχικά να διερευνηθεί ο εντοπισμός του *Neur1* σε δεσμοσώματα και στην κορυφαία επιφάνεια επιθηλιακών κυττάρων θηλαστικών, καθώς και ο συνεντοπισμός του με αυτές τις πρωτεΐνες. Το γεγονός ότι μεταγγραφα του *Neur1* ανιχνεύονται στα επιθήλια κατά την εμβρυϊκή ανάπτυξη του ποντικού [Pavliopoulos *et al*, 2002], αλλά και ο εντοπισμός του *neur* στην κορυφαία επιφάνεια επιθηλιακών κυττάρων προνυμφικών δίσκων δροσόφιλας (Η. Παυλόπουλος, Διδακτορική Διατριβή, Παν/μιο Κρήτης, 2002) είναι ενδείξεις που υποστηρίζουν την αλληλεπίδραση του *Neur1* με αυτές τις πρωτεΐνες.

➤ **Πρωτεΐνες που σχετίζονται με διαδικασίες κυτταρικού πολλαπλασιασμού ή/και προγραμματισμένου κυτταρικού θανάτου**

Εκτός από τις πρωτεΐνες *Dlc-1* και *Numa1*, των οποίων ο ρόλος στην κυτταρική διαίρεση περιγράφεται παραπάνω, η *Lek1* που επίσης αλληλεπιδρά με το *Neur1*, σχετίζεται με τη ρύθμιση της μίτωσης και τη διαδικασία της απόπτωσης. Η αλληλεπίδραση με την *Peg3* συσχετίζει το *Neur1* με αναπτυξιακές διαδικασίες που ρυθμίζουν τη νευρογένεση και τη μυογένεση.

ε. Lek1

Η πρωτεΐνη *Lek1* είναι ομόλογη της ανθρώπινης μιτοσίνης (ή CENP-F). Αν και για την CENP-F είναι γνωστό ότι συνδέεται παροδικά με το εξωτερικό πρωτεϊνικό περίβλημα των κινητοχώρων των χρωμοσωμάτων, για τα ορθόλογα μόρια στον ποντικό (*Lek1*) και στο κοτόπουλο (*CMF1*), υπάρχουν αντικρουόμενες αναφορές για την πρόσδεση στους κινητοχώρους [Goodwin *et al*, 1999/ Yang *et al*, 2003]. Όμως, τα ορθόλογα αυτά μόρια φαίνεται ότι συμμετέχουν στη διαφοροποίηση του μυϊκού συστήματος [Goodwin *et al*, 1999 / Rabon-Pena *et al*, 2000]. Το πρότυπο έκφρασης του *Neur1* επικαλύπτεται σε αρκετά σημεία με αυτό του *Lek1* κατά την ανάπτυξη του ποντικού. Και τα δύο γονίδια εκφράζονται πολύ νωρίς (πριν την E9.5) στο αναπτυσσόμενο έμβρυο. Η έκφραση του *Lek1* περιορίζεται στους ιστούς που διατηρούν έντονη μιτωτική δραστηριότητα (καρδιά και ήπαρ) μέχρι την όψιμη εμβρυϊκή ζωή (E16.5) [Yang *et al*, 2003]. Η έκφραση του *Neur1* είναι ευρύτερη, όμως επικαλύπτεται με αυτήν του *Lek1* στον καρδιακό μυ και στο ήπαρ. Στα ενήλικα ζώα, τα μεταγγραφα του *Lek1* δεν ανιχνεύονται [Yang *et al*, 2003], ενώ το *Neur1* εκφράζεται σε ορισμένους μόνο ιστούς (εγκέφαλο, ήπαρ, σκελετικό μυ και όρχεις) [Pavliopoulos *et al*, 2002].

Η πρωτεΐνη *Lek1* επάγει τη μιτωτική διαδικασία στα κύτταρα που εκφράζεται. Έχει πρόσφατα αναφερθεί η αλληλεπίδραση του καρβοξυτελικού της άκρου με τις πρωτεΐνες της οικογένειας του ρετινοβλαστώματος (p107, p130 και pRb), μέσω του 'pocket domain' που αυτές φέρουν. Συγκεκριμένα, οι Ashe *et al*

(2004) προτείνουν ότι η Lek1 αναστέλλει την αντιμιτωτική δράση αυτών των πρωτεϊνών, παρεμποδίζοντας τις αλληλεπιδράσεις τους με συγκεκριμένους στόχους. Αυτό το μοντέλο εξηγεί και τον φαινότυπο κυττάρων στα οποία έχει κατασταλεί η έκφραση της Lek1 με *antisense* RNA, ο οποίος χαρακτηρίζεται από αναστολή της μίτωσης, διακοπή του κυτταρικού κύκλου στη φάση G1/S και απόπτωση [Ashe *et al*, 2004].

Σύμφωνα με αδημοσίευτες παρατηρήσεις τις ίδιες ομάδας, η πρωτεΐνη Lek1 υφίσταται μια μετα-μεταφραστική τροποποίηση με την οποία αποκόπτεται ένα μικρό καρβοξυτελικό της τμήμα, το οποίο φέρει NLS και εισέρχεται στον πυρήνα. Είναι ενδιαφέρον το γεγονός ότι στο πείραμα των δύο υβριδίων εντοπίσαμε στην Lek1 δύο περιοχές αλληλεπίδρασης με την πρωτεΐνη *Neur1*, (αμινοξέα 466-607 και 1349-1655 της αλληλουχίας XP_129658), οι οποίες βρίσκονται αμινοτελικά του σήματος πυρηνικού εντοπισμού (2647-2663) και της περιοχής αλληλεπίδρασης με Rb (2664-2696), πράγμα που ενδεχομένως σημαίνει ότι η αλληλεπίδραση του *Neur1* με το Lek1 πλήρους μήκους ή το N-τελικό του τμήμα γίνεται στο κυτταρόπλασμα (Εικ.87). Η αλληλεπίδραση της πρωτεΐνης *Neur1* με μηχανισμούς μετα-μεταφραστικής τροποποίησης που προκαλούν πρωτεόλυση των υποστρωμάτων και είσοδό τους στον πυρήνα (ουβικουϊτινιλίωση/SUMOλίωση) πιθανόν να ρυθμίζει αυτή τη διαδικασία. Θα ήταν, επομένως, ενδιαφέρον να διερευνηθεί ο ρόλος της αλληλεπίδρασης των *Neur1* και Lek1 κατά την είσοδο του Lek1 στον πυρήνα. Αυτός θα μπορούσε να είναι ένας μηχανισμός συμμετοχής του *Neur1* στη διαδικασία επαγωγής ή διατήρησης της μίτωσης, εκτός του προτεινόμενου ρόλου του κατά την κυτταρική διαίρεση που επιτελείται μέσω της αλληλεπίδρασης του με τη *Numa1* και την *Dlc-1*. Αξίζει να σημειωθεί και η παρατήρηση ότι το ανθρώπινο ορθόλογο της Lek1 (μιτοσίνη/CENP-F) μεταφέρεται κατά τη μίτωση στους πόλους της ατράκτου, μέσω της κυτταροπλασματικής δυνεΐνης, μια μετακίνηση που παρατηρείται για την *Numa1* και άλλες πρωτεΐνες που ρυθμίζουν τη διαδικασία της κυτταρικής διαίρεσης, υποδηλώνοντας ότι η Lek1 δρα και κατά τη διάρκεια της διαίρεσης [Blagden & Glover, 2003/ Yang *et al*, 2003].

στ. Peg3

Η Peg3 χαρακτηρίζεται από τα 12 μοτίβα *Zn-finger*, τύπου C2HC2, τα οποία φέρει διάσπαρτα στο μόριό της (Εικ.87). Για αυτήν προτείνεται ο ρόλος του ρυθμιστή διαφοροποίησης κατά τη μυογένεση και τη νευρογένεση των θηλαστικών. Το αναπτυξιακό και ιστοειδικό πρότυπο έκφρασής της συμφωνεί με το ρυθμιστικό αυτό ρόλο [Relaix *et al*, 1996], ενώ εμφανίζει επικαλύψεις με αυτό του *Neur1* [Pavliopoulos *et al*, 2002]. Αναλυτικότερα, το γονίδιο *Peg3* εκφράζεται καθολικά στο έμβρυο του ποντικού μέχρι το στάδιο του γαστρίδιου και στη συνέχεια η έκφρασή του περιορίζεται στους σκελετικούς μύες και σε περιοχές του ΚΝΣ. Ο σκελετικός μύς είναι ο μόνος ιστός που παρουσιάζει συνεχώς έντονη έκφραση του γονιδίου κατά την ανάπτυξη και την ενήλικη ζωή. Έμβρυα E8-E9.5 εκφράζουν υψηλά επίπεδα μεταγράφων *Peg3* σε όλο το μεσόδερμα, εκτός της καρδιάς, όπου τα μετάγραφα ανιχνεύονται σε χαμηλά επίπεδα. Από την ηλικία E10, η έκφραση του γονιδίου είναι πολύ ισχυρή στους σωμίτες, συμπεριλαμβανομένου του σκληροτομίου (πρόδρομο χόνδρου και οστών) και του μυοτομίου (πρόδρομο μυών). Ακολουθεί καταστολή της έκφρασης σε όλο το μεσόδερμα εκτός των σκελετικών μυών, στους οποίους η έκφραση συνεχίζεται μέχρι και την ενήλικη ζωή. Έντονη έκφραση του *Peg3* και συνεντοπισμός με τα μετάγραφα του *Neur1* ανιχνεύεται στις προεκβολές των αναπτυσ-

σόμενων άκρων (*limb buds*), στο βλαστικό επιθήλιο του μουστακιού, στο εσωτερικό επιθήλιο του κοχλία και σε εσωτερικά επιθήλια των βρόγχων και των νεφρών, όπου παρατηρείται έκφραση και άλλων πρωτεϊνών-‘θηραμάτων’ του *Neurl* (*Ubc9*, *Dlc-1*, *Numa1*) [Μπαζοπούλου ΜΔ, Διπλωματική Εργασία, Παν/μιο Κρήτης, 2002]. Το πρότυπο έκφρασης του *Peg3* στο ΚΝΣ είναι πιο πολύπλοκο. Αρχικά (E9.5), καταστέλλεται η έκφρασή του στο νευροεκτόδερμα και στη συνέχεια (E13) περιορίζεται στο ‘*floor plate*’ και σε κύτταρα του νωτιαίου μυελού και του εγκεφάλου, ενώ σε ενήλικα ποντίκια ανιχνεύεται ισχυρή έκφραση στον υποθάλαμο, στο χοριοειδές πλέγμα και στην υπόφυση. Η συγκριτική μελέτη του προτύπου έκφρασης των *Neurl* και *Peg3* πραγματοποιήθηκε σε έμβρυα E11.5-E14.5, στα οποία διαπιστώθηκε συνέκφραση των δύο γονιδίων στο νευροεπιθήλιο του νωτιαίου μυελού [Μπαζοπούλου ΜΔ, Διπλωματική Εργασία, Παν/μιο Κρήτης, 2002]. Επιπλέον, οι Timmusk *et al* (2002) διαπιστώνουν ισχυρή έκφραση του *Neurl* σε ιστούς ενηλίκων ποντικών που εκφράζουν *Peg3* (υποθάλαμος, υπόφυση, σκελετικοί μύες). Η διατήρηση της έκφρασης των *Neurl* και *Peg3* στο μυϊκό και νευρικό σύστημα υποδηλώνει ότι ο βιολογικός ρόλος της αλληλεπίδρασής τους σχετίζεται όχι μόνο με την ανάπτυξη αλλά και με λειτουργίες μετά τη διαφοροποίηση των συγκεκριμένων κυτταρικών τύπων.

Τα γονίδια *PEG3* και *NEURL* του ανθρώπου έχουν συσχετιστεί με ογκοκαταστολή, λόγω της διαπίστωσης ότι η έκφρασή τους είναι μειωμένη σε κυτταρικές σειρές γλοιομάτων, σε αντίθεση με τον φυσιολογικό εγκέφαλο, όπου παρατηρούνται υψηλά επίπεδα μεταγράφων [Maegawa *et al*, 2001/Nakamura *et al*, 1998]. Ειδικότερα για το *PEG3* έχει διαπιστωθεί ότι το μητρικό αλληλόμορφο του γονιδίου υφίσταται φυσιολογικά *imprinting* και διερευνάται το ενδεχόμενο η εμφάνιση γλοιομάτων να σχετίζεται με επιγενετικά φαινόμενα υπερμεθυλίωσης των νησιδίων CpG της χρωμοσωμικής περιοχής 19q13.4, στην οποία αυτό χαρτογραφείται [Maegawa *et al*, 2001]. Αν και δεν υπάρχουν σαφείς αποδείξεις για τον ρόλο των πρωτεϊνών *Peg3* και *Neurl* στην ογκογένεση, η υπερέκφραση της *Peg3* σε κυτταρικές σειρές γλοιομάτων προκαλεί απώλεια της ογκογενετικής τους ικανότητας σε ανοσοκατασταλμένα ποντίκια, υποστηρίζοντας τον προτεινόμενο ρόλο του γονιδίου *Peg3* στην ογκοκαταστολή [Kohda *et al*, 2001]. Αντίστοιχη μελέτη υπερέκφρασης δεν έχει πραγματοποιηθεί για το *Neurl*. Όμως, η έλλειψη της λειτουργίας του δεν φαίνεται να προκαλεί εμφάνιση όγκων του ΚΝΣ στα *Neurl*^{-/-} και *Neurl*^{+/-} ζώα [Vollrath *et al*, 2001].

Η πρωτεΐνη *Peg3* ανιχνεύεται κατά κύριο λόγο στον πυρήνα καλλιεργούμενων κυττάρων θηλαστικών, νευρώνων και κυττάρων γλοίας, ενώ ένα σημαντικό ποσοστό της βρίσκεται στο κυτταρόπλασμα [Relaix *et al*, 1996/ Deng & Wu, 2000]. Η παρουσία της επηρεάζεται από τον κυτταρικό κύκλο και συμπίπτει με τη φάση της μίτωσης, γεγονός που συμφωνεί με την παρατήρηση ότι και τα μεταγγραφα του *Peg3* ανιχνεύονται σε υψηλά επίπεδα, ειδικά κατά την τελόφαση της μιτωτικής διαίρεσης. Το γεγονός ότι το γονίδιο *Peg3* αποτελείται από ένα μοναδικό εξόνιο θεωρείται ότι επιτρέπει την έκφρασή του κατά τη μίτωση [Relaix *et al*, 1996]. Αντίστοιχα πειράματα παρακολούθησης της έκφρασης του *Neurl* στις διαφορετικές φάσεις του κυτταρικού κύκλου δεν έχουν πραγματοποιηθεί, όμως από πειράματα ανοσοφθορισμού σε κύτταρα COS7 με τα αντισώματα anti-*Neurl* και anti-CyclinA (προσφορά από Ζ. Λυγερού), διαπιστώσαμε παρουσία του *Neurl* τόσο σε κύτταρα θετικά για κυκλίνη Α (φάσεις S, G2 και αρχικά στάδια μίτωσης), όσο και σε απουσία της κυκλίνης

A (τελικά στάδια μίτωσης και G1). Σε αντίθεση με το *Neur1*, οι *Relaix et al* (1996) παρατηρούν ότι η ένταση της χρώσης του *Peg3* είναι ετερογενής στην ίδια καλλιέργεια κυττάρων και ένα ποσοστό των πυρήνων εμφανίζονται αρνητικοί, χωρίς να έχει εξακριβωθεί η μοριακή βάση αυτής της ρύθμισης. Θα ήταν ενδιαφέρον να διερευνηθεί κατά πόσο η αλληλεπίδρασή του με το *Neur1* συμβάλλει σε αυτή τη ρύθμιση, με δεδομένη και την πιθανή συμμετοχή του *Neur1* στον μηχανισμό ουβικουΐτινιλίωσης και πρωτεόλυσης μέσω του πρωτεασώματος [E. Κουτέλου, αδημοσίευτα αποτελέσματα]. Έχει διαπιστωθεί ότι η *Peg3* αλληλεπιδρά με μέλη της οικογένειας των E3 λιγασών με RING *finger* *Siah* (ομόλογες της δροσοφιλικής *Sina*, *Seven in absentia*) [*Relaix et al*, 2000], χωρίς όμως να έχει αναφερθεί ουβικουΐτινιλίωσή μέσω των μορίων *Siah*. Η αλληλεπίδραση *Peg* και *Siah* φαίνεται να οδηγεί σε απόπτωση. Συγκεκριμένα, με βάση την παρατήρηση ότι η υπερέκφραση του *Siah1a* απουσία της *Peg3* σταματά τον κυτταρικό κύκλο στη φάση G1, ενώ παρουσία της *Peg3* ενεργοποιεί τον αποπτωτικό μηχανισμό που επάγεται από την *p53*, οι *Relaix et al* (2000) προτείνουν ότι η *Peg3* δρα συνεργατικά με την *Siah1a* στο αποπτωτικό μονοπάτι που επάγεται από την *p53*. Επιπλέον, η *Peg3* αλληλεπιδρά και με τον παράγοντα TRAF2 (*TNF Receptor Associated Factor 2*), με τον οποίο δρα συνεργατικά για την ενεργοποίηση του μεταγραφικού παράγοντα NFκB και ενδεχομένως συμμετέχει στη ρύθμιση του κυτταρικού πολλαπλασιασμού, της διαφοροποίησης και του προγραμματισμένου κυτταρικού θανάτου μέσω του σηματοδοτικού μονοπατιού του TNF (*Tumor Necrosis Factor*) [*Relaix et al*, 1998].

Σε συμφωνία με τον ρυθμιστικό ρόλο του γονιδίου *Peg3* σε αναπτυξιακές διαδικασίες απόπτωσης και κυτταρικού πολλαπλασιασμού και το ειδικό πρότυπο έκφρασής του στον μυϊκό και νευρικό ιστό, η απενεργοποίηση της λειτουργίας του προκαλεί βλάβες σε συγκεκριμένες νευρικές συνάψεις και απουσία ή μειωμένο αριθμό νευρώνων, ιδιαίτερα αυτών της περιοχής του υποθαλάμου που παράγουν οκυτοκίνη, οδηγώντας σε διαταραγμένη μητρική συμπεριφορά. Σημειώνεται ότι στην περίπτωση του *Peg3*, η ανικανότητα θηλασμού στα θηλυκά οφείλεται σε ορμονικό αίτιο (έλλειψη οκυτοκίνης), σε αντίθεση με τα *Neur1*^{-/-} ζώα που εμφανίζουν υποπλασμένους γαλακτοφόρους αδένες [Vollrath *et al*, 2001]. Επιπλέον, η υστέρηση στην αύξηση του σωματικού βάρους των *Peg3*^{-/-} ζώων πιθανόν να συσχετίζεται με διαταραχές σε διαδικασίες κυτταρικού πολλαπλασιασμού [Li *et al*, 1999].

Τέλος, ο μοριακός μηχανισμός με τον οποίο η *Peg3* συμβάλλει στην απόφαση του κυττάρου για την επαγωγή απόπτωσης ή '*G1 arrest*' από την *p53* έχει αρχίσει να αποκαλύπτεται, με τη διαπίστωση των Deng & Wu (2000) ότι η *Peg3* προκαλεί τη μετατόπιση της πρωτεΐνης *Bax* από το κυτταρόπλασμα στο εσωτερικό των μιτοχονδρίων. Η *Bax*, της οποίας η έκφραση επάγεται από την *p53*, βρίσκεται στο κυτταρόπλασμα κατά το '*G1 arrest*' και η μετατόπισή της στα μιτοχόνδρια είναι απαραίτητη για την απελευθέρωση του κυτοχρώματος *c* και την ενεργοποίηση των κασπασών, βασικά βήματα για την έναρξη της αποπτωτικής διαδικασίας. Το αποτέλεσμα αυτό στον υποκυτταρικό εντοπισμό της πρωτεΐνης *Bax* πιθανόν συνοδεύεται και με ουβικουΐτινιλίωσή της, εφόσον προηγούμενες μελέτες αναφέρουν ότι η παρουσία αναστολέων του πρωτεασώματος προκαλεί συσσώρευση της *Bax* στα μιτοχόνδρια και απόπτωση [Li & Dou, 2000]. Η συμμετοχή του *Neur1* σε αυτές τις διεργασίες δεν έχει διερευνηθεί, όμως με δεδομένη την ικανότητά του να αυτο-ουβικουΐτινιλιώνεται *in vitro* [E. Κουτέλου, αδημοσίευτα αποτελέσματα], η αλληλεπίδρασή του με την *Peg3* στα κύτ-

ταρα των θηλαστικών θα μπορούσε να το φέρνει σε επαφή με πρωτεΐνες-υποστρώματα, των οποίων ενδεχομένως να επάγει την ουβικουΐτινιλίωση.

➤ **Πρωτεΐνες που συμμετέχουν ή υπόκεινται σε μετα-μεταφραστικές τροποποιήσεις από τον ενζυμικό μηχανισμό ομοιοπολικής σύνδεσης μορίων SUMO**

ζ. Dnmt3a

Η Dnmt3a (*DNA methyl transferase 3a*) είναι μία από τις δύο *de novo* μεθυλ-τρανσφεράσες του DNA των θηλαστικών (Dnmt3a και Dnmt3b), οι οποίες μεθυλιώνουν εκ νέου το γονιδιωματικό DNA των εμβρύων, μετά τη γενική του απομεθυλίωση που ακολουθεί τη γονιμοποίηση [Bestor, 2000]. Η μεθυλίωση του DNA επηρεάζει τη δομή της χρωματίνης και συμβάλλει στην καταστολή της μεταγραφής [Robertson, 2002]. Ο συσχετισμός του *Neur1* με καταστολή της μεταγραφής [Timmusk *et al*, 2002] και η πρόσφατη διαπίστωση των Ling *et al* (2004) ότι η λειτουργία της Dnmt3a στην μεταγραφική καταστολή ρυθμίζεται μέσω τροποποίησής της από SUMO-1, καθιστούν την Dnmt3a ένα από τα πιο ενδιαφέροντα 'θηράματα' του *Neur1*.

Η αλληλεπίδραση των δύο μορίων επιβεβαιώθηκε τόσο *in vitro*, όσο και στα κύτταρα θηλαστικών. Από το αποτέλεσμα του πειράματος δύο υβριδίων στον σακχαρομύκητα φαίνεται ότι αρκεί μια μικρή περιοχή της Dnmt3a (αμινοξέα 406-539, Εικ) για την αλληλεπίδρασή της με το *Neur1*. Η περιοχή αυτή επικαλύπτεται σημαντικά με το *domain* PHD (*Plant Homeodomain*), περιοχή που ευθύνεται και για την αλληλεπίδραση της Dnmt3a με τον μεταγραφικό καταστολέα RP58, την απακετυλάση ιστονών HDAC1 [Fuks *et al*, 2001] και με τις E3 λιγάσες του SUMO-1 PIAS1 και PIASxα [Ling *et al*, 2004]. Πρόσθετες ενδείξεις από τη μελέτη του αναπτυξιακού προτύπου έκφρασης της Dnmt3a στα θηλαστικά [Okano *et al*, 1999] υποδηλώνουν ότι και αυτή η πρωτεΐνη, όπως και το *Neur1*, είναι παρούσα νωρίς κατά την εμβρυογένεση (E7.5), εμφανίζοντας σχεδόν καθολική έκφραση στο έμβρυο.

Οι Ling *et al* (2004) διαπιστώνουν ότι η τροποποίηση της Dnmt3a από SUMO-1 ρυθμίζει και τη λειτουργία της, αναστέλλοντας το κατασταλτικό αποτέλεσμά της στη μεταγραφή ενός γονιδίου αναφοράς. Για την τροποποίηση αυτή δεν έχει βρεθεί ακόμα το μόριο που μπορεί να την ενισχύει *in vitro* και *in vivo* δρώντας ως E3 λιγάση, αλλά τα μέλη της οικογένειας PIAS έχουν αποκλειστεί από αυτόν το ρόλο, παρόλο που συνεντοπίζονται σε μεγάλο βαθμό και αλληλεπιδρούν άμεσα με την Dnmt3a [Ling *et al*, 2004]. Η τροποποίηση των Dnmt3a και b από SUMO-1 είναι σημαντικός τρόπος λεπτής ρύθμισης της λειτουργίας τους και κατ'επέκταση της ιστοειδικής και αναπτυξιακής μεταγραφικής ρύθμισης γονιδίων των θηλαστικών και, επομένως, η ανακάλυψη των μορίων που παίζουν ρυθμιστικό ρόλο σε αυτή τη διαδικασία έχει εξαιρετικό ενδιαφέρον. Το γεγονός ότι το *Neur1* αλληλεπιδρά άμεσα με την Dnmt3a αλλά και με την Ubc9, το μοναδικό γνωστό ένζυμο E2 της SUMOλίωσης, ενδεχομένως σημαίνει ότι ρυθμίζει τη SUMOλίωση της Dnmt3a δρώντας ως E3 λιγάση του SUMO-1. Ο έλεγχος αυτού του μοντέλου με *in vitro* και *in vivo* δοκιμές SUMOλίωσης είναι κρίσιμος για την κατανόηση της λειτουργίας του *Neur1*.

η. Ubc9

Η Ubc9 (*Ubiquitin Conjugating Enzyme 9*) είναι μια πρωτεΐνη 158 αμινοξέων, η οποία φέρει το *domain* UBC (*Ubiquitin Conjugating*), χαρακτηριστικό για όλα τα συζευκτικά ένζυμα που καταλύουν το δεύτερο βήμα (E2) στις διαδικασίες τροποποίησης πρωτεϊνικών υποστρωμάτων από ουβικουΐτινη ή από μόρια ομόλογα της ουβικουΐτινης. Η Ubc9 είναι 'δότης' του SUMO στο τελευταίο στάδιο της αντίδρασης, όπου αυτό μεταφέρεται από την κυστεΐνη του ενεργού της κέντρου (Cys93) σε μια λυσίνη του υποστρώματος (Εικ.89). Σε αντίθεση με την ουβικουΐτινίωση, όπου πολλαπλά E2 χρησιμοποιούνται για την τροποποίηση συγκεκριμένων υποστρωμάτων από ουβικουΐτινη, η Ubc9 είναι το μόνο E2 γνωστό ένζυμο της SUMΟλίωσης. Είναι κατά κύριο λόγο πυρηνική πρωτεΐνη [Seufert *et al*, 1995], ενώ ένα ποσοστό της εντοπίζεται στα κυτταροπλασματικά και στα πυρηνοπλασματικά ινίδια των πυρηνικών πόρων [Zhang *et al*, 2002]. Όπως διαπιστώσαμε, συνεντοπίζεται με την πρωτεΐνη Neur1 σε διακριτές περιοχές της περιφέρειας του πυρήνα και απομονώνεται σε σύμπλοκο με το Neur1 από εκχυλίσματα κυττάρων θηλαστικών. Επιπλέον, η άμεση αλληλεπίδρασή της με την πρωτεΐνη Neur1 επιβεβαιώθηκε *in vitro*.

Η αποδεδειγμένη αλληλεπίδραση των πρωτεϊνών Neur1 και Ubc9, σε συνδυασμό με την πληροφορία από τη συνέκφρασή τους σε αρκετούς ιστούς του αναπτυσσόμενου εμβρύου όπως τα νευροεπιθήλια του προσθγκεφάλου, του οπισθγκεφάλου και του νωτιαίου μυελού, τα νωτιαία γάγγλια, το οσφρητικό επιθήλιο, κύτταρα του φακού κατά την πορεία διαφοροποίησής τους σε κρυσταλλικά και η ρίζα της αναπτυσσόμενης τρίχας [Μπαζοπούλου ΜΔ, Διπλωματική Εργασία, Παν/μιο Κρήτης, 2002] συνεισφέρουν σημαντικά στη λειτουργική διερεύνηση του Neur1 γιατί το συσχετίζουν με τον μηχανισμό τροποποίησης πρωτεϊνών από SUMO και με συγκεκριμένες αναπτυξιακές διαδικασίες. Αυτό όμως που παραμένει άγνωστο είναι ο ακριβής βιοχημικός ρόλος του Neur1 σε αυτόν τον μηχανισμό. Με βάση τον μηχανισμό της SUMΟλίωσης (Εισαγωγή, Εικ.15β), οι πρωτεΐνες που αλληλεπιδρούν με την Ubc9 είναι είτε υποστρώματα ή E3 λιγάσες του SUMO.

Αξίζει να σημειωθεί ότι για την Ubc9 έχουν προταθεί και λειτουργίες χωρίς εμφανή σχέση με τη διαδικασία της SUMΟλίωσης, όπως η συμμετοχή της στην μεταφορά της πρωτεΐνης Vsx-1 (*Visual system homeobox 1*) του *Danio rerio* στον πυρήνα, ανεξάρτητα από τον μηχανισμό SUMΟλίωσης, αλληλεπιδρώντας με το σήμα πυρηνικού εντοπισμού της [Kurtzman & Schechter, 2001] και η αλληλεπίδρασή της με την πρωτεΐνη *dynamin* των θηλαστικών, της οποίας ολιγομερή συμμετέχουν στον μηχανισμό ενδοκύττωσης. Ενώ η *dynamin* αλληλεπιδρά με Ubc9, SUMO και την E3 λιγάση PIAS1, δεν SUMΟλιώνεται, όμως αυτές οι αλληλεπιδράσεις της εμποδίζουν τον ολιγομερισμό και το ρόλο της στην ενδοκύττωση [Mishra *et al*, 2004]. Τέτοιου είδους αλληλεπιδράσεις θεωρούνται είτε εξυπηρετούν τη γειτνίαση της πολυπρωτεϊνικής μηχανής SUMΟλίωσης σε συγκεκριμένες θέσεις, κοντά σε πρωτεΐνες-υποστρώματα, ενώ ταυτόχρονα συμβάλλουν στην αποκάλυψη άγνωστων ακόμα ρόλων των πρωτεϊνών Ubc9 και SUMO στη ρύθμιση βιολογικών διεργασιών.

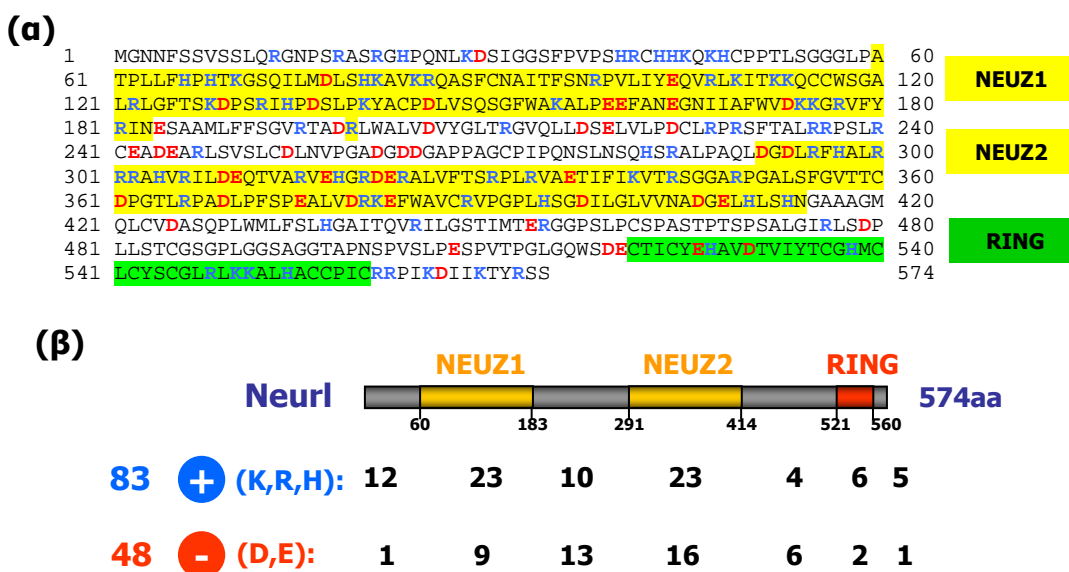
5. Ο ρόλος του RING στις αλληλεπιδράσεις με E2

Είναι γνωστό ότι η παρουσία του RING *finger* συσχετίζεται με τη μετα-μεταφραστική τροποποίηση πολλών πρωτεϊνών από ουβικουϊτίνη και είναι απαραίτητη για την αλληλεπίδρασή τους με 'συζευκτικά' (*conjugating*) ένζυμα E2 της διαδικασίας ουβικουϊτινίωσης (βλ. Εισαγωγή). Η πρωτεΐνη Ubc9 είναι το μοναδικό γνωστό E2 ένζυμο της SUMOλίωσης, για την οποία έχουν αναφερθεί αλληλεπιδράσεις με RING *fingers* (πχ με το RING της πρωτεΐνης PML, *ProMyelocytic Leukemia proto-oncoprotein*, Duprez *et al*, 1999), με εκφυλισμένα RING-like domains των μορίων PIAS (*Protein Inhibitors of Activated STAT*) [Jackson, 2000/ Kahyo *et al*, 2001/ Verger *et al*, 2003], αλλά και με περιοχές πρωτεϊνών που δεν έχουν ομολογία με RING, όπως με την πρωτεΐνη του πυρηνικού πόρου RanBP2 [Saitoh *et al*, 1998/ Pichler *et al*, 2002]. Τα μέχρι στιγμής γνωστά μόρια που αλληλεπιδρούν με Ubc9 είναι είτε υποστρώματα για τροποποίηση από SUMO (πχ η πρωτεΐνη PML), είτε λειτουργούν ως E3 λιγάσες για την τροποποίηση άλλων υποστρωμάτων (πρωτεΐνες PIAS και RanBP2), είτε δεν έχουν γνωστό ρόλο στη διαδικασία SUMOλίωσης (πχ η πρωτεΐνη *dynammin* των θηλαστικών και η Vsx-1 του *Danio rerio*) [Mishra *et al*, 2004/ Kurtzman & Schechter, 2001].

Όπως διαπιστώσαμε, η αλληλεπίδραση του Ubc9 με το Neu1 δεν εξαρτάται αποκλειστικά από το RING του Neu1. Η παρατήρηση αυτή συμφωνεί με τα αποτελέσματα κατευθυνόμενης μεταλλαξιγένεσης και κρυσταλλογραφικής ανάλυσης της δομής συμπλόκων πρωτεϊνών με RING και των αντίστοιχων E2 (πχ της πρωτεΐνης c-Cbl και του UbcH7 ή της RAD5 και του Ubc13), με βάση τα οποία τα συντηρημένα αμινοξέα όχι μόνο του RING αλλά και της ευρύτερης περιοχής του είναι ιδιαίτερα σημαντικά για τις αλληλεπιδράσεις με τα ένζυμα E2 [Zheng *et al*, 2000/ Ulrich, 2003]. Σημαντικό ρόλο σε αυτή την αλληλεπίδραση παίζουν κυρίως φορτισμένα και υδρόφοβα αμινοξέα, που συμμετέχουν στη διαμόρφωση της τυπικής δευτεροταγούς δομής ($\beta\beta\alpha\beta$) των RINGs του τύπου C3HC4, ενώ μια α έλικα που συνήθως προηγείται του RING αλλά και η ενδιάμεση περιοχή που σχηματίζει θηλιά (*loop*) συνεισφέρουν στη αλληλεπίδραση με το E2 [Zheng *et al*, 2000/ Ulrich 2003/ Joazeiro *et al*, 1999/ Meng *et al*, 1999/ Martinez-Noel *et al*, 2001]. Επιπλέον, οι ίδιες μελέτες δείχνουν ότι και τα στοιχεία δευτεροταγούς δομής των ενζύμων E2 που συμμετέχουν στην αλληλεπίδραση με τα RINGs, είναι συντηρημένα μεταξύ των διαφορετικών E2, συμβάλλοντας στη διατήρηση της ομολογίας των επιφανειών αλληλεπίδρασης μεταξύ διαφορετικών ζευγών E2-E3(RING) πρωτεϊνών. Η εξειδίκευση κάθε αλληλεπίδρασης E2-E3 οφείλεται σε μικρές διαφορές που αφορούν κυρίως το βαθμό συμμετοχής συγκεκριμένων αμινοξέων στην αλληλεπίδραση και στο ηλεκτροστατικό δυναμικό των δύο επιφανειών, που καθορίζεται από την κατανομή φορτισμένων αμινοξέων στις επιφάνειες των E2 και E3 που έρχονται σε επαφή [Winn *et al*, 2004/ Winkler *et al*, 2004/ Gervais *et al*, 2001]. Τα θετικό ηλεκτροστατικό δυναμικό που διαφοροποιεί το Ubc9 από τα υπόλοιπα E2 [Winn *et al*, 2004] επιτρέπει την επιλεκτική αλληλεπίδρασή του με το αρνητικά φορτισμένο SUMO [Saitoh *et al*, 1998/ Johnson, 2004] και ενδεχομένως και με άλλες πρωτεΐνες.

Από τις δοκιμές αλληλεπίδρασης των Neu1 και Ubc9 συμπεραίνουμε ότι το RING είναι απαραίτητο αλλά όχι ικανό για την αποτελεσματική αλληλεπίδραση των δύο μορίων. Με βάση και τα παραπάνω στοιχεία από αλληλεπιδράσεις άλλων RINGs με μόρια E2, υποθέτουμε ότι το *domain* NEUZ2 και η περιοχή ανάμεσα στο NEUZ2 και το RING του Neu1, δημιουργούν μαζί με το RING την επιφάνεια αλληλεπίδρασης με το Ubc9

και ίσως και με άλλα E2. Η υπόθεση αυτή υποστηρίζεται από το υψηλό ποσοστό φορτισμένων αμινοξέων στο μόριο Neu1 (Εικ.89), τα περισσότερα από τα οποία είναι συντηρημένα στα ορθόλογα μόρια Neu1 (Εικ.83-86). Είναι αξιοσημείωτο και το γεγονός ότι η πλειονότητα των αρνητικά φορτισμένων αμινοξέων (16/48) βρίσκονται στο NEUZ2, ενώ στο RING υπάρχουν κατά κύριο λόγο θετικά φορτισμένα αμινοξέα, όπως συνηθίζεται στα RINGs των πρωτεϊνών που αλληλεπιδρούν με τα E2 ένζυμα ουβικουϊτινιλίωσης. Πρόσφατα πειράματα έχουν δείξει ότι και η πρωτεΐνη Neu1 του ποντικού, όπως και οι ορθόλογες της δροσόφιλας και του βατράχου, αυτο-ουβικουϊτινιλιώνεται *in vitro* [Ε. Κουτέλου, αδημοσίευτα αποτελέσματα]. Επομένως, τα συντηρημένα αμινοξικά κατάλοιπα της περιοχής του RING πρέπει να καθορίζουν τη λεπτή ρύθμιση αυτής της ουβικουϊτινιλίωσης, διαμορφώνοντας μια τριτοταγή δομή που αλληλεπιδρά εξειδικευμένα με ορισμένα μόνο από τα διαθέσιμα στους διαφορετικούς ιστούς ένζυμα σύζευξης E2. Η συγκριτική μελέτη της αλληλεπίδρασης του Neu1 με το Ubc9 και με τα E2 της ουβικουϊτινιλίωσης είναι πλέον απαραίτητη για την κατανόηση της λειτουργίας του γιατί υποδηλώνει τη συμμετοχή του σε δύο μηχανισμούς μετα-μεταφραστικής τροποποίησης πρωτεϊνών (βλ. παρακάτω). Για τον λεπτομερή χαρακτηρισμό της αλληλεπίδρασης του Neu1(227-574) με το Ubc9, ή με άλλα E2, είναι απαραίτητη η ανάλυση της δομής του συμπλόκου των μορίων και η επιβεβαίωση του μοντέλου με πειράματα κατευθυνόμενης μεταλλαξιγένεσης. Ωστόσο, η πληροφορία που έχουμε από τις *in vitro* και *in vivo* δοκιμές αλληλεπίδρασης που πραγματοποιήθηκαν με ολόκληρα ή ελλειμματικά μόρια Neu1 είναι χρήσιμη στον σχεδιασμό αυτών των πειραμάτων.



Εικ.89 Κατανομή φορτισμένων αμινοξέων στην πρωτεΐνη Neu1 του ποντικού. (α) Με μπλε φαίνονται τα αρνητικά φορτισμένα αμινοξέα (D,E) και με κόκκινο τα θετικά φορτισμένα (K,R,H) στην αλληλουχία αμινοξέων του Neu1. (β) Ο αριθμός θετικά και αρνητικά φορτισμένων αμινοξέων σε κάθε περιοχή του Neu1. Παρατηρείται μεγαλύτερη συγκέντρωση φορτισμένων αμινοξέων στα NEUZ *domains* και στο RING. Τα περισσότερα αρνητικά φορτία βρίσκονται στην περιοχή NEUZ2, ενώ το RING είναι θετικά φορτισμένο.

6. Ο βιολογικός ρόλος του *Neuralized* στα θηλαστικά – προτεινόμενα μοντέλα

Με βάση τις μέχρι τώρα γνώσεις μας σχετικά με τις πρωτεΐνες που αλληλεπιδρούν με το *Neur1*, τον υποκυτταρικό εντοπισμό και το αναπτυξιακό και ιστικό πρότυπο έκφρασής του, προτείνουμε για το *Neur1* τον γενικό ρόλο του 'μεσολαβητή' (*adaptor*), ενός ενδιάμεσου δηλαδή πολυπεπτιδίου που φέρνει σε γεινίαση δύο ή περισσότερες πρωτεΐνες και με αυτό τον τρόπο, διευκολύνει ή/και τροποποιεί τη λειτουργία τους. Αλληλεπιδρώντας με πρωτεΐνες-κινητήρες μέσω της *Dlc-1*, το *Neur1* μπορεί να διευκολύνει τις μετακινήσεις πρωτεϊνών κατά την κυτταρική διαίρεση (π.χ. *Numa1*) και να συμμετέχει στην κίνηση μεμβρανικών κυστιδίων κατά μήκος των μαστιγίων ή των νευραξόνων.

Επιπλέον, ο υποκυτταρικός του εντοπισμός στην περιφέρεια της πυρηνικής μεμβράνης, ο συνεντοπισμός του με τροποποιημένες από *SUMO-1* πρωτεΐνες, οι αλληλεπιδράσεις του με παράγοντες της ενζυμικής μηχανής *SUMO*λίωσης ή/και ουβικουϊτινιλίωσης και η παρουσία του *RING finger*, ενδεχομένως προσδίδουν στο *Neur1* ιδιότητες 'μεσολαβητή' σε μετα-μεταφραστικές τροποποιήσεις, οι οποίες επιτελούνται μέσω των λιγασών *E3*. Τόσο οι λιγάσες *E3* της ουβικουϊτινιλίωσης που φέρουν *RING*, όσο και αυτές της *SUMO*λίωσης, χαρακτηρίζονται ως 'μεσολαβητές', εφόσον δεν έχουν καταλυτική δράση, αλλά επιλέγουν και φέρνουν σε επαφή τα συζευκτικά ένζυμα *E2* με τα υποστρώματά τους, ώστε να διευκολυνθεί η τελική μεταφορά της ουβικουϊτίνης ή του *SUMO* σε αυτά [Robinson & Ardley, 2004/ Johnson, 2004/ Gill, 2004]. Χαρακτηριστικό στοιχείο για τη διερεύνηση του ρόλου του *Neur1* στη *SUMO*λίωση είναι η αλληλεπίδρασή του με την *Dnmt3a*, της οποίας η τροποποίηση από *SUMO-1* δεν επάγεται *in vivo* από τις γνωστές λιγάσες *E3* της οικογένειας *PIAS* με τις οποίες επίσης αλληλεπιδρά [Ling *et al*, 2004]. Το ενδεχόμενο συμμετοχής του *Neur1* σε αυτή την τροποποίηση της *Dnmt3a* σχετίζεται και με τον προτεινόμενο ρόλο του στην μεταγραφική [Timmusk *et al*, 2002]. Ο έλεγχος αυτής της υπόθεσης είναι εξαιρετικά κρίσιμος για την κατανόηση του βιοχημικού ρόλου της πρωτεΐνης *Neur1*. Το πρώτο βήμα ελέγχου του ρόλου του *Neur1* ως *E3* λιγάση του *SUMO* θα είναι η διερεύνηση της ικανότητάς του να αυξάνει τα επίπεδα της τροποποιημένης μορφής της *Dnmt3a* σε κύτταρα θηλαστικών και σε *in vitro* αντιδράσεις *SUMO*λίωσης. Επίσης, θα πρέπει να ελεγχθεί το ενδεχόμενο τροποποίησης και άλλων 'θηραμάτων' του *Neur1* (κυρίως των *Numa1*, *Lek1*, *Dlc-1*) από τον μηχανισμό *SUMO*λίωσης. Μια σημαντική ένδειξη προς αυτήν την κατεύθυνση μπορεί να προκύψει έμμεσα, ελέγχοντας την πιθανή αλληλεπίδραση αυτών των μορίων με την *Ubc9*. Τέλος, με βάση και τον φαινότυπο απώλειας λειτουργίας του *Neur1*, πρέπει να ληφθεί υπόψη το ενδεχόμενο η λειτουργία του στις παραπάνω τροποποιήσεις να είναι 'συνεργατική' της δράσης άλλων *E3* λιγασών ή να υποκαθίσταται πλήρως ή μερικώς από αυτές, γεγονός που θα προσθέσει πολυπλοκότητα στις *in vivo* δοκιμές.

Αν και δεν υπάρχουν αναφορές για *SUMO*λίωση άλλων από τις πρωτεΐνες που αλληλεπιδρούν με το *Neur1*, αυτό δεν αποκλείει ορισμένα από αυτά να είναι υποστρώματα για *SUMO*λίωση. Στην αμινοξική αλληλουχία των *Numa1*, *Lek1*, *Dlc-1*, *Gas8* και *Enr1* ανιχνεύεται *in silico* το χαρακτηριστικό συντηρημένο τετραπεπτιδίο γΚxΕ (γ: υδρόφοβο αμινοξύ, x: οποιοδήποτε αμινοξύ), που συναντάται σε μεγάλο αριθμό χαρακτηρισμένων υποστρωμάτων και περιλαμβάνει τη λυσίνη-στόχο του *SUMO* [Sampson *et al*, 2001]

(Πίνακας 27). Για την Lek1 είναι γνωστό ότι η μετακίνησή της στον πυρήνα ρυθμίζεται μετα-μεταφραστικά, χωρίς να έχει διευκρινιστεί σε μοριακό επίπεδο η φύση αυτής της τροποποίησης [Ashe *et al*, 2004]. Επιπλέον, η δράση των μορίων Numa1, Dlc-1 και Lek1 σχετίζεται άμεσα με τον μηχανισμό της κυτταρικής διαίρεσης, μια λειτουργία που αποδεδειγμένα ρυθμίζεται από τον μηχανισμό SUMΟλίωσης. Η απενεργοποίηση της Ubc9 στον *S. cerevisiae* προκαλεί σταμάτημα του κυτταρικού κύκλου στα αρχικά στάδια της μίτωσης, μετά την εκβλάστηση [Seufert *et al*, 1995/ Johnson & Blobel, 1999], ενώ στα σπονδυλωτά οδηγεί σε προβλήματα κυτοκίνησης κατά την κυτταρική διαίρεση και σε κύτταρα με πολλαπλούς πυρήνες, ενώ αυξάνει σημαντικά το ποσοστό των αποπτωτικών κυττάρων [Hayashi *et al*, 2002]. Αυτοί οι φαινότυποι παρατηρούνται και σε μεταλλαγές των Numa1, Dlc-1 και Lek1. Η προτεινόμενη λειτουργική σχέση των Numa1 και Dlc-1 με την πρωτεΐνη Ubc9 υποστηρίζεται και από τη συγκριτική ανάλυση του αναπτυξιακού προτύπου έκφρασης των αντίστοιχων γονιδίων, στην οποία διαπιστώθηκε συνέκφραση των Ubc9 και Numa1 σε διάφορα επιθήλια, με υψηλότερα επίπεδα μεταγράφων κατά μήκος της επενδυματικής στιβάδας του νωτιαίου μυελού και στην εσωτερική στιβάδα του βλενογόννου επιθηλίου του στομάχου, ενώ συνέκφραση των Ubc9 και Dlc-1 παρατηρήθηκε στο τρίδυμο γάγγλιο και στα νωτιαία γάγγλια. Σε όλους τους παραπάνω ιστούς εκφράζεται έντονα και το Neurl.

Η αλληλεπίδραση του Neurl με την Ubc9 πιθανόν σχετίζεται και με την τοπολογία της Ubc9 κατά την κυτταρική διαίρεση. Βασικά συστατικά του μηχανισμού SUMΟλίωσης, όπως το SUMO, η λιγάση RanBP2 και η πρωτεΐνη-υπόστρωμα RanGAP1 (*Ran GTPase-activating protein*) στρατολογούνται στη μιτωτική άτρακτο και στους κινητοχώρους των χρωμοσωμάτων κατά την κυτταρική διαίρεση [Joseph *et al*, 2002/ Ayaydin & Dasso, 2004]. Ο μηχανισμός και η βιολογική σημασία του φαινομένου δεν έχει εξακριβωθεί, όμως είναι πιθανό να εξυπηρετεί τη SUMΟλίωση συγκεκριμένων πρωτεϊνών, ενδεχομένως και αυτών που αλληλεπιδρούν με τη μιτωτική άτρακτο και τους κινητοχώρους. Η έλλειψη αναφορών για SUMΟλίωση άλλων 'θηραμάτων' του Neurl ενδεχομένως να οφείλεται σε χαμηλά επίπεδα των SUMΟλιωμένων τους μορφών στα κύτταρα των θηλαστικών. Αξίζει να σημειωθεί ότι η ανίχνευση οποιασδήποτε τροποποιημένης από SUMO πρωτεΐνης εκτός ελάχιστων εξαιρέσεων (πχ της πρωτεΐνης RanGAP1 του πυρηνικού πόρου) είναι εξαιρετικά δύσκολη, λόγω του πολύ μικρού ποσοστού των τροποποιημένων μορίων (~1%) και της παρουσίας ισοπεπτιδασών ειδικών για αποSUMΟλίωση [Johnson, 2004/ Gill, 2004]. Εναλλακτικά, το ενδεχόμενο το Neurl να αποτελεί υπόστρωμα του SUMO θα εξηγούσε την αλληλεπίδραση του Neurl με το Ubc9 αλλά και τον πυρηνικό εντοπισμό που παρατηρείται σε μικρό ποσοστό των μορίων Neurl [Timmusk *et al*, 2002 και 'Αποτελέσματα'], εφόσον η SUMΟλίωση αποτελεί μηχανισμό ρύθμισης του πυρηνικού ή υποπυρηνικού εντοπισμού γνωστών υποστρωμάτων του SUMO (βλ. Εισαγωγή). Η απουσία από το Neurl του τυπικού τετραπεπτιδίου γKxE, που αποτελεί θέση SUMΟλίωσης πολλών υποστρωμάτων, δεν αποκλείει το ενδεχόμενο να τροποποιείται από SUMO σε μια 'μη τυπική' θέση, όπως συμβαίνει και με άλλα υποστρώματα (πχ τις πρωτεΐνες Smad4, TEL, PML, Axin, Mdm2, Daxx, CREB και CTBP-2) που αλληλεπιδρούν με Ubc9 [Johnson, 2004]. Όπως έδειξε και η κρυσταλλογραφική ανάλυση του συμπλόκου της Ubc9 με την RanGAP1, επαφές της Ubc9 και με άλλες περιοχές της RanGAP1, εκτός αυτής του τετραπεπτιδίου γKxE, προσδίδουν εξειδίκευση στην επιλογή του υποστρώματος [Bernier-Villamor *et al*, 2002]. Πάντως, η πιθανότητα το Neurl να

SUMOλιώνεται δεν αποκλείει τον ρόλο του ως λιγάση E3 του SUMO. Αντίστοιχα και η RanBP2 (*Ran Binding Protein 2*), γνωστή E3 που ευθύνεται για τη SUMOλίωση της RanGAP1, αποτελεί υπόστρωμα του SUMO-1 [Pichler *et al*, 2002/ Saitoh *et al*, 1998]. Η ενδεχόμενη τροποποίηση του Neurl με SUMO πιθανόν να σχετίζεται με τη ρύθμιση του υποκυτταρικού του εντοπισμού, της πυρηνοκυτταροπλασματικής του διακίνησης αλλά και των αλληλεπιδράσεων του με άλλες πρωτεΐνες. Χαρακτηριστικό είναι το παράδειγμα της RanGAP1, της οποίας η SUMOλίωση ευθύνεται για τον εντοπισμό της στους πυρηνικούς πόρους και στη μιτωτική άτρακτο [Mahajan *et al*, 1997/ Joseph *et al*, 2002]. Για τον έλεγχο αυτής της υπόθεσης απαιτείται η ανασύσταση του μηχανισμού SUMOλίωσης *in vitro*, χρησιμοποιώντας βακτηριακά εκφρασμένες πρωτεΐνες E1, E2, SUMO και Neurl, καθώς και η ανίχνευση της SUMOλιωμένης μορφής του Neurl στα κύτταρα παρουσία αναστολέων των ισοπεπτιδασών του SUMO.

Πίνακας 27. Πρόβλεψη θέσεων SUMOλίωσης

Πρωτεΐνη	Λυσίνη-στόχος	Αλληλουχία θέσης σουμολίωσης	Βαθμός
Numa1 (AAH49791, 2094aa)	K1083	LKIQL VKKE KEHPA	0.93
Lek1 (XP_129658, 2818aa)	K2419 K1729 K687	SELGL IKLE KVALV RELTR IKSE KANIE EELLQ IKGE VEENL	0.94 0.94 0.94
Gas8 (NP_061343, 489aa)	K92	NNLAE VKAE GTVVM	0.93
Dlc-1 (AAH08106, 89 aa)	K43	DIAAH IKKE FDKKY	0.94
Evpl (Q9D952, 2035aa)	K1082 K1468	VVKEV VKVE KDLEM IMEEV VKLE KDPDL	0.93 0.93
Dnmt3a (NP_031898, 908aa)	K573	AAQAA IKED PWNCY	0.94
Vsx-1 (NP_571408, 344aa)	K187	REMLA MKTE LPEDR	0.80
GLUD1 (NP_005262, 558aa)	K457	YGRLT FKYE RDSNY	0.85

Σημ: Αποτελέσματα πρόβλεψης θέσεων SUMOλίωσης με το πρόγραμμα SUMOplot για τις πρωτεΐνες που αλληλεπιδρούν με το Neurl. Φαίνονται οι προβλεπόμενες θέσεις των 'θηραμάτων' του Neurl, με βαθμό >0.91. Σε γκριζό υπόβαθρο, τα αποτελέσματα της ανάλυσης για τις πρωτεΐνες Vsx-1 και GLUD1 που δεν SUMOλιώνονται (αρνητικός έλεγχος της αξιολόγησης των θέσεων SUMOλίωσης). Η Dnmt3a, γνωστό υπόστρωμα του SUMO-1, αναγνωρίζεται από το πρόγραμμα. Στις παρενθέσεις (πρώτη στήλη), ο κωδικός πρόσβασης στην Genbank και ο αριθμός αμινοξέων (aa) κάθε πρωτεΐνης.

Επίσης, δεν μπορεί να αποκλειστεί η πιθανότητα το Neurl αλληλεπιδρώντας με την Ubc9 να οδηγείται στον πυρήνα. Είναι γνωστό ότι η πρωτεΐνη Ubc9 εισέρχεται στον πυρήνα αλληλεπιδρώντας με την Importin 13 [Mingot *et al*, 2001]. Η αλληλεπίδραση του Neurl με αυτήν πιθανόν να σχετίζεται με την παρουσία μικρού ποσοστού των μορίων του στον πυρήνα, όπως συμβαίνει και για την πρωτεΐνη Vsx-1 [Kurtzman & Schechter, 2001]. Όπως διαπιστώσαμε, η πρωτεΐνη Neurl σχηματίζει έναν δακτύλιο γύρω από τον πυρήνα, ενώ η Ubc9 είναι κυρίως πυρηνική (Εικ.60). Οι δύο πρωτεΐνες συνεντοπίζονται σε διακριτά

σημεία του πυρηνικού φακέλου και του πυρηνοπλάσματος, παρατήρηση που υποστηρίζει την παραπάνω υπόθεση. Αξίζει να τονιστεί ότι η πυρηνοκυτταροπλασματική διακίνηση του Neu1 ενισχύει την πιθανότητα αυτό να συμμετέχει στον μηχανισμό SUMΟλίωσης ως υπόστρωμα ή/και ως E3 λιγάση, εφόσον η ρύθμιση της μετακίνησης προς και από τον πυρήνα είναι μία από τις βασικές συνέπειες της τροποποίησης πρωτεϊνών από SUMO [Johnson, 2004].

Τέλος, η πρωτεΐνη Neu1 έχει συσχετιστεί και με τον μηχανισμό ουβικουϊτινιλίωσης. Όπως δείχνουν πρόσφατα πειράματα *in vitro* ουβικουϊτινιλίωσης του Neu1 των θηλαστικών [Ε. Κουτέλου, αδημοσίευτα αποτελέσματα], σε συμφωνία με προηγούμενες μελέτες των ορθόλογων μορίων της *Drosophila* και του *Xenopus* [Yeh *et al*, 2001/ Lai *et al*, 2001/ Deblandre *et al*, 2001] (βλ. Εισαγωγή), το Neu1 λειτουργεί ως E3 λιγάση ουβικουϊτίνης, επάγοντας τουλάχιστον την αυτο-ουβικουϊτινιλίωσή του. Η αυτο-ουβικουϊτινιλίωση είναι σύνθετος φαινόμενο μεταξύ των E3 λιγασών ουβικουϊτίνης [Robinson & Ardley, 2004] και ενδεχομένως αποτελεί μηχανισμό ρύθμισης των επιπέδων του Neu1 στα κύτταρα. Προς το παρόν δεν υπάρχουν στοιχεία για τις πρωτεΐνες-υποστρώματα των θηλαστικών, των οποίων η ουβικουϊτινιλίωση πιθανόν ρυθμίζεται μέσω του Neu1. Θα ήταν ενδιαφέρον να διερευνηθεί το ενδεχόμενο ρύθμισης από το πρωτεάσωμα των πρωτεϊνών που ταυτοποιήθηκαν σε αυτή την εργασία ως 'θηράματα' του Neu1. Ειδικότερα για τις Lek1, Numa1 και Peg3, τα δεδομένα που υπάρχουν για τη ρύθμιση και τη λειτουργία τους υποστηρίζουν ένα τέτοιο μοντέλο. Συγκεκριμένα, η Lek1 υφίσταται μερική πρωτεόλυση για να εισέλθει στον πυρήνα [Ashe *et al*, 2004], τα επίπεδα της Numa1 επηρεάζονται από τις διαφορετικές φάσεις του κυτταρικού κύκλου και μειώνονται κατά την απόπτωση [Gueth-Hallonet *et al*, 1997/ Taimen *et al*, 2000/ Taimen & Kallajoki, 2003], ενώ η Peg3 ανιχνεύεται μόνο στη φάση της μίτωσης [Relaix *et al*, 1996] και αλληλεπιδρά με τις λιγάσες ουβικουϊτίνης της οικογένειας Siah [Relaix *et al*, 2000]. Η άμεση αλληλεπίδραση και η συνέκφρασή τους σε αρκετούς ιστούς με το Neu1 καθιστά αυτές τις πρωτεΐνες σημαντικά υποψήφια υποστρώματα που ενδεχομένως ελέγχονται μετα-μεταφραστικά μέσω του Neu1.

Ο προτεινόμενος ρόλος ως E3 λιγάση ουβικουϊτίνης δεν αποκλείει τη συμμετοχή του Neu1 και στον μηχανισμό SUMΟλίωσης. Αντίθετα, φαίνεται ότι το Neu1 είναι ικανό να αλληλεπιδρά τόσο με την πρωτεΐνη Ubc9 (βλ. Αποτελέσματα), όσο και με ένζυμα E2 ουβικουϊτινιλίωσης [Ε. Κουτέλου, Μ. Κοκκινάκη, αδημοσίευτες παρατηρήσεις] και πιθανόν συμμετέχει και στους δύο μηχανισμούς. Στο επίπεδο των E3 λιγασών, ο μηχανισμός της τροποποίησης πρωτεϊνικών υποστρωμάτων με ουβικουϊτίνη εμφανίζει ομολογία με τροποποιήσεις από μόρια ομόλογα της ουβικουϊτίνης. Συγκεκριμένα, για την πρωτεΐνη Mdm2, η οποία φέρει RING *finger* και λειτουργεί ως E3 λιγάση για την ουβικουϊτινιλίωση της p53, πρόσφατα διαπιστώθηκε ο ρόλος της ως E3 λιγάση και για την τροποποίηση της p53 με το ομόλογο της ουβικουϊτίνης μόριο Nedd8 [Xirodimas *et al*, 2004]. Η ταυτοποίηση μορίων όπως το Neu1 που δρουν και στους δύο αυτούς μηχανισμούς είναι δυνατόν να αποκαλύψει νέα στοιχεία για την λεπτή ρύθμιση αυτών των μετα-μεταφραστικών τροποποιήσεων πρωτεϊνών.

7. Συμπεράσματα και προοπτικές

Τα συμπεράσματα που προκύπτουν από αυτή την εργασία συνοψίζονται ως εξής:

1. Η περιοχή ανάμεσα στους γενετικούς δείκτες D10S541-D10S583 του ανθρώπινου χρωμοσώματος 10 έχει μήκος 4.38Mb και περιέχει 51 λειτουργικά γονίδια από τα οποία 34 κωδικοποιούν πρωτεΐνη.
2. Το γονίδιο *Neur1* του ποντικού:
 - είναι το ορθόλογο του ανθρώπινου γονιδίου *NEURL*, καταλαμβάνει 79Kb γονιδιωματικής αλληλουχίας & αποτελείται από 6 εξόνια
 - παράγει δύο εναλλακτικά μετάγραφα που κωδικοποιούν τις ετερογενείς στο αμινοτελικό άκρο, ισομορφές *Neur1*, μήκους 574 και 557 αμινοξέων
 - εμφανίζει μεγαλύτερη συντήρηση στα εξόνια που κωδικοποιούν τις περιοχές NEUZ1,2 και RING
3. Η πρωτεΐνη *Neur1* αλληλεπιδρά και συνεντοπίζεται με πρωτεΐνες που:
 - σχετίζονται με τον κυτταροσκελετό ή αποτελούν συστατικά ινιδίων
 - ρυθμίζουν τη διαδικασία της κυτταρικής διαίρεσης και της απόπτωσης
 - συμμετέχουν ή υπόκεινται σε μετα-μεταφραστικές τροποποιήσεις
4. Οι περιοχές NEUZ1 και NEUZ2 είναι αναγκαίες για όλες τις αλληλεπιδράσεις της πρωτεΐνης, ενώ το RING finger συνεισφέρει στην αλληλεπίδραση του *Neur1* με την *Ubc9*.

Η εργασία αυτή συνέβαλλε στην κατασκευή του φυσικού χάρτη υψηλής ευκρίνειας του χρωμοσώματος 10 του ανθρώπου αλλά και στη μελέτη της λειτουργίας του γονιδίου *Neur1* των θηλαστικών. Ο φυσικός χάρτης των κλώνων PAC και BAC που δημιουργήθηκε, χρησιμοποιήθηκε για την αλληλούχηση της αντίστοιχης χρωμοσωμικής περιοχής και την ταυτοποίηση των γονιδίων που περιέχονται σε αυτήν. Η πλήρης νουκλεοτιδική αλληλουχία και τα αποτελέσματα του *'gene annotation'* αυτής της περιοχής είναι διαθέσιμα στη βάση δεδομένων *Ensembl*. Πρόσθετες πληροφορίες για τη ρύθμιση της έκφρασης καθενός από αυτά τα γονίδια ή για την εξελικτική τους ιστορία, μπορούν να προκύψουν μελετώντας την οργάνωση των κωδικών τους αλληλουχιών σε σχέση με μη κωδικές και εξελικτικά συντηρημένες, αλληλουχίες. Η συγκριτική γονιδιωματική ανάλυση είναι δυνατόν να αποκαλύψει περιοχές πιθανών ρυθμιστικών στοιχείων και να επιταχύνει την πειραματική επιβεβαίωση του ρόλου τους στη γονιδιακή έκφραση. Στα πλαίσια αυτά, η συγκριτική ανάλυση της γονιδιωματικής οργάνωσης των γονιδίων *Neur1* και *NEURL*, εξυπηρετεί τον ίδιο σκοπό.

Η ταυτοποίηση των πρωτεϊνών που αλληλεπιδρούν άμεσα με την πρωτεΐνη *Neur1* συμβάλλει σημαντικά στη μελέτη του βιολογικού της ρόλου στα θηλαστικά γιατί την συσχετίζει με μεγάλο εύρος βιολογικών διεργασιών: την κυτταρική διαίρεση, την ενδοκυττάρια κίνηση μέσω πρωτεϊνών-κινητήρων και τον ενζυμικό μηχανισμό μετα-μεταφραστικής τροποποίησης πρωτεϊνών από μόρια SUMO. Τα αποτελέσματά μας για το *Neur1* συμφωνούν με τον χαρακτηρισμό του ως 'μεσολαβητή', που αποδίδεται στις πρωτεΐνες με

RING [Borden, 2000/ Jackson *et al*, 2000/ Joazeiro & Weissman, 2000]. Το γεγονός ότι μεταξύ των πρωτεϊνών που αλληλεπιδρούν άμεσα με το *Neur1* δεν ανιχνεύσαμε πρωτεΐνες με χαρακτηρισμένο ρόλο στο μονοπάτι σηματοδότησης του υποδοχέα Notch (πχ τους δεσμευτές Delta των θηλαστικών), δεν διαφοροποιεί αναγκαστικά το ρόλο του *Neur1* των θηλαστικών από αυτόν του *neur* της δροσόφιλας στο συγκεκριμένο μονοπάτι. Η απουσία του δεσμευτή Delta δικαιολογείται από το γεγονός ότι η αλληλεπίδραση του *Neur1* με αυτόν *in vivo*, είναι πιθανότατα έμμεση όπως δείχθηκε με πείραμα δύο υβριδίων και πειράματα ανοσοκατακρήμνισης [E. Κουτέλου, αδημοσίευτα αποτελέσματα]. Η συντήρηση αυτού του σηματοδοτικού μονοπατιού κατά την ανάπτυξη των ζωικών οργανισμών, αλλά και η συγκριτική ανάλυση του προτύπου έκφρασης του *Neur1* και των συνδετών του Notch στα θηλαστικά *Dll1*, *Dll4*, *Jag1* και *Jag2* [E. Κουτέλου, Μεταπτυχιακή Διατριβή, Παν/μιο Κρήτης, 2001] υποστηρίζουν τη συμμετοχή του *Neur1* σε αναπτυξιακές διαδικασίες που ελέγχονται από το Notch. Ειδικότερα για τη νευρογένεση και την σπερματογένεση, διαδικασίες που επηρεάζονται από την έλλειψη λειτουργικού *Neur1* στα θηλαστικά [Vollrath *et al*, 2001/ Ruan *et al*, 2001], υπάρχουν πολλαπλές αναφορές που καταδεικνύουν τη σημασία της σηματοδότησης Notch [Hayashi *et al*, 2001/ Shen *et al*, 2002/ Chambers *et al*, 2001]. Η αλληλεπίδραση του *Neur1* με ορισμένες από τις πρωτεΐνες που ταυτοποιήσαμε, όπως η *Numa1* και η *Dlc-1*, σε συγκεκριμένους κυτταρικούς τύπους κατά την ανάπτυξη των θηλαστικών, θα μπορούσε να συμβάλει στη ρύθμιση της σηματοδότησης του υποδοχέα Notch. Ιδιαίτερο ενδιαφέρον προκαλεί το γεγονός ότι και οι δύο αυτές πρωτεΐνες συμμετέχουν στον καθορισμό του ισημερινού επιπέδου κατά την ασύμμετρη κυτταρική διαίρεση που αποτελεί βασικό μηχανισμό κυτταρικής διαφοροποίησης κατά τη νευρογένεση των σπονδυλωτών [Rogiers & Jan, 2004] και τη ρύθμιση της σηματοδότησης του Notch [Shen *et al*, 2002]. Με δεδομένη την αλληλεπίδραση του *Neur1* με τις *Numa1* και *Dlc-1* στο κεντρόσωμα αλλά και την ασύμμετρη κατανομή του δροσοφιλικού *neur* κατά τις διαιρέσεις των πρόδρομων κυττάρων αισθητηρίων οργάνων [Le Borgne & Schweisguth, 2003a], η διερεύνηση του ρόλου του *Neur1* στην ασύμμετρη κυτταρική διαίρεση μπορεί να βοηθήσει στην κατανόηση της λειτουργίας του στη σηματοδότηση Notch των θηλαστικών.

Τέλος, ο προτεινόμενος ρόλος του *Neur1* ως E3 λιγάση καθορίζεται από την ικανότητά του να αλληλεπιδρά με διαφορετικά μόρια E2 (SUMOλίωσης και ουβικουϊτινιλίωσης) αλλά και με άλλες πρωτεΐνες, διευκολύνοντας ενδεχομένως την τροποποίησή τους μέσω των δύο μηχανισμών. Η ταυτοποίηση των πρωτεϊνών που αλληλεπιδρούν με το *Neur1* συμβάλλει σημαντικά στον έλεγχο αυτού του μοντέλου, προσφέροντας μια σειρά από πιθανά υποστρώματα που θα δοκιμασθούν σε *in vitro* και *in vivo* πειράματα SUMOλίωσης και ουβικουϊτινιλίωσης.

ΒΙΒΛΙΟΓΡΑΦΙΑ

ΒΙΒΛΙΟΓΡΑΦΙΑ*

- Abaza A, Soleilhac JM, Westendorf J, Piel M, Crevel I, Roux A, Pirollet F (2003). M phase phosphoprotein 1 is a human plus-end-directed kinesin-related protein required for cytokinesis. *J Biol Chem.*, 278, (30): 27844-52.
- Agatep R, Kirkpatrick RD, Parchaliuk DL, Woods RA, Gietz RD (1998). Transformation of *Saccharomyces cerevisiae* by the lithium acetate/single-stranded carrier DNA/polyethylene glycol protocol. *Technical Tips Online* (<http://tto.trends.com>), 1, (51): P01525.
- Aguilar RC, Wendland B. (2003). Ubiquitin: not just for proteasomes anymore. *Curr Opin Cell Biol.*, 15, (2): 184-90.
- Alonso C, Miskin J, Hernaez B, Fernandez-Zapatero P, Soto L, Canto C, Rodriguez-Crespo I, Dixon L, Escribano JM (2001). African swine fever virus protein p54 interacts with the microtubular motor complex through direct binding to light-chain dynein. *J Virol.*, 75, (20): 9819-27.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. *J Mol Biol*, 215, (3):403-10.
- Anand R, Riley JH, Butler R, Smith JC, Markham AF (1990). A 3.5 genome equivalent multi access YAC library: construction, characterisation, screening and storage. *Nucleic Acids Res.*, 18, (8): 1951-6.
- Antonarakis SE (1994). Genome linkage scanning: systematic or intelligent? *Nat Genet.*, 8, (3): 211-2.
- Artavanis-Tsakonas S, Delidakis C, Fehon RG (1991). The Notch locus and the cell biology of neuroblast segregation. *Annu Rev Cell Biol.*, 7, (427-52):
- Artavanis-Tsakonas S, Matsuno K, Fortini ME (1995). Notch signaling. *Science*, 268, (5208): 225-32.
- Artavanis-Tsakonas S, Rand MD, Lake RJ (1999). Notch signaling: cell fate control and signal integration in development. *Science*, 284, (5415): 770-6.
- Ashe M, Pabon-Pena L, Dees E, Price KL, Bader D (2004). LEK1 is a potential inhibitor of pocket protein-mediated cellular processes. *J Biol Chem.*, 279, (1): 664-76.
- Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. (2000). *Saccharomyces cerevisiae*; *Current Protocols in Molecular Biology*, Unit 13, John Wiley & Sons, Inc., New York.
- Ayaydin F, Dasso M (2004). Distinct in vivo dynamics of vertebrate SUMO paralogues. *Mol Biol Cell*, 15, (12): 5208-18.
- Bailey AM, Posakony JW (1995). Suppressor of hairless directly activates transcription of enhancer of split complex genes in response to Notch receptor activity. *Genes Dev.*, 9, (21): 2609-22.
- Batzoglou S, Pachter L, Mesirov JP, Berger B, Lander ES. (2000). Human and mouse gene structure: comparative analysis and application to exon prediction. *Genome Res.*, 10, (7): 950-8.
- Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2003). GenBank. *Nucleic Acids Res.*, 31, (1): 23-7.
- Bentley DR *et al.* (2001). The physical maps for sequencing human chromosomes 1, 6, 9, 10, 13, 20 and X. *Nature*, 409, (6822): 942-3.
- Bernier-Villamor V, Sampson DA, Matunis MJ, Lima CD (2002). Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. *Cell*, 108, (3): 345-56.

* Τα άρθρα με περισσότερους από 15 συγγραφείς αναφέρονται με το όνομα του πρώτου συγγραφέα, πχ Lander ES *et al.* (2001).

- Bestor TH (2000). The DNA methyltransferases of mammals. *Hum Mol Genet.*, 9, (16): 2395-402.
- Bettenhausen B, Hrabe de Angelis M, Simon D, Guenet JL, Gossler A (1995). Transient and restricted expression during mouse embryogenesis of Dll1, a murine gene closely related to Drosophila Delta. *Development*, 121, (8): 2407-18.
- Bies J, Markus J, Wolff L. (2002). Covalent attachment of the SUMO-1 protein to the negative regulatory domain of the c-Myb transcription factor modifies its stability and transactivation capacity. *J Biol Chem.*, 277, (11): 8999-9009.
- Bingham S, Chaudhari S, Vanderlaan G, Itoh M, Chitnis A, Chandrasekhar A (2003). Neurogenic phenotype of mind bomb mutants leads to severe patterning defects in the zebrafish hindbrain. *Dev Dyn.*, 228, (3): 451-63.
- Blagden SP, Glover DM (2003). Polar expeditions--provisioning the centrosome for mitosis. *Nat Cell Biol.*, 5, (6): 505-11.
- Blair SS (2000). Notch signaling: Fringe really is a glycosyltransferase. *Curr Biol.*, 10, (16): R608-12.
- Blattner FR *et al.* (1997). The complete genome sequence of Escherichia coli K-12. *Science.*, 277, (5331): 1453-74.
- Bloom K (2001). Nuclear Migration: Cortical anchors for cytoplasmic dynein. *Cur Biol*, 11, (8): R326-9.
- Borden KL (2000). RING domains: master builders of molecular scaffolds? *J Mol Biol.*, 295, (5): 1103-12.
- Boulianne GL, de la Concha A, Campos-Ortega JA, Jan LY, Jan YN (1991). The Drosophila neurogenic gene neuralized encodes a novel protein and is expressed in precursors of larval and adult neurons. *EMBO J.*, 10, (10): 2975-83.
- Bradley A, Evans M, Kaufman MH, Robertson E (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature*, 309, (5965): 255-6.
- Bradley A (2002). Mining the mouse genome. *Nature*, 420, (6915): 512-4.
- Bray-Ward P, Bowlus C, Choi J, Paslier DL, Weissenbach J, Gruen JR. (1996). FISH-Mapped CEPH YACs spanning 0 to 46cM on human chromosome 6. *Genomics.*, 36, (1): 104-11.
- Brenner S, Elgar G, Sandford R, Macrae A, Venkatesh B, Aparicio S. (1993). Characterization of the pufferfish (Fugu) genome as a compact model vertebrate genome. *Nature.*, 366, (6452): 265-8.
- Burge CB Karlin S. (1998). Finding the genes in genomic DNA. *Curr Opin Struct Biol.*, 8, (3): 346-54.
- Burkhard P, Stetefeld J, Strelkov SV (2001). Coiled coils: a highly versatile protein folding motif. *Trends Cell Biol.*, 11, (2): 82-8.
- Burset M, Guigo R. (1996). Evaluation of gene structure prediction programs. *Genomics.*, 34, (3): 353-67.
- Camargo AA *et al.* (2001). The contribution of 700,000 ORF sequence tags to the definition of the human transcriptome. *Proc Natl Acad Sci U S A.*, 98, (21): 12103-8.
- Campos-Ortega JA (1988). Cellular interactions during early neurogenesis of Drosophila melanogaster. *Trends Neurosci.*, 11,(9):400-5.
- Carter S, Urbe S, Clague MJ. (2004). The met receptor degradation pathway: requirement for Lys48-linked polyubiquitin independent of proteasome activity. *J Biol Chem.*, 279, (51): 52835-9.
- Cayouette M, Whitmore AV, Jeffery G, Raff M (2001). Asymmetric segregation of Numb in retinal development and the influence of the pigmented epithelium. *J Neurosci.*, 21, (15): 5643-51.
- Cayouette M, Raff M (2002). Asymmetric segregation of Numb: a mechanism for neural specification from Drosophila to mammals. *Nat Neurosci.*, 5, (12): 1265-9.
- Chambers CB, Peng Y, Nguyen H, Gaiano N, Fishell G, Nye JS. (2001). Spatiotemporal selectivity of response to Notch1 signals in mammalian forebrain precursors. *Development*, 128, (5): 689-702.
- Chen W, Casey Corliss D (2004). Three modules of zebrafish Mind bomb work cooperatively to promote Delta ubiquitination and endocytosis. *Dev Biol.*, 267, (2): 361-73.

- Chitnis A, Henrique D, Lewis J, Ish-Horowicz D, Kintner C (1995). Primary neurogenesis in *Xenopus* embryos regulated by a homologue of the *Drosophila* neurogenic gene Delta. *Nature*, 375, (6534): 761-6.
- Clark MS, Smith SF, Elgar G. (2001). Use of the Japanese pufferfish (*Fugu rubripes*) in comparative genomics. *Mar Biotechnol* (NY). 3, (Supplement 1): S130-40.
- Cleveland DW (1995). NuMA: a protein involved in nuclear structure, spindle assembly, and nuclear re-formation. *Trends Cell Biol.*, 5, (2): 60-4.
- Cohen SN, Chang AC, Hsu L (1972). Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* by R-factor DNA. *Proc Natl Acad Sci U S A*, 69, (8): 2110-4.
- Compton DA, Cleveland DW (1994). NuMA, a nuclear protein involved in mitosis and nuclear reformation. *Curr Opin Cell Biol.*, 6, (3): 343-6.
- Cornell M, Evans DA, Mann R, Fostier M, Flasz M, Monthatong M, Artavanis-Tsakonas S, Baron M (1999). The *Drosophila melanogaster* Suppressor of deltex gene, a regulator of the Notch receptor signaling pathway, is an E3 class ubiquitin ligase. *Genetics*, 152, (2): 567-76.
- Craig JM, Bickmore WA (1994). The distribution of CpG islands in mammalian chromosomes. *Nat Genet.*, 7, (3): 376-82.
- Crepieux P, Kwon H, Leclerc N, Spencer W, Richard S, Lin R, Hiscott J (1997). I kappaB alpha physically interacts with a cytoskeleton-associated protein through its signal response domain. *Mol Cell Biol.*, 17, (12): 7375-85.
- Das AK, Cohen PTW, Barford D (1998). The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions. *EMBO J*, 15, (1192-1199):
- de Celis JF, Bray S, Garcia-Bellido A (1997). Notch signalling regulates veinlet expression and establishes boundaries between veins and interveins in the *Drosophila* wing. *Development*, 124, (10): 1919-28.
- de la Concha A, Dietrich U, Weigel D, Campos-Ortega JA (1988). Functional interactions of neurogenic genes of *Drosophila melanogaster*. *Genetics*, 118: 499-508.
- Deblandre GA, Wettstein DA, Koyano-Nakagawa N, Kintner C (1999). A two-step mechanism generates the spacing pattern of the ciliated cells in the skin of *Xenopus* embryos. *Development*, 126, (21): 4715-28.
- Deblandre GA, Lai EC, Kintner C (2001). *Xenopus* neuralized is a ubiquitin ligase that interacts with XDelta1 and regulates Notch signaling. *Dev Cell*, 1, (6): 795-806.
- Deloukas P, Schuler GD, Gyapay G, et al. (1998). A physical map of 30,000 human genes. *Science.*, 282, (5389): 744-6.
- Deloukas P *et al.* (2001). The DNA sequence and comparative analysis of human chromosome 20. *Nature.*, 414, (6866): 865-71.
- Deloukas P *et al.* (2004). The DNA sequence and comparative analysis of human chromosome 10. *Nature.*, 429, (6990): 375-81.
- Deng Y, Wu X (2000). Peg3/Pw1 promotes p53-mediated apoptosis by inducing Bax translocation from cytosol to mitochondria. *Proc Natl Acad Sci U S A.*, 97, (22): 12050-5.
- Dermitzakis ET, Reymond A, Scamuffa N, Ucla C, Kirkness E, Rossier C, Antonarakis SE. (2003). Evolutionary discrimination of mammalian conserved non-genic sequences (CNGs). *Science.*, 302, (5647): 1033-5.
- Dib C *et al.*(1996). A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature*, 380(6570): 152-4.
- Dick T, Ray K, Salz HK, Chia W (1996). Cytoplasmic dynein (*ddlc1*) mutations cause morphogenetic defects and apoptotic cell death in *Drosophila melanogaster*. *Mol Cell Biol.*, 16, (5): 1966-77.
- DiColandrea T, Karashima T, Maatta A, Watt FM (2000). Subcellular distribution of envoplakin and periplakin: insights into their role as precursors of the epidermal cornified envelope. *J Cell Biol.*, 151, (3): 573-86.
- Dower WJ, Miller JF, Ragsdale CW (1988). High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res.*, 16, (13): 6127-45.

- Du Q, Taylor L, Compton DA, Macara IG (2002). LGN blocks the ability of NuMA to bind and stabilize microtubules. A mechanism for mitotic spindle assembly regulation. *Curr Biol.*, 12, (22): 1928-33.
- Du Q, Macara IG (2004). Mammalian Pins is a conformational switch that links NuMA to heterotrimeric G proteins. *Cell*, 119, (4): 503-16.
- Dunham I, Shimizu N, Roe BA, et al. (1999). The DNA sequence of human chromosome 22. *Nature.*, 402, (6761): 489-95.
- Dunwoodie SL, Henrique D, Harrison SM, Beddington RS (1997). Mouse Dll3: a novel divergent Delta gene which may complement the function of other Delta homologues during early pattern formation in the mouse embryo. *Development*, 124, (16): 3065-76.
- Dupre S Urban-Grimal D, Haguenaer-Tsapis R (2004). Ubiquitin and endocytic internalization in yeast and animal cells. *Biochim Biophys Acta.*, 1695, (1-3): 89-111.
- Duprez E, Saurin AJ, Desterro JM, Lallemand-Breitenbach V, Howe K, Boddy MN, Solomon E, de The H, Hay RT, Freemont PS (1999). SUMO-1 modification of the acute promyelocytic leukaemia protein PML: implications for nuclear localisation. *J Cell Sci.*, 112, (Pt 3): 381-93.
- Epstein E, Sela-Brown A, Ringel I, Kilav R, King SM, Benashski SE, Yisraeli JK, Silver J, Naveh-Many T (2000). Dynein light chain binding to a 3'-untranslated sequence mediates parathyroid hormone mRNA association with microtubules. *J Clin Invest.*, 105, (4): 505-12.
- Espreafico EM, Coling DE, Tsakraklides V, Krogh K, Wolenski JS, Kalinec G, Kachar B (1998). Localization of myosin-V in the centrosome. *Proc Natl Acad Sci U S A*, 95, (15): 8636-41.
- Evans LL, Hammer J, Bridgman PC (1997). Subcellular localization of myosin V in nerve growth cones and outgrowth from dilute-lethal neurons. *J Cell Sci.*, 110, (Pt 4): 439-49.
- Fan J, Zhang Q, Tochio H, Li M, Zhang M (2001). Structural basis of diverse sequence-dependent target recognition by the 8 kDa dynein light chain. *J Mol Biol.*, 306, (1): 97-108.
- Faulkner NE Dujardin DL, Tai CY, Vaughan KT, O'Connell CB, Wang Y, Vallee RB (2000). A role for the lissencephaly gene LIS1 in mitosis and cytoplasmic dynein function. *Nat Cell Biol.*, 2, (11): 784-91.
- Ferhat L, Cook C, Kuriyama R, Baas PW (1998). The nuclear/mitotic apparatus protein NuMA is a component of the somatodendritic microtubule arrays of the neuron. *J Neurocytol.*, 27, (12): 887-99.
- Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM, et al. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science.*, 269, (5223): 496-512.
- Frangioni JV Neel BG (1993). Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. *Anal Biochem.*, 210, (1): 179-87.
- Fuks F, Burgers WA, Godin N, Kasai M, Kouzarides T (2001). Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. *EMBO J.*, 20, (10): 2536-44.
- Gaglio T, Saredi A, Compton DA (1995). NuMA is required for the organization of microtubules into aster-like mitotic arrays. *J Cell Biol.*, 131, (3): 693-708.
- Gervais V, Busso D, Wasielewski E, Poterszman A, Egly JM, Thierry JC, Kieffer B (2001). Solution structure of the N-terminal domain of the human TFIIF MAT1 subunit: new insights into the RING finger family. *J Biol Chem.*, 276, (10): 7457-64.
- Ghysen A, Dambly-Chaudiere C, Jan LY, Jan YN (1993). Cell interactions and gene interactions in peripheral neurogenesis. *Genes Dev.*, 7, (5): 723-33.
- Gill G (2004). SUMO and ubiquitin in the nucleus: different functions, similar mechanisms? *Genes Dev.*, 18, (17): 2046-59.
- Gilligan P, Brenner S, Venkatesh B. (2002). Fugu and human sequence comparison identifies novel human genes and conserved non-coding sequences. *Gene.*, 294, (1-2): 35-44.

- Goodwin RL, Pabon-Pena LM, Foster GC, Bader D (1999). The cloning and analysis of LEK1 identifies variations in the LEK/centromere protein F/mitosin gene family. *J Biol Chem.*, 274, (26): 18597-604.
- Gopal S, Schroeder M, Pieper U, Sczyrba A, Aytekin-Kurban G, Bekiranov S, Fajardo JE, Eswar N, Sanchez R, Sali A, Gaasterland T. (2001). Homology-based annotation yields 1,042 new candidate genes in the *Drosophila melanogaster* genome. *Nat Genet.*, 27, (3): 337-40.
- Graham FL, van der Eb AJ (1973). Transformation of rat cells by DNA of human adenovirus 5. *Virology*, 54, (2): 536-9.
- Greenwald I (1998). LIN-12/Notch signaling: lessons from worms and flies. *Genes Dev.*, 12, (12): 1751-62.
- Grovdal LM, Stang E, Sorkin A, Madshus IH. (2004). Direct interaction of Cbl with pTyr 1045 of the EGF receptor (EGFR) is required to sort the EGFR to lysosomes for degradation. *Exp Cell Res.*, 300, (2): 388-95.
- Grunstein M, Hogness DS (1975). Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. *Proc Natl Acad Sci U S A*, 72, (10): 3961-5.
- Gueth-Hallonet C, Weber K, Osborn M (1997). Cleavage of the nuclear matrix protein NuMA during apoptosis. *Exp Cell Res*, 233,(1):21-4.
- Guthrie C, Fink GR. (1991). *Guide to Yeast Genetics and Molecular Biology; Methods in Enzymology*, vol:169, Academic Press, San Diego.
- Harborth J, Wang J, Gueth-Hallonet C, Weber K, Osborn M (1999). Self assembly of NuMA: multiarm oligomers as structural units of a nuclear lattice. *EMBO J.*, 18, (6): 1689-700.
- Hartenstein V, Posakony JW (1990). A dual function of the Notch gene in *Drosophila* sensillum development. *Dev Biol.*,142, (1): 13-30.
- Hattori M, Fujiyama A, Taylor TD, Watanabe H; Chromosome 21 mapping and sequencing consortium. (2000). The DNA sequence of human chromosome 21. *Nature.*, 405, (6784): 311-9.
- Hayashi T, Kageyama Y, Ishizaka K, Xia G, Kihara K, Oshima H (2001). Requirement of Notch 1 and its ligand jagged 2 expressions for spermatogenesis in rat and human testes. *J Androl.*, 22, (6): 999-1011.
- Hayashi T, Seki M, Maeda D, Wang W, Kawabe Y, Seki T, Saitoh H, Fukagawa T, Yagi H, Enomoto T (2002). Ubc9 is essential for viability of higher eukaryotic cells. *Exp Cell Res.*, 280, (2): 212-21.
- Heilig R *et al.* (2003). The DNA sequence and analysis of human chromosome 14. *Nature.*, 421, (6923): 601-7.
- Hershko A, Ciechanover A (1998). The ubiquitin system. *Annu Rev Biochem.*, 67, (425-79):
- Herzig RP, Andersson U, Scarpulla RC (2000). Dynein light chain interacts with NRF-1 and EWG, structurally and functionally related transcription factors from humans and *drosophila*. *J Cell Sci.*, 113, (Pt 23): 4263-73.
- Hicke L, Dunn R. (2003). Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins. *Annu Rev Cell Dev Biol.*, 19: 141-72
- Hillier LD, Lennon G, Becker M, Bonaldo MF, et al. (1996). Generation and analysis of 280,000 human expressed sequence tags. *Genome Res.*, 6, (9): 807-28.
- Hillier LW *et al.* (2003). The DNA sequence of human chromosome 7. *Nature.*, 424, (6945): 157-64.
- Hirokawa N (1998). Kinesin and dynein superfamily proteins and the mechanism of organelle transport. *Science*, 279, (5350): 519-26.
- Hollenberg SM, Sternglanz R, Cheng PF, Weintraub H (1995). Identification of a new family of tissue-specific basic helix-loop-helix proteins with a two-hybrid system. *Mol Cell Biol.*, 15, (7): 3813-22.
- Holt RA *et al.* (2002). The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science.*, 298, (5591): 129-49.
- Hood JK, Silver PA (1999). In or out? Regulating nuclear transport. *Curr Opin Cell Biol.*, 11, (2): 241-7.

- Hu G, Zhang S, Vidal M, Baer JL, Xu T, Fearon ER. (1997). Mammalian homologs of seven in absentia regulate DCC via the ubiquitin-proteasome pathway. *Genes Dev.*, 11, (20): 2701-14.
- Hubbard T *et al.* (2002). The Ensembl genome database project. *Nucleic Acids Res.*, 30, (1): 38-41.
- Hudson TJ, Stein LD, Gerety SS, Ma J, Castle AB, Silva J, Slonim DK, Baptista R, Kruglyak L, Xu SH, et al. (1995). An STS-based map of the human genome. *Science.*, 270, (5244): 1945-54.
- Huppert SS, Jacobsen TL, Muskavitch MA (1997). Feedback regulation is central to Delta-Notch signalling required for *Drosophila* wing vein morphogenesis. *Development*, 124, (17): 3283-91.
- Istrail S *et al.* (2004). Whole-genome shotgun assembly and comparison of human genome assemblies. *Proc Natl Acad Sci U S A*, 101, (7): 1916-21.
- Itoh M, Kim CH, Palardy G, Oda T, Jiang YJ, Maust D, Yeo SY, Lorick K, Wright GJ, Ariza-McNaughton L, Weissman AM, Lewis J, Chandrasekharappa SC, Chitnis AB (2003). Mind bomb is a ubiquitin ligase that is essential for efficient activation of Notch signaling by Delta. *Dev Cell.*, 4, (1): 67-82.
- Jackson PK, Eldridge AG, Freed E, Furstenenthal L, Hsu JY, Kaiser BK, Reimann JD (2000). The lore of the RINGs: substrate recognition and catalysis by ubiquitin ligases. *Trends Cell Biol.*, 10, (10): 429-39.
- Jacob Y, Badrane H, Ceccaldi PE, Tordo N (2000). Cytoplasmic dynein LC8 interacts with lyssavirus phosphoprotein. *J Virol.*, 74, (21): 10217-22.
- Jaffrey SR, Snyder SH (1996) PIN: an associated protein inhibitor of neuronal nitric oxide synthase. *Science*, 274, (5288): 774-7.
- James P, Halladay J, Craig EA (1996). Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics*, 144, (4): 1425-36.
- Jennings B, Preiss A, Delidakis C, Bray S (1994). The Notch signalling pathway is required for Enhancer of split bHLH protein expression during neurogenesis in the *Drosophila* embryo. *Development*, 120, (12): 3537-48.
- Jin Y, Blue EK, Dixon S, Shao Z, Gallagher PJ (2002). A death-associated protein kinase (DAPK)-interacting protein, DIP-1, is an E3 ubiquitin ligase that promotes tumor necrosis factor-induced apoptosis and regulates the cellular levels of DAPK. *J Biol Chem.*, 277, (49): 46980-6.
- Joazeiro CA, Wing SS, Huang H, Levenson JD, Hunter T, Liu YC (1999). The tyrosine kinase negative regulator c-Cbl as a RING-type, E2-dependent ubiquitin-protein ligase. *Science*, 286, (5438): 309-12.
- Joazeiro CA, Weissman AM (2000). RING finger proteins: mediators of ubiquitin ligase activity. *Cell*, 102, (5): 549-52.
- Johnson ES, Blobel G (1999). Cell cycle-regulated attachment of the ubiquitin-related protein SUMO to the yeast septins. *J Cell Biol.*, 147, (5): 981-94.
- Johnson ES (2004). Protein modification by SUMO. *Annu Rev Biochem.*, 73, (355-82):
- Joseph J, Tan SH, Karpova TS, McNally JG, Dasso M (2002). SUMO-1 targets RanGAP1 to kinetochores and mitotic spindles. *J Cell Biol.*, 156, (4): 595-602.
- Kahyo T, Nishida T, Yasuda H (2001). Involvement of PIAS1 in the sumoylation of tumor suppressor p53. *Mol Cell*, 8, (3): 713-8.
- Kaiser FJ, Tavassoli K, Van den Bemd GJ, Chang GT, Horsthemke B, Moroy T, Ludecke HJ (2003). Nuclear interaction of the dynein light chain LC8a with the TRPS1 transcription factor suppresses the transcriptional repression activity of TRPS1. *Hum Mol Genet.*, 12, (11): 1349-58.
- Kalinin AE, Idler WW, Marekov LN, McPhie P, Bowers B, Steinert PM, Steven AC (2004). Co-assembly of envoplakin and periplakin into oligomers and Ca(2+)-dependent vesicle binding: implications for cornified cell envelope formation in stratified squamous epithelia. *J Biol Chem.*, 279, (21): 22773-80.
- Karki S, Holzbaaur EL(1999). Cytoplasmic dynein and dynactin in cell division and intracellular transport. *Curr Opin Cell Biol*, 11, (1)45-53.

- Kasai H, Isono S, Kitakawa M, Mineno J, Akiyama H, Kurnit DM, Berg DE, Isono K. (1992). Efficient large-scale sequencing of the *Escherichia coli* genome: implementation of a transposon- and PCR-based strategy for the analysis of ordered lambda phage clones. *Nucleic Acids Res.*, 20, (24): 6509-15.
- King SM, Barbaresse E, Dillman JF III, Patel-King RS, Carson JH, Pfister KK (1996). Brain cytoplasmic and flagellar outer arm dyneins share a highly conserved Mr 8,000 light chain. *J Biol Chem.*, 271, (32): 19358-66.
- King SM (2000). The dynein microtubule motor. *Biochim Biophys Acta*, 1496, (1): 60-75.
- King SM (2003). Organization and regulation of the dynein microtubule motor. *Cell Biol Int.*, 27, (3): 213-5.
- Knoblich JA (2001). Asymmetric cell division during animal development. *Nat Rev Mol Cell Biol.*, 2, (1): 11-20.
- Kohda T, Asai A, Kuroiwa Y, Kobayashi S, Aisaka K, Nagashima G, Yoshida MC, Kondo Y, Kagiya N, Kirino T, Kaneko-Ishino T, Ishino F (2001). Tumour suppressor activity of human imprinted gene PEG3 in a glioma cell line. *Genes Cells*, 6, (3): 237-47.
- Kracklauer MP, Schmidt C (2003). At the crossroads of SUMO and NF-kappaB. *Mol Cancer*, 2, (1): 39.
- Kumar S, Tamura K, Jakobsen IB, Nei M (2001). MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics*, 17, (12): 1244-5.
- Kurtzman AL, Schechter N (2001). Ubc9 interacts with a nuclear localization signal and mediates nuclear localization of the paired-like homeobox protein Vsx-1 independent of SUMO-1 modification. *Proc Natl Acad Sci U S A.*, 98, (10): 5602-7.
- Lai EC, Deblandre GA, Kintner C, Rubin GM (2001). *Drosophila* neuralized is a ubiquitin ligase that promotes the internalization and degradation of delta. *Dev Cell*, 1, (6): 783-94.
- Lai EC, Rubin GM (2001a). Neuralized is essential for a subset of Notch pathway-dependent cell fate decisions during *Drosophila* eye development. *Proc Natl Acad Sci U S A*, 98, (10): 5637-42.
- Lai EC, Rubin GM (2001b). neuralized functions cell-autonomously to regulate a subset of notch-dependent processes during adult *Drosophila* development. *Dev Biol.*, 231, (1): 217-33.
- Lai EC (2002). Protein degradation: four E3s for the notch pathway. *Curr Biol.*, 12, (2): R74-8.
- Lallemand-Breitenbach V, Zhu J, Puvion F, Koken M, Honore N, Doubeikovsky A, Duprez E, Pandolfi PP, Puvion E, Freemont P, de Thé H (2001). Role of promyelocytic leukemia (PML) sumolation in nuclear body formation, 11S proteasome recruitment, and As2O3-induced PML or PML/retinoic acid receptor alpha degradation. *J Exp Med.*, 193, (12): 1361-71.
- Lander ES, Linton LM, Birren B; International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. *Nature.*, 409, (6822): 860-921.
- Le Borgne R, Schweisguth F (2003a). Unequal segregation of Neuralized biases Notch activation during asymmetric cell division. *Dev Cell*, 5, (1): 139-48.
- Le Borgne R, Schweisguth F (2003b). Notch signaling: endocytosis makes delta signal better. *Curr Biol.*, 13, (7): R273-5.
- Lehmann R, Jimenez F, Dietrich U, Campos-Ortega, JA (1983). On the phenotype and development of mutants of early neurogenesis in *Drosophila melanogaster*. *Roux's Arch. Dev. Biol.*, 192: 62-74.
- Leung CL, Liem RK, Parry DA, Green KJ (2001). The plakin family. *J Cell Sci.*, 114, (Pt 19): 3409-10.
- Levkowitz G, Waterman H, Ettenberg SA, Katz M, Tsygankov AY, Alroy I, Lavi S, Iwai K, Reiss Y, Ciechanover A, Lipkowitz S, Yarden Y (1999). Ubiquitin ligase activity and tyrosine phosphorylation underlie suppression of growth factor signaling by c-Cbl/Sli-1. *Mol Cell.*, 4, (6): 1029-40.
- Li B, Dou QP (2000). Bax degradation by the ubiquitin/proteasome-dependent pathway: involvement in tumor survival and progression. *Proc Natl Acad Sci U S A*, 97, (8): 3850-5.
- Li L, Keverne EB, Aparicio SA, Ishino F, Barton SC, Surani MA (1999). Regulation of maternal behavior and offspring growth by paternally expressed Peg3. *Science*, 284, (5412): 330-3.

- Lindsell CE, Shawber CJ, Boulter J, Weinmaster G (1995). Jagged: a mammalian ligand that activates Notch1. *Cell*, 80, (6): 909-17.
- Ling Y, Sankpal UT, Robertson AK, McNally JG, Karpova T, Robertson KD (2004). Modification of de novo DNA methyltransferase 3a (Dnmt3a) by SUMO-1 modulates its interaction with histone deacetylases (HDACs) and its capacity to repress transcription. *Nucleic Acids Res.*, 32, (2): 598-610.
- Lohse P, Lohse P, Chahrokh-Zadeh S, Seidel D (1997). Human lysosomal acid lipase/cholesteryl ester hydrolase and human gastric lipase: site-directed mutagenesis of Cys227 and Cys236 results in substrate-dependent reduction of enzymatic activity. *J Lipid Res.*, 38, (9): 1896-905.
- Lopez-Schier H, St Johnston D (2001). Delta signaling from the germ line controls the proliferation and differentiation of the somatic follicle cells during *Drosophila* oogenesis. *Genes Dev.*, 15, (11): 1393-405.
- Lorick KL, Jensen JP, Fang S, Ong AM, Hatakeyama S, Weissman AM. (1999). RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proc Natl Acad Sci U S A.*, 96, (20): 11364-9.
- Lund EG, Kerr TA, Sakai J, Li WP, Russell DW (1998). cDNA cloning of mouse and human cholesterol 25-hydroxylases, polytopic membrane proteins that synthesize a potent oxysterol regulator of lipid metabolism. *J Biol Chem.*, 273, (51): 34316-27.
- Maatta A, Ruhrberg C, Watt FM (2000). Structure and regulation of the envoplakin gene. *J Biol Chem.*, 275, (26): 19857-65.
- Maegawa S, Yoshioka H, Itaba N, Kubota N, Nishihara S, Shirayoshi Y, Nanba E, Oshimura M (2001). Epigenetic silencing of PEG3 gene expression in human glioma cell lines. *Mol Carcinog.*, 31, (1): 1-9.
- Mahajan R, Delphin C, Guan T, Gerace L, Melchior F (1997). A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell*, 88, (1): 97-107.
- Martinez-Noel G, Muller U, Harbers K (2001). Identification of molecular determinants required for interaction of ubiquitin-conjugating enzymes and RING finger proteins. *Eur J Biochem.*, 268, (22): 5912-9.
- Marzella R, Kokkinaki MA, Kapsetaki M, Ricco A, Argyrokastritis A, Kamakari S, Sarafidou T, Archidiacono N, Roussou A, Pasparaki A, Rocchi M, Moschonas NK. (1997). Map integration at human chromosome 10: molecular and cytogenetic analysis of a chromosome-specific somatic cell hybrid panel and genomic clones, based on a well-supported genetic map. *Cytogenet Cell Genet.*, 79,(3-4):257-65.
- Matunis MJ, Wu J, Blobel G. (1998). SUMO-1 modification and its role in targeting the Ran GTPase-activating protein, RanGAP1, to the nuclear pore complex. *J Cell Biol.*, 140, (3): 499-509.
- McCabe JB, Berthiaume LG (2001). N-terminal protein acylation confers localization to cholesterol, sphingolipid-enriched membranes but not to lipid rafts/caveolae. *Mol Biol Cell*, 12, (11): 3601-17.
- Meng W, Sawadkisol S, Burakoff SJ, Eck MJ (1999). Structure of the amino-terminal domain of Cbl complexed to its binding site on ZAP-70 kinase. *Nature*, 398, (6722): 84-90.
- Merdes A, Ramyar K, Vechio JD, Cleveland DW (1996). A complex of NuMA and cytoplasmic dynein is essential for mitotic spindle assembly. *Cell*, 87, (3): 447-58.
- Merdes A, Heald R, Samejima K, Earnshaw WC, Cleveland DW (2000). Formation of spindle poles by dynein/dynactin-dependent transport of NuMA. *J Cell Biol.*, 149, (4): 851-62.
- Mingot JM, Kostka S, Kraft R, Hartmann E, Gorlich D (2001). Importin 13: a novel mediator of nuclear import and export. *EMBO J.*, 20, (14): 3685-94.
- Mishra RK, Jatiani SS, Kumar A, Simhadri VR, Hosur RV, Mittal R (2004). Dynamin interacts with members of the sumoylation machinery. *J Biol Chem.*, 279, (30): 31445-54.
- Modrek B, Resch A, Grasso C, Lee C. (2001). Genome-wide detection of alternative splicing in expressed sequences of human genes. *Nucleic Acids Res.*, 29, (13): 2850-9.
- Muller S, Matunis MJ, Dejean A (1998). Conjugation with the ubiquitin-related modifier SUMO-1 regulates the partitioning of PML within the nucleus. *EMBO J.*, 17, (1): 61-70.

- Muller S, Berger M, Lehembre F, Seeler JS, Haupt Y, Dejean A. (2000). c-Jun and p53 activity is modulated by SUMO-1 modification. *J Biol Chem.*, 275, (18): 13321-9.
- Muller U. (1999). Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis. *Mech Dev*, 82 (1-2):3-21.
- Mungall AJ *et al.* (2003). The DNA sequence and analysis of human chromosome 6. *Nature*, 425 (6960): 805-11.
- Muratani M, Tansey WP. (2003). How the ubiquitin-proteasome system controls transcription. *Nat Rev Mol Cell Biol.*, 4, (3): 192-201.
- Myers EW, Sutton GG, Smith HO, Adams MD, Venter JC. (2002). On the sequencing and assembly of the human genome. *Proc Natl Acad Sci U S A.*, 99, (7): 4145-6.
- Naisbitt S, Valtschanoff J, Allison DW, Sala C, Kim E, Craig AM, Weinberg RJ, Sheng M (2000). Interaction of the postsynaptic density-95/guanylate kinase domain-associated protein complex with a light chain of myosin-V and dynein. *J Neurosci.*, 20, (12): 4524-34.
- Nakamura H, Yoshida M, Tsuiki H, Ito K, Ueno M, Nakao M, Oka K, Tada M, Kochi M, Kuratsu J, Ushio Y, Saya H (1998). Identification of a human homolog of the *Drosophila* neuralized gene within the 10q25.1 malignant astrocytoma deletion region. *Oncogene*, 16, (8): 1009-19.
- Nastasi T, Bongiovanni A, Campos Y, Mann L, Toy JN, Bostrom J, Rottier R, Hahn C, Conaway JW, Harris AJ, D'Azzo A. (2004). Ozz-E3, a muscle-specific ubiquitin ligase, regulates beta-catenin degradation during myogenesis. *Dev Cell.*, 6, (2): 269-82.
- Nie J, McGill MA, Dermer M, Dho SE, Wolting CD, McClade CJ (2002). LNX functions as a RING type E3 ubiquitin ligase that targets the cell fate determinant Numb for ubiquitin-dependent degradation. *EMBO J.*, 21, (1-2): 93-102.
- Nirenberg M, Caskey T, Marshall R, Brimacombe R, Kellogg D, Doctor B, Hatfield D, Levin J, Rottman F, Pestka S, Wilcox M, Anderson F (1966). The RNA code and protein synthesis. *Cold Spring Harb Symp Quant Biol.*, 31: 11-24.
- Oda H, Stockert RJ, Collins C, Wang H, Novikoff PM, Satir P, Wolkoff AW (1995). Interaction of the microtubule cytoskeleton with endocytic vesicles and cytoplasmic dynein in cultured rat hepatocytes. *J Biol Chem.*, 270, (25): 15242-9.
- Okano M, Bell DW, Haber DA, Li E (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*, 99, (3): 247-57.
- Okazaki Y, *et al* ;FANTOM Consortium; RIKEN Genome Exploration Research Group Phase I & II Team. (2002). Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature.*, 420, (6915): 563-73.
- Osoegawa K, Mammoser AG, Wu C, Frengen E, Zeng C, Catanese JJ, de Jong PJ (2001). A bacterial artificial chromosome library for sequencing the complete human genome. *Genome Res.*, 11, (3): 483-96.
- Pabon-Pena LM, Goodwin RL, Cise LJ, Bader D (2000). Analysis of CMF1 reveals a bone morphogenetic protein-independent component of the cardiomyogenic pathway. *J Biol Chem.*, 275, (28): 21453-9.
- Parks AL, Klueg KM, Stout JR, Muskavitch MA (2000). Ligand endocytosis drives receptor dissociation and activation in the Notch pathway. *Development*, 127, (7): 1373-85.
- Patanjali SR, Parimoo S, Weissman SM. (1991). Construction of a uniform-abundance (normalized) cDNA library. *Proc Natl Acad Sci U S A.*, 88, (5): 1943-7.
- Pavlopoulos E, Pitsouli C, Klueg KM, Muskavitch MA, Moschonas NK, Delidakis C (2001). neuralized Encodes a peripheral membrane protein involved in delta signaling and endocytosis. *Dev Cell*, 1, (6): 807-16.
- Pavlopoulos E, Kokkinaki M, Koutelou E, Mitsiadis TA, Prinios P, Delidakis C, Kilpatrick MW, Tsiouras P, Moschonas NK (2002). Cloning, chromosomal organization and expression analysis of *Neurl*, the mouse homolog of *Drosophila melanogaster* neuralized gene. *Biochim Biophys Acta*, 1574, (3): 375-82.
- Peng YF, Mandai K, Sakisaka T, Okabe N, Yamamoto Y, Yokoyama S, Mizoguchi A, Shiozaki H, Monden M, Takai Y (2000). Ankyrcorbin: a novel actin cytoskeleton-associated protein. *Genes Cells*, 5, (12): 1001-8.

- Penn SG, Rank DR, Hanzel DK, Barker DL. (2000). Mining the human genome using microarrays of open reading frames. *Nat Genet.*, 26, (3): 315-8.
- Pichler A, Gast A, Seeler JS, Dejean A, Melchior F (2002). The nucleoporin RanBP2 has SUMO1 E3 ligase activity. *Cell*, 108,(1): 109-20.
- Price BD, Chang Z, Smith R, Bockheim S, Laughon A (1993). The *Drosophila* neuralized gene encodes a C3HC4 zinc finger. *EMBO J.*, 12, (6): 2411-8.
- Price CM, Pettijohn DE (1986). Redistribution of the nuclear mitotic apparatus protein (NuMA) during mitosis and nuclear assembly. *Exp. Cell Res*, 166, (295-311):
- Puthalakath H, Huang DC, O'Reilly LA, King SM, Strasser A (1999). The proapoptotic activity of the Bcl-2 family member Bim is regulated by interaction with the dynein motor complex. *Mol Cell*, 3, (3): 287-96.
- Qiu L, Joazeiro C, Fang N, Wang HY, Elly C, Altman Y, Fang D, Hunter T, Liu YC (2000). Recognition and ubiquitination of Notch by Itch, a hect-type E3 ubiquitin ligase. *J Biol Chem.*, 275, (46): 35734-7.
- Raux H, Flamand A, Blondel D (2000). Interaction of the rabies virus P protein with the LC8 dynein light chain. *J Virol.*, 74, (21):10212-6.
- Read MA, Brownell JE, Gladysheva TB, Hottelot M, Parent LA, Coggins MB, Pierce JW, Podust VN, Luo RS, Chau V, Palombella VJ. (2000). Ned8 modification of cul-1 activates SCF(betaTrCP)-dependent ubiquitination of I kappaBalpha. *Mol Cell Biol.*, 20, (7): 2326-33.
- Reboul J *et al.* (2001). Open-reading-frame sequence tags (OSTs) support the existence of at least 17,300 genes in *C. elegans*. *Nat Genet.*, 27, (3): 332-6.
- Rechsteiner M, Rogers SW (1996). PEST sequences and regulation by proteolysis. *Trends Biochem Sci.*, 21, (7): 267-71.
- Reck-Peterson SL, Provance DW Jr, Mooseker MS, Mercer JA (2000). Class V myosins. *Biochim Biophys Acta*, 1496, (1): 36-51.
- Redonnet-Vernhet I, Chatelut M, Basile JP, Salvayre R, Levade T (1997). Cholesteryl ester storage disease: relationship between molecular defects and in situ activity of lysosomal acid lipase. *Biochem Mol Med.*, 62, (1): 42-9.
- Relaix F, Weng X, Marazzi G, Yang E, Copeland N, Jenkins N, Spence SE, Sassoon D (1996). Pw1, a novel zinc finger gene implicated in the myogenic and neuronal lineages. *Dev Biol.*, 177, (2): 383-96.
- Relaix F, Wei XJ, Wu X, Sassoon DA (1998). Peg3/Pw1 is an imprinted gene involved in the TNF-NFkappaB signal transduction pathway. *Nat Genet.*, 18, (3): 287-91.
- Relaix F, Wei X, Li W, Pan J, Lin Y, Bowtell DD, Sassoon DA, Wu X (2000). Pw1/Peg3 is a potential cell death mediator and cooperates with Siah1a in p53-mediated apoptosis. *Proc Natl Acad Sci U S A*, 97, (5): 2105-10.
- Riordan JR, Rommens JM, Kerem B, Alon N, Rozmahel R, Grzelczak Z, Zielenski J, Lok S, Plavsic N, Chou JL, et al. (1989). Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. *Science.*, 245, (4922): 1066-73.
- Robertson KD (2002). DNA methylation and chromatin - unraveling the tangled web. *Oncogene*, 21, (35): 5361-79.
- Robinson PA, Ardley HC (2004). Ubiquitin-protein ligases. *J Cell Sci.*, 117, (Pt 22): 5191-4.
- Rodriguez MS, Desterro JM, Lain S, Midgley CA, Lane DP, Hay RT. (1999). SUMO-1 modification activates the transcriptional response of p53. *EMBO J.*, 18, (22): 6455-61.
- Roegiers F, Jan YN (2004). Asymmetric cell division. *Curr Opin Cell Biol.*, 16, (2): 195-205.
- Roest Crollius H *et al.* (2000). Estimate of human gene number provided by genome-wide analysis using *Tetraodon nigroviridis* DNA sequence. *Nat Genet.*, 25, (2): 235-8.
- Rosenfeld MA *et al.* (1992). In vivo transfer of the human cystic fibrosis transmembrane conductance regulator gene to the airway epithelium. *Cell.*, 68, (1): 143-55.

- Ruan Y, Tecott L, Jiang MM, Jan LY, Jan YN (2001). Ethanol hypersensitivity and olfactory discrimination defect in mice lacking a homolog of *Drosophila* neuralized. *Proc Natl Acad Sci U S A*, 98, (17): 9907-12.
- Rui L, Yuan M, Frantz D, Shoelson S, White MF. (2002). SOCS-1 and SOCS-3 block insulin signaling by ubiquitin-mediated degradation of IRS1 and IRS2. *J Biol Chem.*, 277, (44): 42394-8.
- Rupp G, Porter ME (2003). A subunit of the dynein regulatory complex in *Chlamydomonas* is a homologue of a growth arrest-specific gene product. *J Cell Biol.*, 162, (1): 47-57.
- Saitoh H, Sparrow DB, Shiomi T, Pu RT, Nishimoto T, Mohun TJ, Dasso M (1998). Ubc9p and the conjugation of SUMO-1 to RanGAP1 and RanBP2. *Curr Biol.*, 8, (2): 121-4.
- Sambrook J, Fritsch EF, Maniatis T. (1989). *Molecular cloning : A laboratory manual*. Cold Spring Harbor Laboratory Press.
- Sampson DA, Wang M, Matunis MJ (2001). The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. *J Biol Chem.*, 276, (24): 21664-9.
- Sanger F, Nicklen S, Coulson AR. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A.*, 74, (12): 5463-7.
- Sanger F, Coulson AR, Hong GF, Hill DF, Petersen GB. (1982). Nucleotide sequence of bacteriophage lambda DNA. *J Mol Biol.*, 162, (4): 729-73.
- Sauer B, Henderson N. (1988). Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1. *Proc Natl Acad Sci U S A.*, 85, (14): 5166-70.
- Sbodio JI, Lodish HF, Chi NW (2002). Tankyrase-2 oligomerizes with tankyrase-1 and binds to both TRF1 (telomere-repeat-binding factor 1) and IRAP (insulin-responsive aminopeptidase). *Biochem J.*, 361, (Pt 3): 451-9.
- Schneider C, King RM, Philipson L (1988). Genes specifically expressed at growth arrest of mammalian cells. *Cell*, 54, (6): 787-93.
- Schwartz DC, Cantor C (1984). Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell*, 37, (1): 67-75.
- Schweisguth F (1999). Dominant-negative mutation in the beta2 and beta6 proteasome subunit genes affect alternative cell fate decisions in the *Drosophila* sense organ lineage. *Proc Natl Acad Sci U S A*, 96, (20): 11382-6.
- Seufert W, Futcher B, Jentsch S (1995). Role of a ubiquitin-conjugating enzyme in degradation of S- and M-phase cyclins. *Nature*, 373, (6509): 78-81.
- Sharp PM, Matassi G (1994). Codon usage and genome evolution. *Curr Opin Genet Dev.*, 4, (6): 851-60.
- Shawber C, Boulter J, Lindsell CE, Weinmaster G (1996). Jagged2: a serrate-like gene expressed during rat embryogenesis. *Dev Biol.*, 180, (1): 370-6.
- Shcherbik N, Haines DS. (2004). Ub on the move. *J Cell Biochem.*, 93, (1): 11-9.
- Shen Q, Zhong W, Jan YN, Temple S (2002). Asymmetric Numb distribution is critical for asymmetric cell division of mouse cerebral cortical stem cells and neuroblasts. *Development*, 129, (20): 4843-53.
- Shoemaker DD *et al.* (2001). Experimental annotation of the human genome using microarray technology. *Nature.*, 409, (6822): 922-7.
- Shutter JR, Scully S, Fan W, Richards WG, Kitajewski J, Deblandre GA, Kintner CR, Stark KL (2000). Dll4, a novel Notch ligand expressed in arterial endothelium. *Genes Dev.*, 14, (11): 1313-8.
- Skaletsky H *et al.* (2003). The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature.*, 423, (6942): 825-37.
- Smith JB, Nguyen TT, Hughes HJ, Herschman HR, Widney DP, Bui KC, Rovai LE. (2002). Glucocorticoid-attenuated response genes induced in the lung during endotoxemia. *Am J Physiol Lung Cell Mol Physiol.*, 283, (3): L636-47.

- Soares MB, Bonaldo MF, Jelene P, Su L, Lawton L, Efstratiadis A. (1994). Construction and characterization of a normalized cDNA library. *Proc Natl Acad Sci U S A.*, 91, (20): 9228-32.
- Southern EM (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol.*, 98, (3):503-17.
- Starke HR, Yan JY, Zhang JZ, Muhlegger K, Effgen K, Dovichi NJ. (1994). Internal fluorescence labeling with fluorescent deoxynucleotides in two-label peak-height encoded DNA sequencing by capillary electrophoresis. *Nucleic Acids Res.*, 22, (19): 3997-4001.
- Strausberg RL, Feingold EA, Grouse LH; Mammalian Gene Collection Program Team. (2002). Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc Natl Acad Sci U S A.*, 99, (26): 16899-903.
- Struhl G, Adachi A (1998). Nuclear access and action of notch in vivo. *Cell*, 93, (4): 649-60.
- Su HL, Li SS (2002). Molecular features of human ubiquitin-like SUMO genes and their encoded proteins. *Gene*, 296, (1-2): 65-73.
- Tabb JS, Molyneaux BJ, Cohen DL, Kuznetsov S, Langford G (1998) Transport of ER vesicles on actin filaments in neurons by myosin V. *J Cell Sci.*, 111, (Pt 21): 3221-34.
- Taimen P, Viljamaa M, Kallajoki M (2000). Preferential expression of NuMA in the nuclei of proliferating cells. *Exp Cell Res.*, 256, (1): 140-9.
- Taimen P, Kallajoki M (2003). NuMA and nuclear lamins behave differently in Fas-mediated apoptosis. *J Cell Sci.*, 116, (Pt 3): 571-83.
- Taimen P, Parvinen M, Osborn M, Kallajoki M. (2004). NuMA in rat testis--evidence for roles in proliferative activity and meiotic cell division. *Exp Cell Res.*, 298, (2): 512-20.
- Tang CJ, Hu HM, Tang TK (2004). NuMA expression and function in mouse oocytes and early embryos. *J Biomed Sci.*, 11, (3): 370-6.
- Tang TK, Tang CJ, Chao YJ, Wu CW (1994). Nuclear mitotic apparatus protein (NuMA): spindle association, nuclear targeting and differential subcellular localization of various NuMA isoforms. *J Cell Sci.*, 107, (Pt 6): 1389-402.
- The International HapMap Consortium (2003). The International HapMap Project. *Nature.*, 426, (6968): 789-96.
- Timmusk T, Palm K, Belluardo N, Mudo G, Neuman T (2002). Dendritic localization of mammalian neuralized mRNA encoding a protein with transcription repression activities. *Mol Cell Neurosci.*, 20, (4): 649-68.
- Tyers M, Willems AR. (1999). One ring to rule a superfamily of E3 ubiquitin ligases. *Science.*, 284, (5414): 601, 03-4.
- Tyers M, Jorgensen P (2000). Proteolysis and the cell cycle: with this RING I do thee destroy. *Curr Opin Genet Dev.*, 10, (1): 54-64.
- Ulrich HD (2003). Protein-protein interactions within an E2-RING finger complex. Implications for ubiquitin-dependent DNA damage repair. *J Biol Chem.*, 278, (9): 7051-8.
- Venkatesh B, Erdmann MV, Brenner S. (2001). Molecular synapomorphies resolve evolutionary relationships of extant jawed vertebrates. *Proc Natl Acad Sci U S A.*, 98, (20): 11382-7.
- Venter JC *et al.* (2001). The sequence of the human genome. *Science*, 291, (5507):1304-51.
- Verger A, Perdomo J, Crossley M (2003). Modification with SUMO. A role in transcriptional regulation. *EMBO Rep.*, 4, (2): 137-42.
- Viard I, Wehrli P, Bullani R, Schneider P, Holler N, Salomon D, Hunziker T, Saurat JH, Tschopp J, French LE (1998). Inhibition of toxic epidermal necrolysis by blockade of CD95 with human intravenous immunoglobulin. *Science*, 282, (5388): 490-3.

- Vollrath B, Pudney J, Asa S, Leder P, Fitzgerald K (2001). Isolation of a murine homologue of the *Drosophila* neuralized gene, a gene required for axonemal integrity in spermatozoa and terminal maturation of the mammary gland. *Mol Cell Biol.*, 21, (21): 7481-94.
- Wakamatsu Y, Maynard TM, Jones SU, Weston JA (1999). NUMB localizes in the basal cortex of mitotic avian neuroepithelial cells and modulates neuronal differentiation by binding to NOTCH-1. *Neuron*, 23, (1): 71-81.
- Waterston RH, Lindblad-Toh K, Birney E,....., Lander ES; Mouse Genome Sequencing Consortium. (2002a). Initial sequencing and comparative analysis of the mouse genome. *Nature.*, 420, (6915): 520-62.
- Waterston RH, Lander ES, Sulston JE. (2002b). On the sequencing of the human genome. *Proc Natl Acad Sci U S A.*, 99, (6): 3712-6.
- Waterston RH, Lander ES, Sulston JE. (2003). More on the sequencing of the human genome. *Proc Natl Acad Sci U S A.*, 100, (6): 3022-4.
- Wathelet M, Moutschen S, Defilippi P, Cravador A, Collet M, Huez G, Content J (1986). Molecular cloning, full-length sequence and preliminary characterization of a 56-kDa protein induced by human interferons. *Eur J Biochem.*, 155, (1): 11-7.
- Wathelet MG, Clauss IM, Nols CB, Content J, Huez GA (1987). New inducers revealed by the promoter sequence analysis of two interferon-activated human genes. *Eur J Biochem.*, 169, (2): 313-21.
- Watson JD, Crick FH (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature*, 171, (4356): 737-8.
- Watson JD, Cook-Deegan RM. (1991). Origins of the Human Genome Project. *FASEB J.*, 5, (1): 8-11.
- Weber JL, Myers EW. (1997). Human whole-genome shotgun sequencing. *Genome Res.*, 7, (5): 401-9.
- Wettstein DA, Turner DL, Kintner C (1997). The *Xenopus* homolog of *Drosophila* Suppressor of Hairless mediates Notch signaling during primary neurogenesis. *Development*, 124, (3): 693-702.
- Whitehead CM, Rattner JB (1998). Expanding the role of HsEg5 within the mitotic and post-mitotic phases of the cell cycle. *J Cell Sci.*, 111, (Pt 17): 2551-61.
- Whitmore SA *et al.* (1998). Characterization and screening for mutations of the growth arrest-specific 11 (GAS11) and C16orf3 genes at 16q24.3 in breast cancer. *Genomics*, 52, (3): 325-31.
- Wiggins AK, Shen PJ, Gundlach AL (2003). Neuronal-NOS adaptor protein expression after spreading depression: implications for NO production and ischemic tolerance. *J Neurochem.*, 87, (6): 1368-80.
- Wilson MJ, Salata MW, Susalka SJ, Pfister KK (2001). Light chains of mammalian cytoplasmic dynein: identification and characterization of a family of LC8 light chains. *Cell Motil Cytoskeleton*, 49, (4): 229-40.
- Winkler GS, Albert TK, Dominguez C, Legtenberg YI, Boelens R, Timmers HT (2004). An altered-specificity ubiquitin-conjugating enzyme/ubiquitin-protein ligase pair. *J Mol Biol.*, 337, (1): 157-65.
- Winn PJ, Religa TL, Battey JN, Banerjee A, Wade RC (2004). Determinants of functionality in the ubiquitin conjugating enzyme family. *Structure (Camb)*, 12, (9): 1563-74.
- Wu G, Lyapina S, Das I, Li J, Gurney M, Pauley A, Chui I, Deshaies RJ, Kitajewski J (2001). SEL-10 is an inhibitor of notch signaling that targets notch for ubiquitin-mediated protein degradation. *Mol Cell Biol.*, 21, (21): 7403-15.
- Xirodimas DP, Saville MK, Bourdon JC, Hay RT, Lane DP (2004). Mdm2-mediated NEDD8 conjugation of p53 inhibits its transcriptional activity. *Cell*, 118, (1): 83-97.
- Yang CH, Lambie EJ, Snyder M (1992). NuMA: an unusually longcoiled-coil related protein in the mammalian nucleus. *J. Cell Biol*, 116, (6): 1303-17.
- Yang ZY, Guo J, Li N, Qian M, Wang SN, Zhu XL (2003). Mitosin/CENP-F is a conserved kinetochore protein subjected to cytoplasmic dynein-mediated poleward transport. *Cell Res.*, 13, (4): 275-83.

- Ye S, Pang H, Gu YY, Hua J, Chen XG, Bao CD, Wang Y, Zhang W, Qian J, Tsao BP, Hahn BH, Chen SL, Rao ZH, Shen N (2003). Protein interaction for an interferon-inducible systemic lupus associated gene, IFIT1. *Rheumatology (Oxford)*, 42, (10): 1155-63.
- Yeh E, Zhou L, Rudzik N, Boulianne GL (2000). Neuralized functions cell autonomously to regulate *Drosophila* sense organ development. *EMBO J.*, 19, (17): 4827-37.
- Yeh E, Dermer M, Commisso C, Zhou L, McGlade CJ, Boulianne GL (2001). Neuralized functions as an E3 ubiquitin ligase during *Drosophila* development. *Curr Biol.*, 11, (21): 1675-9.
- Yeh SD, Chen YJ, Chang AC, Ray R, She BR, Lee WS, Chiang HS, Cohen SN, Lin-Chao S (2002). Isolation and properties of Gas8, a growth arrest-specific gene regulated during male gametogenesis to produce a protein associated with the sperm motility apparatus. *J Biol Chem.*, 277, (8): 6311-7.
- Zeng C (2000). NuMA: a nuclear protein involved in mitotic centrosome function. *Microsc Res Tech.*, 49, (5): 467-77.
- Zhang H, Saitoh H, Matunis MJ (2002). Enzymes of the SUMO modification pathway localize to filaments of the nuclear pore complex. *Mol Cell Biol.*, 22, (18): 6498-508.
- Zhang Z, Gerstein M (2004). Large-scale analysis of pseudogenes in the human genome. *Curr Opin Genet Dev.*, 14, (4): 328-35.
- Zheng N, Wang P, Jeffrey PD, Pavletich NP (2000). Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. *Cell*, 102, (4): 533-9.
- Zhong W, Feder JN, Jiang MM, Jan LY, Jan YN (1996). Asymmetric localization of a mammalian numb homolog during mouse cortical neurogenesis. *Neuron*, 17, (1): 43-53.
- Zhong W, Jiang MM, Schonemann MD, Meneses JJ, Pedersen RA, Jan LY, Jan YN (2000). Mouse numb is an essential gene involved in cortical neurogenesis. *Proc Natl Acad Sci U S A*, 97, (12): 6844-9.

ΔΗΜΟΣΙΕΥΣΕΙΣ

The DNA sequence and comparative analysis of human chromosome 10

P. Deloukas¹, M. E. Earthrowl¹, D. V. Grafham¹, M. Rubenfield^{2,3}, L. French¹, C. A. Steward¹, S. K. Sims¹, M. C. Jones¹, S. Searle¹, C. Scott¹, K. Howe¹, S. E. Hunt¹, T. D. Andrews¹, J. G. R. Gilbert¹, D. Swarbreck¹, J. L. Ashurst¹, A. Taylor¹, J. Battles², C. P. Bird¹, R. Ainscough¹, J. P. Almeida¹, R. I. S. Ashwell¹, K. D. Ambrose¹, A. K. Babbage¹, C. L. Bagguley¹, J. Bailey¹, R. Banerjee¹, K. Bates¹, H. Beasley¹, S. Bray-Allen¹, A. J. Brown¹, J. Y. Brown¹, D. C. Burford¹, W. Burrill¹, J. Burton¹, P. Cahill², D. Camire², N. P. Carter¹, J. C. Chapman¹, S. Y. Clark¹, G. Clarke¹, C. M. Cleve¹, S. Clegg¹, N. Corby¹, A. Coulson¹, P. Dhami¹, I. Dutta¹, M. Dunn¹, L. Faulkner¹, A. Frankish¹, J. A. Frankland¹, P. Garner¹, J. Garnett¹, S. Gribble¹, C. Griffiths¹, R. Grocock¹, E. Gustafson^{2,3}, S. Hammond¹, J. L. Harley¹, E. Hart¹, P. D. Heath¹, T. P. Ho², B. Hopkins¹, J. Horne², P. J. Howden¹, E. Huckle¹, C. Hynds², C. Johnson¹, D. Johnson¹, A. Kana², M. Kay¹, A. M. Kimberley¹, J. K. Kershaw¹, M. Kokkinaki⁴, G. K. Laird¹, S. Lawlor¹, H. M. Lee², D. A. Leongamornlert¹, G. Laird¹, C. Lloyd¹, D. M. Lloyd¹, J. Loveland¹, J. Lovell¹, S. McLaren¹, K. E. McLay¹, A. McMurray¹, M. Mashreghi-Mohammadi¹, L. Matthews¹, S. Milne¹, T. Nickerson¹, M. Nguyen², E. Overton-Larty¹, S. A. Palmer¹, A. V. Pearce¹, A. I. Peck¹, S. Pelan¹, B. Phillimore¹, K. Porter¹, C. M. Rice¹, A. Rogosin^{2,3}, M. T. Ross¹, T. Sarafidou⁴, H. K. Sehra¹, R. Shownkeen¹, C. D. Skuce¹, M. Smith¹, L. Standing², N. Sycamore¹, J. Tester¹, A. Thorpe¹, W. Torcasso², A. Tracey¹, A. Tromans¹, J. Tsolas^{2,3}, M. Wall¹, J. Walsh², H. Wang², K. Weinstock², A. P. West¹, D. L. Willey¹, S. L. Whitehead¹, L. Wilming¹, P. W. Wray¹, L. Young¹, Y. Chen⁵, R. C. Lovering⁶, N. K. Moschonas⁴, R. Siebert⁷, K. Fichtel², D. Bentley¹, R. Durbin¹, T. Hubbard¹, L. Doucette-Stamm^{2,3}, S. Beck¹, D. R. Smith^{2,3} & J. Rogers¹

¹The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK

²Genome Therapeutics Corporation, 100 Beaver Street, Waltham, Massachusetts 02453, USA

³Agencourt Bioscience Corporation, 100 Cummings Center, Beverly, Massachusetts 01915, USA

⁴Department of Biology, University of Crete & Institute of Molecular Biology and Biotechnology, Foundation of Research and Technology, PO Box 2208, 71409 Heraklion, Crete, Greece

⁵European Bioinformatics Institute (EBI), Wellcome Trust Genome Campus, Hinxton CB10 1SD, UK

⁶HUGO Gene Nomenclature Committee, Department of Biology, University College London, London NW1 2HE, UK

⁷Institute of Human Genetics, University Hospital Schleswig-Holstein Campus Kiel, Schwanenweg 24, D-24105 Kiel, Germany

The finished sequence of human chromosome 10 comprises a total of 131,666,441 base pairs. It represents 99.4% of the euchromatic DNA and includes one megabase of heterochromatic sequence within the pericentromeric region of the short and long arm of the chromosome. Sequence annotation revealed 1,357 genes, of which 816 are protein coding, and 430 are pseudogenes. We observed widespread occurrence of overlapping coding genes (either strand) and identified 67 antisense transcripts. Our analysis suggests that both inter- and intrachromosomal segmental duplications have impacted on the gene count on chromosome 10. Multispecies comparative analysis indicated that we can readily annotate the protein-coding genes with current resources. We estimate that over 95% of all coding exons were identified in this study. Assessment of single base changes between the human chromosome 10 and chimpanzee sequence revealed nonsense mutations in only 21 coding genes with respect to the human sequence.

We report here the final clone map and sequence assembly of human chromosome 10. This metacentric chromosome accounts for 4.5% of the genome and is best known for harbouring the *PTEN* tumour suppressor gene and the *RET* proto-oncogene.

With the human genome sequence in hand, the task ahead is to identify the different units of genetic information embedded in the sequence and understand their function both at the molecular and cellular level. In this study we address the former by reporting a comprehensive annotation of manually inspected gene structures and their correlation to sequence variation and other features of the genomic sequence. The annotation process is assessed by comparative analysis to the genome sequence of two rodents, *Mus musculus* and *Rattus norvegicus*, and three fishes, *Tetraodon nigroviridis*, *Fugu rubripes* and *Danio rerio*. Finally, we report our preliminary findings on the distribution of single base differences along human chromosome 10 in comparison to the chimpanzee genome.

The clone map and finished sequence

A clone map spanning the euchromatic regions of the short (p) and long (q) arm of human chromosome 10 was assembled by restriction fingerprint and sequence-tagged site (STS) content analysis¹. We identified clones by screening approximately 85 genomic equivalents of P1-derived artificial chromosome (PAC), bacterial

artificial chromosome (BAC), yeast artificial chromosome (YAC), cosmid and fosmid libraries. The tiling path consists of 1,144 minimally overlapping clones (Supplementary Table S1) organized into 12 contigs (Table 1). Contig-1340 spans the entire p arm and harbours the boundary between euchromatin and heterochromatin with the proximal 250 kilobases (kb) extending into pericentromeric satellite repeats. In a detailed study of this region two further clones carrying satellite 3 sequences, contig-2069, were identified and mapped by pulse field gel electrophoresis (PFGE) ~50 kb proximal to contig-1340 (ref. 2). Similarly, contig-43 harbours the q-arm boundary with the proximal 240 kb composed of satellite repeats³. In addition, clone RP11-745D9, 94% of which consists of α -satellite repeats, was arbitrarily placed proximal to contig-43 as it was suggested to map to human chromosome 10 (E. Eichler, personal communication). A 9.75-megabase (Mb) PFGE map spanning the chromosome 10 centromere⁴ places the core α -satellite block (D10Z1) ~0.2 Mb distal to contig-102 and 1 Mb proximal to contig-43.

In contrast with the p arm, nine gaps remain in the clone map of the q arm. Five of them are clustered within a ~4-Mb region (Table 1 and Fig. 1 (rollfold); see also Supplementary Fig. S1 for a detailed view). Our inability to walk across these five gaps was due to the extensive segmental duplications in this region (Fig. 2);

however, we obtained a size estimate by fluorescent *in situ* hybridization (FISH) of RP11-172C24 (AL512595) and RP11-13E1 (AC013284) on metaphase chromosomes. We sized the remaining four gaps by FISH with clones immediately flanking each gap to extended DNA fibres. No gap was estimated to be larger than 50 kb in size. Altogether the euchromatic gaps account for no more than 840 kb (Table 1). Finally, we defined the location of both telomeres. Clone RP11-631M1 (AL713922) ends ~20 kb away from the telomeric repeats of the p arm based on the telomeric half-YAC XX-YAC22O3 (<http://www.wistar.upenn.edu/Riethman/>). At the end of the q arm (qtel), clone XX-YAC2136 (BX322534) contains part of the telomeric repeat block.

Each clone of the tiling path was subjected to random subcloning and sequencing at either the Genome Therapeutics Corporation (GENE) or the Sanger Institute—the initial draft sequence of a few clones was carried out by other centres that are credited in the corresponding submissions to the EMBL/GenBank/DBJ databases. We finished clones according to the international finishing standard (<http://genome.wustl.edu/Overview/g16stand.php>). Of the 1,144 clones in the human chromosome 10 tiling path, 221 and 913 were finished at GENE and the Sanger Institute, respectively, and three elsewhere (Supplementary Fig. S1). The remaining seven clones show persistent deletion of internal fragments. In total, we finished 131,666,441 base pairs (bp) in 18 sequence contigs; euchromatic coverage is estimated at 99.4%. Sequence accuracy was estimated as described in ref. 5 and found to exceed 99.99%. The sequence assembly comprises all known chromosome 10 messenger RNAs (RefSeq set) and STS markers from available radiation hybrid⁶ and genetic maps^{7,8} (T. Furey, personal communication). In addition, the integrity of the sequence map was independently assessed at the University of California, Santa Cruz, by alignment of fosmid and BAC paired end sequences (<http://genome.cse.ucsc.edu/>). Table 1 lists the size of each sequence contig, with the largest one spanning 44,693,577 bp.

Table 1 Clone and sequence contigs on chromosome 10

Clone contig	Sequence contig	Size (bp)	Gap size estimate (bp)
Telomere	—	—	20,000
1340	BX32259-BX276080	5,574,992	10,000
	AL365356-BX255924	12,337,510	10,000
	AL928729-AL133216	20,793,917	—
Gap	—	—	50,000
2069	AL133173-AL590623	286,100	—
Gap	—	—	50,000
Centromere	—	—	2,380,000
Gap	—	—	50,000
102	BX322613	191,752	—
Gap	—	—	50,000
43	AL031601-AC012044	3,830,268	—
Gap*	—	—	ND*
4000	AL831769-BX649215	952,201	—
Gap*	—	—	ND*
3003	AL450388-AL603965	263,307	—
Gap*	—	—	ND*
101	AL954360	163,321	—
Gap*	—	—	ND*
14	BX547991-AL731733	989,826	—
Gap*	—	—	ND*
15	AL603966-AL731572	1,941,839	10,000
	AL672187	211,435	10,000
	AL589822-AL132656	30,068,948	—
Gap	—	—	10,000
16	AC068139-AL731667	44,693,577	10,000
	AL606510-AL772134	2,696,534	—
Gap	—	—	50,000
17	BX470155-AL607044	4,615,046	—
Gap	—	—	10,000
2493 (includes telomere)	BX294094-BX511297	246,123	10,000
	AL732395-BX322534	1,809,745	—

*Cumulative gap size estimate of 750,000 for all five of the indicated gaps together. ND, not determined.

The gene and protein index

The Sanger Institute has established a standardized annotation pipeline (outlined at <http://vega.sanger.ac.uk/>) in which gene structures are drawn on the basis of human interpretation of the combined supportive evidence generated during sequence analysis. Annotation of the human chromosome 10 sequence resulted in a total of 1,787 gene structures that we then classified, as described in ref. 9, into: (1) 654 ‘known’ genes; (2) 162 ‘novel genes’; (3) 219 ‘novel transcripts’; (4) 322 ‘putative genes’; and (5) 430 ‘pseudogenes’. Pseudogenes were further subdivided into processed (371) and unprocessed (59).

Excluding the pseudogenes, human chromosome 10 is a chromosome with an average gene density (10.4 genes Mb⁻¹). The 1,357 genes span 66,309,730 bp in total (mean 51,335 bp per gene). Therefore, 50.6% of the analysed sequence is transcribed, matching the figure reported for chromosome 22 (51%), which is gene-rich, but appearing elevated in comparison with chromosomes 6, 7, 14 and 20 (42.2%, 46.5%, 43.6% and 42.4%, respectively), which have gene densities similar to chromosome 10. The latter suggests that the human chromosome 10 genes have on average a larger genomic span than those on chromosomes 6, 7, 14 and 20. Gene size along human chromosome 10 varies enormously, with the two extremes being *CTNNA3* (1,776,209 bp) and *CALML5* (859 bp). Exons account for only 2.3% of the sequence and the mean exon size is 313 bp. The longest and shortest exons annotated in this study have a length of 9,763 (*SH3MD1*) and 3 bp (*CDH23*), respectively. *CDH23* is also the gene with most exons (69) on this chromosome. Table 2 summarizes the features of each gene class.

Alternative splicing is a major contributor to the complexity of the human transcriptome. We annotated a total of 4,204 transcripts for 1,357 gene structures (Table 2). No splice variants were annotated on the basis of alternative polyadenylation sites. Approximately 73% of the protein-coding genes have more than one transcript and 5.8 on average. For *ADD3* we annotated 22 variants. Note that the use of partial expressed sequences (for example, expressed sequence tags (ESTs)) may result in the annotation of more than one transcript per splice variant. Given this caveat, our analysis suggests a significantly higher level of alternative splicing compared with previous estimates¹⁰. Approximately 50% of the 3,456 transcripts (known and novel) do not seem to encode a protein. Annotation of these transcripts is largely based on ESTs, many of which may correspond to aberrant transcripts. Their precise role is largely unknown but they may be part of the machinery of transcriptional regulation (for example, via nonsense-mediated decay). Nevertheless, there are 1,837 transcripts with an open reading frame (ORF). Of the 342 genes with at least two transcripts having a complete ORF (that is, possess both a 5' and a 3' untranslated region (UTR)), 312 encode at least two distinct peptides.

Identification of transcription start sites and promoter regions remains a challenge in the annotation process. First, we scanned the human chromosome 10 unmasked sequence and predicted a total of

Figure 1 The sequence map of chromosome 10 and its features (see rollfold). The current sequence assembly (v1.1) and that used in the analysis presented in this study (v1.0) are available at <http://www.sanger.ac.uk/HGP/Chr10>. Tracks from top to bottom are: (1) the scale bar (in Mb); (2) the sequence map of human chromosome 10 represented by a black solid bar interrupted by clone and sequence gaps (grey); (3) syntenic blocks in the mouse (top track) and the rat (bottom track) where blocks are colour-coded per chromosome and labelled with the chromosome number and coordinates (Mb) (for example, 2: 3.1–11.3 Mb; unordered sequence contigs are tagged as random); (4) predicted CpG islands (brown); (5) regions of sequence homology to *Fugu* (blue), zebrafish (dark blue) and *Tetraodon* (dark pink); and (6) protein-coding genes. Gene names are in dark blue for the known and black for the novel CDS.

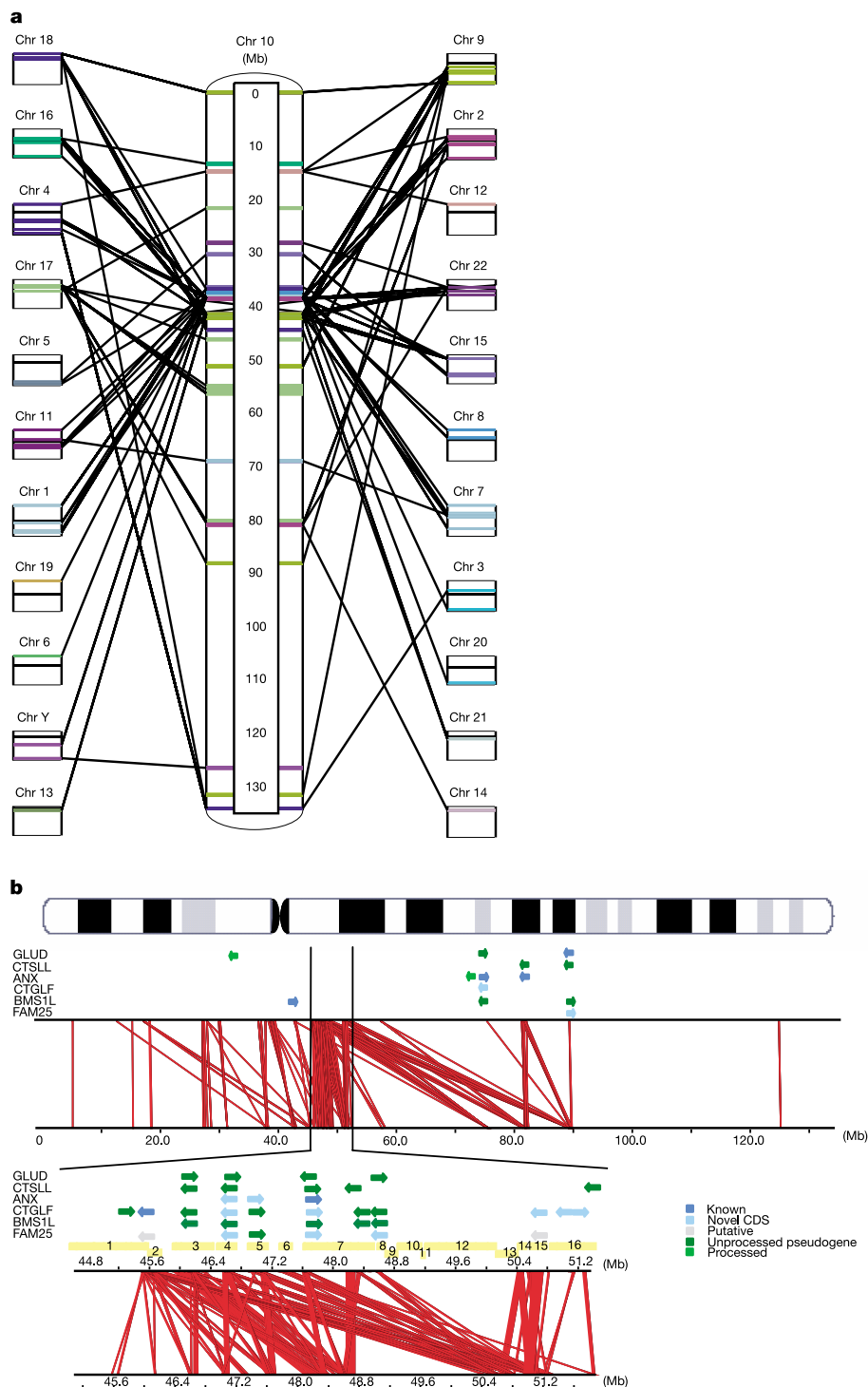


Figure 2 Chromosome 10 inter- and intrasegmental duplications. **a**, Interchromosomal duplications across chromosome 10 showing blocks of 10 kb or greater. Duplicated regions are colour-coded per chromosome and indicated as lines (thickness is proportional to physical distance). Each chromosome other than 10 is represented by an open black rectangle with a black line representing the centromere. **b**, 10q11:10q22:10q23.1:10q23.3 intrasegmental duplications. Top row, ideogram of chromosome 10; second row, positions of members of the six main gene families outside 10q11 (colour-coded per gene class; arrows indicate transcription); third row, intrachromosomal duplications across the whole chromosome showing blocks of 10 kb or greater; bottom row, enlarged view of the 10q11 region. Yellow bars represent sequence (bottom row for sequences submitted in reverse orientation) contigs between AL358394 and AL589794 (complete clone list in Supplementary Fig. S1). From left to right the

members (where parentheses indicate members only appearing in the enlarged section) of each family are: GLUD family, *GLUDP5*, (*GLUDP7*, *GLUDP8*, *GLUDP6*, *GLUDP2*), *GLUDP3*, *GLUD1*; CTSLL family, (*CTSLL5*, *CTSLL7*, *CTSLL2*, *CTSLL3*, *CTSLL4*), *CTSLL6*, *CTSLL1*; ANXA family, (*ANXA8L1*, *ANXA8L2*, *ANXA8*), *ANXA2P3*, *ANXA7*, *ANXA11*; CTGLF family, (*CTGLF10P*, *CTGLF1*, *CTGLF13P*, *CTGLF7*, *CTGLF11P*, *CTGLF6*, *CTGLF9P*, *CTGLF12P*, *CTGLF5*, *CTGLF4*, *CTGLF3*), *CTGLF2*; BMS1L family, *BMS1L*, (*BMSILP1*, *BMSILP2*, *BMSILP6*, *BMSILP5*, *BMSILP7*), *BMSILP4*, *BMSILP3*; FAM25 family, (*FAM25E*, *FAM25B*, *FAM25HP*, *FAM25G*, *FAM25C*, *FAM25D*), *FAM25A*. *ANXA7* and *ANXA11* were included owing to their proximity to the 10q22 and 10q23.1 locus, respectively. Seven *CTGLF1* paralogues (Supplementary Table S3) were annotated as novel genes for consistency but probably represent expressed pseudogenes.

Table 2 Structural characteristics of annotated gene structures on chromosome 10

Gene structure class	Number of genes	Total transcribed length (bp)	Mean gene length (bp)	Mean exon length (bp)	Mean transcripts per gene	Mean exons per gene
Known genes	654	56,312,218	86,870	316	4.76	11.02
Novel CDS	162	3,615,155	22,435	344	2.12	4.96
Total protein coding	816	59,482,121	74,078	320	4.24	9.81
Novel transcript	219	6,123,588	29,166	272	1.77	4.06
Putative genes	322	2,798,239	8,780	287	1.12	2.15
Processed pseudogenes	371	328,131	887	638	1.00	1.19
Unprocessed	59	982,662	16,655	176	1.00	6.95
Total structures	1,787	67,353,917	39,717	322	2.59	5.84

Note that transcribed length is not additive because some genes overlap. CDS, coding sequence.

1,025 CpG islands (Fig. 1), which are known to be associated with the 5' end of an estimated 60% of human genes¹¹. We then used Eponine¹² to predict transcription start sites and FirstEF¹³ to predict regions that encompass the promoter and 5' exon. In total, FirstEF and Eponine predicted 1,801 (rank = 1, score ≥0.8) and 2,800 features, respectively (Supplementary Fig. S1). Notably, 62% of FirstEF and 96% of Eponine features directly overlapped CpG islands, suggesting a heavy bias towards this feature in both algorithms. The distribution of CpG islands, FirstEF and Eponine hits was also assessed relative to the first exon of each of the 4,635 annotated transcripts using a window of 5,000 bp upstream and 1,000 bp downstream of the exon. Table 3 summarizes the results obtained per gene class. For example, in the 'known' class, 1,544 (49.6%) and 1,124 (36.1%) transcripts had a FirstEF and Eponine feature, respectively. Not surprisingly, 89.9% of FirstEF and 96.7% of Eponine transcripts were also associated with a CpG island. Note that Eponine predicts multiple transcription start sites per transcript (4.24 on average).

Regulation of gene expression by antisense transcription is a recognized mechanism with examples reported in species ranging from bacteria to mammals^{14–16}. We observed widespread occurrence of overlapping coding genes (either strand) in human chromosome 10 (101 pairs in total). In 38 cases one of the genes is fully contained within an intron of the other, typically on the opposite strand (34). For example, the second intron of the splice variant LIPA-004 encompasses four members of the IFIT gene cluster (Supplementary Table S3 and Fig. S2) and appears to be transcribed from the same bidirectional promoter as *IFIT5* (IFIT cluster member). Interestingly, the LIPA-001 and -002 variants do not overlap any IFIT gene, whereas LIPA-003 and LIPA-005 both overlap with *IFIT2* and *IFIT4* (Supplementary Fig. S2). Mutations in *LIPA* can cause Wolman and cholesteryl ester storage disease (Online Mendelian Inheritance in Man (OMIM) 278000). Among partially overlapping pairs (opposite strands), 34% involve the respective 5' exons, which is indicative in each case of a bidirectional promoter. We also searched for non-coding transcripts located on the opposite strand of coding genes. There are 67 antisense transcripts overlapping 63 coding genes. The two most common patterns observed were intragenic with partial overlap of one exon, and partial overlap with the most 5' exon of the coding gene. For *ZNF32*, we found two antisense transcripts (*ZNF32OS1* and *ZNF32OS2*).

Finally, we looked at the distribution of known protein domains

in both human chromosome 10 (this study) and the whole genome (Ensembl v.17.33.1) proteome using InterProScan. At the gene level, 70.6% of the human chromosome 10 peptides have at least one InterPro match with a Pfam domain and 32% are multidomain (1.37 distinct InterPros on average). Supplementary Table S2 shows the top 24 domains in chromosome 10 alongside their genome-wide ranking, suggesting that this chromosome is enriched in peptides with a lipase (IPR000734), aldo/keto reductase (IPR001395) or alpha/beta hydrolase (IPR000073) domain. BLASTP analysis (*e*-values <10⁻¹⁵) showed that all six genes encoding the peptides carrying the IPR001395 domain are clustered (AKRIC cluster) at position 4.8–5.3 Mb, whereas there are two lipase clusters, LIP (90.0–91.0 Mb; six members) and PNLIP (117.9–118.1 Mb; four members). In total, we found 42 gene clusters along human chromosome 10 (Supplementary Table S3).

Genomic landscape

The average G+C and repeat content of human chromosome 10 are 41.58% and 43.66%, respectively. The distribution of the main classes of repeats (Supplementary Table S4), the G+C and CpG density plots all seem to follow the known genome-wide trends. For example, the G+C content fluctuates along the chromosome, peaks at the qtel and is positively correlated to gene density (Supplementary Fig. S1). Large genes tend to be located adjacent to or within gene-poor regions, for example *PRKG1* and *PCDH15* (interval 52.0–59.0 Mb), whereas regions of high gene density seem enriched in short interspersed elements (SINES).

Segmental duplications are an important feature of the genomic landscape, being an integral part of the evolutionary machinery. Figure 2a illustrates the interchromosomal duplications along human chromosome 10 (see also ref. 17 and <http://humanparalogy.cwru.edu/SDD>). Figure 2b shows the extensive segmental duplications within a 5 Mb region at 10q11 with sequences further dispersed at 10q22, 10q23.1 and 10q23.2. Using the draft genome sequence Crosier and colleagues¹⁸ reported that 10q11 has been subject to multiple rounds of local duplications in at least the last 30–40 million years (Myr); they also showed that a 10q11:10q23 paracentric inversion occurred after the divergence of orang-utan from other great apes and hypothesized that the 10q22 locus resulted from chromosome-specific duplicative transposition. Bryce and colleagues¹⁹ characterized three cathepsin-L-like paralogues, which are expressed pseudogenes, and reported FISH signals

Table 3 Correlation of CpG island, FirstEF and Eponine features with annotated transcripts

Gene class	All Transcripts		FirstEF transcripts		Eponine transcripts	
	Complete	Incomplete	Complete	Incomplete	Complete	Incomplete
Known	1,490 (580)	1,623 (1,018)	593 (531)	951 (877)	421 (412)	703 (682)
Novel CDS	130 (57)	213 (75)	45 (43)	62 (57)	38 (38)	43 (41)
Novel transcript	388 (119)	0 (0)	120 (105)	0 (0)	83 (79)	0 (0)
Processed pseudogene	371 (32)	0 (0)	29 (22)	0 (0)	21 (19)	0 (0)
Unprocessed pseudogene	59 (11)	1 (2)	14 (10)	1 (1)	9 (8)	0 (0)
Putative	360 (74)	0 (0)	87 (64)	0 (0)	57 (50)	0 (0)

Values in parentheses are for transcripts associated with a CpG island.

at 10q11, 10q22 and 10q23, a pattern previously seen with the *GLUD* paralogues²⁰ (Fig. 2b). The duplication of the *CTSL* locus between chromosomes 9 and 10 was estimated to have occurred some 40 Myr ago¹⁹. We identified four and three additional *CTSL* and *GLUD* pseudogenes, respectively (Fig. 2b), consistent with the local duplication events involving *BMS1L* (known as *KIAA0187*)¹⁸ and *CTGLF1* (this study). In total, we annotated 7 *BMS1L* and 11 *CTGLF1* paralogues (Fig. 2b; functional genes in dark blue). Retro-position of a truncated *KIAA1099* (*CENTG2*; chromosome 2) mRNA gave rise to a processed pseudogene on chromosome 10 (ref. 18). However, this pseudogene forms the 3' exon of *CTGLF1*, suggesting that this gene resulted from a fusion between an ancestral gene and the *CENTG2* pseudogene, and the retroposition event predated all segmental duplications. Note that there is also evidence of transcripts combining exons of *CTGLF1* and *BMS1L* paralogues. Interestingly, we predicted a novel gene, *FAM25A* (based on mouse complementary DNA AK008614 and without similarity to any known protein), with seven human chromosome 10 paralogues (Fig. 2b) that follow the pattern of *GLUD*, *BMS1L*, *CTGLF* and *CTSL*. In addition to these five types of low copy repeats, the segmental duplications in 10q11:10q22:10q23 seem to have impacted on the number of genes on this chromosome. Notably, 31% of all the functionally related gene clusters are located within 10q11 and 10q23 (Supplementary Fig. S1 and Table S3).

The average recombination rate across the chromosome is 1.32 cM Mb⁻¹ (Fig. 3, sex average). Note that ref. 7 used the draft human chromosome 10 sequence (inflated by ~8%) and thus obtained a lower figure (1.21 cM Mb⁻¹). The rate of male recombination is higher than the female rate near the telomeres, whereas between D10S211 and D10S575 the female rate is higher (Fig. 3). This comparison also indicates the presence of two female-specific recombination hotspots (Fig. 3, arrows). As expected, pericentromeric regions display a low rate of recombination, more than twofold below the chromosome average. In particular, the region between D10S1247 and D10S1783 has a rate of 0.3 cM Mb⁻¹ and contains the only human chromosome 10 recombination 'desert'²¹. The region of extensive segmental duplications at 10q11 shows a low

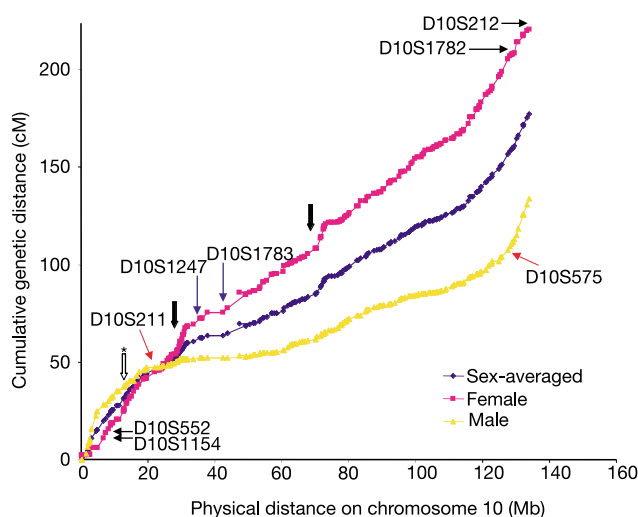


Figure 3 Alignment of the deCODE genetic map of chromosome 10 to the physical map from the telomeric end of the short arm to the telomeric end of the long arm. The position of each genetic marker on the female, male and sex-averaged genetic map is indicated. Female-specific recombination hotspots are indicated by thick arrows (left, D10S1732–D10S208; right, D10S599–D10S676). The location of markers flanking recombination deserts (blue arrows) and jungles (black arrows) is shown. The asterisk indicates the location of the refuted jungle (D10S189–D10S1728).

rate of recombination (0.72 cM Mb⁻¹). Thus, meiotic recombination is unlikely to have been the driving force in generating these duplications. Our analysis confirmed the recombination 'jungle' between D10S1782–D10S1651, which we extended to D10S212 (3.4 cM Mb⁻¹); we refuted the one between D10S189 and D10S1728 (2.3 cM Mb⁻¹), and identified a new one between D10S1154 and D10S552 (4.55 cM Mb⁻¹).

Comparative sequence analysis

Many of the functional units in present-day vertebrate genomes have been conserved through evolution. Coding regions are highly conserved across all vertebrates, whereas non-coding regions are conserved among more closely related species. The genome sequences of three fishes (*Tetraodon*, *Fugu* and zebrafish) and two rodents (mouse and rat) were publicly available at the time of analysis. We compared the sequence of human chromosome 10 to each of the above species and searched for conserved regions with coding potential—distinguishing functional non-coding elements above background sequence conservation requires the use of additional genomes²². We then correlated the obtained hits in each species with the annotated genes. As expected, each of the fish genomes showed higher specificity (>0.82; that is, 82% of sequences conserved in fish overlapped human annotated exons) than the rodents (>0.29), whereas the highest specificity was obtained using all genomes together (0.96). Sensitivity was higher using a rodent genome (0.7; that is, 70% of all annotated exons had matches in rodent) than a fish genome (0.4).

Of the 1,787 gene structures (16,765 exons), 84% (78% of exons) had at least one exon supported by a conserved region in one of the other genomes and 52% (32% of exons) in all genomes. Note that the figures given for exons are underestimated owing to lower sequence conservation in the untranslated regions. Protein-coding genes are highly conserved (98.5%; 85% of exons). In contrast, 61% (29% of exons) and 45% (27% of exons) of novel transcripts and putative genes, respectively, have at least one match. Furthermore, only 21% of novel transcripts and 8% of putative genes show conservation with a fish (91% for protein-coding genes). Typically, novel transcripts were annotated on the basis of solid experimental evidence (that is, human mRNA) and may represent either genes that have evolved more rapidly or non-coding RNAs.

On the basis of specificity, regions conserved in all six species can serve as a measure of completeness of the gene annotation process that occurred independently of the comparative analysis. We found 5,604 such evolutionarily conserved regions (ECRs) of which 5,292 mapped inside annotated exons (including pseudogenes). Of the remaining 312 ECRs, 142 were intergenic and 170 intragenic. On inspection, we found 79 of these ECRs with supportive evidence to annotate a missed exon, most of which were part of a pseudogene (79%). The remaining 233 ECRs provide the basis to estimate that we have annotated at least 95.8% of all conserved coding exons on human chromosome 10. This is a conservative estimate as 131 of these ECRs are located in introns and may represent conserved non-exonic sequences. Interestingly, 54 (41%) of them are associated with just four genes: *C10orf11* (26; also known as *CDA017*), *EBF3* (20), *TCF7L2* (5) and *PAX2* (3). All but *C10orf11* (unknown function) are transcription factors. Figure 4 shows a MultiPip-Maker²³ alignment of the orthologous *EBF3* loci and the relative position of ECRs in the human gene. Note that sequence identity is often higher in ECRs than in exons.

Sequence variation

During the human chromosome 10 project we discovered 35,882 single nucleotide polymorphisms (SNPs) by sequence alignment in regions of clone overlaps. In total, we mapped 143,364 SNPs (dbSNP release 115) to the chromosome 10 sequence. Supplementary Fig. S1 shows the density plots for randomly discovered²⁴ and all SNPs across the chromosome.

There are 5,864 (4.1%) exonic and 65,973 (46%) intronic SNPs. Of the 1,821 SNPs in coding exons 984 are non-synonymous. *MSMB* has the most polymorphic coding region with 43 SNPs kb⁻¹; it encodes a protein with inhibin-like activity and its expression is decreased in prostate cancer²⁵.

We also considered 729,553 human–chimpanzee single base differences (SBDs) remapped on the current assembly of human chromosome 10. These were high-confidence sequence differences originally identified by aligning 14 million shotgun reads of the chimpanzee genome, generated jointly by the Whitehead Institute and Washington University Genome Centers, to the human genome sequence assembly (NCBI build 31). We first removed all human–chimpanzee SBDs that co-localized with known human SNPs. Supplementary Fig. S1 shows the density plot of the remaining 703,338 SBDs. Of those, 55.3% are intergenic, 42.9% intronic and 1.8% exonic. The highest density of human–chimpanzee SBDs, fourfold greater than the average level, was observed in a 200-kb gene-poor region at 19.43–19.63 Mb. We then examined the 12,710 human–chimpanzee SBDs that lie in exons of the 816 human coding genes. Of those, 3,972 were in coding regions and can be subdivided further into 2,273 synonymous, 1,678 non-synonymous and 21 nonsense with respect to the human sequence. For each gene we calculated the rate of evolution of non-synonymous (K_a) and synonymous (K_s) changes, and the ratio K_a/K_s , which provides a measure of evolutionary selection. Supplementary Table S5 lists the 1,413 transcripts with at least one coding human–chimpanzee SBD sorted on the K_a/K_s value. There are only 29 transcripts (21 genes) that have a K_a/K_s value ≥ 1 , whereas there are 484 without non-synonymous SBDs. Note that several caveats apply in this type of analysis owing to the incomplete nature of both the chimpanzee data and the list of human SNPs; we used the number of intronic human–chimpanzee SBDs per base in comparison to the chromosome average of 0.005 as a possible estimate of coverage. The gene with most non-synonymous human–chimpanzee SBDs is *MKI67*, an antigen identified by monoclonal antibody Ki-67, which appears to be fast evolving in humans ($K_a/K_s = 1.038507$; SNP data). The expression pattern of *MKI67* in gastric and other cancers is under investigation as this gene is expressed in proliferating cells. Interestingly, a nonsense human–chimpanzee SBD is present in both of its coding transcripts. Among the 21 genes with nonsense human–chimpanzee SBDs notable examples are the serotonin receptor *HTR7* (the neurotransmitter serotonin is thought to be involved

in cognition and behaviour), *PSAP* (prosaposin; involved in variant Gaucher's disease and metachromatic leukodystrophy) and the developmental gene *NODAL*.

Medical implications

At the time of writing there were 85 disease loci reported on human chromosome 10 (<http://www.ncbi.nlm.nih.gov/omim/>), a 47% increase since 1999 (ref. 26). Several of these loci account for multiple disease phenotypes caused by mutations in the same gene; notable examples are *FGFR2* (OMIM 176943), *PTEN* (OMIM 601728) and the proto-oncogene *RET* (OMIM 164761). Since *PTEN* was first shown to be mutated in brain, breast and prostate cancers²⁷, there has been an explosion of reported mutations (110 germline and 332 somatic mutations)²⁸ and disease phenotypes²⁹. Human chromosome 10 harbours several other genes involved in tumorigenesis; for example, deregulation of *TLX1*, *NFKB2* or *BMI1* caused by chromosomal translocations or amplifications has been detected in lymphatic neoplasms. Mapping of allelic imbalances and functional studies suggest the presence of additional tumour-related genes. The finished and annotated sequence is key in the process of cloning these and other hitherto unidentified disease-associated genes.

The prompt release of both the clone and sequence map resources throughout the project has accelerated the cloning of many disease-causing genes. To this end we recently showed as part of the European ADLTE consortium that mutations in the *LGII* gene cause autosomal dominant lateral temporal epilepsy³⁰. Notably, we found that the *FRA10A* folate-sensitive fragile site is located close to *LGII* and its expression is associated with the expansion of a polymorphic CGG repeat located at the 5' UTR of *FRA10AC1*, a gene encoding a novel nuclear protein³¹. There are seven fragile sites mapping to human chromosome 10 (ref. 32).

The challenge ahead is to unravel the molecular basis of common disease. An increasing number of susceptibility loci for complex diseases is being mapped to human chromosome 10, including metabolic diseases such as type I diabetes (IDDM10 and IDDM17), psychiatric disorders such as schizophrenia, or neurodegenerative illnesses such as Alzheimer's disease^{26,32}. In a case control study of morbidly obese and healthy individuals Boutin and colleagues³³ identified a SNP in the *GAD2* gene that increases the risk for obesity as well as a protective haplotype. Studies so far have mainly focused on candidate genes. The construction of a comprehensive haplotype

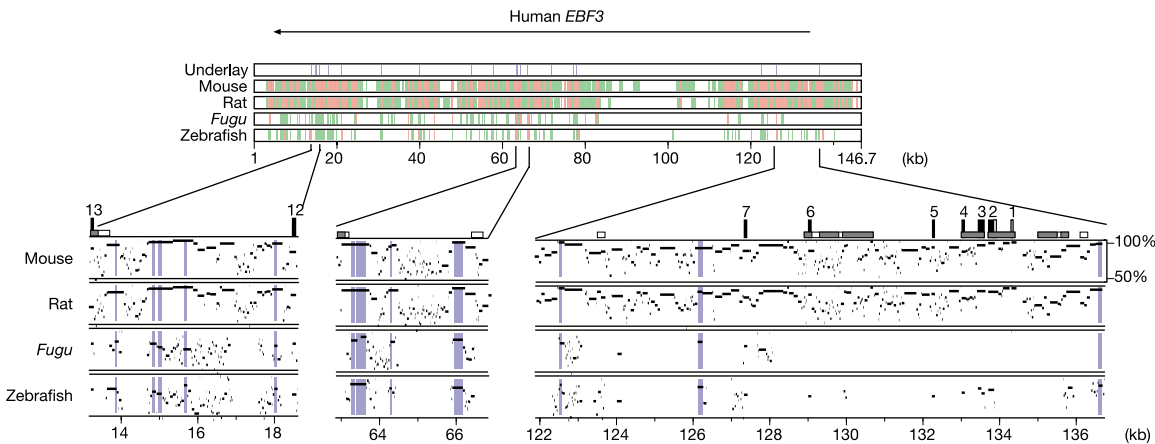


Figure 4 Multispecies alignment of orthologous *EBF3* loci. The human early B-cell factor 3 (*EBF3*) gene is represented by the arrow at the top. Alignments are displayed using MultiPipMaker²³. In the top panel, the first track shows the location of the ECRs (blue lines) across the human locus, whereas the following four tracks show regions conserved in mouse, rat, *Fugu* and zebrafish, respectively (green, aligned regions; orange, aligned

regions with at least 70% nucleotide identity and no gap over 100 bp). The bottom panel shows a detailed view of the three regions with the highest number of ECRs. Vertical black and grey numbered boxes represent coding and UTR exons, respectively. The scale at the right indicates the percentage of sequence identity. Physical distance is given in kilobases (kb).

map of the human genome is well underway³⁴, making it possible to undertake a systematic approach to scanning the genome for associations to disease-related and other phenotypes. □

Methods

The mapping and sequencing methods used in the assembly of the bacterial clone and sequence map of chromosome 10, respectively, as well as the tools of the gene annotation pipeline are described in refs 1 and 35 (see ref. 36 for detailed protocols). Manual annotation of gene structures followed the guidelines agreed in the human annotation workshop (HAWK; <http://www.sanger.ac.uk/HGP/havana/hawk.shtml>), whereas gene symbols were approved where possible by the HUGO Gene Nomenclature Committee³⁷. Protein translations were analysed with InterProScan (<http://www.ebi.ac.uk/interpro/scan.html>), which was run via the Ensembl protein annotation pipeline, to obtain Pfam, Prosite, Prints and Profiles domain matches.

Alignments for inter- and intrachromosomal duplications were performed with WU-BLASTN (<http://blast.wustl.edu>) using the current sequence assembly of chromosome 10 and the NCBI34 build for the rest of the genome. All sequences were repeat masked with RepeatMasker (<http://repeatmasker.genome.washington.edu>) and low-quality alignments (e -value $> 10^{-30}$, sequence identity $< 90\%$, length < 80 bp) were discarded. For intrachromosomal duplications, self matches were discarded. For interchromosomal duplications, the sequence was split into 400-kb segments. Adjacent matches in the same orientation were joined together as described by ref. 38. Only blocks of 10 kb or greater were retained.

The following sequence assemblies were used for comparative analysis: *M. musculus* NCBI build 30 (Mouse Genome Sequencing Consortium; http://www.ensembl.org/Mus_musculus/resources.html); *R. norvegicus* version 2.0 (Rat Genome Sequencing Consortium; <http://www.hgsc.bcm.tmc.edu/projects/rat>); *D. rerio* Assembly version 1 (Sanger Institute; http://www.sanger.ac.uk/Projects/D_rerio); *F. rubripes* version 2 (International Fugu Genome Consortium; <http://www.fugu-sg.org/project/info.html>); and *T. nigroviridis* version 6 (Genoscope and Whitehead Institute for Genome Research; <http://www.genoscope.cns.fr/externe/tetraodon/Ressource.html>). The repeat-masked sequence of chromosome 10 was aligned to the mouse and rat genome sequences using BLASTZ³⁹ and the resulting matches were post-processed by axtBest and subsetAxt (W. J. Kent; <http://www.soe.ucsc.edu/~kent/src>) as described elsewhere³⁵. Alignment of the three fish genome sequences to chromosome 10 was performed with WU-TBLASTX using the scoring matrix, parameters and filtering strategy applied by Exofish⁴⁰. Overlapping alignments to different sequences were merged to produce contiguous regions of sequence conservation, analogous to the ECRs or 'Ecores', reported by Exofish.

SNPs in sequence overlaps were identified using a modification of the SSAHA software⁴¹. The chromosome 10 SNPs (dbSNP release 115) were mapped to the sequence assembly of this chromosome (this study) first with SSAHA and then Cross-Match. Of the approximately 14 million chimpanzee reads (<http://www.genome.gov/11008056>) mapped onto the human sequence assembly (NCBI build 31), those mapping to chromosome 10 were remapped to our sequence assembly and used to identify human-chimpanzee SBDs.

Received 17 December 2003; accepted 9 March 2004; doi:10.1038/nature02462.

1. Bentley, D. R. *et al.* The physical maps for sequencing human chromosomes 1, 6, 9, 10, 13, 20 and X. *Nature* **409**, 942–943 (2001).
2. Guy, J. *et al.* Genomic sequence and transcriptional profile of the boundary between pericentromeric satellites and genes on human chromosome arm 10p. *Genome Res.* **13**, 159–172 (2003).
3. Guy, J. *et al.* Genomic sequence and transcriptional profile of the boundary between pericentromeric satellites and genes on human chromosome arm 10q. *Hum. Mol. Genet.* **9**, 2029–2042 (2000).
4. Jackson, M. S., See, C. G., Mulligan, L. M. & Lauffart, B. F. A 9.75-Mb map across the centromere of human chromosome 10. *Genomics* **33**, 258–270 (1996).
5. Felsenfeld, A., Peterson, J., Schloss, J. & Guyer, M. Assessing the quality of the DNA sequence from the Human Genome Project. *Genome Res.* **9**, 1–4 (1999).
6. Deloukas, P. *et al.* A physical map of 30,000 human genes. *Science* **282**, 744–746 (1998).
7. Kong, A. *et al.* A high-resolution recombination map of the human genome. *Nature Genet.* **31**, 241–247 (2002).
8. Broman, K. W., Murray, J. C., Sheffield, V. C., White, R. L. & Weber, J. L. Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am. J. Hum. Genet.* **63**, 861–869 (1998).
9. Deloukas, P. *et al.* The DNA sequence and comparative analysis of human chromosome 20. *Nature* **414**, 865–871 (2001).
10. International Human Genome Sequencing Consortium, Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
11. Antequera, F. & Bird, A. Number of CpG islands and genes in human and mouse. *Proc. Natl Acad. Sci. USA* **90**, 11995–11999 (1993).
12. Down, T. A. & Hubbard, T. J. Computational detection and location of transcription start sites in mammalian genomic DNA. *Genome Res.* **12**, 458–461 (2002).
13. Davuluri, R. V., Grosse, I. & Zhang, M. Q. Computational identification of promoters and first exons in the human genome. *Nature Genet.* **29**, 412–417 (2001).

14. Wagner, E. G. & Simons, R. W. Antisense RNA control in bacteria, phages, and plasmids. *Annu. Rev. Microbiol.* **48**, 713–742 (1994).
15. Eddy, S. R. Non-coding RNA genes and the modern RNA world. *Nature Rev. Genet.* **2**, 919–929 (2001).
16. Kiyosawa, H. *et al.* Antisense transcripts with FANTOM2 clone set and their implications for gene regulation. *Genome Res.* **13**, 1324–1334 (2003).
17. Bailey, J. A. *et al.* Recent segmental duplications in the human genome. *Science* **297**, 1003–1007 (2002).
18. Crosier, M. *et al.* Human paralogs of KIAA0187 were created through independent pericentromeric-directed and chromosome-specific duplication mechanisms. *Genome Res.* **12**, 67–80 (2002).
19. Bryce, S. D. *et al.* A novel family of cathepsin L-like (CTSLL) sequences on human chromosome 10q and related transcripts. *Genomics* **24**, 568–576 (1994).
20. Deloukas, P., Dauwerse, J. G., Moschonas, N. K., van Ommen, G. J. & van Loon, A. P. Three human glutamate dehydrogenase genes (GLUD1, GLUDP2, and GLUDP3) are located on chromosome 10q, but are not closely physically linked. *Genomics* **17**, 676–681 (1993).
21. Yu, A. *et al.* Comparison of human genetic and sequence-based physical maps. *Nature* **409**, 951–953 (2001).
22. Thomas, J. W. *et al.* Comparative analyses of multi-species sequences from targeted genomic regions. *Nature* **424**, 788–793 (2003).
23. Schwartz, S. *et al.* PipMaker—a web server for aligning two genomic DNA sequences. *Genome Res.* **10**, 577–586 (2000).
24. Sachidanandam, R. *et al.* A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409**, 928–933 (2001).
25. Vanaja, D. K., Chevillat, J. C., Iturria, S. J. & Young, C. Y. Transcriptional silencing of zinc finger protein 185 identified by expression profiling is associated with prostate cancer progression. *Cancer Res.* **63**, 3877–3882 (2003).
26. Deloukas, P., French, L., Meitinger, T. & Moschonas, N. K. Report of the third international workshop on human chromosome 10 mapping and sequencing 1999. *Cytogenet. Cell Genet.* **90**, 1–12 (2000).
27. Li, J. *et al.* PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *Science* **275**, 1943–1946 (1997).
28. Bonneau, D. & Longy, M. Mutations of the human PTEN gene. *Hum. Mutat.* **16**, 109–122 (2000).
29. Eng, C. PTEN: one gene, many syndromes. *Hum. Mutat.* **22**, 183–198 (2003).
30. Morante-Redolat, J. M. *et al.* Mutations in the LGI1/Epitempin gene on 10q24 cause autosomal dominant lateral temporal epilepsy. *Hum. Mol. Genet.* **11**, 1119–1128 (2002).
31. Sarafidou, T. *et al.* Folate-sensitive fragile site *FRA10A* is due to an expansion of a CGG-repeat in a novel gene *FRA10AC1*, encoding a nuclear protein. *Genomics* (in the press).
32. Moschonas, N. K. in *Encyclopedia of the Human Genome* Vol. 1, 618–625 (Nature Publishing Group, London, 2003).
33. Boutin, P. *et al.* *GAD2* on chromosome 10p12 is a candidate gene for human obesity. *PLoS Biol.* **1**, 1–11 (2003).
34. The International HapMap Consortium. The International HapMap project. *Nature* **426**, 789–796 (2003).
35. Mungall, A. J. *et al.* The DNA sequence and analysis of human chromosome 6. *Nature* **425**, 805–811 (2003).
36. Dunham, I. (ed.) *Genome Mapping and Sequencing* (Horizon Scientific, Wymondham, UK).
37. Wain, H. M., Lush, M., Ducluzeau, F. & Povey, S. Genew: the human gene nomenclature database. *Nucleic Acids Res.* **30**, 169–171 (2002).
38. Cheung, J. *et al.* Genome-wide detection of segmental duplications and potential assembly errors in the human genome sequence. *Genome Biol.* **4**, R25 (2003).
39. Schwartz, S. *et al.* Human–mouse alignments with BLASTZ. *Genome Res.* **13**, 103–107 (2003).
40. Roest Crollius, H. *et al.* Estimate of human gene number provided by genome-wide analysis using *Tetraodon nigroviridis* DNA sequence. *Nature Genet.* **25**, 235–238 (2000).
41. Ning, Z., Cox, A. J. & Mullikin, J. C. SSAHA: a fast search method for large DNA databases. *Genome Res.* **11**, 1725–1729 (2001).

Supplementary Information accompanies the paper on www.nature.com/nature.

Acknowledgements We thank the *Fugu*, Mouse and Rat Genome Sequencing Consortia; Genoscope and Whitehead Institute for Genome Research (*Tetraodon* genome data); the members of the zebrafish project at the Sanger Institute; the groups at the Washington University Genome Sequencing Center and the Whitehead Institute for early access to the chimpanzee whole-genome shotgun data; the laboratories of H. Riethman and M. S. Jackson for the provision of telomeric and pericentromeric clones, respectively; H. M. Wain, E. A. Bruford, C. C. Talbot, V. K. Khodiyar, M. J. Lush, M. W. Wright and S. Povey for assigning official gene symbols; the EMBL, GenBank and Ensembl database groups; I. Dunham for critical reading of the manuscript; and the Wellcome Trust for financial support.

Competing interests statement The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to P.D. (panos@sanger.ac.uk). All DNA sequences reported in this study have been deposited in the EMBL or GenBank databases and the accession numbers can be found in Supplementary Fig. S1.

Short sequence-paper

Cloning, chromosomal organization and expression analysis of *Neurl*¹, the mouse homolog of *Drosophila melanogaster neuralized* gene

Elias Pavlopoulos^{a,b}, Maria Kokkinaki^{a,b}, Evangelia Koutelou^a,
Thymios A. Mitsiadis^c, Panagiotis Prinos^{d,2}, Christos Delidakis^{a,b},
Michael W. Kilpatrick^d, Petros Tsipouras^d, Nicholas K. Moschonas^{a,b,*}

^a Department of Biology, University of Crete, P.O. Box 2208, Heraklion, 714 09, Crete, Greece

^b Institute of Molecular Biology and Biotechnology, Foundation of Research and Technology-Hellas, P.O. Box 1527, Heraklion, 711 10, Crete, Greece

^c IMEB EA 2198, Faculté d'Odontologie, Université de la Méditerranée, Marseille, France

^d Department of Pediatrics, University of Connecticut Health Center, Farmington, CT 06030, USA

Received 2 August 2001; received in revised form 1 October 2001; accepted 2 October 2001

Abstract

The *Drosophila neuralized* (*neur*) gene belongs to the neurogenic group of genes involved in regulating cell–cell interactions required for neural precursor development. *neur* mutant phenotypes include strong overcommitment to neural fates at the expense of epidermal fates. The human *neuralized* homolog (*NEURL*) has been recently determined and found to map to chromosome 10q25.1 within the region frequently deleted in malignant astrocytomas. Because of its potential importance in developmental processes, we analyzed the structure of the mouse homolog, *Neurl*, and its expression pattern in embryonic tissues. *Neurl* activity is detected from early developmental stages in several tissues and organs including neural tissues, limbs, the skeletal system, sense organs and internal organs undergoing epithelial–mesenchymal interactions. *Neurl* encodes a polypeptide associated with the plasma membrane but also detected in the cytoplasm. Similarly to the *Drosophila* gene, mammalian *neuralized* may code for an important regulatory factor. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Neurogenic gene; NEUZ domain; RING finger domain; Mouse embryo development; Malignant astrocytoma

Drosophila neuralized encodes a RING finger polypeptide and belongs to the neurogenic genes which also include *Notch*, *Delta*, *Enhancer of split*, *big brain*, and *mastermind* [1,2]. These genes are involved in cell fate determination of, among others, the neurogenic regions of the *Drosophila* embryo [3] by mediating lateral inhibition, a process whereby neural precursor cells inhibit neighboring cells from adopting a neural fate [4,5]. Loss-of-function mutations of any of the neurogenic genes are associated with hyperplasia of the embryonic nervous sys-

tem [6]. *neur* represents a strong neurogenic mutant. Recent studies have suggested that *neur* participates in *Notch-Delta* signaling during *Drosophila* peripheral neurogenesis [7,8]. The human homolog, *NEURL*, has been identified and mapped within the region of loss of heterozygosity at chromosome 10q25.1, the most frequent chromosomal aberration found in malignant astrocytomas [9] and references therein). Compared to normal brain tissue, *NEURL* was found to be expressed at very low levels in astrocytomas and several glioma cell lines examined [9]. Among *Drosophila*, *Caenorhabditis elegans* ([10]; Acc. No. U40945) and human *neuralized* genes [9], extensive sequence homology is observed at the carboxy-terminal C₃HC₄ RING finger domain. Plenty of regulatory proteins containing this domain are involved in important biological processes including oncogenesis and development (e.g. [11,12]). Recent evidence suggests that proteins containing this domain participate in protein interactions and regulate cellular protein levels serving as ubiquitin E3 ligases [13–15]. Based on *Drosophila neur* functional role in

* Corresponding author. Fax: +30-81-394-408.

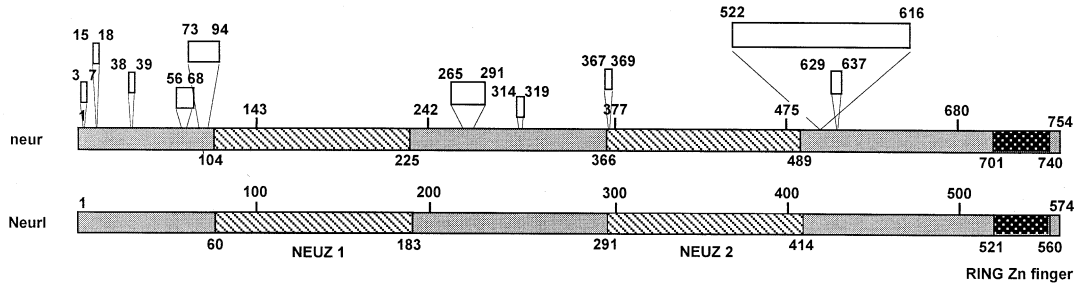
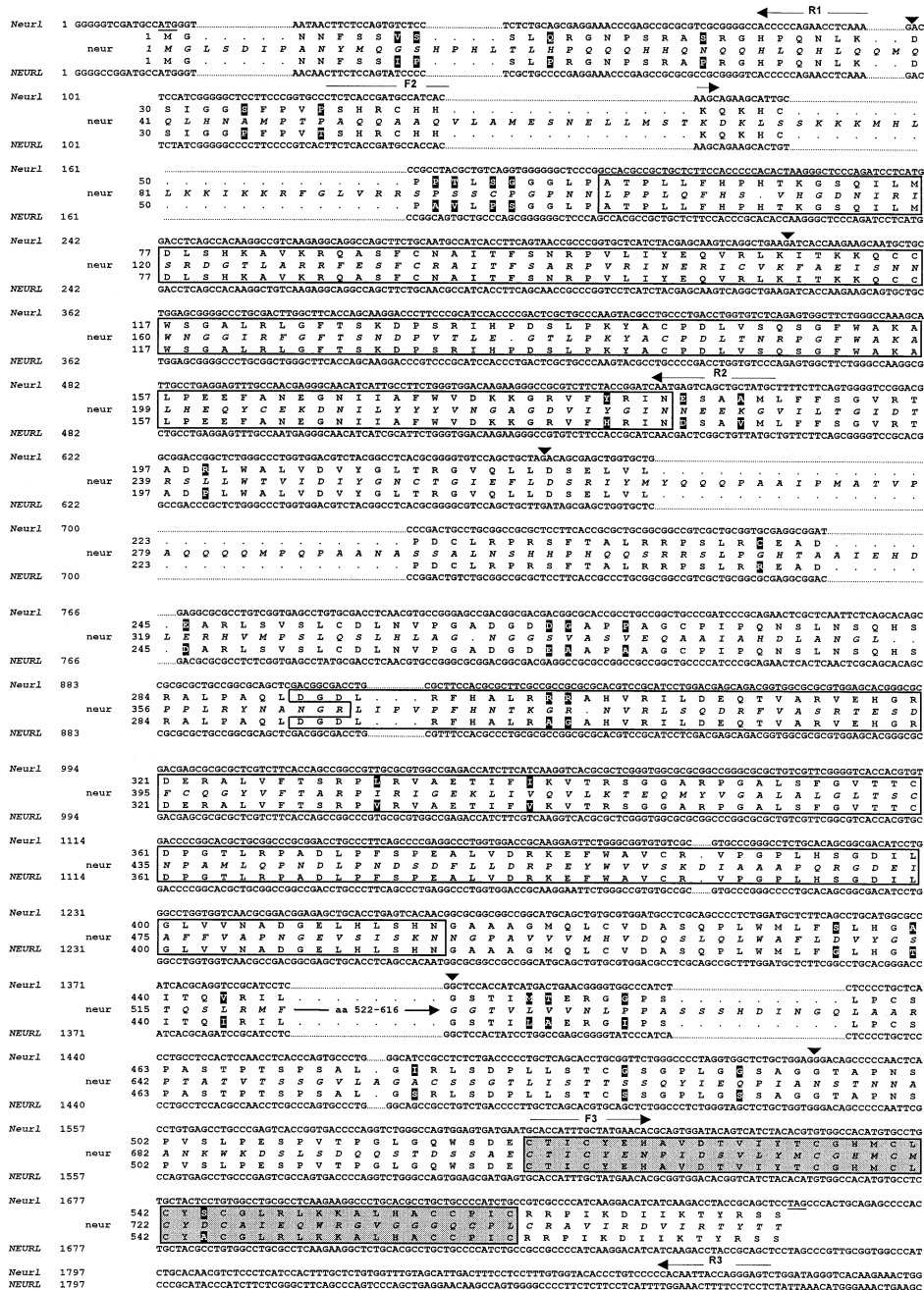
E-mail address: moschon@imbb.forth.gr (N.K. Moschonas).

¹ The mouse *Neurl* cDNA sequence has been submitted to GenBank with Acc. No.: NM_021360; *Neurl* genomic DNA sequences are available under GenBank Acc. Nos. AG271919–23. *NEURL* (*Drosophila neuralized*-like) is a gene symbol approved by the HUGO/GDB Nomenclature Committee.

² Present address: Laboratory of Cellular and Molecular Biology, Clinical Research Institut of Montreal, 110 Pine Avenue West, Montreal, QC, Canada H2W 1R7.

neurogenesis, it has been hypothesized that NEURL may be involved in cell fate determination of the nervous system, and its inactivation may be associated with progression of malignant neuroectodermal tumors in human [9].

Because of its potential contribution in important developmental processes in mammals, we isolated and analyzed the mouse *neur* homolog, *Neurl*. Here, we present its structure and genomic organization, the subcellular localization



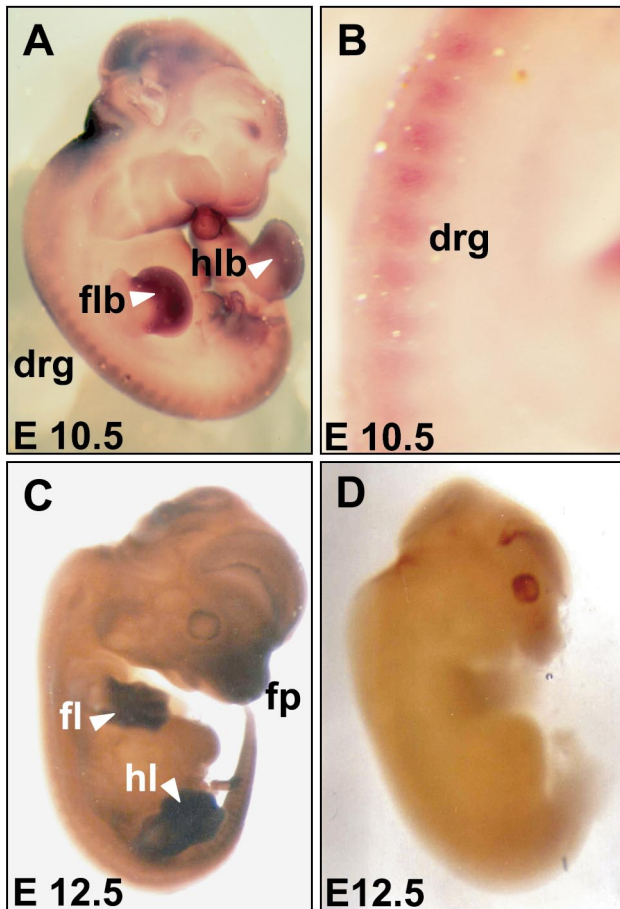


Fig. 2. Whole-mount in situ hybridization to a *Neurl* DIG-labeled probe. Lateral view of E10.5–11.0 and E12.5 BALB/C mouse embryos. *Neurl* transcripts are clearly detected in the forelimb and the hindlimb bud (flb and hlb, respectively, A), and the dorsal root ganglia (drg; B, zoom in of A). At E12.5 (C), *Neurl* activity is detected in the facial processes (fp) and the distal parts of the developing limbs (fl, forelimb; hl, hindlimb). The *Neurl* antisense riboprobe was synthesized by T3 RNA polymerase (Promega) from a 1275 bp *SmaI/EcoRI Neurl* cDNA template cloned in pBS KSII/+ and linearized at the *SmaI* site. D is a control E12.5 embryo hybridized to the sense *Neurl* riboprobe.

and the expression profile during mouse embryo development.

An IMAGE human EST (#179548) was used to probe a C57BL/6 mouse E14.5 λ ExCell *EcoRI* CIP cDNA library (inserts containing *EcoRI*–*NotI* adaptors; a generous gift of R. Wehr and P. Gruss) under standard conditions. Overlapping restriction fragments of a positive 2433 bp long cDNA clone, λ *Neurl*, were appropriately subcloned

in pBluescript (pBS), sequenced on both strands on an ABI 373 sequencer and evaluated in silico. Sequence alignment to the *NEURL* cDNA (GenBank Acc. Nos.: NM_004210, [9] and AF029729; the latter corresponding to a 3634 bp long *NEURL* cDNA, unpublished data of our group) and in silico analysis (see also, below) revealed a nearly full-length, 1710 bp long, open reading frame (i.e. 570 aa) followed by 723 bp of a 3'-untranslated region (UTR) (Fig. 1). To generate the missing part of the coding sequence, a *Neurl* primer (R1; Fig. 1) and a *NEURL*-specific primer (5'-GCAGCTCCTGCCCGGCCT-3'; AF029729) located at the 5'-UTR of the human gene (data not shown) assumed to hybridize to the respective region of the mouse gene, were employed to reverse transcribe and PCR amplify (RT-PCR) an E14.5 BALB/C mouse brain poly A⁺ RNA sample (80 ng) pre-treated by DNase I. The RT DNA product was amplified with denaturation at 94°C (1 min), annealing at 61°C (45 s) and extension at 72°C (2 min). Indeed, the 140 bp product included the missing *Neurl* translation initiation region as well as the adjacent 55 bp of the putative 5'-UTR of the gene (GenBank Acc. No. NM_021360). As suggested by a recently submitted *Neurl* cap-trapper selected cDNA clone ([16]; GenBank Acc. No. AK010787), the transcription initiation site extends 261 bp further upstream, thus making 316 bp the overall size of *Neurl* 5'-UTR (not shown).

Alignment between *NEURL* and the composite *Neurl* cDNA sequence suggested a high degree of homology at both the nucleotide and the cDNA-deduced polypeptide level (Fig. 1). The length of the two putative polypeptides is the same, i.e. 574 aa, sharing 97% sequence similarity and 94% identity. Thirty-three amino acid changes may cause slight variation of their physicochemical properties, however, they possibly do not interfere with their functional characteristics. Furthermore, comparison between mouse/human and *Drosophila neuralized* suggested a protein sequence identity of about 33% and a similarity of 57% determined throughout the molecules (Fig. 1). This is not an uncommon finding when mammalian and insect homologous polypeptides are compared [17]. Although the molecules share the same overall architecture (see below), there is a 180 aa difference between the mammalian and the *Drosophila* molecule (Fig. 1). The extra *Drosophila* sequences may reflect additional functional characteristics of the insect molecule or, simply, represent unstructured extra chain fragments linking adjacent polypeptide do-

Fig. 1. Upper part: Nucleotide (nt) and derived amino acid (aa) sequence of *Neurl* aligned to human *NEURL* gene and *D. melanogaster* neur protein. *Neurl* is equivalent to the recently reported *Neurl1* [26]. The alignment was done by using either the 'GAP' or the 'BEST FIT' software (GCG Wisconsin Package, Accelrys). The proposed translation initiation codon [9] and the stop codon are underlined. The exon boundaries have remained the same in mouse and human; they are indicated by vertical arrow heads. Boxed are the neuralized-specific domains, NEUZ (open) and the RING finger (gray). Amino acid differences between *Neurl* and *NEURL* are shown in black boxes. The *D. melanogaster* neur protein sequence is shown in italics. Arrows indicate the position and orientation of synthetic oligonucleotides used as PCR primers for structural (R1, F2/R2) or expression pattern (F3/R3) analysis. Lower part: Comparative schematic presentation of the mouse (*Neurl*) with the *Drosophila* (*neur*) proteins. The overall organization has remained the same. Extra *Drosophila* protein regions are indicated by open boxes. Numbers show aa positions.

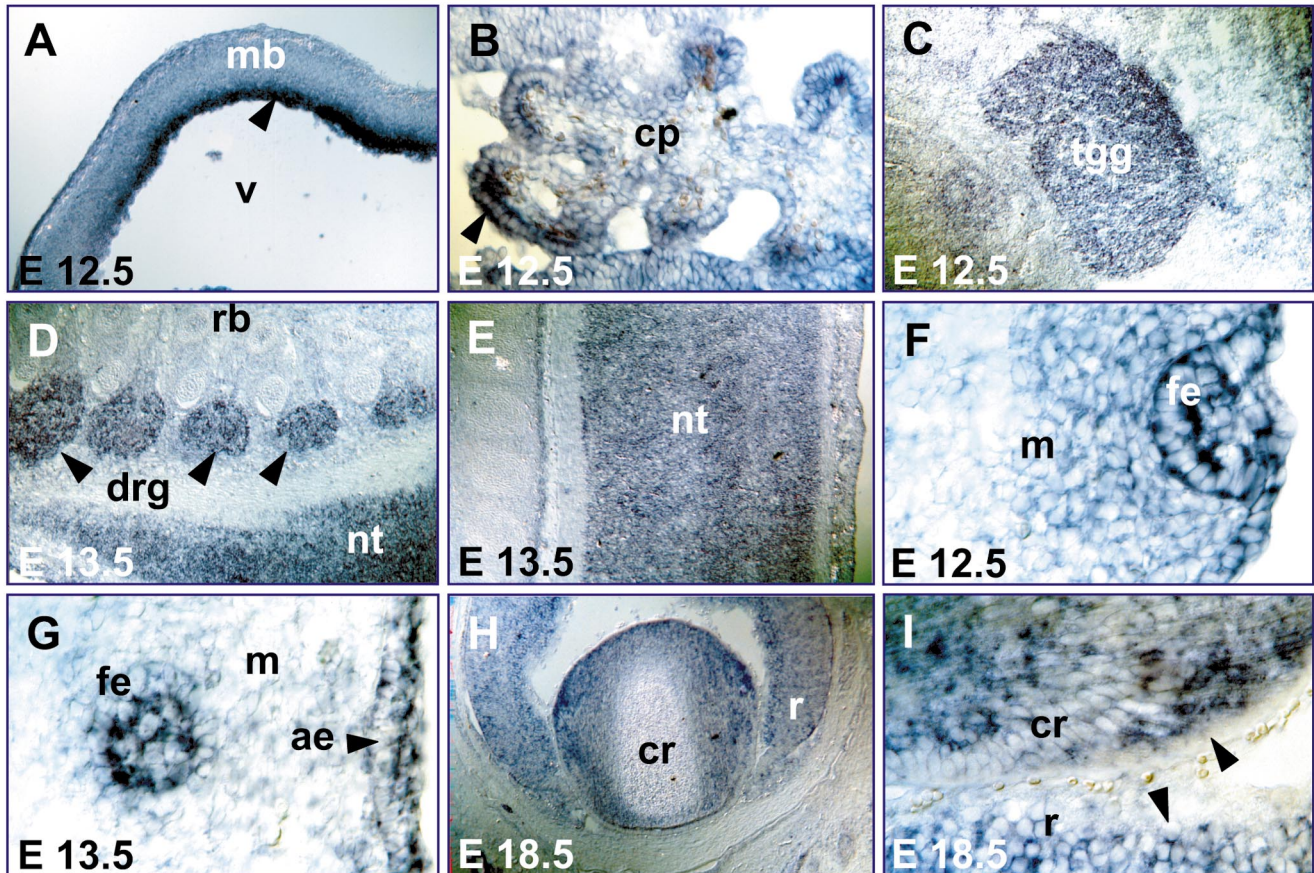


Fig. 3. Representative cryosection in situ hybridization panels of E12.5, E13.5 and E18.5 Swiss mouse embryos. *Neurl* transcripts are detected in the ventricular zone of the midbrain (mb), A; the choroid plexus (cp) epithelium, B; the trigeminal ganglion (tgg), C; the dorsal root ganglia (drg), D; the neural tube (nt), E; the whisker follicular epithelium (fe), F and G; and the eye, retina (r) and crystalline or lens fibers (cr), H and I (detail of H). Additional abbreviations: v, third ventricle; rb, rib; m, mesenchyme; ae, aboral epithelium. All sections are sagittal.

mains. Analysis of the putative *Neurl* polypeptide sequence using the SMART server tools (smart.embl-heidelberg.de/), identified a carboxy-terminal C₃HC₄ RING finger, a domain suggested to be involved in DNA–protein, RNA–protein and protein–protein interactions which is characteristic of a large variety of regulatory proteins (e.g. [15,18]), and two NEUZ domains, in the same order as in human, *Drosophila* and *C. elegans* homologs. NEUZ, a recently introduced domain of unknown function, was firstly identified as an internal neuralized-specific repeat in NEURL [9]. Interestingly, two novel cDNA-deduced protein groups have been predicted as containing NEUZ domains but no RING fingers (SMART server). The first consists of two possibly orthologous genes, i.e. *Drosophila melanogaster* Q9VUC2 and human Q9H0B0, containing four and six NEUZ domains, respectively. The second includes a *D. melanogaster* gene, Q9W113, and two highly homologous mammalian genes, i.e. human Q9BRO9 and mouse Q9D0S4, all containing a NEUZ domain accompanied by a SOCs ('suppressor of cytokine signaling') domain. Sequence similarity among the above and neuralized proteins is limited to NEUZ (not shown) suggesting that it is a well-conserved protein domain and that neuralized

belongs to a functionally diverged NEUZ superfamily with members both in vertebrate and invertebrate species.

Neurl gene locus was mapped to mouse chromosome 19, between *Pitx3* and *D19Mit1* [19], a region syntenic to human chromosome 10q25.1 [20] associated with several genetic defects of the nervous system and developmental malformations (<http://www3.ncbi.nlm.nih.gov/OMIM/>). To investigate whether *Neurl* exists in a single or multiple copies in the mouse genome, we performed standard genomic DNA Southern analysis using three non-overlapping *Neurl* cDNA probes sequentially hybridized, at high stringency, to a nylon blot containing several lanes of BALB/C mouse DNA digests. The simple hybridization pattern detected, is consistent with the existence of a unique *Neurl* gene (data not shown). To determine the genomic organization of *Neurl*, we screened a mouse (strain 129SV) liver λFIXII genomic library (Stratagene) using as probes three cDNA fragments, i.e. the almost full-length 2433 bp *Neurl* cDNA together with a 231 bp *EcoRI/BstYI* and a 424 bp *NotI/SacII* fragment corresponding to the 5'- and the 3'-end of the gene, respectively. Three positive clones representing more than 52 kb covered the entire *Neurl*-specific genomic region (data

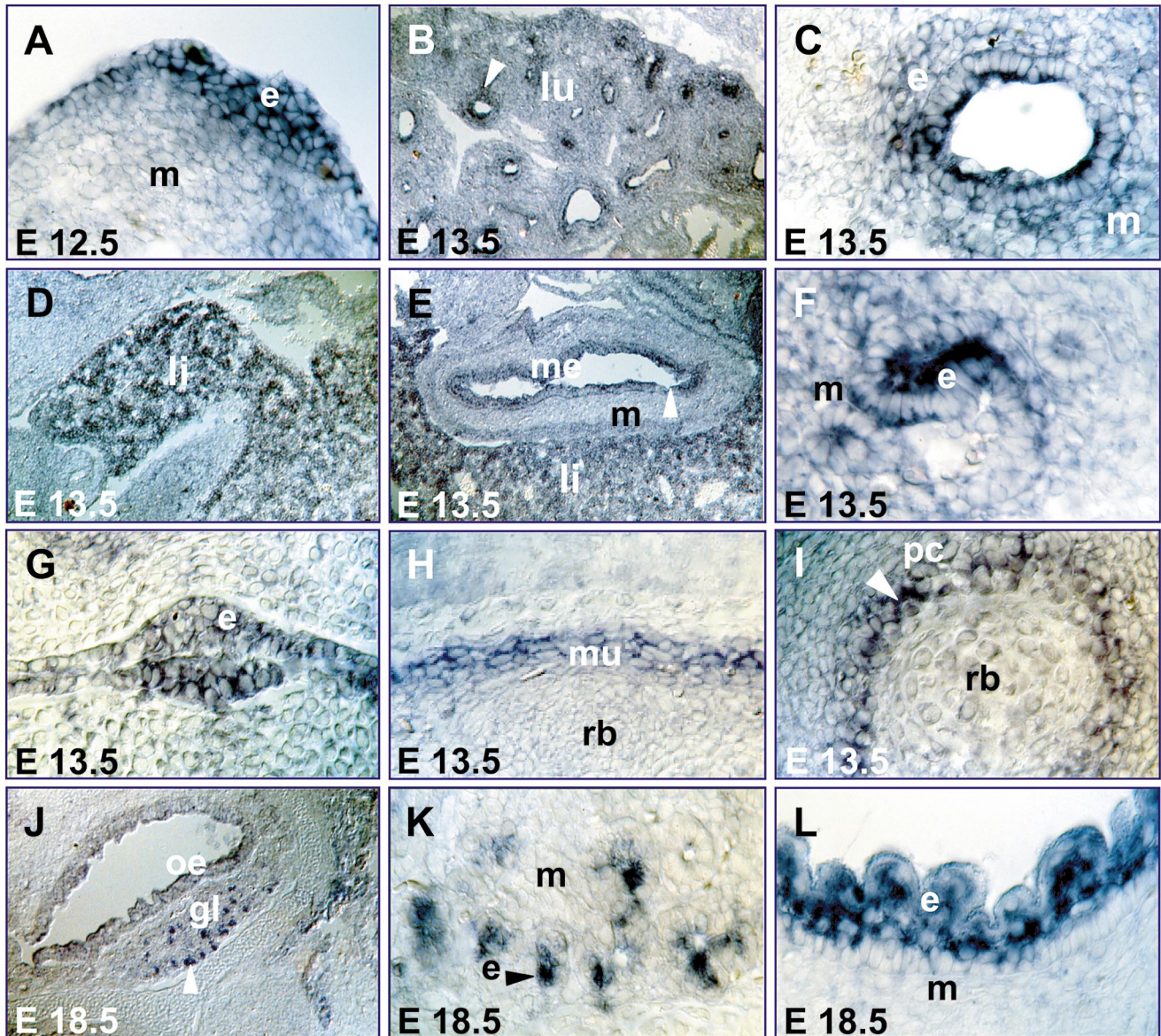


Fig. 4. Representative cryosection in situ hybridization panels of E12.5, E13.5 and E18.5 mouse embryos. *Neurl* transcripts are detected in limb epithelium (e), A; lung (lu) alveolar epithelium, B and C (detail of B); liver (li) parenchyma, D; mucosa epithelium (me) of stomach, E; kidney tubular epithelium, F; epithelium (e) of mesonephric (Wolffian) urogenital duct, G; muscles (mu), H; perichondrium (pc) of the ribs (rb), I; nasal olfactory epithelium (oe) and gland (gl) epithelium (e), J and K (detail of J); facial skin epithelium, L. Additional abbreviation: m, mesenchyme. All sections are sagittal.

not shown). Exon/intron boundaries, intron approximate sizes and the full sequence of the second intron were determined by Southern analysis and appropriate subcloning and sequencing (GenBank Acc. Nos. AG271919–23). *Neurl* coding sequence is embedded in a region of about 21 kb (not shown) and consists of six exons (Fig. 1). Mapping analysis suggested an intron of about 10 kb interrupting the 3'-UTR of *Neurl* (data not shown). In silico analysis of *NEURL* cDNA and genomic sequence (GenBank Acc. Nos. AL139339 and AL121929) permitted us to determine the respective structure of the human homolog. Comparison with *Neurl* suggested that the two genes share the same exon/intron organization; however the relative sizes of most of the introns differ significantly (not shown).

Furthermore, *NEURL* 3'-UTR is not interrupted by an intron suggesting an increased structural divergence in the non-translated regions of the two genes.

Expression analysis of *NEURL* suggested that the gene is transcriptionally active in normal human brain and at much lower levels in malignant astrocytoma tissues [9]. Investigation of *Neurl* expression pattern by RNA in situ hybridization on mouse embryo tissues permitted us to identify sites potentially requiring *Neurl* function. A digoxigenin (DIG)-labeled *Neurl* antisense riboprobe was used under conditions previously described [21]. *Neurl* expression was detected at specific sites and cell layers in several organs of the developing embryo. In situ hybridization on whole-mount E10.5 and E12.5 embryos, indicated expres-

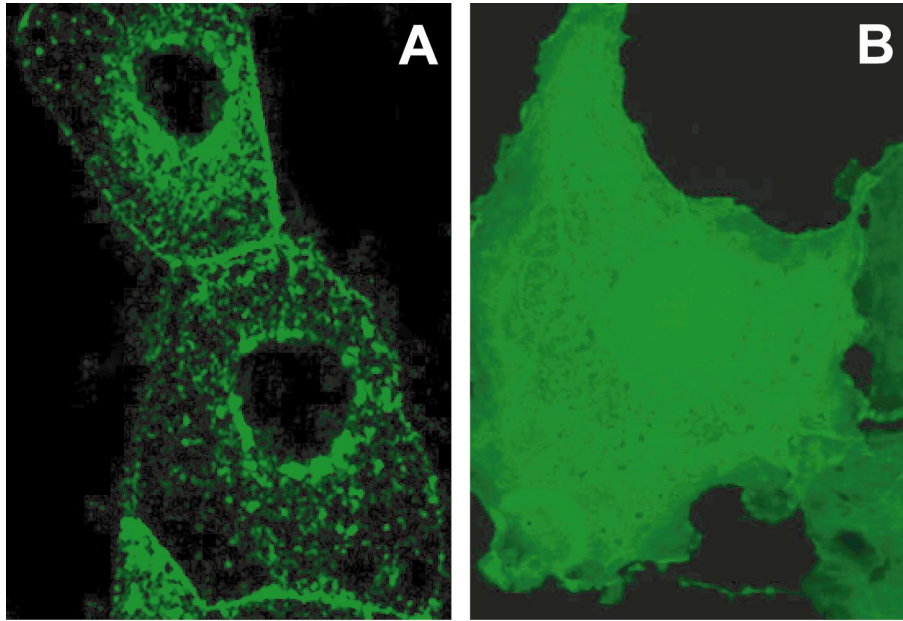


Fig. 5. Confocal fluorescence optical sections of transfected COS-7 cells showing the subcellular localization of GFP tag Neurl chimeras. (A) Transfection performed by pEGFP/Neurl, generated by direct insertion of the 2083 bp *EcoRI* (at λ ExCell adaptor)/*KpnI* fragment of λ Neurl to the *EcoRI/KpnI* linearized pEGFPC1 vector. About 2×10^5 COS-7 cells were transfected with 0.5–1.0 μ g of plasmid DNA using the calcium phosphate co-precipitation method [22] and examined 24–36 h after transfection. (B) Transfection performed by pEGFP/ Δ Neurl, made from pEGFP/Neurl by excision of a 677 bp *NotI* (at λ ExCell adaptor)/*NotI* (at nt 693; Fig. 1) fragment.

sion of *Neurl* in the developing limb buds, dorsal root ganglia and facial processes (Fig. 2A–C). Furthermore, analysis of cryosections from E12.5, E13.5 and E18.5 embryos (Figs. 3 and 4) revealed *Neurl* transcripts in neural tissues, such as the ventricular (proliferative) zone of the midbrain (Fig. 3A), the choroid plexus epithelium (Fig. 3B), the neural tube (Fig. 3E), the trigeminal ganglion (Fig. 3C), the dorsal root ganglia (Fig. 3D) and sense organs, i.e. the retina and eye lens cells (Fig. 3H, I), the olfactory and gland epithelium of the nose (Fig. 4J, K) and the whisker cells (Fig. 3F, G) undergoing epithelial–mesenchymal interactions. *Neurl* is also expressed at distinct cell layers of internal organs, i.e. lung alveolar epithelium (Fig. 4B, C), liver parenchyma (Fig. 4D), stomach mucosa epithelium (Fig. 4E), kidney tubular epithelium (Fig. 4F), mesonephric (Wolffian) urogenital duct epithelium (Fig. 4G), limb epithelium (Fig. 4A), muscle (Fig. 4H), perichondrium of the ribs (Fig. 4I) and in facial skin epithelium (Fig. 4L). In addition, RT-PCR applying the F3/R3 set of primers (Fig. 1) to DNA-free poly A⁺ RNA showed that *Neurl* is also expressed in the adult brain, testis and liver as well as in embryonal mouse NIH 3T3 and teratocarcinoma P19 cells (not shown). Similarly to mouse, in situ hybridization to 6 weeks human embryo sections, showed *NEURL* expression throughout the skin, the differentiating cartilaginous elements of the limb, rib and vertebrae, as well as in liver parenchyma, lung and muscle (R. Kosher, C. Dealy and P. Tspouras, unpublished observations).

Finally, to assess the subcellular localization of Neurl,

we proceeded in transient expression of a green fluorescent protein (GFP) Neurl chimera in African green monkey COS-7 cells. The construct, pEGFP/Neurl, contained the enhanced GFP open reading frame (Clontech) combined at its 3'-end with the almost full-length *Neurl* translated region (Neurl aa 5–574). The GFP tag Neurl was found associated with the plasma membrane as well as in the cytoplasm localized to punctate structures (Fig. 5A). The same accumulation pattern was obtained when the GFP tag was placed at the C-terminal region of the chimera suggesting that GFP possibly does not interfere with Neurl localization. In similar transfection experiments using mouse neuroblastoma Neuro2a cells, Ruan et al. [26] found Neurl throughout the cytoplasm. The difference in the accumulation pattern observed may be due to different cell types employed. Notably, *Drosophila* epitope-tagged *neuralized* protein has been recently shown by in vivo experiments to be primarily associated with the plasma membrane, also localized to punctate structures [7,8]. Since there are no recognizable signal peptide or transmembrane domains in either the mammalian or the *Drosophila* molecules (SMART search), it is likely that Neurl will have a peripheral membrane association. Biochemical experiments are underway to elucidate the basis of Neurl association with the membrane and the nature of the punctate structures in the cytoplasm. The latter may be relevant to the current generalized concept about RING proteins function, i.e. their ability to form large macromolecular scaffolds required for various complex molecular or enzymatic processes taking place in the cell [14]. To

get a better understanding on *Neurl* subcellular localization in relation to its structure, we extended our analysis to a series of truncated *Neurl* GFP tag derivatives from either the N- or the C-terminus. Our data revealed that the absence of 226 aa from the *Neurl* N-terminus (chimeric plasmid pEGFP/ Δ *Neurl*) results in the abolishment of the characteristic accumulation pattern and drives the molecule to nucleo-cytoplasmic diffusion (Fig. 5B) reminiscent of the control GFP accumulation pattern [23]. Interestingly, expression of a reciprocal *Neurl* (aa 1–269)/GFP tag chimera containing the first NEUZ domain resulted in a pattern comparable to that of the intact molecule (not shown). This suggests the existence of a necessary and sufficient element for the proper subcellular accumulation of *Neurl* within its N-terminal 269 aa.

In conclusion, we have analyzed the structure and genomic organization of the mouse *Neurl* homolog and compared its genomic organization to the human counterpart. Orthologous *neuralized* genes encoding a subgroup of RING proteins may, also, be considered as members of a novel gene family characterized by the presence of one or multiple NEUZ domains, the biological role of which remains to be elucidated. The gene is active from early embryonic stages, in several tissues (mostly epithelia) in a variety of sense and internal organs as well as in the nervous and skeletal system. The *Drosophila* homolog seems to be intimately associated with the *Notch-Delta* signaling [3,7,8,24]. Interestingly, in the developing mouse embryo, the expression pattern of *Delta1(Dll1)*, a homolog of *Drosophila Delta* gene, includes several sites of activity, e.g. nasal olfactory and gland epithelia, lung respiratory epithelia, skeletal muscles, whiskers, central nervous system [25], also exhibiting *Neurl* activity. Although extensive experiments are needed for a definitive proof, it is tempting to suggest a possible involvement of *Neurl* in *Notch*-dependent processes during fetal mammalian development.

The authors wish to thank Drs. R.A. Kosher and C. Dealy for permitting the use of unpublished observations, Dr. G. Chalepakis for useful discussions and A. Pasparki for technical assistance. This work was supported by IMBB (grant #56), University of Crete (grant #1199), and the Greek Ministry of Development (#236/234/603) (N.K.M.), HD-22610 and Coles Family Foundation (P.T.) and the Association pour la Recherche sur le Cancer (T.A.M.).

References

- [1] J.A. Campos-Ortega, Cellular interactions during early neurogenesis of *Drosophila melanogaster*, *Trends Neurosci.* 11 (1988) 400–405.
- [2] R. Lehmann, F. Jimenez, U. Dietrich, J.A. Campos-Ortega, On the phenotype and development of mutants of early neurogenesis in *Drosophila melanogaster*, *Roux's Arch. Dev. Biol.* 192 (1983) 62–74.
- [3] A. de la Concha, U. Dietrich, D. Weigel, J.A. Campos-Ortega, Functional interactions of neurogenic genes of *Drosophila melanogaster*, *Genetics* 118 (1988) 499–508.
- [4] S. Artavanis-Tsakonas, P. Simpson, Choosing a cell fate: a view from the Notch locus, *Trends Genet.* 7 (1991) 403–408.
- [5] S. Artavanis-Tsakonas, M.D. Rand, R.J. Lake, Notch signaling cell fate control and signal integration in development, *Science* 284 (1999) 770–776.
- [6] S. Artavanis-Tsakonas, K. Matsuno, M.E. Fortini, Notch signaling, *Science* 268 (1995) 225–232.
- [7] E.C. Lai, G.M. Rubin, neuralized functions cell-autonomously to regulate a subset of notch-dependent processes during adult *Drosophila* development, *Dev. Biol.* 231 (2001) 217–233.
- [8] E. Yeh, L. Zhou, N. Rudzik, G.L. Boulianne, Neuralized functions cell autonomously to regulate *Drosophila* sense organ development, *EMBO J.* 19 (2000) 4827–4837.
- [9] H. Nakamura, M. Yoshida, H. Tsuiki, K. Ito, M. Ueno, M. Nakao, K. Oka, M. Tada, M. Kochi, J. Kuratsu, Y. Ushio, H. Saya, Identification of a human homolog of the *Drosophila* neuralized gene within the 10q25.1 malignant astrocytoma deletion region, *Oncogene* 16 (1998) 1009–1019.
- [10] R. Wilson, R. Ainscough, K. Anderson, C. Baynes, M. Berks, J. Bonfield, J. Burton, M. Connell, T. Copesey, J. Cooper et al., 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*, *Nature* 368 (1994) 32–38.
- [11] Y. Miki, J. Swensen, D. Shattuck-Eidens, P.A. Futreal, K. Harshman, S. Tavtigian, Q. Liu, C. Cochran, L.M. Bennett, W. Ding, A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1, *Science* 266 (1994) 66–71.
- [12] J.I. Bardos, A.J. Saurin, C. Tissot, E. Duprez, P.S. Freemont, HPC3 is a new human polycomb orthologue that interacts and associates with RING1 and Bmi1 and has transcriptional repression properties, *J. Biol. Chem.* 275 (2000) 28785–28792.
- [13] K.L. Lorick, J.P. Jensen, S. Fang, A.M. Ong, S. Hatakeyama, A.M. Weissman, RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination, *Proc. Natl. Acad. Sci. USA* 96 (1999) 11364–11369.
- [14] K.L. Borden, RING domains: master builders of molecular scaffolds?, *J. Mol. Biol.* 295 (2000) 1103–1112.
- [15] P.K. Jackson, A.G. Eldridge, E. Freed, L. Furstenthal, J.Y. Hsu, B.K. Kaiser, J.D. Reimann, The lore of the RINGs: substrate recognition and catalysis by ubiquitin ligases, *Trends Cell Biol.* 10 (2000) 429–439.
- [16] P. Carninci, Y. Shibata, N. Hayatsu, Y. Sugahara, K. Shibata, M. Itoh, H. Konno, Y. Okazaki, M. Muramatsu, Y. Hayashizaki, Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes, *Genome Res.* 10 (2000) 1617–1630.
- [17] Y. Sasai, R. Kageyama, Y. Tagawa, R. Shigemoto, S. Nakanishi, Two mammalian helix-loop-helix factors structurally related to *Drosophila* hairy and Enhancer of split, *Genes Dev.* 6 (1992) 2620–2634.
- [18] D.P. Satiijn, M.J. Gunster, J. van der Vlag, K.M. Hamer, W. Schul, M.J. Alkema, A.J. Saurin, P.S. Freemont, R. vanDriel, A.P. Otte, RING1 is associated with the polycomb group protein complex and acts as a transcriptional repressor, *Mol. Cell. Biol.* 17 (1997) 4105–4113.
- [19] G.M. Hansen, D. Tackles, C. Schwartz, M.J. Justice, A mouse chromosome 19 genetic map including the Lvis viral insertion site, *Genomics* 56 (1999) 228–231.
- [20] S. Banfi, G. Borsani, E. Rossi, L. Bernard, A. Guffanti, F. Rubboli, A. Marchitello, S. Giglio, E. Coluccia, M. Zollo, O. Zuffardi, A. Ballabio, Identification and mapping of human cDNAs homologous to *Drosophila* mutant genes through EST database searching, *Nat. Genet.* 13 (1996) 167–174.
- [21] T.A. Mitsiadis, E. Hirsinger, U. Lendahl, C. Goridis, Delta-notch signaling in odontogenesis: correlation with cytodifferentiation and evidence for feedback regulation, *Dev. Biol.* 204 (1998) 420–431.

- [22] F.L. Graham, A.J. van der Eb, Transformation of rat cells by DNA of human adenovirus 5, *Virology* 54 (1973) 536–539.
- [23] M.A. Powers, D.J. Forbes, Cytosolic factors in nuclear transport: what's importin?, *Cell* 79 (1994) 931–934.
- [24] E.C. Lai, G.M. Rubin, Neuralized is essential for a subset of Notch pathway-dependent cell fate decisions during *Drosophila* eye development, *Proc. Natl. Acad. Sci. USA* 98 (2001) 5637–5642.
- [25] J. Beckers, A. Clark, K. Wunsch, M. Hrabe De Angelis, A. Gossler, Expression of the mouse Delta1 gene during organogenesis and fetal development, *Mech. Dev.* 84 (1999) 165–168.
- [26] Y. Ruan, L. Tecott, M.M. Jiang, L.Y. Jan, Y.N. Jan, Ethanol hypersensitivity and olfactory discrimination defect in mice lacking a homolog of *Drosophila* neuralized, *Proc. Natl. Acad. Sci. USA* 14 (98) (2001) 9907–9912.

Folate-sensitive fragile site *FRA10A* is due to an expansion of a CGG repeat in a novel gene, *FRA10AC1*, encoding a nuclear protein ☆☆☆

Theologia Sarafidou,^{a,1} Christina Kahl,^{b,1} Isabel Martinez-Garay,^{c,1} Marie Mangelsdorf,^{d,1} Stefan Gesk,^b Elizabeth Baker,^d Maria Kokkinaki,^a Polly Talley,^e Edna L. Maltby,^e Lisa French,^f Lana Harder,^b Bernd Hinzmann,^g Carlo Nobile,^h Kathy Richkind,ⁱ Merran Finnis,^d Panagiotis Deloukas,^f Grant R. Sutherland,^d Kerstin Kutsche,^c Nicholas K. Moschonas,^a Reiner Siebert,^b and Jozef Gécz^{d,*}

European Collaborative Consortium for the Study of ADLTE²

^aDepartment of Biology, University of Crete, and Institute of Molecular Biology and Biotechnology(IMBB), Foundation of Research and Technology (FORTH-GR), P.O. Box 2208, 714 09 Heraklion, Crete, Greece

^bInstitute of Human Genetics, University Hospital Schleswig-Holstein Campus Kiel, Schwanenweg 24, D-24105 Kiel, Germany

^cInstitute of Human Genetics, University Hospital Hamburg-Eppendorf, Butenfeld 42, D-22529 Hamburg, Germany

^dDepartment of Genetic Medicine, Women's and Children's Hospital, 72 King William Road, North Adelaide, SA 5006 Adelaide, Australia, and Department of Paediatrics, Adelaide University, Adelaide, Australia

^eNorth Trent Cytogenetics Department, Sheffield Children's NHS Trust, Sheffield S10 2TH, UK

^fThe Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK

^gmetaGen Pharmaceuticals GmbH, Oudenarder Strasse 16, D-13347 Berlin, Germany

^hCNR Institute of Neurosciences, Section of Padua, Viale G. Colombo 3, 35121 Padua, Italy

ⁱGenzyme, Genzyme Genetics, Santa Fe, NM 87505, USA

Received 8 October 2003; accepted 31 December 2003

Available online 23 April 2004

Abstract

Fragile sites appear visually as nonstaining gaps on chromosomes that are inducible by specific cell culture conditions. Expansion of CGG/CCG repeats has been shown to be the molecular basis of all five folate-sensitive fragile sites characterized molecularly so far, i.e., *FRAXA*, *FRAXE*, *FRAXF*, *FRA11B*, and *FRA16A*. In the present study we have refined the localization of the *FRA10A* folate-sensitive fragile site by fluorescence in situ hybridization. Sequence analysis of a BAC clone spanning *FRA10A* identified a single, imperfect, but polymorphic CGG repeat that is part of a CpG island in the 5'UTR of a novel gene named *FRA10AC1*. The number of CGG repeats varied in the population from 8 to 13. Expansions exceeding 200 repeat units were methylated in all *FRA10A* fragile site carriers tested. The *FRA10AC1* gene consists of 19 exons and is transcribed in the centromeric direction from the *FRA10A* repeat. The major transcript of ~1450 nt is ubiquitously expressed and codes for a highly conserved protein, FRA10AC1, of unknown function. Several splice variants leading to alternative 3' ends were identified (particularly in testis). These give rise to FRA10AC1 proteins with altered COOH-termini. Immunofluorescence analysis of full-length, recombinant EGFP-tagged FRA10AC1 protein showed that it was present exclusively in the nucleoplasm. We show that the expression of *FRA10A*, in parallel to the other cloned folate-sensitive fragile sites, is caused by an expansion and subsequent methylation of an unstable CGG

☆ The gene symbol *FRA10AC1* was approved by the HUGO Gene Nomenclature Committee for the gene name fragile site, folic acid type, rare, fra(10)(q23.3) candidate 1, alias *C10orf4*.

☆☆ Supplementary data for this article may be found on ScienceDirect.

* Corresponding author. Fax: +61-8-8161-7342.

E-mail address: jozef.gecz@adelaide.edu.au (J. Gécz).

¹ These authors contributed equally to this work.

² A list of members of the Consortium is given under Acknowledgments.

trinucleotide repeat. Taking advantage of three cSNPs within the *FRA10A* gene we demonstrate that one allele of the gene is not transcribed in a *FRA10A* carrier. Our data also suggest that in the heterozygous state *FRA10A* is likely a benign folate-sensitive fragile site. © 2004 Elsevier Inc. All rights reserved.

Keywords: Folate-sensitive fragile site; Triplet repeat expansion; Novel gene; Nuclear protein

Chromosomal fragile sites are loci that show gaps or breaks in metaphase spreads of cells grown under specific cell culture conditions. In general, common and rare fragile sites are distinguished. To date, more than 80 common fragile sites have been identified in the human genome [1]. Some of these, like the most frequently seen fragile site, *FRA3B* in 3p14.2, have been proposed to be regions of genomic instability implicated in, e.g., breakage events leading to recurrent chromosomal aberrations in cancer cells [for review see 2 and 3]. Although some of these common fragile sites were sequenced, no recognizable sequence elements have been identified that would be sufficient to confer fragility to these chromosomal sites.

According to the culture conditions under which fragile sites are induced, rare fragile sites are classified into folate-sensitive and distamycin A- or BrdU-requiring fragile sites. Twenty-two folate-sensitive fragile sites have been cytogenetically described, the best known being *FRAXA* in Xq27.3. Expression of *FRAXA* is associated with the fragile X syndrome (MIM309550), the most frequent familial form of mental retardation [4]. Similarly, expression of another rare gonosomal folate-sensitive fragile site, *FRAXE*, leads to mild to borderline mental retardation (MIM309548) [5]. To date, two autosomal folate-sensitive fragile sites, *FRA11B* and *FRA16A*, have been cloned [6,7], but only the former has been associ-

ated with a human disease, Jacobsen syndrome, which is caused by terminal deletion of part of the long arm of chromosome 11 [8].

The common molecular basis of the currently characterized rare fragile sites is the expansion of repetitive DNA elements. Expression of the distamycin A-inducible *FRA16B* site in 16q22 and of the BrdU-requiring *FRA10B* site in 10q25 is caused by expansion of 33- and 42-bp AT-rich minisatellite repeats, respectively [9,10]. In contrast, expression of all five folate-sensitive rare fragile sites cloned so far, i.e., *FRAXA*, *FRAXE*, and *FRAXF* on the X chromosome as well as *FRA16A* in 16p13.11 and *FRA11B* in 11q23.3, is uniformly based on expansion of CCG/CGG-trinucleotide repeats. The majority of characterized folate-sensitive fragile site repeats are located in the regulatory or 5' untranslated regions (UTRs) of transcribed genes, probably with the exception of *FRA16A* for which an associated gene has not yet been reported [7]. All of these repeats follow the paradigm of expansion and hypermethylation with subsequent gene silencing as a result [for review see 11]. The CGG-repeat expansion is the only known molecular mechanism underlying expression of folate-sensitive fragile sites. It is yet unknown why this kind of genomic instability leading to chromosomal fragility affects only a very limited subset of all CGG-repeats within the human genome.

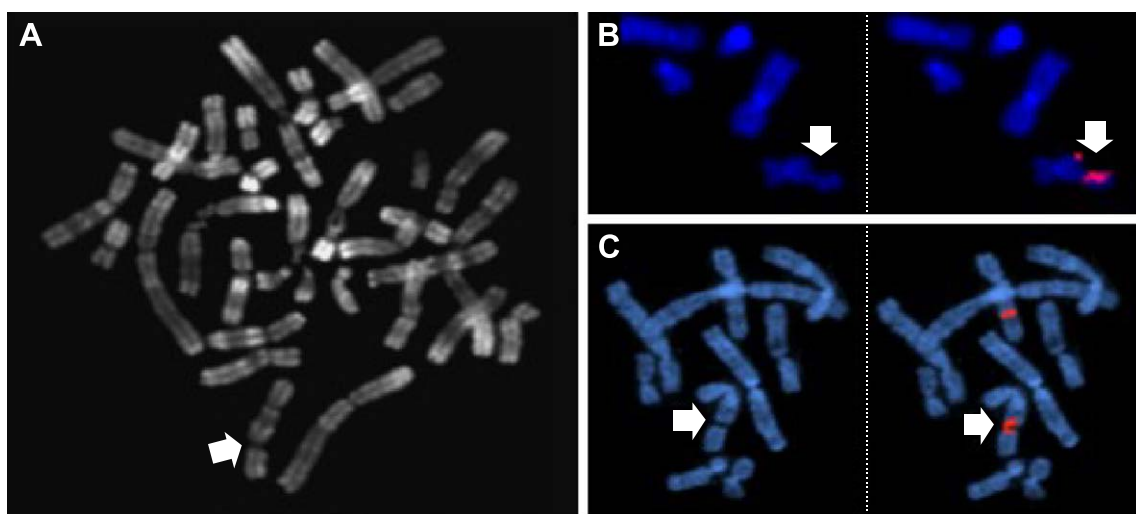


Fig. 1. Cytogenetic and FISH-mapping of the *FRA10A* site. (A) Fluorescence R-banded metaphase of case I expressing the *FRA10A* as a clear gap (arrow) in chromosomal band 10q23.3. (B) Partial metaphase of case I stained with DAPI (left) and hybridized with YAC 912C7 (right). The red signal from the YAC probe is split due to expression of the *FRA10A* site (arrow) and can be seen on both the centromeric and the telomeric fragment, indicating that YAC 912C7 spans the fragile site. (C) Partial metaphase of case III stained with DAPI (left) and hybridized with BAC 437J2 (right), which produces a signal proximal and distal to the gap, indicating that this clone is spanning the *FRA10A* site (arrow).

To learn more about the molecular basis of rare fragile site expression, we have characterized *FRA10A* in 10q23.3, the most prevalent among the rare folate-sensitive autosomal fragile sites in the human genome [12]. It has been estimated that about 1 in 500 individuals manifest *FRA10A*. Similar to the other autosomal folate-sensitive fragile sites, no distinct phenotype has been associated with expression of *FRA10A*. Nevertheless, some studies have proposed that fragile site expression in 10q23–q24 might be associated with mental retardation [12,13], tumorigenesis [14,15], or neurological disorders, including schizophrenia [16]. It is likely that these associations can be attributed to ascertainment bias. The 10q23–q24 chromosomal region has also been shown to contain tumor suppressor genes [17,18]; however their relation to the *FRA10A* is not clear.

In the present work we have refined the localization of the *FRA10A* fragile site and identified its molecular basis. We show that expression of *FRA10A* is caused by an

expansion of a polymorphic CGG repeat within the 5'UTR of a ubiquitously expressed gene, *FRA10AC1*, encoding a novel nuclear protein.

Results

Fluorescence in situ hybridization (FISH) mapping of the FRA10A site

To narrow down the locus containing *FRA10A*, a sequential FISH approach was applied to metaphase spreads of the index patient (case I). In this patient, cytogenetic expression of *FRA10A* resulted in chromosomal breakage on one homolog at 10q23.3 in two-thirds of cells (Fig. 1A). In a first step, double-color FISH was performed with YAC clones from a recently established contig of 10q23–q24 using YAC 790H10 as an internal control that hybridizes to the *MXI1* gene in 10q25 ([18,19] and R.S., unpublished

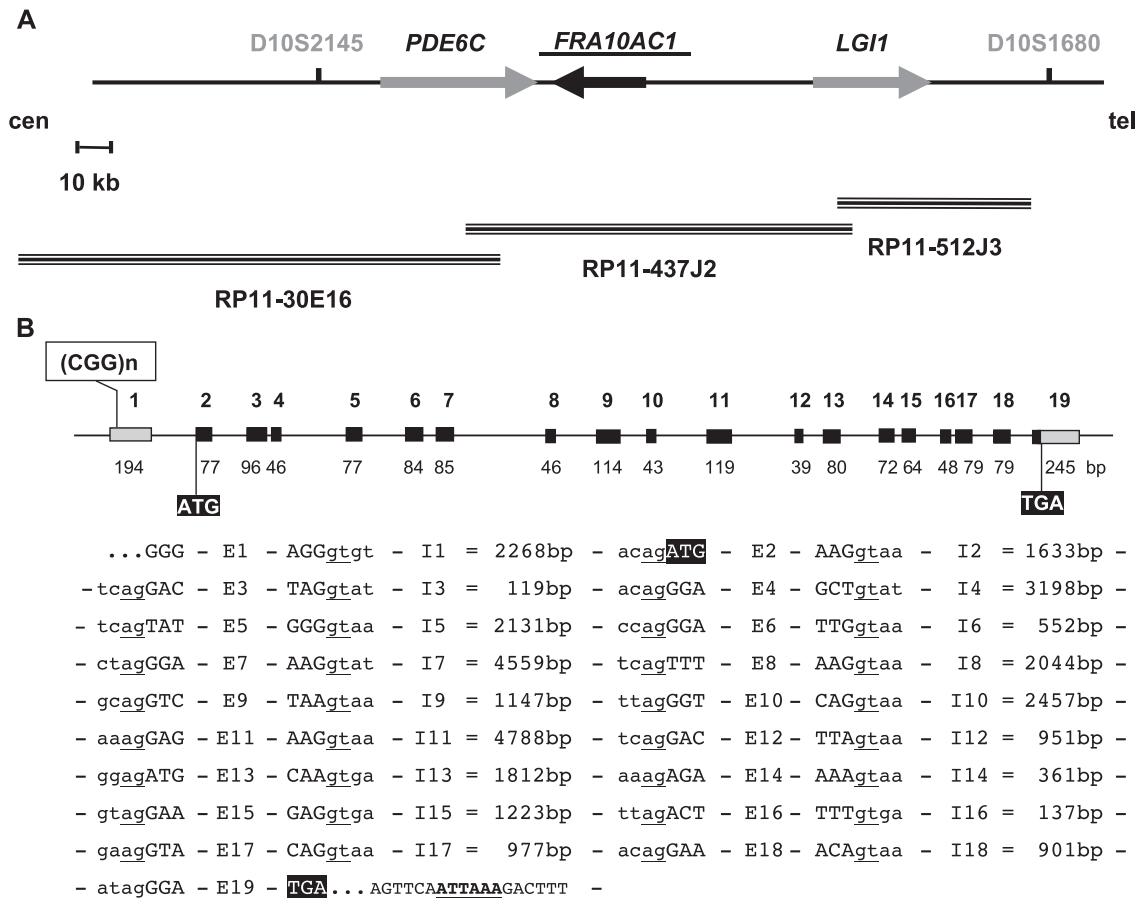


Fig. 2. Mapping of the *FRA10A* locus and genomic organization of the *FRA10AC1* gene. (A) Schematic representation of the genomic region containing the genes *FRA10AC1*, *PDE6C*, and *LGI1* as well as two microsatellite markers. Genes are indicated by large arrows. The 5' → 3' orientation of each gene is given by the orientation of the arrow. BAC clones are represented by black lines. (B) Exon–intron structure of *FRA10AC1*. In the upper part, the genomic structure of *FRA10AC1* is shown schematically (not drawn to scale). Exons are indicated by boxes. Coding regions of exons are shown in black, whereas noncoding regions are in gray. The number of each exon is depicted above and the size below the exon. Start as well as stop codons and the (CGG)_n repeat are indicated. In the lower part, partial sequences of the exon–intron boundaries and the intron sizes are given. Exon sequences are given in uppercase letters and intron sequences in lowercase. Exon is abbreviated by “E” and intron by “I”. The ATG and TGA codons are shown in black boxes. The conserved “ag” and “gt” bases at the intron boundaries are underlined. The polyadenylation signal ATAAA is in bold and underlined.

data). YAC clones 905F4 and 912C7 showed signals distal and proximal to the fragile site, indicating that these clones span the fragile site in the index case (Fig. 1B and Supplementary Table S1).

Based on the available physical map of 10q23–q24 that is partially covered by the inserts of the breakpoint-spanning YACs 905F4 and 912C7, contiguous BAC clones of the chromosome 10 sequence tile path [12] were selected to delineate the fragile site more precisely. FISH analysis showed that BAC clones RP11-341B18 and RP11-30E16 hybridized proximally, whereas RP11-562K6, -237M6, -76P2, -121A3, -162K11, and -397J18 were distal to the fragile site. Clones RP11-568I23, -437J2, -512J3, -306O10, and -367C3 spanned the fragile site in the index case (Supplementary Table S2).

To verify mapping of the *FRA10A* fragile site, metaphase spreads of three additional *FRA10A* carriers (cases II–IV) were analyzed by FISH (Supplementary Table S3). As in the index case, BAC 437J2 was shown to span the fragile site in all of them (Fig. 1C and Supplementary Table S3).

Sequence analysis of the *FRA10A* locus

FISH mapping refined the *FRA10A* locus to an ~50-kb region. Interestingly, BAC clone RP11-437J2 spans the entire *FRA10A* site (EMBL Accession No. AL157396). The total length of the RP11-437J2 clone insert is 162,840 bp. The true right end of this clone is at position 128,796 in the reported sequence (AL157396). The true left end of RP11-437J2 is at position 129,920 of AL356214 (RP11-30E16). Database searches identified the 3' end of the *PDE6C* gene encoding phosphodiesterase 6C as well as the first two 5' exons of the gene encoding the leucine-rich, glioma inactivated 1 protein, *LGII* (Fig. 2), within this *FRA10A* spanning BAC. A single (CGG)₈ repeat was identified in the very central part of the sequence (kb 54.809–54.832 of AL157396). Adjacent to it lies an incomplete and invariable CCA repeat. Both repeats are part of a CpG island of a novel gene, named *FRA10AC1* (fragile site, folic acid type, rare, fra(10) (q23.3) candidate 1), encoding a protein with unknown function (see below).

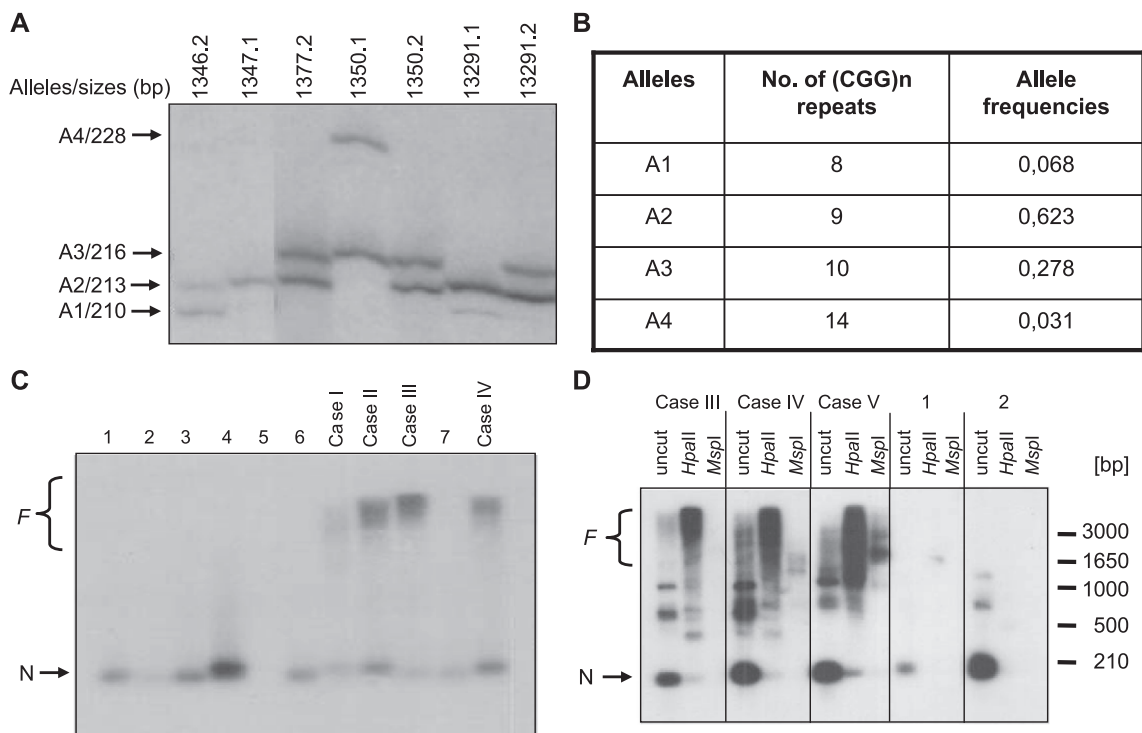


Fig. 3. CGG-repeat polymorphism in *FRA10A* carriers and controls. (A) Analysis of the CGG polymorphism in the 5'UTR of *FRA10AC1* in individuals of the CEPH families. Autoradiogram showing the four alleles (A1–A4) of the repeat detected by PCR. Seven representative genotypes are shown. (B) Allele distribution of the polymorphic *FRA10AC1* repeats in 81 unrelated individuals of the CEPH families. (C) Analysis of the (CGG)_n trinucleotide repeat in controls (lanes 1–7) and individuals expressing the *FRA10A* site (cases I, II, III, IV, and V). In addition to a normal allele (labeled with “N”) individuals who cytogenetically express the fragile site show PCR fragments of larger size (labeled with “F”), indicating repeat expansion. (D) Methylation analysis of the *FRA10A* CGG repeats. DNA from *FRA10A* heterozygous individuals (cases III, IV, and V) and controls (1 and 2) were digested with *HpaII* or *MspI* restriction enzyme, prior to PCR. The (CGG)_n repeat was then PCR-amplified from the restricted and unrestricted genomic DNA (uncut). Hybridization signals were obtained only for the expanded (CGG)_n repeats in individuals expressing *FRA10A* after digestion with *HpaII*, indicating that the region around the (CGG)_n repeat is hypermethylated. Approximate sizes of DNA marker bands are indicated on the right.

Analysis of the CGG repeat in controls and carriers of the *FRA10A* site

PCR-based amplification of the identified CGG repeat in the 5'UTR of *FRA10AC1* in various individuals revealed this repeat to be polymorphic. Testing of 81 unrelated individuals from 40 CEPH families identified four alleles (A1–A4) with repeat size $n = 8, 9, 10,$ and 14 . The most common allele, A2, contained nine CGG repeats (Figs. 3A and 3B). In contrast, individuals cytogenetically expressing the fragile site showed much larger PCR fragments (detected as a smear of higher molecular weight products), indicating expansion of the repeat (Fig. 3C). One normal CGG allele was detected in all *FRA10A* carriers, which is consistent with the heterozygous status of the repeat expansion. To test whether this expanded CGG repeat is subject to methylation (as is the case for the other cloned folate-sensitive fragile sites), we digested genomic DNA from *FRA10A* carriers and controls with a methylation-sensitive (*HpaII*) or -insensitive (*MspI*) restriction enzyme. Following digestion, PCR, Southern blot, and hybridization with a GCC/GGC probe, the expanded CGG repeats were detected only from the undigested or *HpaII*-digested genomic DNA and not from the *MspI*-digested DNA of *FRA10A* carriers (cases III, IV, and V in Fig. 3D). These data indicate that the expanded CGG repeat is hypermethylated. Control individuals showed the presence of normal allele only. In all experiments, the expanded alleles presented as a smear rather as a distinct

band, suggesting some difficulties for the *Taq* polymerase to amplify across the expanded CGG repeat. The size of the expanded alleles in *FRA10A* carriers exceeded that of the normal alleles by at least 600–800 bp (up to several kilobases). Thus, the number of CGG repeat units in *FRA10A* can be estimated to be at least 200.

Full-length cDNA sequence and genomic organization of the *FRA10AC1* gene

By combining in silico data with results obtained from RT-PCRs, we assembled the full-length *FRA10AC1* cDNA sequence of 1344 bp that contains an open reading frame encoding a protein of 315 amino acids (GenBank Accession No. AJ431721). The genomic organization of *FRA10AC1* was established by BLASTN searches of the complete cDNA sequence with the genomic sequence of clone RP11-437J2 (GenBank Accession No. AL157396). In total, this novel gene consists of 19 exons (Fig. 2B). Nonetheless, 4 exons (exons 13, 14, 15, and 16) are alternatively spliced in (see below), and exon 17 has been identified at least once, e.g., in the EST clone GKBBABE07 (GenBank Accession No. AV682434). *FRA10AC1* spans a region of 33 kb and is transcribed from telomere to centromere (Fig. 2A). All intron boundaries follow the AG/GT rule of the consensus splice site. The (CGG) n repeat is embedded in exon 1, within the putative 5'UTR of the *FRA10AC1* gene, and is part of the CpG island. The ATG codon at position 195 (in

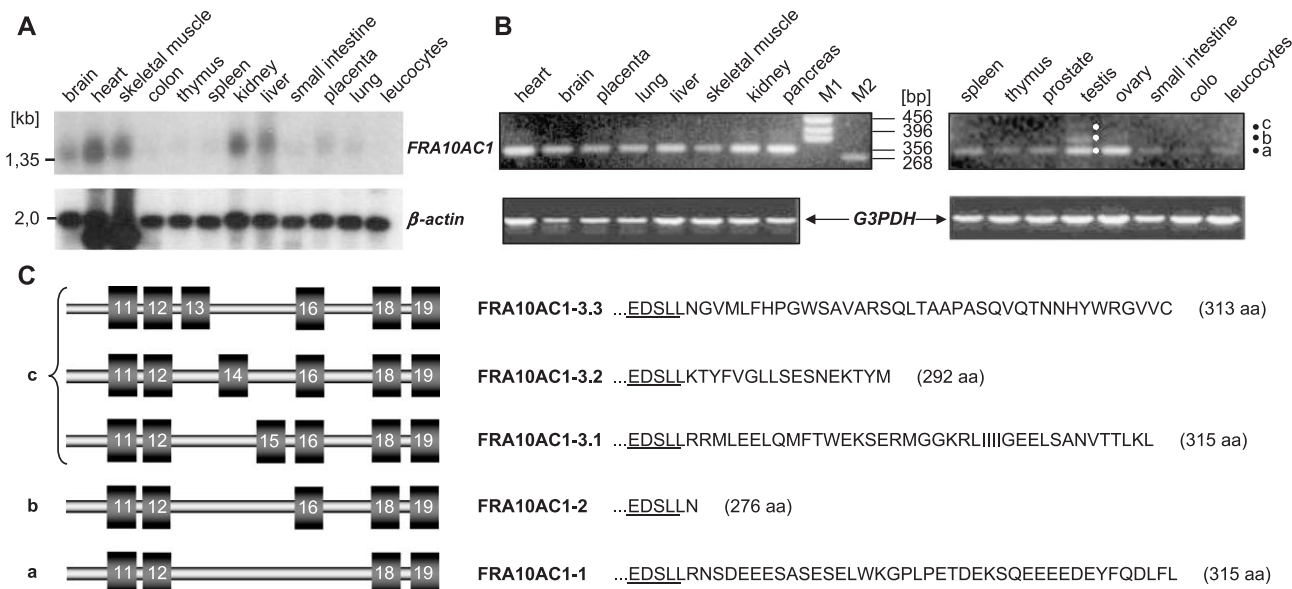


Fig. 4. Expression pattern of *FRA10AC1* in various human tissues and identification of alternative splicing variants. (A) Northern blot analysis of *FRA10AC1*. Hybridization of a *FRA10AC1* exon 19-containing probe to a human multiple-tissue Northern blot containing total RNA. A single transcript of approximately 1450 nt was detectable in numerous tissues (top). The 1.35-kb band of the marker is indicated. The same blot was reprobbed with a β -actin probe as control (bottom). (B) RT-PCRs with human MTC cDNA panels. *FRA10AC1*-specific PCR product spanning exons 11 to 19 was amplified from all tissues tested (top). White and black dots indicate the three different transcript variants “a”, “b”, and “c”. pBluescript/*HinfI* (M1) and λ EcoRV (M2) were used as DNA size markers; sizes of marker bands are indicated. Control RT-PCR amplification of a *G3PDH* product was done from the same cDNA samples (bottom). (C) Identification of alternative splice variants of the *FRA10AC1* gene. Schematic representation of the alternatively spliced *FRA10AC1* variants a, b, and c (left) and their derived amino acid sequences (right). Exons are indicated by dark gray boxes and are numbered. Introns are depicted as thick gray lines. Amino acids encoded by exon 12, common to all *FRA10AC1* variants, are underlined.

exon 2) was considered the most likely translation start site conforming to the Kozak rule. The translation termination codon, TGA, was identified in-frame with this start codon at position 1140 bp in the cDNA sequence (in exon 19). By using the ExPASy proteomics tools, the encoded protein was calculated to have a theoretical pI of 8.25 and a calculated molecular mass of 37.548 kDa. Its most basic region lies between amino acids 200 and 275 (theoretical $pI \sim 10.0$), whereas the carboxy-terminus (aa 280–315) represents the most acidic region ($pI \sim 3.70$). Various serine residues, particularly at the C-terminus (aa 250–313), are potential sites for phosphorylation [20]. By database searches using the Pfam and Smart software tools, we found no significant similarities of FRA10AC1 to any known protein domains or repeats. However, a bipartite nuclear localization signal (NLS; at aa positions 223–238 and 245–262), which is embedded within a larger lysine-rich region (aa 129–269), was identified.

Expression analysis of FRA10AC1

For Northern blot analysis, we used a FRA10AC1-specific probe containing the first 616 bp of the 5' coding region and a probe covering part of exon 19. We obtained identical hybridization patterns for both probes. A single transcript of approximately 1450 bp, which corresponds well to the estimated cDNA size, was detected in all adult tissues analyzed (Fig. 4A), indicating ubiquitous expression of FRA10AC1. Strong expression of FRA10AC1 was found in brain, heart, skeletal muscle, kidney, and liver, whereas low expression was found in colon, thymus, spleen, small intestine, placenta, and lung. In leukocytes, the transcript is barely detectable (Fig. 4A). To investigate further the expression profile of the FRA10AC1 gene, we employed RT-PCR on normalized first-strand cDNAs of a panel of human tissues using primers spanning exons 11 to 19. Amplification was done under nonsaturating conditions allowing semiquantitative estimation of the relative FRA10AC1 transcript levels. In general, ubiquitous expression of FRA10AC1 was confirmed (Fig. 4B).

In addition to the tissues tested by Northern blot, FRA10AC1-specific transcripts were detected in pancreas, testis, ovary, and, at lower levels, prostate and leukocytes (Fig. 4B). In the majority of the tissues examined, a fainter and larger band ("b" in Fig. 4B) was found in addition to the abundant FRA10AC1 transcript ("a" in Fig. 4B) indicating the presence of a splice variant. Moreover, in ovarian and particularly in testicular tissue, a third band with slower mobility (variant "c" in Fig. 4B) appeared. Sequence analysis showed that these transcript variants are due to an alternative splicing of exons 13–16 (Fig. 4C). In the abundant variant a, exon 12 is spliced directly to exon 18 (GenBank Accession No. AJ431721). The ubiquitously expressed larger transcript variant b contains the 48-bp exon 16 spliced between exons 12 and 18 (GenBank Accession No. AJ431725). The largest amplicon, c, turned out to be a

mixture of three different FRA10AC1 transcript variants, in which exon 13, 14, or 15 is individually spliced in, in addition to exon 16 (GenBank Accession Nos. AJ431722, AJ431723, and AJ431724). Comparison of the derived amino acid sequences of the five splicing variants showed that the encoded C-terminal regions are different (Fig. 4C).

cSNP/RT-PCR analysis of FRA10AC1 in a FRA10A carrier individual

To investigate that one allele of the FRA10AC1 gene is transcriptionally silenced by the methylation at the FRA10A fragile site, we performed cSNP/RT-PCR on cDNA from one FRA10A carrier (case I, Fig. 3). First, at least three cSNPs were detected within the FRA10AC1 open reading frame by analysis of available ESTs (Hs.303727): one at position 47 in exon 2 (c.47G → A, 17 ESTs with the G and 11 with the A allele), the second at position 90 in exon 3 (c.90T → C, 22 ESTs with T and 8 with C), and the third at position 233 in exon 5 (c.233G → C, 21 ESTs with G and 9 with C). Second, FRA10A carriers for whom we had cell lines available were tested for these three cSNPs and one was found informative with genotype g.47A–g.90T/C–g.233C/G (case I). When the exon 2/5 RT-PCR product (primers available upon request) of this FRA10A carrier was sequenced, only one allele was detected: c.47A–c.90C–c.233G (see Fig. 5). This result demonstrated that one FRA10AC1 allele (c.47A–c.90T–c.233C in this case) was not transcribed. That the silenced allele is the one with the FRA10A fragile site is highly likely but was not formally established (due to the lack of available family material).

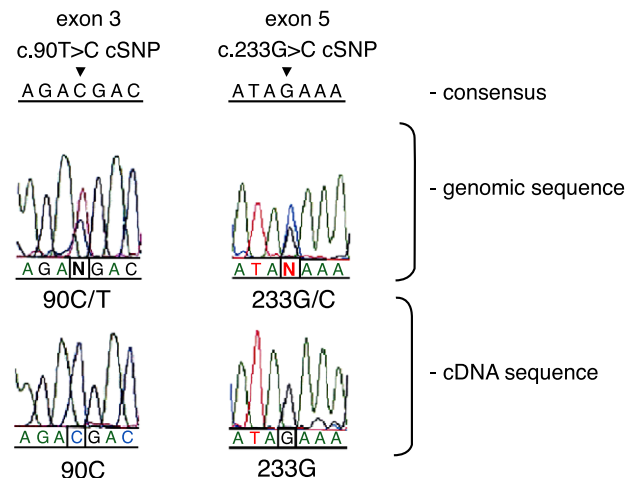


Fig. 5. cSNP/RT-PCR on FRA10A carrier individual. Partial genomic and cDNA sequence of exon 3 (c.90T → C) and exon 5 (c.233G → C) cSNPs of the FRA10AC1 gene of a FRA10A carrier individual (case I) is shown. While the genomic sequence shows heterozygosity for T/C and G/C alleles at both positions, the cDNA sequence clearly shows the presence of only one allele, c.90C–c.233G. This individual was not informative for the c.47G → A cSNP. The consensus FRA10AC1 sequence (GenBank Accession No. AJ431721) is also shown.

The *FRA10AC1* protein is located to the nucleus

To examine the subcellular localization of the *FRA10AC1* protein and, thereby, the functional potential of the predicted NLS, COS-7 cells were transiently transfected with an enhanced green fluorescent protein (EGFP)–*FRA10AC1* expression plasmid. The cloned *FRA10AC1* cDNA corresponds to the most abundant splice variant a (Fig. 4C). Upon transient transfection the green fluorescence was diffusely detected in the nucleus (Figs. 6A and B-a). However, in about 10% of the transfected cells a punctate pattern was found (Fig. 6B-d). Coimmunostaining of cells with lamin A/C antibody, which marks

the nuclear envelope, suggested that *FRA10AC1* is located exclusively in the nucleoplasm (Fig. 6B-f).

FRA10AC1 is a conserved protein

We used the human *FRA10AC1* cDNA sequence to search for ortholog ESTs in the databases. Several overlapping ESTs were found. Three of them (GenBank Accession Nos. BQ939866, AA840143, and BE946023) were selected for the assembly of the mouse *FRA10AC1* ortholog (size 1380 bp; GenBank Accession No. BN000292). *Fra10Ac1* maps to mouse chromosome 19C2–C3, between the genes *Pde6c* and *Lgi1*, in a

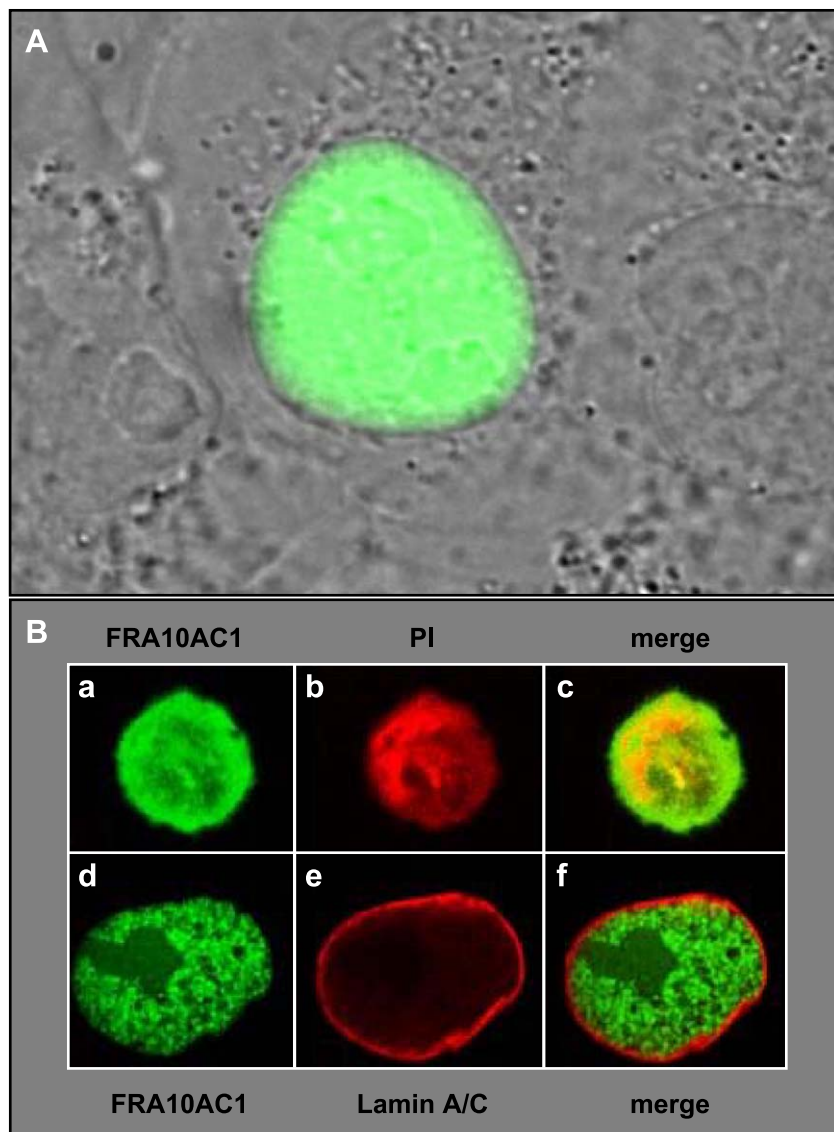


Fig. 6. Subcellular localization of *FRA10AC1*. (A) Subcellular localization of the EGFP-tagged *FRA10AC1* fusion protein (GATEWAY compatible) detected by fluorescence microscopy. Representative COS-7 cells from transient transfections were photographed. More than 95% of transfected cells showed identical patterns of subcellular localization. (B) The EGFP-tagged *FRA10AC1* protein is exclusively detected in the nucleoplasm of transfected COS-7 cells. (a) and (d) show fluorescence photomicrographs of nuclei with the green signal of *FRA10AC1*. (b) Propidium iodide staining of the same cell as shown in (a). (c) Merge of (a) and (b). (e) Labeling of the nuclear envelope of the same cell as shown in (d) with a lamin A/C antibody. (f) Merge of (d) and (e).

region syntenic to human chromosome 10q23–q24. *Fra10Ac1* encodes a protein of the same size (315 aa) compared to human and both proteins share a sequence identity of 86% and similarity of 92% (Fig. 7A). BLASTP analysis revealed that numerous FRA10AC1 orthologs exist, in both animal and plant species (Fig. 7B). The FRA10AC1 proteins identified in plants (*Arabidopsis thaliana* and *Pimpinella brachycarpa*) and *Caenorhabditis elegans* show about 35% identity to the human protein compared to 58% in insects (*Anopheles gambiae*) (Fig. 7B). These data suggest that a common ancestor gene had existed before plants and animals diverged. It is interesting to note that the central portion (aa 105–240 of human FRA10AC1; Fig. 7B) of the

protein is 45–67% identical among all orthologs analyzed, suggesting a highly conserved domain with an important function.

Discussion

In the present study we have elucidated the molecular basis of the most prevalent autosomal folate-sensitive fragile site, *FRA10A* [12], to be an unstable CGG repeat. The *FRA10A* repeat is similar to those identified at the other five folate-sensitive fragile sites, *FRAXA* [21,22], *FRAXE* [23], *FRAXF* [24,25], *FRA11B* [8], and *FRA16A* [7], that have been cloned so far. Overall, the *FRA10A*

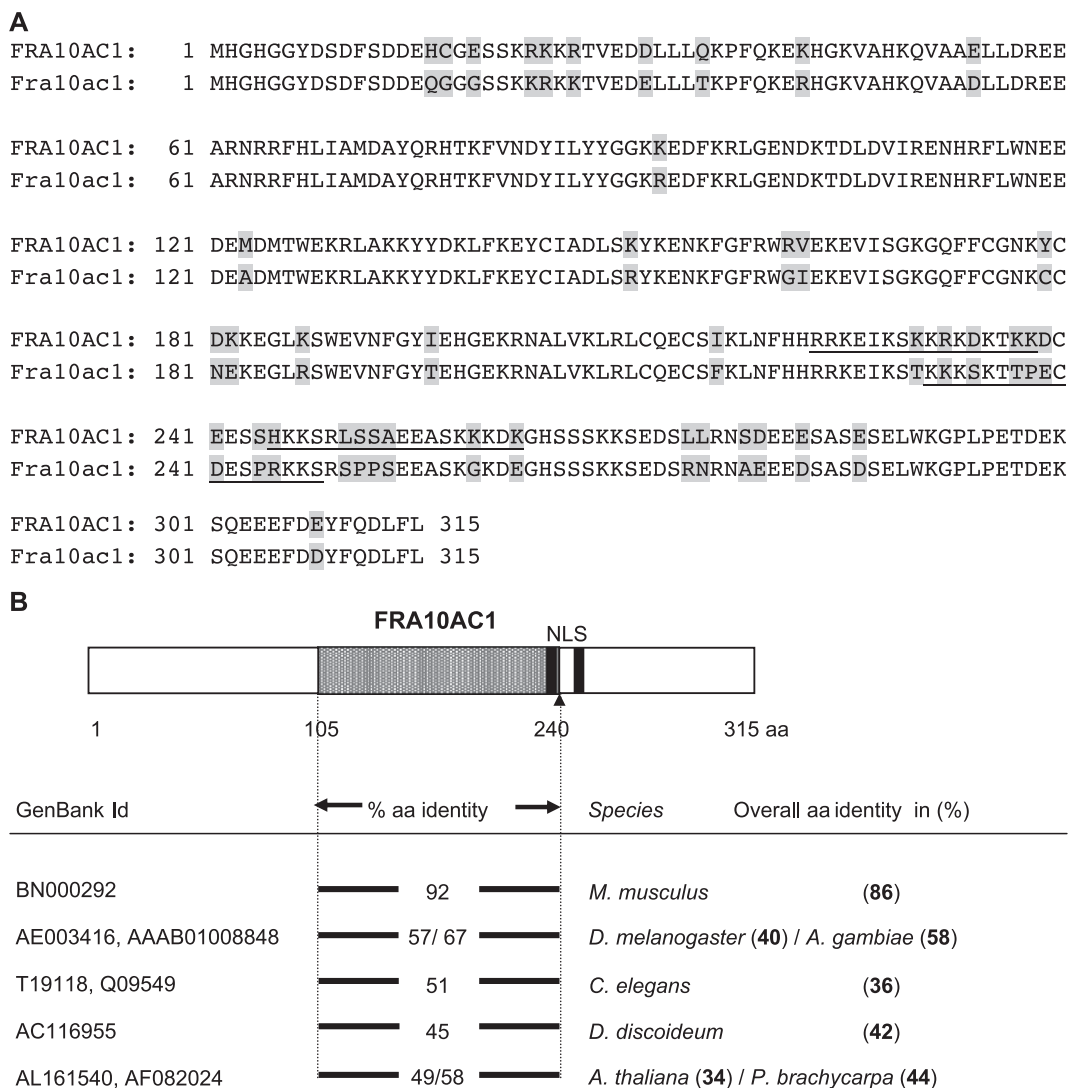


Fig. 7. Conservation of the FRA10AC1 protein. (A) Amino acid sequence alignment of the predicted human and mouse FRA10AC1 polypeptides. Amino acids that differ between human and mouse are highlighted in gray. The bipartite nuclear localization signals (NLS) are underlined. Numbers at the left and right indicate amino acid positions. (B) Schematic representation and conservation of the human FRA10AC1 protein. Top: The central conserved region (aa 105–240) is depicted as a gray box. The bipartite NLS (aa 223–238 and 245–262) is indicated by two black bars. Bottom: On the left, GenBank accession numbers of the orthologs are given. Amino acid identity within the conserved central part of FRA10AC1 is given in the middle, whereas the overall protein identity is shown on the right.

CGG repeat does not differ from the other known folate-sensitive fragile site-associated CCG/CGG repeats. The repeat is normally polymorphic with four alleles of 8, 9, 10, and 14 repeat units. In *FRA10A* carriers the repeat expands beyond 200 repeats (600–800 bp) and in that size it becomes also methylated. It appears to be somatically stable, although this was not rigorously tested. Among the 81 individuals (162 chromosomes) tested so far we have not seen any premutations.

Similar to the other folate-sensitive fragile sites, perhaps with the exception of *FRA16A* [7], *FRA10A* is associated with a transcription unit. To investigate that the expansion and methylation of the CGG repeat in the 5'UTR of the *FRA10AC1* gene results in the loss of its expression we tested RNA of one *FRA10A* carrier. For this purpose we took advantage of the presence of three cSNPs in the *FRA10AC1* mRNA. While genomic PCR showed heterozygosity for cSNPs at positions 90 (c.90T → C) and 233 (c.233G → C), only one *FRA10AC1* mRNA allele (c.90C and c.233G) has been detected by RT-PCR. This result clearly demonstrated that one allele of the *FRA10AC1* gene is not transcribed in this individual. It is highly likely that this *FRA10AC1* allele is transcriptionally silenced as a direct consequence of the presence of the *FRA10A* fragile site on this chromosome. To our knowledge this region of chromosome 10 is not imprinted. Whether the lack of one *FRA10AC1* mRNA allele expression is compensated for at the posttranscriptional level was not investigated.

While for the three X-linked folate-sensitive fragile sites the association with a particular phenotype is more (*FRAXA* and *FRAXE*) or less (*FRAXF*) resolved, the situation is not clear-cut for the two autosomal sites, *FRA11B* and *FRA16A*. Jones et al. [26] characterized an additional eight CCG repeats from the 40-Mb region of Jacobsen patient deletion breakpoints. These authors concluded that the nonrandom clustering of chromosome deletion breakpoints with CCG repeats suggests that they may play an important role in a common mechanism of chromosome breakage, perhaps including also the *FRA11B* CCG repeat. The great majority of Jacobsen syndrome patients, however, do not have their chromosome 11 deletion breakpoint at or near *FRA11B* [27]. As for the *FRA16A* site [28], there is no phenotype associated with this fragile site reported yet.

Carrier individuals of *FRA10A* analyzed as part of this study do have a clinical phenotype broadly defined as learning difficulties or developmental delay. However, this phenotype most likely cannot be attributed to the expression of the fragile site and likely haploinsufficiency of the *FRA10AC1* gene expression, but rather to an ascertainment bias (these individuals would have been primarily screened for *FRAXA*). Additional evidence supporting ascertainment bias as a likely cause comes (where available) from the evaluation of the parents of *FRA10A* carriers who did carry the fragile site, but did not display any clinical phenotype similar to that found in their

offspring. On the other hand it is intriguing that lack of one copy of the *FRA10AC1* gene does not give any obvious phenotype. The *FRA10AC1* gene is highly conserved and widely expressed and has no obvious paralogs, which could indicate some redundancy of its function. Recently, Shaw et al. [29] identified the *FAM11A* gene, whose expression was silenced as a consequence of the presence of the *FRAXF* fragile site. Only one individual with *FRAXF* was analyzed for *FAM11A* expression, which was extinguished. Although this individual did not present with any obvious *FRAXF*-related phenotype it is too early to speculate that the lack of *FAM11A* does not cause any discernible phenotype [29]. This also applies for *FRA10A* for which additional cases together with further functional studies of the *FRA10AC1* gene will probably help to resolve this issue. Similar early controversy about the existence of an associated phenotype, debated for the *FRAXE* site and mental retardation, has later been resolved by postulating a minor and/or partly redundant function for the *FMR2* gene [for review see 5]. For *FRA10AC1* it is not clear yet whether haploinsufficiency would have any pathologic consequences in tissues where the gene is expressed. It is interesting to observe that the highest *FRA10AC1* expression is seen in metabolically active tissues like heart, skeletal muscle, brain, kidney, and liver. Whether the single copy of the *FRA10AC1* gene in *FRA10A* carriers can be sufficiently up-regulated in times of high demand needs to be addressed.

Contrary to the fact that *FRA10A* is one of the most frequently seen folate-sensitive fragile sites so far, to our knowledge, there is no *FRA10A* homozygote reported. Such a case would be very desirable (if not in utero lethal) in addressing the function and associated phenotype (if any) of the *FRA10AC1* gene. However, this could be achieved by manipulating *FRA10AC1* expression in any of the model organisms with *FRA10AC1* orthologs, including fly, worm, or mouse.

The *FRA10A* fragile site lies about 50 kb from the 5' end of the *LGII* gene, which was recently implicated in autosomal-dominant partial epilepsy with auditory features [30,31]. Moreover, down-regulation of *LGII* has been shown during malignant progression of astrocytic gliomas due to a hitherto unknown mechanism [32]. Although the presence of some long-range regulatory elements of the *LGII* gene affected by the expression of the *FRA10A* fragile site can be speculated, such hypothesis has not been tested. Neither an increased incidence of tumors nor such type of epilepsy has been noticed in *FRA10A* carriers. There does not seem to be any other gene present in the 50-kb region containing *FRA10AC1* and *LGII* (Fig. 2A).

In conclusion, *FRA10A* represents another unstable CCG trinucleotide repeat-triggered folate-sensitive fragile site. Although no phenotype can be associated with the fragile site as yet, it is intriguing that the *FRA10A*-associated *FRA10AC1* gene is highly conserved and, thus,

should serve an important function. Identification and characterization of the *FRA10A* site and its associated gene will allow the elucidation of any possible genotype–phenotype correlation in the future.

Materials and methods

Chromosome analysis

All cytogenetic studies were performed after obtaining informed consent as part of a diagnostic workup. In the index patient (case I), chromosome analysis was initiated because of speech delay and autistic features and revealed expression of *FRA10A* in 66% of the cells. The expression of the fragile site in 10q23 resulted in chromosomal breakage in the majority of the patient's cells, leading to an acentric fragment of the distal part of 10q, which was gained or lost in part of the metaphases. Another patient (case II), a 1-week-old boy, who presented with mildly dysmorphic features, expressed *FRA10A* in 5 of 20 metaphases. Three other unrelated *FRA10A*-positive individuals (cases III, IV, and V) were detected when tested for the cause of their developmental delay (case V) or developmental delay in another family member (cases III and IV). In two instances the index case (case III and case IV) did not present with *FRA10A*, while one of the parents did. Follow-up studies on the parents and/or siblings of the remaining *FRA10A*-positive patients showed that the *FRA10A* site was inherited from one of the parents, in each case, who was not affected. Chromosome analyses from peripheral blood cells suitable for the demonstration of folate-sensitive fragile sites as well as harvesting, slide preparation, and G or R banding were performed according to standard methods.

YAC and BAC clones

YAC clones were derived from the CEPH-mega-YAC library. A detailed STS-based map of 10q23–q24 containing the YAC clones used in the present study has been published previously [19,33]. BAC clones were derived from the RPCI-11 library (<http://www.chori.org/bacpac/home.htm>). They were part of the BAC contig covering chromosome region 10q23–q24, which was assembled by means of STS-content mapping and fingerprinting (<http://www.sanger.ac.uk/cgi-bin/humace/fpcwebmap.cgi?mode=map&map=bac.10.1.html>) [34,35]. Shotgun sequencing and finishing of clones from this contig including the clone RP11-437J2 (GenBank Accession No. AL157396) was performed as described previously [36].

FISH

BAC and YAC DNA was prepared using the Perfect-Prep Plasmid Maxi Kit (Eppendorf) or the Nucleon MiY

kit (Amersham Pharmacia). Indirect or direct labeling of BAC DNA and *Alu*-PCR products from YACs was performed with the BioPrime DNA labeling system (Invitrogen) or by nick translation. Indirect and direct FISH to metaphase spreads was performed as described previously [37]. Chromosomes were counterstained with DAPI.

Northern blot and RT-PCR analyses

The expression pattern of *FRA10AC1* in human tissues was studied using multiple-tissue Northern blots (BD Biosciences Clontech). The probes used were a 129-bp PCR product representing bp 1124–1252 of exon 19 and a 617-bp PCR product flanking the ATG start codon. The exon 19 probe was radiolabeled directly by addition of [α - 32 P]dCTP (Amersham) into the PCR. Fifty nanograms of DNA from a human fetal liver cDNA library (Invitrogen) was amplified using primers F2 (5'-TCAG-GATTTGTTTCTATGAGAC-3') and R2 (5'-CCACACAAGCTGTAAGTTCAGCA-3'). The 5' coding region probe was amplified with primers N19F (5'-GATGCATGGTTCATGGAGGCTATG-3') and N19R (5'-CAAGTGCATTTCTTCTCACCATG-3') and radiolabeled by random priming with hexanucleotides (Invitrogen; Amersham Biosciences) and [α - 32 P]dCTP (Amersham Biosciences). Blots were hybridized according to the manufacturer's instructions, washed at high stringency, and exposed to X-ray film. To control for RNA integrity and loading amount, the blots were rehybridized with an α - 32 P-labeled β -actin probe as described in the protocol provided. For the analysis of the *FRA10AC1* expression profile, primers F3 (5'-GCAGAAGAGGCTCCAA-3') and R2 were employed in PCRs using cDNA samples (1 ng) from a multiple-tissue panel of normalized first-strand cDNAs (MTC, BD Biosciences Clontech). PCR products were separated by 1.5% agarose gel electrophoresis, visualized using ethidium bromide, and sequence analyzed either by direct sequencing or after subcloning into vector pGEM-T Easy (Promega).

cSNP/RT-PCR analysis of *FRA10AC1* in a *FRA10A* carrier individual

Putative cSNPs within the *FRA10AC1* open reading frame were identified by sequence comparison of available ESTs (Hs.303727). The presence of cSNPs in *FRA10A* carriers was tested by sequencing of their genomic DNA. The following primers were used: for c.47G \rightarrow A cSNP, E2Fi, 5'-CCTGGACAATTGGCAACATAG-3', and E2Ri, 5'-CATTCACTAGAGCAATTGTCTTC-3', 254-bp amplicon; for c.90T \rightarrow C cSNP, E3Fi, 5'-ATCTCATTAAAGATCAGT-GAGTC-3', and E3Ri, 5'-AATCCGAAGCAACT-TAAAC-CAAG-3', 283-bp amplicon; and for c.233G \rightarrow C cSNP, E5Fi, 5'-GACACCAAACAGCTCCTAAGAG-3', and E5Ri, 5'-CATAACCCACATTCCCTTCTATG-3', 170-bp ampli-

con. cDNA was prepared from EBV-transformed lymphoblastoid cell lines of *FRA10A* carriers using standard techniques and tested for the presence of the cSNPs by direct sequencing of an RT-PCR product encompassing exons 2–5. The following primers were used: E2F, 5'-ATGCATGGT-CATGGAGGCTATG-3', and E5R, 5'-CCCAAACGCTT-GAAGTCTTCT-3', 296-bp amplicon. Dye terminator (Big Dye Terminator) sequencing was performed according to the instructions of the supplier of the kit (Perkin–Elmer). Exon PCR products were purified using the UltraClean PCR Clean-up DNA purification kit (MoBio Laboratories, Inc.). All PCR products were sequenced in both forward and reverse directions.

Construction of green fluorescence protein-tagged *FRA10AC1* cDNA expression plasmids

A cDNA fragment corresponding to the 315-aa-long *FRA10AC1* protein was obtained by PCR using as template 1 ng of human testis first-strand cDNA (MTC, BD Biosciences Clontech) with primers F, 5'-CGGGATCCGAAAGGATG CATGGTCATG-3' (*Bam*HI site and ATG start codon underlined), and R4, 5'-CGGAATTCGAAACAAATCCTGAAAATACTC-3' (*Eco*RI site underlined). The 964-bp agarose gel-purified PCR product was cloned into pGEM-T Easy (Promega), sequence verified, excised as a *Bam*HI/*Eco*RI fragment, and ligated to pEGFP-N1 (BD Biosciences Clontech) digested with the same enzymes. The resulting pEGFP construct encoded a GFP-tagged *FRA10AC1* polypeptide. In an alternative approach, an amplicon of the *FRA10AC1* ORF compatible with the GATEWAY cloning system (Invitrogen) was generated from fetal brain cDNA with primers N19C3F (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTGCATGGT-CATGGAGGCTATGATTC-3') and N19C3R (5'-GGGGACCACTTTGTACAAGAAAGCTGGGTGT-TACTATAGAAACAAATCCTGAAAATACTCATC-3') using cDNA Advantage polymerase (BD Biosciences Clontech). PCR products were purified according to the instructions. Plasmid pEGFP-C3 (BD Biosciences Clontech) was digested with *Sma*I and purified, and 5' phosphates were removed by alkaline phosphatase. After a second purification step by phenol/chloroform extraction, the plasmid was ligated to the GATEWAY Reading Frame Cassette C (RfC.1) of the GATEWAY Vector Conversion System (Invitrogen) according to the manufacturer's instructions. The clones obtained represent GATEWAY-compatible "destination vectors" that were sequenced for integrity and correct orientation of the cassette with primers 5'attR1 (5'-CACATTATACGAGCCGGAAGCAT-3') and 3'attR2 (5'-CAGTGTGCCGGTCTCCGTTATCG-3'). These vectors were propagated in *Escherichia coli* host DB3.1 (Invitrogen) to compensate for expression of the *ccdB* gene present on the cassette. Construction of the EGFP-*FRA10AC1* plasmid with the GATEWAY system expressing EGFP fused to the N-terminus was carried out as previously described [38].

Cell culture, transfection, and immunocytochemistry

COS-7 cells were grown in DMEM containing 10% fetal calf serum and gentamycin. About 3×10^4 cells grown on glass coverslips were transfected with 0.5–1.0 μ g of GFP-tagged *FRA10AC1* plasmid DNA using the calcium phosphate coprecipitation method [39]. Thirty-six hours after transfection, cells were washed with $1 \times$ phosphate-buffered saline (PBS) and fixed for 10 min with 4% paraformaldehyde followed by three washes with PBS. After fixation, cells were treated for 5 min with 20 mM glycine followed by three washes with PBS. Cells were permeabilized with blocking buffer (0.2% Triton X, 2 mM $MgCl_2$, 1% fish gelatin in PBS) for 5 min and incubated with a lamin A/C mouse monoclonal antibody (Santa Cruz Biotechnology) for 1 h at room temperature, then washed three times in blocking buffer, and incubated for 1 h with fluorescence-conjugated secondary antibody, Alexa Fluor 568 goat anti-mouse IgG conjugate (Molecular Probes). Primary and secondary antibodies were used at dilutions of 1:500 and 1:1000, respectively. After incubation, cells were washed three times with blocking buffer and once with PBS. Cell staining with propidium iodide was accomplished after treatment of the cells with RNase A for 30 min at room temperature. Processed cells were visualized with a Leica SP confocal microscope.

For transient expression of the EGFP-*FRA10AC1* GATEWAY-compatible plasmid, COS-7 cells were transfected and analyzed as described in Martinez-Garay et al. [38].

PCR analysis of the *FRA10AC1* CGG repeat

Genotyping of the *FRA10AC1* CGG repeat in 81 unrelated individuals, parents, or grandparents of the 40 CEPH families was performed by amplification of 100 ng of genomic DNA using 1 U of *Taq* polymerase (Minotech) with primers F1 (5'-CCATTTAGATGTAACACGTG-3') and R1 (5'-GTCGTTTCCTTCTTTCCC-3'). Primer F1 was added to the reaction after 5' end-labeling with [α - 32 P]dATP. PCR products were separated on a denaturing acrylamide–urea DNA sequencing gel and visualized by X-ray autoradiography.

To amplify across the CGG repeat in *FRA10A* carriers, primers 5'-TCGGACCGGAAGTGGTGCCC-3' and 5'-TCCCTGTGTGCCAAGTTCGC-3' were applied in a 20- μ l reaction containing 100 ng of genomic DNA; $1 \times$ buffer (Expand Long Template Buffer 2; Roche); 200 μ M each dCTP, dTTP, and dATP; 500 μ M 7-deaza dGTP; 10% (v/v) DMSO; 20 pmol of each oligonucleotide; and 0.4 μ l enzyme mix (Expand Long Template; Roche). Ten microliters of each PCR was run on a 1.5% agarose gel with $1 \times$ TBE. The gel was then blotted on a nylon membrane (Amersham) and hybridized with a GCC/GGC oligonucleotide probe.

Labeling and hybridization using the GCC/GGC probe

This was done essentially as described for the poly(AC)_n/(GT)_n probe by Thompson et al. [40]. Briefly, reagents as for random priming were used; however, the random hexamers were omitted. Instead, two complementary synthetic CCG oligonucleotides (30-mers; 10 copies of the CGG/CCG repeat) were used. They prime off of each other to generate the probe. Fifty nanograms of each oligonucleotide was used in the labeling reaction followed by conventional hybridization at 65°C and washing at 42–65°C. After a final stringent wash (0.1× SSC at 65°C for 20 min), the X-ray films were developed after an overnight exposure at room temperature. For the PCR studies to assess the methylation status of the expanded *FRA10AC1* CGG repeat, 1 μg of genomic DNA was digested with *HpaII* or *MspI* restriction enzyme before PCR amplification. Undigested and digested DNA was PCR-amplified using *FRA10AC1* flanking oligonucleotides (see above), run on a 1.5% agarose gel, blotted, and hybridized using the GCC/GGC probe.

Sequence analysis

Analysis of DNA and protein sequences was performed using BLAST (<http://www.ncbi.nlm.nih.gov/blast/>), SMART (<http://smart.embl-heidelberg.de/>), <http://www.ebi.ac.uk/interpro/scan.html>, Pfam (<http://www.sanger.ac.uk/Software/Pfam/search.shtml>), Ensembl (<http://www.ensembl.org/>), GoldenPath (<http://genome.ucsc.edu/>), NetPhos (www.cbs.dtu.dk/services/NetPhos), and ExpASY (<http://us.expasy.org/tools/#proteome>) servers.

Acknowledgments

This study was supported by grants from the Deutsche Krebshilfe (10-1643-Si1), the Hensel-Stiftung Kiel, the National Health and Medical Research Council of Australia, the Deutsche Forschungsgemeinschaft (GRK336), and the Greek Ministry of Development–Secretariat for Research and Technology (No. 236/234/603 and 6th Session/37) to N.K.M. We thank A. Pasparki (IMBB-FORTH) and Dorit Schuster for excellent technical assistance. Other members of the European Collaborative Consortium for the Study of ADLTE are J. Galan, A. Gorostidi, L. Kluwe, A. López de Munain, J.F. Martí Massó, V.F. Mautner, R. Michelucci, J.M. Morante-Redolat, J. Perez-Tur, J.J. Poza, J.F. Prud'homme, A. Rosenthal, A. Sáenz, E. Staub, U. Stephani, and C.A. Tassinari.

References

- [1] G.R. Sutherland, F. Hecht, Fragile Sites on Human Chromosomes, Oxford Univ. Press, New York, 1985.
- [2] Y. Pekarsky, A. Palamarchuk, K. Huebner, C.M. Croce, *FHIT* as

- tumor suppressor: mechanisms and therapeutic opportunities, *Cancer Biol. Ther.* 1 (2002) 232–236.
- [3] M.F. Arlt, A.M. Casper, T.W. Glover, Common fragile sites, *Cytogenet. Genome Res.* 100 (2003) 92–100.
- [4] B. Bardoni, J.L. Mandel, G.S. Fisch, *FMR1* gene and fragile X syndrome, *Am. J. Med. Genet.* 97 (2000) 153–163.
- [5] J. Gécz, The *FMR2* gene, *FRAXE* and non-specific X-linked mental retardation: clinical and molecular aspects, *Ann. Hum. Genet.* 64 (2000) 95–106.
- [6] C. Jones, et al., Physical linkage of the fragile site *FRA11B* and a Jacobsen syndrome chromosome deletion breakpoint in 11q23.3, *Hum. Mol. Genet.* 3 (1994) 2123–2130.
- [7] J.K. Nancarrow, et al., Implications of *FRA16A* structure for the mechanism of chromosomal fragile site genesis, *Science* 264 (1994) 1938–1941.
- [8] C. Jones, et al., Association of a chromosome deletion syndrome with a fragile site within the proto-oncogene *CBL2*, *Nature* 376 (1995) 145–149.
- [9] S. Yu, et al., Human chromosomal fragile site *FRA16B* is an amplified AT-rich minisatellite repeat, *Cell* 88 (1997) 367–374.
- [10] D.R. Hewett, et al., *FRA10B* structure reveals common elements in repeat expansion and chromosomal fragile site genesis, *Mol. Cell* 1 (1998) 773–781.
- [11] G.R. Sutherland, Rare fragile sites, *Cytogenet. Genome Res.* 100 (2003) 77–84.
- [12] P. Petit, J.P. Fryns, H. van den Berghe, F. Hecht, Population cytogenetics of autosomal fragile sites, *Clin. Genet.* 29 (1986) 96–100.
- [13] A. Mavrou, M. Syrou, C. Tsenghi, C. Metaxotou, Autosomal folate sensitive fragile sites in normal and mentally retarded individuals in Greece, *Am. J. Med. Genet.* 38 (1991) 437–439.
- [14] L. Lacombe, et al., Microsatellite instability and deletion analysis of chromosome 10 in human prostate cancer, *Int. J. Cancer* 69 (1996) 110–113.
- [15] S.K. Kim, et al., Identification of two distinct tumor-suppressor loci on the long arm of chromosome 10 in small cell lung cancer, *Oncogene* 17 (1998) 1749–1753.
- [16] B.J. Mowry, et al., Second stage of a genome scan of schizophrenia: study of five positive regions in an expanded sample, *Am. J. Med. Genet.* 96 (2000) 864–869.
- [17] O.B. Chernova, R.P. Somerville, J.K. Cowell, A novel gene, *LGII*, from 10q24 is rearranged and downregulated in malignant brain tumors, *Oncogene* 17 (1998) 2873–2881.
- [18] R. Siebert, et al., Deletions in the long arm of chromosome 10 in lymphomas with t(14;18): a pathogenetic role of the tumor suppressor genes *PTEN/MMAC1* and *MXI1*? *Blood* 92 (1998) 4487–4489.
- [19] C. Nobile, et al., A refined physical and EST map spanning 7.4 Mb of 10q24, a region involved in neurological disorders, *Mamm. Genome* 9 (1998) 835–837.
- [20] N. Blom, S. Gammeltoft, S. Brunak, Sequence- and structure-based prediction of eukaryotic protein phosphorylation sites, *J. Mol. Biol.* 294 (1999) 1351–1362.
- [21] E.J. Kremer, et al., Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)_n, *Science* 252 (1991) 1711–1714.
- [22] A.J. Verkerk, et al., Identification of a gene (*FMR-1*) containing a CCG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome, *Cell* 65 (1991) 905–914.
- [23] S.J.L. Knight, et al., Trinucleotide repeat amplification and hypermethylation of a CpG island in *FRAXE* mental retardation, *Cell* 74 (1993) 127–134.
- [24] J.E. Parrish, et al., Isolation of a GCC repeat showing expansion in *FRAXF*, a fragile site distal to *FRAXA* and *FRAXE*, *Nat. Genet.* 8 (1994) 229–235.
- [25] R.J. Ritchie, et al., The cloning of *FRAXF*: trinucleotide repeat expansion and methylation at a third fragile site in distal Xqter, *Hum. Mol. Genet.* 3 (1994) 2115–2121.
- [26] C. Jones, et al., Co-localisation of CCG repeats and chromosome

- deletion breakpoints in Jacobsen syndrome: evidence for a common mechanism of chromosome breakage, *Hum. Mol. Genet.* 9 (2000) 1201–1208.
- [27] R.C. Michaelis, et al., Most Jacobsen syndrome deletion breakpoints occur distal to *FRA11B*, *Am. J. Med. Genet.* 76 (1998) 222–228.
- [28] J.K. Nancarrow, et al., Molecular basis of p(CCG)_n repeat instability at the *FRA16A* fragile site locus, *Hum. Mol. Genet.* 4 (1995) 367–372.
- [29] M.A. Shaw, P. Chiurazzi, D.R. Romain, G. Neri, J. Géczy, A novel gene, *FAM11A*, associated with the *FRAXF* CpG island is transcriptionally silent in *FRAXF* full mutation, *Eur. J. Hum. Genet.* 10 (2002) 767–772.
- [30] S. Kalachikov, et al., Mutations in *LGII* cause autosomal-dominant partial epilepsy with auditory features, *Nat. Genet.* 30 (2002) 335–341.
- [31] J.M. Morante-Redolat, et al., Mutations in the *LGII*/Epitempin gene on 10q24 cause autosomal dominant lateral temporal epilepsy, *Hum. Mol. Genet.* 11 (2002) 1119–1128.
- [32] R. Besleaga, M. Montesinos-Rongen, J. Perez-Tur, R. Siebert, M. Deckert, Expression of the *LGII* gene product in astrocytic gliomas: downregulation with malignant progression, *Virchows Arch.* 443 (2003) 561–564.
- [33] I.C. Gray, et al., An integrated physical and genetic map spanning chromosome band 10q24, *Genomics* 43 (1997) 85–88.
- [34] D.R. Bentley, et al., The physical maps for sequencing human chromosomes 1, 6, 9, 10, 13, 20 and X, *Nature* 409 (2001) 942–943.
- [35] J.D. McPherson, et al., A physical map of the human genome, *Nature* 409 (2001) 934–941.
- [36] E. Lander, et al., Initial sequencing and analysis of the human genome, *Nature* 409 (2001) 860–921.
- [37] B. Schlegelberger, S. Metzke, S. Harder, R. Zühlke-Jenisch, Y. Zhang, R. Siebert, Classical and molecular cytogenetics of tumor cells, in: R.D. Wegner (Ed.), *Diagnostic Cytogenetics*, Springer-Verlag, Berlin, 1999, pp. 151–185.
- [38] I. Martinez-Garay, et al., A new gene family (*FAM9*) of low-copy repeats in Xp22.3 expressed exclusively in testis: implications for recombinations in this region, *Genomics* 80 (2002) 259–267.
- [39] F.L. Graham, A.J. van der Eb, A new technique for the assay of infectivity of human adenovirus 5 DNA, *Virology* 52 (1973) 456–467.
- [40] A.D. Thompson, et al., Isolation and characterisation of (AC)_n microsatellite genetic markers from human chromosome 16, *Genomics* 13 (1992) 402–408.

Map integration at human chromosome 10: molecular and cytogenetic analysis of a chromosome-specific somatic cell hybrid panel and genomic clones, based on a well-supported genetic map

R. Marzella,^a M.A. Kokkinaki,^{b,c} M. Kapsetaki,^b A. Ricco,^a A. Argyrokastritis,^b S. Kamakari,^b T. Sarafidou,^{b,c} N. Archidiacono,^a A. Roussou,^c A. Paspaparakis,^b M. Rocchi,^a and N.K. Moschonas^{b,c}

^aIstituto di Genetica, Bari (Italy),

^bInstitute of Molecular Biology and Biotechnology, Foundation of Research and Technology, and

^cDepartment of Biology, University of Crete, Crete (Greece)

Abstract. Well-characterized, chromosome-specific somatic cell hybrid panels are powerful tools for the analysis of the human genome. We have characterized a panel of human × hamster somatic cell hybrids retaining fragments of human chromosome 10 by fluorescence in situ hybridization and associated them to genetic markers. Most of the hybrids were generated by the radiation-reduction method, starting from a chromosome 10-specific monochromosomal hybrid, whereas some were collected from hybrids retaining chromosome 10-specific fragments as a result of spontaneous in vitro rearrangements. PCR was used to score the retention of 57 microsatellite markers evenly distributed along a well-supported framework genet-

ic map containing 149 loci uniquely placed at 69 anchor points (odds exceeding 1,000:1), with an average spacing of 2.8 cM. As an additional resource for genomic studies involving human chromosome 10, we report the cytogenetic localization of a series of YAC and PAC clones recognized by at least one genetic marker. Somatic cell hybrids provide a powerful source of partial chromosome paints useful for detailed clinical cytogenetic and primate chromosome evolution investigations. Furthermore, correlation of the above physical, genetic, and cytogenetic data contribute to an emerging consensus map of human chromosome 10.

Human chromosome 10 is a medium-size, submetacentric chromosome corresponding to about 4.6% of the length of the autosomes. To date, more than 40 genetic diseases have been associated with chromosome 10 loci; about 165 genes and more than 880 partial and full-length cDNAs (<http://www.ncbi.nlm.gov/SCIENCE96/>)

have been mapped to this chromosome. Several disease-causing genes have been identified on chromosome 10 and are under intensive investigation, including the RET proto-oncogene responsible for multiple endocrine neoplasia and Hirschsprung disease (see, e.g., Eng et al., 1997) and fibroblast growth factor receptor 2, which has been associated with Apert, Jackson-Weiss, Pfeiffer, Crouzon, and Beare-Stevenson cutis gyrata syndromes (Jabs et al., 1994; Reardon et al., 1994; Rutland et al., 1995; Wilkie et al., 1995; Przylepa et al., 1996). However, a number of other disease-related loci on chromosome 10, including several involved in developmental abnormalities, cancer, and neurological diseases (e.g., split hand/foot malformation type 3 [Nunes et al., 1995], DiGeorge II and velocardiofacial syndromes [Daw et al., 1996], endometrial carcinoma [Nagase et al., 1997], and infantile onset of spi-

Supported in part by the European Gene Mapping Project (EC EUROGEN), a Telethon grant, and AIRC, and by the Hellenic Secretariat for Research and Technology. R.M. and M.A.K. have equally contributed to the project.

Received 21 July 1997; revision accepted 12 January 1998.

Request reprints from Dr. Nicholas K. Moschonas, Institute of Molecular Biology and Biotechnology, Foundation of Research and Technology, PO Box 1527, Heraklion 711 10, Crete (Greece); telephone: 30-81-394-357; fax: 30-81-394-408; e-mail: moschon@imbb.forth.gr.

nocerebellar atrophy [Nikali et al., 1997]), still remain unidentified.

Positional cloning of disease-related genes and advanced molecular and cytogenetic diagnostic tools depend upon detailed integrated chromosome maps and the availability of well-characterized probes with precisely defined genetic, cytogenetic, and physical locations. Such resources, recently used for the construction of high-density, whole genome genetic and physical maps, include collections of microsatellite markers (Dib et al., 1996), YAC clones (Chumakov et al., 1995), and radiation hybrids (e.g., Hudson et al., 1995; Gyapay et al., 1996). To date, available chromosome 10-specific resources include λ phage (Deaven et al., 1986) and cosmid flow-sorted libraries (Ma et al., 1996) and a large collection of cytogenetically mapped YAC and PAC clones (Moschonas et al., 1996; Meitinger et al., 1997).

The radiation hybrid technique (Goss and Harris, 1975) was designed initially as a tool for physical mapping by ordering genetic markers according to the frequency of radiation-induced breakages between them. The numerous highly informative genetic markers available today permit detailed characterization of radiation hybrids, leading to adequate integration of the genetic and physical maps of human chromosomes. Some laboratories have produced chromosome-specific radiation hybrids derived from a lethally irradiated human \times rodent monochromosomal hybrid and fused to a second rodent cell line (Francke et al., 1994). In the present study, we report the construction of a well-supported genetic linkage map of human chromosome 10 and the detailed cytogenetic and molecular characterization of a chromosome 10-specific panel of 21 somatic cell hybrids retaining specific fragments of chromosome 10 by fluorescence in situ hybridization (FISH) and by the use of a set of microsatellite markers evenly distributed along the genetic map. As an example of the applications that such a panel may have in cytogenetic studies, we report the characterization of a paracentric inversion that differentiates human chromosome 10 and its corresponding homolog in the orangutan.

Materials and methods

Characterization of novel microsatellites and YAC library screening

Two Charon 21A-phage libraries (LA10NS01 and LL10NS01), constructed from a chromosome 10-specific human \times hamster somatic cell hybrid (ATCC), were screened with $5'$ - 32 P-labeled (GATA) $_{10}$, (AAAT) $_{10}$, or (CA) $_{20}$ oligonucleotides. A monochromosomal somatic cell hybrid panel (Corriell Cell Repositories Mapping Panel #2) was used in Southern blots to confirm the chromosomal origin of the positive clones. DNA sequencing was performed with either universal primers or complementary primers corresponding to the repeat, using conventional methodologies. Primers flanking each repeat were designed and employed in PCRs using as templates 80 Centre d'Etude du Polymorphisme Humain (CEPH) parental DNA samples. Genotyping of the 40 CEPH families was performed either manually or by using a Pharmacia A.L.F. DNA sequencer. Genotypic data evaluation was performed using ALP version 1.00 software (A. Wright and A. Brown, personal communication). DNA sequencing and genotypic data for nine markers were submitted to EMBL/GenBank database and GDB, respectively. Primer pairs corresponding to each marker were used for PCR screening of the ICI YAC library (supplied by the United Kingdom Human Gene Mapping Project [UK HGMP] Resource Center) in three successive steps, according to the instructions of the Resource Center.

Genetic linkage analysis

Genotypic data generated locally for 36 markers (D10S249, D10S92, D10S33, D10S2314, D10S179, D10S189, D10S547, D10S191, D10S211, D10S89, D10S197, D10S193, GLUDP5, D10S141, ZNF22, GLUDP2, D10S220, D10S589, D10S210, D10S535, D10S109, D10S201, D10S608, D10S2311, D10S2316, D10S192, D10S597, D10S2330, D10S2312, D10S190, D10S186, D10S212, D10S2318, D10S2309, D10S2317, and D10S2310) by screening the 40 CEPH families (Kapssetaki et al., 1994; Cox et al., 1996; this study) were submitted electronically to the European Gene Mapping (EUROGEM) ftp server (Cox et al., 1996) and subsequently passed on to CEPH. For the construction of a genetic map of human chromosome 10, our genotypic data were combined with data (for 117 markers) retrieved from the CEPH version 6 and CHLC version 2 databases, as described by Cox et al. (1996). Linkage analysis was performed by the CRI-MAP version 2.4 program essentially as described by Spurr et al. (1994), using a criterion of at least 1,000:1 odds. Error checking was performed by the CROMPIC option of the program, and potential data errors for markers genotyped locally were checked by retyping or removed from the analysis according to certain guidelines (Cox et al., 1996).

Cell lines

The normal human \times hamster somatic cell hybrids retaining fragments of human chromosome 10 were first reported by Antonacci et al. (1995). Hybrid GM10926D, which retained this chromosome as the only human contribution, was obtained from the Coriell Institute. The orangutan (*Pongo pygmaeus pygmaeus*, PPY) lymphoblastoid cell line was kindly donated by Dr. R.E. Bontrop (The Netherlands).

Radiation hybrids

A 75-cm² flask containing a subconfluent culture of the GM10926D cell hybrid was gamma-irradiated at low dosage (5,000 rad) to obtain hybrids with medium-to-large fragments that would be particularly suited to cytogenetic analysis. After exposure, the cell line was fused to a thymidine kinase-deficient hamster cell line (B14-150, Flow Laboratories) and selected in HAT medium. The hybrid clones were transferred to 24-well plates and subjected to trypsinization at near confluence. An aliquot was seeded in an eight-well chamber slide (Nunc), and the remaining cells were kept in the well. The chamber slides were harvested and hybridized in situ with biotin-labeled total human DNA. The signal was detected using Cy3-conjugated avidin (Amersham). FISH analysis provided information on the presence of human material, on its approximate size, and, particularly, on the homogeneity of the clone. Clones with less than 50% positive nuclei were discarded. Each positive clone was then expanded in a 75-cm² flask for liquid nitrogen storage and in a 150-cm² flask for DNA extraction. Hybrid PCP3B10 was a gift from P. Goodfellow (Cambridge).

Alu-PCR probes

PCR amplification of cell hybrid DNA templates was performed using the following *Alu* primers: 5'-GGA TTA CAG GYR TGA GCC A-3' and 5'-RCC AYT GCA CTC CAG CCT G-3' (where Y = C/T and R = A/G) (Liu et al., 1993). From 50 to 100 ng of DNA was amplified in a volume of 50 μ l. Following an initial denaturation step of 5 min at 95°C, 30 cycles of amplification (1 min at 95°C, 1 min at 65°C, and 4 min at 72°C) were performed, with a final extension step of 10 min at 72°C. A series of conventional *Alu* primers was used for the preparation of YAC-specific probes (Nelson et al., 1989). A mixture of *Alu*-PCR products corresponding to ICI YAC-specific sequences were used as probes for filter hybridization screening of a gridded PAC library (Ioannou et al., 1994), supplied by the UK HGMP Resource Center. Interpretation of the hybridization results was performed according to the scheme provided by the Resource Center.

FISH analysis of the hybrids and the YAC clones

Human metaphase spreads were obtained from PHA-stimulated peripheral blood lymphocytes from a normal human donor, hybridized in situ with *Alu*-PCR products biotinylated by nick translation and detected by Cy3-conjugated avidin. Hybridization was performed essentially as described by Archidiacono et al. (1994). Briefly, 400 ng of labeled DNA from each hybrid were used for each slide; hybridization was performed at 37°C in 2 \times SSC, 50% (v/v) formamide, 10% (w/v) dextran sulfate, 5 μ g Cot-I DNA (Boehringer Mannheim), and 3 μ g sonicated salmon sperm DNA, in a volume of 10 μ l. Posthybridization washing was carried out at 42°C in 2 \times SSC, 50%

formamide (three washes), followed by three additional washes in $0.1 \times \text{SSC}$ at 60°C . In situ hybridization of YAC probes was performed similarly, using 300–500 ng of biotin-labeled yeast DNA. Alternatively, chromosome spreads banded with 4',6-diamidino-2-phenylindole (DAPI) were hybridized to biotin-labeled yeast DNA or *Alu*-PCR products of the clones, essentially as described by Knapp et al. (1995). The signal was detected with FITC-conjugated avidin. Chromosome 10 was identified by cytological banding using DAPI and by cohybridization to a control plasmid clone, D10Z1, which detected chromosome 10-specific repetitive centromeric sequences (ATCC).

Digital imaging

Digital images were obtained using a Leica DMRXA epifluorescence microscope equipped with a CCD camera (Princeton Instruments). The fluorescent Cy3, FITC, and DAPI signals, detected with specific filters, were recorded separately as grayscale images. Pseudocoloring and merging of images were performed using Adobe Photoshop software.

Results

Isolation, characterization, and FISH mapping of novel human chromosome 10 microsatellite markers

To isolate novel polymorphic microsatellites of human chromosome 10, we screened two chromosome 10-specific phage libraries. Screening of about 100,000 plaques, yielded 36 candidate marker loci of the relatively frequent $(\text{CA})_n$ and the rare $(\text{GATA})_n$ and $(\text{AAAT})_n$ repeat type (26, 7, and 3 clones, respectively). The human origin of the positives was tested by differential hybridization to human vs. hamster genomic DNA and their human chromosome 10 origin, by Southern analysis of a somatic cell hybrid panel. Nine microsatellites, viz., D10S2311, D10S2312, D10S2314, and D10S2330 of the $(\text{GATA})_n$ repeat type, D10S2309 and D10S2310 of the $(\text{AAAT})_n$ type, and D10S2316, D10S2317, and D10S2318 of the $(\text{GT})_n$ repeat type, were selected for further analysis. To evaluate their genetic informativeness, we genotyped the 80 CEPH parental DNA samples. Allele numbers ranged from 5 to 11. The observed heterozygosities varied from 0.19 to 0.81. Codominant segregation was observed in all CEPH families genotyped. To determine the cytogenetic localization of the markers, ICI YACs corresponding to each of the markers, were isolated by PCR screening (a total of 21 YACs), and the corresponding *Alu*-PCR products were used in FISH experiments (Table 1). One marker (D10S2314) was mapped to the subtelomeric region of 10p, another one (D10S2318) to 10p12, and the rest to 10q23→q25.

Construction of a genetic map of human chromosome 10

A chromosome 10 genetic linkage map (Fig. 1) was constructed by employing 153 polymorphic loci with an average heterozygosity of 0.7. Genotypic data retrieved from the CEPH and CHLC databases (for 32 and 85 markers, respectively), were combined with data (for 36 markers) generated locally. The map (EUROGEM map) has a total sex-average length of 194.7 cM (male length, 148.2 cM; female length, 243.3 cM) and consists of 69 framework markers (viz., 40 haplotype systems and 29 individual markers) ordered with a likelihood support of at least 1,000:1 (Fig. 1). The average marker spacing on the sex-average map is 2.8 cM. The map extends from D10S249 (pter) to D10S6 (qter). Increased recombination was observed on the distal long arm of chromosome 10 in females and on the

distal short arm in males (Fig. 1). This finding is not unusual for several human chromosomes (Spurr et al., 1994); however, the molecular and biological etiology of this phenomenon remains to be resolved. Four markers which could not be ordered uniquely with odds of greater than 1,000:1 (D10S2318, D10S2309, D10S2317, and D10S2310) were positioned approximately.

FISH and molecular characterization of the somatic cell hybrid panel

Figure 2 summarizes the results of the cytogenetic and molecular characterization of a set of somatic cell hybrids retaining fragments of human chromosome 10. Since the distribution of *Alu* repeats along the human chromosomes correlates with G-negative bands (Korenberg and Rykowski, 1988), inter-*Alu* PCR products do not produce a homogeneous staining but, rather, an R-banding pattern (Baldini and Ward, 1991). Consequently, FISH characterization of fragments with breakpoints occurring in very pale regions could be inaccurate. In addition, small chromosomal deletions are not detectable by FISH because of the chromatin condensation in metaphase chromosomes. Thus, in order to characterize the hybrids at the molecular level and to overcome the limits of the FISH method, we employed 57 microsatellite markers chosen from those used for the construction of the genetic map. Selection of the markers was based on their informativeness and their nearly even distribution on the genetic map. A relatively larger number of markers mapped to the disease-gene-rich region 10q23→q25 was employed. PCR screening of 21 hybrid DNA samples was performed at least twice. In parallel, a positive (human DNA) and a negative (hamster DNA) control was always used. Some regions detected by the markers were not detected by FISH and vice versa, as indicated in Fig. 2. To enrich the physical and cytogenetic mapping resources available for human chromosome 10, we also characterized by FISH a set of CEPH YACs, in addition to those from the ICI library (Table 1). Table 1 also includes markers found by PCR to recognize these YACs, so that linkage and physical mapping data can be correlated. All but one YAC (886-F-3) recognized at least one human chromosome 10 marker. In addition, for YACs mapped to the disease-gene-rich 10q23→q25 region, we isolated several PAC clones recognized by at least one marker participating in our genetic map.

A paradigm of partial chromosome paints use in primate cytogenetic evolution

Partial chromosome paints (PCPs) generated from these hybrids are a powerful tool for cytogenetic investigations, particular for investigating the cytogenetic evolution of primates (Muller et al., 1996). Figure 3 presents an example of the combined use of PCPs and YACs in the fine characterization of a paracentric inversion present in orangutan species. It has been shown that the long arm of orangutan chromosome 7, which is syntenic to human chromosome 10, exhibits a banding pattern different from that of human 10q, suggesting the existence of a paracentric inversion (Yunis and Prakash, 1982). For a detailed characterization of this inversion, we performed cohybridization experiments, using appropriate human chromo-

Table 1. Regional assignment of human chromosome 10-specific YAC and PAC clones and the corresponding markers/STSs

YACs	Markers	Chromosome location	YACs	Markers	Chromosome location ^a	PACs
912-B-3	D10S249	10pter	821-D-2	D10S541	10q23	
946-G-12	D10S533	10p15	12HA1	D10S2309	10q23	
	D10S559			D10S1753		181-E18, 109-L90, 31-J24, 48-P11
	D10S602		24DE9	D10S2309	10q23	
	D10S1745			D10S1753		
	D10S1716		647-G-1	D10S564	10q23	
	D10S1706		8EH2	D10S2311	10q23.3→q24	
10GA2	D10S2314	10pter		D10S564		
10GA7	D10S2314	10pter	10BA1	D10S2311	10q23.3→q24	172-M13, 63-O24, 27-H17, 241-O11
81E1	D10S2314	10pter		D10S1753		23-G1, 135-E5
809-F-9	D10S1729	10p15		D10S564		27-H17, 241-O11, 82-C17, 105-N10, 216-P14, 74-P12
	D10S591		32GE7	D10S2311	10q23.3→q24	
	D10S552			D10S564		
851-H-9	D10S547	10p14	37IF9	D10S2311	10q23.3→q24	
747-H-8	D10S547	10p14		D10S564		
	D10S585			D10S1753		
808-A-2	D10S223	10p14	11DG4	D10S2316	10q23	
	D10S1725		22FG5	D10S2316	10q23	
886-F-3		10p13	3HH7	D10S2317	10q23	
959-G-10	D10S191	10p12	878-F-11	D10S566	10q24	
958-C-9	D10S191	10p12		D10S530		
33GB2	D10S2318	10p12	27BC9	D10S597		30-C6
36HC3	D10S2318	10p12		D10S2330	10q24	30-C6, 171-A21, 171-F5, 75-B20, 119-I13, 317-G21
940-F-9	D10S548	10p12	35GE1	D10S597		30-C6
934-E-11	D10S1660	10p12		D10S2330	10q24	30-C6, 171-A21, 171-F5, 75-B20, 119-I13, 317-G21
	D10S1747		27DD2	D10S2312	10q24→q25.1	32-B12, 320-F13, 59-L4, 241-L17, 84-E23
	D10S1749		14CB8	D10S2312	10q24→q25.1	
875-B-4	D10S197	10p12	19FC12	D10S2312	10q24→q25.1	
	D10S1775		24AC1	D10S2312	10q24→q25.1	
751-D-5	D10S1771	10p11.23→p12.1	22DC2	D10S2310	10q25.1	
	D10S588		966-H-9	D10S543	10q25	
953-E-3	D10S1732	10p11	935-F-11	D10S1748	10q25	
	D10S1684			D10S1657		
	D10S204			D10S544		
	D10S213			D10S545		
	D10S224		937-A-6	D10S209	10q25	
686-D-11	D10S224	10p11	934-C-4	D10S575	10q26	
945-A-12	D10S224	10p11		D10S214		
	D10S193		932-F-11	D10S1727	10q26	
	D10S1674			D10S217		
754-D-10	D10S1781	10p11	<u>799-E-5</u>		10q11.2/ 10q22→q23	
798-C-7	D10S208				10q23/10q11.2	
837-B-5	D10S578	10p11	<u>738-B-12</u>	D10S1687	10q23/10q11.2	
895-D-3	D10S1783	10q11	<u>889-C-10</u>	D10S558	10p15/5p15	
	D10S604			D10S594		
933-A-3	D10S196	10q11	<u>764-E-4</u>	D10S203	10p12/8p23	
	D10S225		<u>760-C-2</u>	D10S593	10p11.2/1q43/ 19p13.1	
	D10S1724				10q21/12q21?	
885-D-10	D10S568	10q11	<u>928-B-3</u>	D10S1767	10q21/12q21?	
905-G-8	D10S599	10q21	<u>665-D-2</u>	D10S556	10q22/7q34	
781-F-5	D10S210	10q21	<u>955-A-4</u>	D10S607	10q23/Yq11.22	
779-E-7	D10S1688	10q21	<u>968-E-6</u>	D10S541	10q23/6q21	
876-B-5	D10S218	10q22	<u>24DE4</u>	D10S2309	10q23/Chr 6	
	D10S188		<u>9CH9</u>	D10S2316	10q23/Chr 4	
804-H-2	D10S202	10q22	<u>649-H-1</u>		10q24/ 2q35→q36.3	
807-D-4	D10S551	10q23			10q24/ centromeres	
	D10S1686		<u>10AE2</u>	D10S2330	10q24→q25.1/ Chr 2	
	D10S1689		<u>38HG5</u>	D10S2312	10q26.1/ multiple sites	
			<u>771-A-11</u>			

^a Fifteen YACs (shown underscored in the lower right half of the table) mapped to more than one chromosome and possibly represent chimeric clones. FISH images of most of the YACs can be viewed at <http://bioserver.uniba.it/fish/rocchi/welcome.html>.

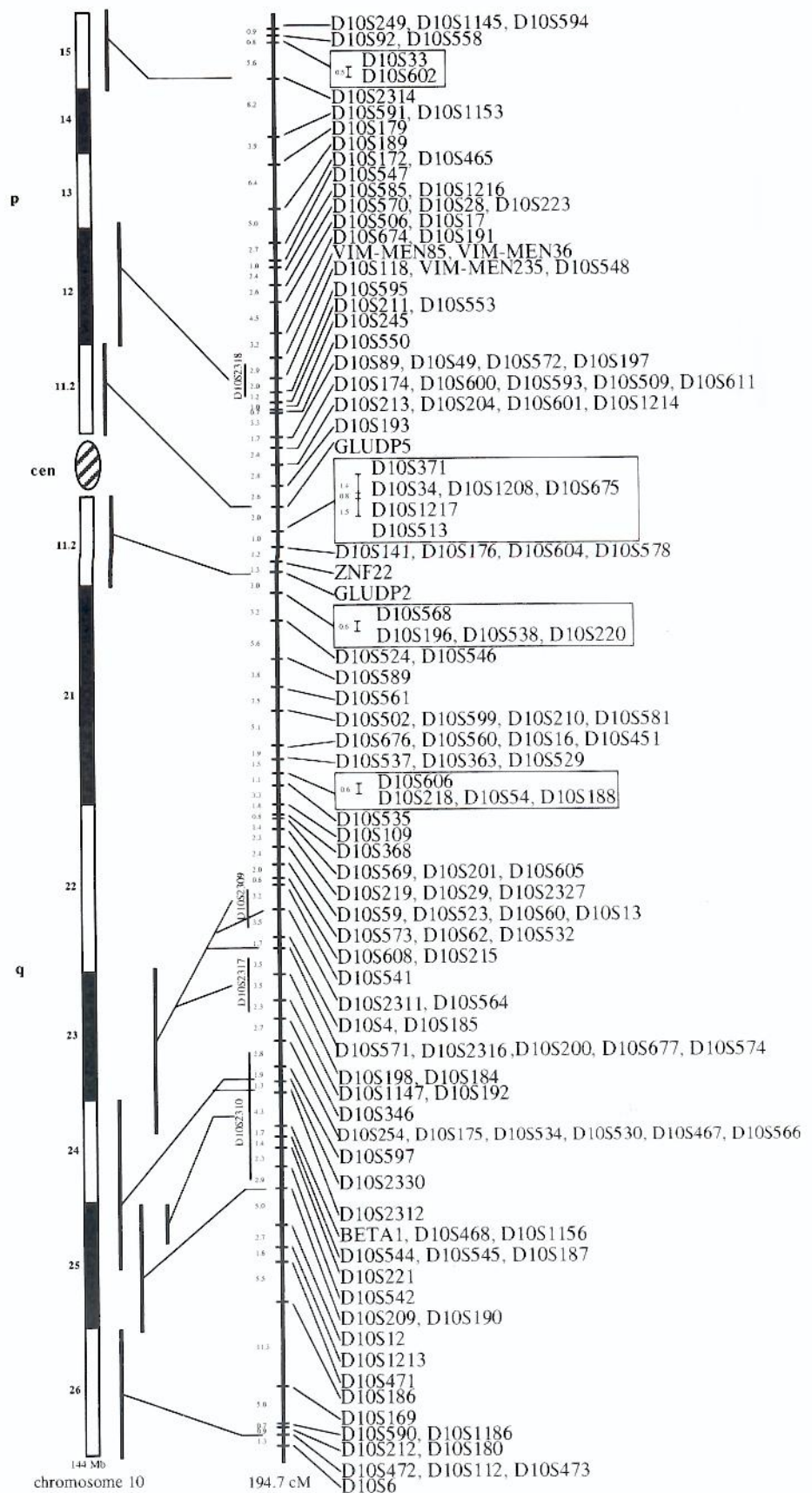


Fig. 1. A framework genetic linkage map of human chromosome 10. The sex-average map and the telomeric regions of the female and male maps, where significantly skewed recombination distances were determined, are presented. The distances between markers are shown to scale. The support intervals of four markers that were placed approximately are indicated by vertical lines. Markers haplotyped together are shown on the same line or as framed groups. The thin vertical lines to the left of framed markers indicate groups of loci for which the order could not be resolved with odds exceeding 1,000:1. The diagram at the left is an ideogram of chromosome 10 showing the cytogenetic localization of 13 of the markers used (Deloukas et al., 1993; Tzimogiorgis et al., 1993; this study). An earlier version of this map is presented in GDB (GDB No. 4584969). The male and female maps are accessible at: <http://www.imbb.forth.gr/papers/malemap.jpg> and <http://www.imbb.forth.gr/papers/femalemap.jpg>, respectively.

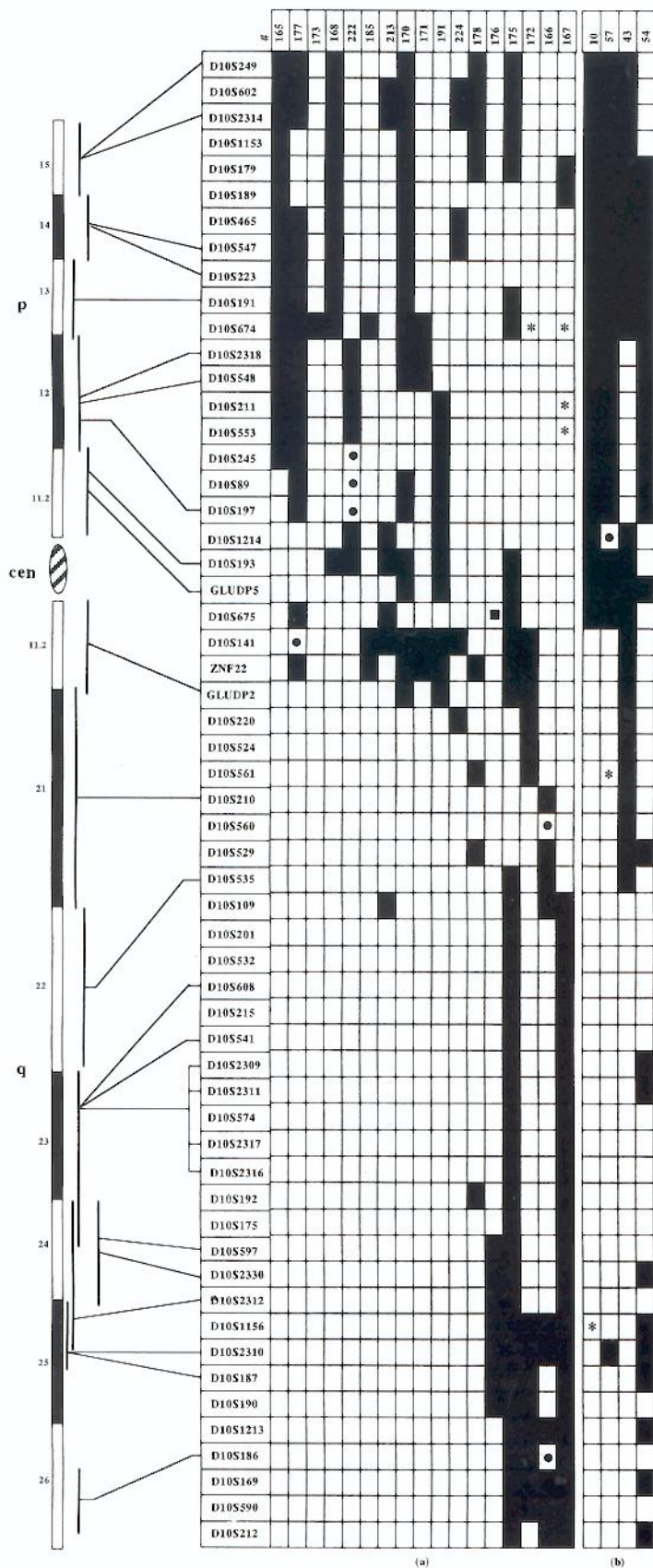


Fig. 2. FISH analysis and molecular characterization of 21 somatic cell hybrids retaining fragments of human chromosome 10, based on the use of 57 microsatellite markers. Shaded boxes indicate retention of the respective marker and detection of chromosomal segments by FISH. Information on the discrepancies between molecular characterization and FISH analysis is also included. Chromosome 10 segments recognized by the markers only are indicated by an asterisk (*), a segment detected only by FISH is marked by a small filled square, and molecularly characterized microdeletions which do not produce a FISH signal are marked by a dot. (a) Radiation hybrids. (b) Somatic cell hybrids with spontaneous chromosomal rearrangements containing additional human cytogenetic material. The diagram at the left is an ideogram of chromosome 10 showing the cytogenetic localization of 25 of the markers.

some 10 PCPs and YACs. Shown in Fig. 3a is the result of cohybridization of PCP 167, specific for 10q22→qter, with PCP 3B10, specific for 10q11.2 and 10q21.3. The pattern observed in the orangutan clearly indicates the presence of a paracentric inversion with breakpoints at 10q22.1 and 10q11.23. The result of cohybridization of PCP 3B10 and YAC 928B3, which maps to 10q21.1, confirmed the previous observation (Fig. 3b). Several YACs were used to refine the breakpoints. The most informative result was obtained when YAC 933A3 was applied (Fig. 3c). This YAC was clearly disrupted by the inversion, thereby allowing a very precise identification of the rearrangement.

Discussion

In this work, we characterized a somatic cell hybrid panel retaining fragments of human chromosome 10. Hybrids were mapped by FISH and subsequently screened with a group of microsatellite markers distributed along a well-supported genetic map of this chromosome. The map consists of a mixture of novel as well as previously employed microsatellite and fragment length polymorphic markers, thus providing a useful resource for comparisons of genetic distances and marker order. In regard to the Généthon chromosome 10 map (Dib et al., 1996), our genetic map contains 66 markers in common (43% of the markers we employed). With a few minor exceptions, the order of the markers is the same in both maps. This observation confirms the location of the markers along chromosome 10 and supports the credibility of our map. Only a small number of minor order conflicts, within short genetic distances, were observed. For example, within a comparable 3-cM interval, we have placed D10S190 and D10S209 distal to D10S542 (Fig. 1), whereas in the Généthon map, D10S209 is placed distal to D10S542 and D10S190. Similarly, D10S578 and D10S604 are separated by 2 cM in the Généthon map, whereas in our map they co-localize with D10S141 and D10S176. These conflicts may be due to (1) differences in the approach to map construction, (2) the relative amount of genotypic data used, (3) the overall marker content of the two maps, or (4) unchecked genotyping errors. Three reference markers (D10S33, D10S1214, and D10S12) mapped to more than one interval on the Généthon map have been placed at unique positions on our map. Within a map segment of comparable genetic distance (~33 cM) in both maps framed by markers D10S541 and D10S187 at 10q23.3→q25, we have placed seven novel markers, viz., D10S2311, D10S2316, D10S2330, D10S2312, D10S2309, D10S2317, and D10S2310. Interestingly, five of these markers are of the rare tetranucleotide repeat type. Although the sample is small, this could indicate a nonrandom distribution of such markers along the chromosomes. Alternatively, it could indicate a biased representation of the libraries employed for their isolation. Within the disease-gene-rich 10q23→q25 region, still unidentified disease-related genes exist (Meitinger et al. 1997). The high density of markers in the region may be useful for the positional cloning of, as yet, unknown disease-causing genes. The largest gap (11.3 cM) on the genetic map is between D10S186 and D10S169 on the dis-

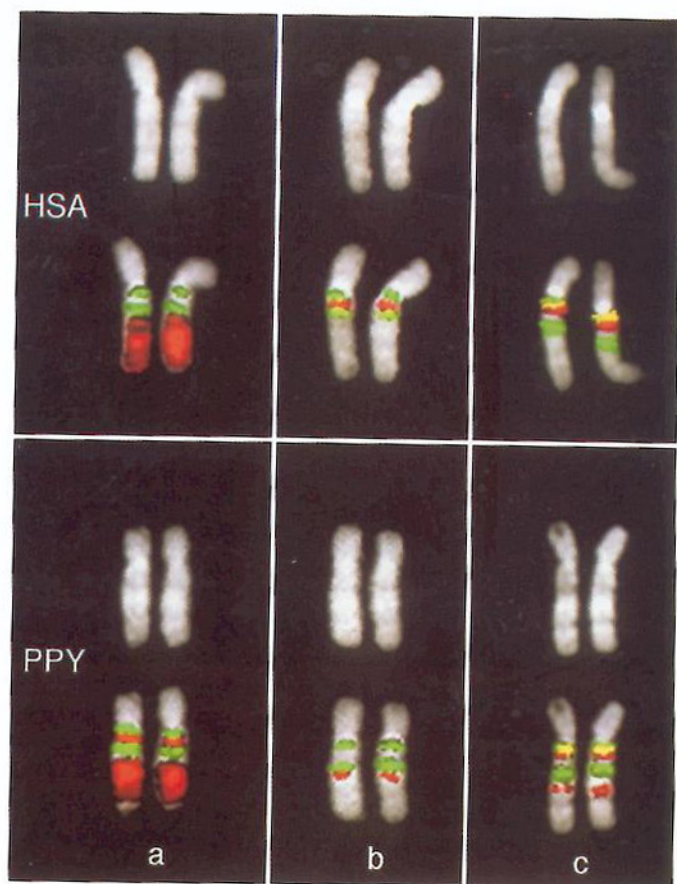


Fig. 3. Conclusive co-hybridization experiments to investigate the rearrangement differentiating the long arm of human chromosome 10 (HSA 10) and its orangutan counterpart (PPY 7). (a) Cohybridization of partial chromosome paint (PCP) 167 (red signal) with PCP 3B10 (green signal). (b, c) Co-hybridization of PCP 3B10 with YAC 928B3 (b) and YAC 933A3 (c). The DAPI-banded chromosomes are shown without hybridization signals (upper row) to display their morphology better. Additional information on the cytogenetic applications of PCPs may be found at <http://bioserver.uni-ba.it/fish/rocchi/welcome.html>.

tal portion of 10q. D10S186 maps to 10q26 (MIT database, Fig. 2). Since D10S169 is distal to D10S186, we conclude that it also maps to 10q26. The observed distance gap could be due to the fact that 10q26 is a subtelomeric region exhibiting a high rate of recombination. However, it should be mentioned that these markers were found to be physically separated in two hybrids (54 and 166), as suggested by both FISH and molecular analysis (Fig. 2).

Highly informative microsatellite markers are the most useful STSs involved in linkage analysis and in somatic cell hybrid characterization. Linkage analysis and molecular screening of the hybrid panel, permitted the dissection of chromosome 10 into 49 regions at an average genetic distance of about 3.9 cM, reflecting an average physical distance of about 2.9 Mb (size of HC10: 144 Mb). A combination of the two approaches provided better resolution for subregional mapping of sequences located on chromosome 10. Accordingly, markers D10S191 and D10S674, positioned together by linkage analysis, were

found physically to be separated in five hybrids, viz., 173, 185, 171, 172, and 167 (Fig. 2). In contrast, other markers grouped together at the genetic map, viz., D10S211/D10S553, D10S89/D10S197, and D10S608/D10S215, were found to be retained together in all hybrids scored, possibly suggesting their close physical linkage. Furthermore, molecular analysis of the hybrids permitted more precise mapping of the markers positioned approximately by linkage analysis. Thus, D10S2310, mapped to an 18.6-cM interval between D10S254 and D10S209 (Fig. 1), was found to be closely linked to D10S1156, as deduced from hybrids 176, 166, and 54 (Fig. 2). Similarly, the marker pattern of hybrids 222 and 171 suggested a close physical linkage of D10S2318 with D10S548.

FISH and molecular characterization of the hybrids revealed complementary and mutually supporting mapping information. FISH is inadequate for detecting microdeletions (present, as expected, mainly in radiation hybrids) because of chromatin condensation and because spreading of the fluorescent signals tends to fill small gaps; however, it is very efficient in detecting small chromosomal fragments. In contrast, the molecular approach offers a more detailed profile of the hybrid. In fact, several microdeletions that were not obvious in the FISH studies were detected by scoring for microsatellite markers. For example, hybrid 222, mapped by FISH to 10p12→p11.2, showed an internal microdeletion between D10S553 and D10S1214. When comparing FISH and PCR screening results, it should be taken into account that the ability of the molecular approach to detect small deletions or remote fragments depends upon both the number and spacing of the markers used for the analysis and the degree of representation of the particular sequence in the hybrid. Chromosome fragments produced by spontaneous rearrangements in cell hybrids are considered as having been created by a single breakage event, thus consisting of an uninterrupted chromosomal segment. Indeed, such hybrids in our panel appear with a considerably smaller number of internal microdeletions (Fig. 2b). These hybrids may be particularly useful in mapping studies, especially when radiation hybrid experiments produce conflicting results. Furthermore, they can be used for generating PCPs via *Alu*-PCR. PCPs offer an alternative approach for basic or clinical research and for diagnostic studies. They are useful for detecting chromosome rearrangements in tumors (Mitelman et al., 1997), where the classical techniques frequently fail due to the poor morphology of cancer cell chromosomes, or in studies of mammalian and, particularly, primate chromosome evolution (Muller et al., 1996). A combination of appropriate human chromosome 10 hybrids and YAC clones were employed for the refinement of cytogenetic and physical mapping of the paracentric inversion distinguishing the orangutan chromosome homolog from human chromosome 10.

As an additional resource for high-resolution map integration and cytogenetic investigations, we characterized 57 human chromosome 10-specific YACs. The physical order of the YAC clones, as determined by FISH, was found to correlate well with the MIT and the EUROGEN genetic linkage map data (results not shown). The majority of the YACs were recognized by markers used for the analysis of the cell hybrid panel, thus enabling the integrated use of YACs and PCPs derived from

the hybrids. In some cases, groups of YACs were recognized by the same marker, e.g., 10GA2, 10GA7, and 8IE1 by D10S2314 and 27DD2, 14CB8, 19FC12, and 24AC1 by D10S2312. *Alu*-PCR fingerprinting and pulsed-field gel electrophoresis restriction analysis revealed that these genomic clones comprise contigs with sizes ranging from 500 to 800 kb (M.A. Kokkinaki et al., unpublished data). Furthermore, markers grouped together in the genetic map turned out to be located, in fact, within short physical distances. Thus, D10S2311 and D10S564 (Fig. 1) were identified on a single YAC (10BA1). YACs 12HA1 and 37IF9, overlapping with 10BA1, contained two additional markers, D10S1753 and D10S2309, suggesting the close physical linkage of all four markers. To improve the resolution of the determined contig, several PAC clones were isolated. *Alu*-PCR fingerprinting and restriction analysis resulted in the selection of 16 PACs (Table 1) consisting a well-defined contig of about 700 kb in size, at a resolution of 50–80 kb (M.A. Kokkinaki et al., unpublished results). In another case, two YACs (27BC9 and 35GE1) and two PACs (171-A21 and 30-C6) revealed the close physical linkage of D10S2330 and D10S597, thus invalidating the estimated genetic distance (1.9 cM) of the two markers (Fig. 1). Approaches as those described, upgrade the integration of the genetic, cytogenetic, and physical maps of human chromosome 10. The molecular analysis of the panel facilitates the choice of appropriate genomic sequences to be used as additional tools for refining the molecular structure of cytogenetic rearrangements. This analysis, combined with the characterization of chromosome 10-specific genomic clones, permit the precise placement of the hybrids to the integrated chromosome 10 map, contribute to resources available for mapping this chromosome, and provide a powerful tool for various analytical cytogenetic methodologies and positional cloning procedures.

Acknowledgements

We are particularly grateful to the YAC Screening Center, DIBIT-HSR, Milan. We thank J. Attwood for assistance with the linkage map building. CEPH DNA filters and samples were provided to one of us (N.K.M.) by Génethon (Paris) in the context of the EUROGEN consortium.

References

- Antonacci R, Marzella R, Finelli P, Lonoce A, Forabosco A, Archidiacono N, Rocchi M: A panel of subchromosomal painting libraries representing over 300 regions of the human genome. *Cytogenet Cell Genet* 68:25–32 (1995).
- Archidiacono N, Antonacci R, Forabosco A, Rocchi M: Preparation of human chromosome painting probes from somatic cell hybrids, in Choo KHA (ed): *Methods in Molecular Biology: In situ Hybridization Protocols*, pp 1–13 (Humana Press, Ottowa 1994).
- Baldini A, Ward DC: In situ hybridisations of human chromosomes with *Alu*-PCR products: a simultaneous karyotype for gene mapping studies. *Genomics* 9:770–774 (1991).
- Chumakov IM, Rigault P, LeGall I, Bellané-Chantelot C, Billault A, Guillou S, Soularue P, Guasconi G, Poullier E, Gros I, Belova M, Sambucy JL, Susini L, Gervy P, Glibert F, Beauvais S, Bui H, Massart C, De Tand MF, Dukasz F, Lekoulant S, Ougen P, Perrot V, Saumier M, Soravito C, Bahouayila R, Cohen-Akenine A, Berillot E, Bertrand S, Codani JJ, Caterina D, Georges I, Laeroix B, Lucotte G, Sabbatou M, Schmit C, Sanguoat M, Tubacher E, Dib C, Fauré S, Fizames C, Gyapay G, Millasseau P, Nguyen S, Muselet D, Vignal A, Morissette J, Menninger J, Lieman J, Desai T, Banks A, Bray-Ward P, Ward D, Hudson T, Gerety S, Foote S, Stein L, Page DC, Lander ES, Weissenbach J, LePaslier D, Cohen D: A YAC contig map of the human genome. *Nature* 377(suppl):175–297 (1995).
- Cox SA, Attwood J, Bryant SP, Bains R, Povey S, Rebello M, Kapssetaki M, Moschonas NK, Grzeschik KH, Otto M, Dixon M, Sudworth HE, Kooy RF, Wright A, Teague P, Terrenato L, Vergnaud G, Monfouilloux S, Weissenbach J, Alibert O, Dib C, Faure S, Bakker E, Pearson NM, Vossen RHAM, Gal A, Mueller-Myhsok, Cann HM, Spurr NK: European Gene Mapping Project (EUROGEM): breakpoint panels for human chromosomes based on the CEPH reference families. *Ann Hum Genet* 60:447–486 (1996).
- Daw SC, Taylor C, Kraman M, Call K, Mao J, Schuffenhauer S, Meitinger T, Lipson T, Goodship J, Scrambler P: A common region of 10p deleted in DiGeorge and velocardiofacial syndromes. *Nature Genet* 13:458–460 (1996).
- Deaven LL, Van Dilla MA, Barthold MF, Carrano AV, Cram LS, Fuscoe JC, Gray JW, Hildebrand CE, Moyzis RK, Perlman J: Construction of human chromosome-specific DNA libraries from flow-sorted chromosomes. *Cold Spring Harb Symp Quant Biol* 51:159–167 (1986).
- Deloukas P, Dauwerse JG, Moschonas NK, van Ommen G-JB, van Loon APGM: Three human glutamate dehydrogenase genes (GLUD1, GLUDP2 and GLUDP3) are located on chromosome 10q, but are not closely physically linked. *Genomics* 17:676–681 (1993).
- Dib C, Fauré S, Fizames C, Samson D, Drouot N, Vignal A, Millasseau P, Marc S, Hazan J, Seboun E, Lathrop M, Gyapay G, Morissette J, Weissenbach J: A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* 380:152–154 (1996).
- Eng C, Mulligan LM: Mutations of the RET proto-oncogene in the multiple endocrine neoplasia type 2 syndromes, related sporadic tumours, and Hirschsprung disease. *Hum Mutat* 9:97–109 (1997).
- Francke U, Chang E, Comeau K, Geigle E-M, Giacalone J, Li X, Luna J, Moon A, Welch S, Wilgenbus P: A radiation hybrid map of human chromosome 18. *Cytogenet Cell Genet* 66:196–213 (1994).
- Goss SJ, Harris H: New method for mapping genes in human chromosomes. *Nature* 255:680–684 (1975).
- Gyapay G, Schmitt K, Fizames C, Jones H, Vega-Czarny N, Spillet D, Muselet D, Prud'Homme JF, Dib C, Auffray C, Morissette J, Weissenbach J, Goodfellow PN: A radiation hybrid map of the human genome. *Hum Molec Genet* 5:339–346 (1996).
- Hudson TJ, Stein LD, Gerety SS, Ma J, Castle AB, Silva JS, Slonim DK, Baptista R, Kruglyak L, Xu SH, Hu X, Colbert AME, Rosenberg C, Reeve-Daly MP, Rozen S, Hui L, Wu X, Vestergaard C, Wilson KM, Bae JS, Maitra S, Ganiatsas S, Evans CA, DeAngelis MM, Ingalls KA, Nahf RW, Horton LT, Anderson MO, Collymore AJ, Ye W, Kouyoumjian V, Zemsteva IS, Tam J, Devine R, Courtney DF, Renaud MT, Nguyen H, O'Connor TJ, Fizames C, Fauré S, Gyapay G, Dib C, Morissette J, Orlin JB, Birren BW, Goodman N, Weissenbach J, Hawkins TL, Foote S, Page DC, Lander ES: An STS-based map of the human genome. *Science* 270:1945–1954 (1995).
- Ioannou PA, Amemiya CT, Garnes J, Kroisel PM, Shizuya H, Chen C, Batzer MA, de Jong PJ: A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nature Genet* 6:84–89 (1994).
- Jabs EW, Li X, Scott AF, Meyers G, Chen W, Eccles M, Mao J, Chamas LR, Jackson CE, Jaye M: Jackson-Weiss and Crouzon syndromes are allelic with mutations in fibroblast growth factor receptor 2. *Nature Genet* 8:275–279 (1994).
- Kapssetaki M, Kokkinaki M, Angelicheva D, Lubyova B, Mavraki H, Argyrokastritis A, Vergnaud G, Ferguson-Smith M, Rubinsztein DC, Lush M, Moschonas NK: The EUROGEM map of human chromosome 10. *Eur J Hum Genet* 2:222–223 (1994).
- Knapp RD, Smith LC, Baldini A: Eigenanalysis of DAPI-stained chromosomes: tools and strategies toward computer-assisted analysis of FISH experiments. *Cytogenet Cell Genet* 69:81–86 (1995).
- Korenberg JR, Rykowski MC: Human genome organization: *Alu*, Lines and the molecular structure of metaphase chromosome bands. *Cell* 53:391–400 (1988).
- Liu P, Siciliano J, Seong D, Craig J, Zhao Y, de Jong PJ, Siciliano MJ: Dual *Alu* polymerase chain reaction primers and conditions for isolation of human chromosome painting probes from hybrid cells. *Cancer Genet Cytogenet* 65:93–99 (1993).
- Ma NSF, Zheng C, Benchekroun Y, Deaven LL, Longmire JL, Moir DT, Mao J: Characterization of a flow-sorted human chromosome 10 cosmid library by FISH. *Cytogenet Cell Genet* 74:266–271 (1996).
- Meitinger T, Scharfe C, Call K, Moschonas N: Report of the second international workshop on human chromosome 10 mapping 1997. *Cytogenet Cell Genet* 78:183–196 (1997).
- Mitelman F, Mertens F, Johansson B: A breakpoint map of recurrent chromosomal rearrangements in human neoplasia. *Nature Genet* 15(suppl):417–474 (1997).
- Moschonas NK, Spurr NK, Mao J: Report of the first international workshop on human chromosome 10 mapping 1995. *Cytogenet Cell Genet* 72:99–112 (1996).
- Muller S, Koehler U, Wienberg J, Marzella R, Finelli P, Antonacci R, Rocchi M, Archidiacono N: Comparative fluorescence in situ hybridization mapping of primate chromosomes with *Alu*-PCR generated probes from human/rodent somatic cell hybrids. *Chrom Res* 4:38–42 (1996).
- Nagase S, Yamakawa H, Sato S, Yajima A, Horii A: Identification of a 790-kilobase region of common allelic loss in chromosome 10q25-q26 in human endometrial cancer. *Cancer Res* 57:1630–1633 (1997).
- Nelson DL, Ledbetter SA, Corbo L, Victoria MF, Ramirez-Solis R, Webster TD, Ledbetter DH, Caskey CT: *Alu*-polymerase chain reaction: a method for rapid isolation of human-specific sequences from complex DNA sources. *Proc Natl Acad Sci, USA* 86:6686–6690 (1989).
- Nikali K, Isoomppi J, Lonnqvist T, Mao JI, Suomalainen A, Peltonen L: Toward cloning of a novel ataxia gene: refined assignment and physical map of the IOSCA locus (SCA8) on 10q24. *Genomics* 39:185–191 (1997).
- Nunes ME, Schutt G, Kapur RP, Luthardt F, Kukolich M, Byers P, Evans JP: A second autosomal split hand/split foot locus maps to chromosome 10q24-q25. *Hum Molec Genet* 4:2165–2170 (1995).
- Przyłępa KA, Paznekas W, Zhang M, Golabi M, Bias W, Bamshad MJ, Carey JC, Hall BD, Stevenson R, Orlow SJ, Cohen Jr MM, Jabs EW: Fibroblast growth factor receptor 2 mutations in Beare-Stevenson cutis gyrata syndrome. *Nature Genet* 13:492–494 (1996).
- Reardon W, Winter RM, Pulleyn LJ, Jones BM, Malcolm S: Mutation in the fibroblast growth factor receptor 2 gene cause Crouzon syndrome. *Nature Genet* 8:98–103 (1994).
- Rutland P, Pulleyn LJ, Reardon W, Baraitser M, Hayward R, Jones B, Malcolm S, Winter RM, Oldridge M, Slaney SF, Poole MD, Wilkie AOM: Identical mutations in the FGFR2 gene cause both Pfeiffer and Crouzon syndrome phenotypes. *Nature Genet* 9:173–176 (1995).
- Spurr NK, Bryant SP, Attwood J, Nyberg K, Cox SA, Mills A, Bains R, Warne D, Cullin L, Povey S, Seboun JM, Weissenbach J, Cann HM, Lathrop M, Dausset J, Marcadet-Troton A, Cohen D: European Gene Mapping Project (EUROGEM): genetic maps based on the CEPH reference families. *Eur J Hum Genet* 2:193–252 (1994).
- Tzimogiorgis G, Leversha M, Chroniary K, Goulielmos G, Sargent CA, Ferguson-Smith M, Moschonas NK: Structure and expression analysis of a novel member of the human glutamate dehydrogenase (GLUD) gene family mapped to chromosome 10p11.2. *Hum Genet* 91:433–438 (1993).
- Wilkie AOM, Slaney SF, Oldridge M, Poole MD, Ashworth GJ, Hockley AD, Hayward RD, David DJ, Pulleyn LJ, Rutland P, Malcolm S, Winter RM, Reardon W: Apert syndrome results from localized mutations of FGFR2 and is allelic with Crouzon syndrome. *Nature Genet* 9:165–172 (1995).
- Yunis JJ, Prakash O: The origin of man: a chromosomal pictorial legacy. *Science* 215:1525–1530 (1982).