



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

**ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ
ΒΙΟΙΑΤΡΙΚΗ**

**Γενετικοί πολυμορφισμοί των microRNAs και συσχέτισή τους με
ανθρώπινες νόσους. Συστηματική ανασκόπηση και σύνοψη πεδίου.**

Παμπάλου Ανδρομάχη

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Υπεύθυνοι

Μπράλιου Γεωργία

Διδάσκουσα σύμφωνα με το Π.Δ. 407/80

Μπάγκος Παντελής

Αναπληρωτής Καθηγητής

Λαμία, 2017



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΒΙΟΙΑΤΡΙΚΗ

**Γενετικοί πολυμορφισμοί των microRNAs και συσχέτισή τους με
ανθρώπινες νόσους. Συστηματική ανασκόπηση και σύνοψη
πεδίου.**

Παμπάλου Ανδρομάχη

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Επιβλέπων

Μπράλιου Γεωργία

Διδάσκουσα σύμφωνα με το Π.Δ. 407/80

Μπάγκος Παντελής

Αναπληρωτής Καθηγητής

Λαμία, 2017

Με ατομική μου ευθύνη και γνωρίζοντας τις κυρώσεις ⁽¹⁾, που προβλέπονται από της διατάξεις της παρ. 6 του άρθρου 22 του Ν. 1599/1986, δηλώνω ότι:

1. Δεν παραθέτω κομμάτια βιβλίων ή άρθρων ή εργασιών άλλων αυτολεξεί **χωρίς να τα περικλείω σε εισαγωγικά** και χωρίς να αναφέρω το συγγραφέα, τη χρονολογία, τη σελίδα. Η αυτολεξεί παράθεση χωρίς εισαγωγικά χωρίς αναφορά στην πηγή, είναι λογοκλοπή. Πέραν της αυτολεξεί παράθεσης, λογοκλοπή θεωρείται και η παράφραση εδαφίων από έργα άλλων, συμπεριλαμβανομένων και έργων συμφοιτητών μου, καθώς και η παράθεση στοιχείων που άλλοι συνέλεξαν ή επεξεργάστηκαν, χωρίς αναφορά στην πηγή. Αναφέρω πάντοτε με πληρότητα την πηγή κάτω από τον πίνακα ή σχέδιο, όπως στα παραθέματα.
2. Δέχομαι ότι η αυτολεξεί **παράθεση χωρίς εισαγωγικά**, ακόμα κι αν συνοδεύεται από αναφορά στην πηγή σε κάποιο άλλο σημείο του κειμένου ή στο τέλος του, είναι αντιγραφή. Η αναφορά στην πηγή στο τέλος π.χ. μιας παραγράφου ή μιας σελίδας, δεν δικαιολογεί συρραφή εδαφίων έργου άλλου συγγραφέα, έστω και παραφρασμένων, και παρουσίασή τους ως δική μου εργασία.
3. Δέχομαι ότι υπάρχει επίσης περιορισμός στο μέγεθος και στη συχνότητα των παραθεμάτων που μπορώ να εντάξω στην εργασία μου εντός εισαγωγικών. Κάθε μεγάλο παράθεμα (π.χ. σε πίνακα ή πλαίσιο, κλπ), προϋποθέτει ειδικές ρυθμίσεις, και όταν δημοσιεύεται προϋποθέτει την άδεια του συγγραφέα ή του εκδότη. Το ίδιο και οι πίνακες και τα σχέδια
4. Δέχομαι όλες τις συνέπειες σε περίπτωση λογοκλοπής ή αντιγραφής.

Ημερομηνία:/...../20.....

Η Δηλ.

(Υπογραφή)

(1) «Όποιος εν γνώσει του δηλώνει ψευδή γεγονότα ή αρνείται ή αποκρύπτει τα αληθινά με έγγραφη υπεύθυνη δήλωση του άρθρου 8 παρ. 4 Ν. 1599/1986 τιμωρείται με φυλάκιση τουλάχιστον τριών μηνών. Εάν ο υπαίτιος αυτών των πράξεων σκόπευε να προσπορίσει στον εαυτόν του ή σε άλλον περιουσιακό όφελος βλάπτοντας τρίτον ή σκόπευε να βλάψει άλλον, τιμωρείται με κάθειρξη μέχρι 10 ετών.

Γενετικοί πολυμορφισμοί των microRNAs και συσχέτισή τους με ανθρώπινες νόσους. Συστηματική ανασκόπηση και σύνοψη πεδίου.

Παμπάλου Ανδρομάχη

Τριμελής Επιτροπή:

Μπράλιου Γεωργία, Διδάσκουσα σύμφωνα με το Π.Δ. 407/80

Μπάγκος Παντελής, Αναπληρωτής Καθηγητής

Ευριπίδης Μάρκου, Επίκουρος Καθηγητής

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τους επιβλέποντες καθηγητές μου, κ. Μπράλιου Γεωργία και κ. Μπάγκο Παντελή, για την καθοδήγηση και βοήθεια που μου πρόσφεραν σε όλη αυτή την πορεία. Επιπλέον είμαι ευγνώμων στην οικογένεια μου, τους φίλους μου και σε όλους όσους με στηρίζουν καθημερινά και με προτρέπουν να κυνηγάω τους στόχους μου.

Περιεχόμενα

Περίληψη

Abstract

1. Εισαγωγή

1.1. Σκοπός Πτυχιακής Εργασίας

1.2. microRNAs

1.2.1. Ιστορική αναδρομή

1.2.2. Ονοματολογία

1.2.3. Βιογένεση των microRNA

1.2.4. Λειτουργία των microRNA

1.2.4.1. Αποικοδόμηση mRNA

1.2.4.2. Καταστολή της μετάφρασης του mRNA στόχου

1.2.4.3. Ρ-Σωματίδια (P-Bodies)

1.3. Γενετικοί πολυμορφισμοί και SNPs

1.3.1. Πολυμορφισμοί

1.3.2. Single Nucleotide Polymorphism SNP

1.3.3. SNP και microRNA

1.3.3.1. SNPs σε γονίδια που εμπλέκονται στη βιογένεση και επεξεργασία του miRNA

1.3.3.2. SNPs σε γονίδια miRNA

1.3.3.3. SNPs σε θέσεις δέσμησης miRNA σε γονίδια-στόχους

1.4. Μετα-ανάλυση και Συστηματική Ανασκόπηση

1.4.1. Μετα-ανάλυση

1.4.1.1. Ιστορική αναδρομή

1.4.1.2. Η μεθοδολογία της μετα-ανάλυσης

1.4.1.3. Μεμονωμένα ή Συγκεντρωτικά Δεδομένα

1.4.1.4. Αναζήτηση Βιβλιογραφίας

1.4.1.5. Στατιστική Ανάλυση

1.4.1.6. Ετερογένεια

1.4.1.7. Παρουσίαση Αποτελεσμάτων

1.4.2. Συστηματική ανασκόπηση

1.4.2.1. Στάδια συστηματικής ανασκόπησης

1.4.2.2. Πλεονεκτήματα συστηματικής ανασκόπησης

1.4.2.3. Μειονεκτήματα συστηματικής ανασκόπησης

2. Μεθοδολογία

2.1.Εργαλεία

2.2.Βήματα

2.2.1. Επιλογή άρθρων και μετα-αναλύσεων

2.2.2. Καταγραφή δεδομένων

2.2.3. Επεξεργασία καταγραφών

2.2.4. Υπολογισμός στατιστικών τιμών

3. Αποτελέσματα

3.1.Περιγραφική Στατιστική

3.2.Έλεγχοι γενετικής συσχέτισης των miR-SNPs με ασθένειες

3.3.Δημιουργία, Ανάλυση και Αξιολόγηση δικτύου γενετικής συσχέτισης

3.3.1. Αξιολόγηση με δεδομένα από MIRNASNP2

3.3.2. Αποτελέσματα από ανάλυση με Diana tools

3.3.3. Αξιολόγηση με δεδομένα από Kegg

3.3.4. Δημιουργία δικτύων συσχέτισης miRNA:genes και miRNA:pathways

4. Συμπεράσματα

5. Βιβλιογραφία

Περίληψη

Τα microRNAs (miRNAs) είναι μη κωδικά RNAs μήκους 21-25 νουκλεοτιδίων που ελέγχουν την έκφραση γονιδίων με δέσμευση την δέσμευση τους σε RNA στόχους. Έχει βρεθεί ότι η ύπαρξη ενός μονονουκλεοτιδικού πολυμορφισμού (SNP) σε κάποιο στάδιο βιογένεσης ή ωρίμανσης miRNA διαταράσσει την κανονική λειτουργία του microRNA, το οποίο με τη σειρά του μπορεί να επηρεάσει την έκφραση αρκετών γονιδίων και να συμβάλει σε ασθένειες. Την τελευταία δεκαετία πολλές μετα-αναλύσεις έχουν διερευνήσει αυτή τη σύνδεση. Σε αυτή την εργασία, επιχειρήσαμε να προσδιορίσουμε όλα τα miR-SNPs που σχετίζονται με ανθρώπινες ασθένειες, αφού εξετάσαμε όλες τις διαθέσιμες μετα-αναλύσεις στη βιβλιογραφία.

Μια εκτενής βιβλιογραφική αναζήτηση πραγματοποιήθηκε στη βάση δεδομένων PubMed προκειμένου να βρεθεί "μετα-ανάλυση" σε SNPs στο "microRNA ή miRNA". Για κάθε έρευνα καταγράψαμε διάφορα πεδία όπως ο αριθμός των μελετών και των συμμετεχόντων, μεμονωμένες ομάδες εθνικότητας (Ασιάτες, Καυκάσιοι και Μικτοί Πληθυσμοί), το στατιστικό μοντέλο για το οποίο αναφέρθηκε μαζί με τους συντελεστές απόδοσης και τα διαστήματα εμπιστοσύνης 95%, οι τιμές της ετερογένειας και τεστ για στατιστικές προκαταλήψεις. Κάθε ζεύγος SNP:ασθένεια που βρέθηκε συσχέτιση, διερευνήθηκε στη βάση δεδομένων MIRNASNP2 προκειμένου να συγκριθούν τα αποτελέσματα. Επιπλέον, το miRNA στο οποίο βρέθηκαν SNPs που σχετίζονταν με ασθένειες αναλύθηκε με τα εργαλεία Diana. Τέλος, δημιουργήθηκε ένα δίκτυο που συνδέει τα miR-SNPs και τις ασθένειες χρησιμοποιώντας το εργαλείο Cytoscape.

Abstract

MicroRNAs (miRNAs) are 21–25-nucleotide-long, noncoding RNAs that control the expression of genes by binding to RNA targets. The existence of a SNP at some stage of miRNA biogenesis or maturation has been found to disturb the normal function of the microRNA, which in turn can affect the expression of several genes and contribute to diseases. The last decade many meta-analyses have investigated this connection. In this work, we attempted to determine all the miR-SNPs associated with human diseases after reviewing all the available meta-analyses in the literature.

A comprehensive literature search was performed in the PubMed database in order to find “meta-analysis” on SNPs in “microRNA or miRNA”. For each paper we recorded several fields such as number of studies and participants, individual groups (Asians, Caucasians and Mixed Populations), statistical model reported along with Odds Ratios and 95% Confidence Intervals, heterogeneity and bias statistics. Each pair of SNP:disease which we consider associated, were investigated at MIRNASNP2 database in order to compare the results. In addition the miRNA in which SNPs related to diseases had been found were processed with Diana tools. Finally a network that connects miR-SNPs and diseases were created by using Cytoscape tool.

1. Εισαγωγή

1.1 Σκοπός Πτυχιακής Εργασίας

Στόχος της παρούσας μελέτης είναι η συγκέντρωση και ανασκόπηση όλων των διαθέσιμων δεδομένων από μετα-αναλύσεις, οι οποίες διερευνούν την συσχέτιση διαφόρων SNPs σε microRNAs με ανθρώπινες ασθένειες. Επειδή Υπάρχει πληθώρα τέτοιων μετα-αναλύσεων, ήταν απαραίτητο να υπάρξει μια σύνοψη στο συγκεκριμένο πεδίο ώστε ο μελετητής να μπορεί να έχει μια εμπειριστατωμένη εικόνα για γενετικές συσχετίσεις miR-SNPs και ασθενειών. Αυτή η εικόνα θα μπορούσε να χρησιμοποιηθεί ως οδηγός αναζήτησης συσχετίσεων ώστε να προάγει την πειραματική έρευνα, π.χ. δίνοντας δυνατότητα για δημιουργία γρηγορών διαγνωστικών/προγνωστικών τεστ.

Δόθηκε λοιπόν σημασία στη κατασκευή εποπτικών τρόπων απόδοσης αυτών των συσχετίσεων, μια και από την έρευνά μας βρέθηκε ότι η πληροφορία για κάθε SNP, για κάθε miRNA και για κάθε ασθένεια κατέληγε σε πλήθος πινάκων και ιστογραμμάτων, που δύσκολα μπορούσε να γίνει πλήρως αντιληπτή.

Για αυτό το λόγο κατασκευάστηκε δίκτυο συσχέτισης για τα δεδομένα SNPs σε microRNAs και τις αντίστοιχες ανθρώπινες ασθένειες, που βρέθηκε να έχουν θετική συσχέτιση στις μετα-αναλύσεις. Ένα βιολογικό δίκτυο συσχέτισης αναπαριστά ευδιάκριτα και κατανοητά σύνθετες συσχετίσεις μεταξύ ιατροβιολογικών παραγόντων. Ακολουθώντας τις συνδέσεις που φανερώνουν οι ακμές του δικτύου, σκοπός ήταν η να μπορεί να εξάγεται εύκολα και γρήγορα συμπέρασμα για τους κόμβους του δικτύου (SNP, miRNA ή ασθένεια).

Καθώς στόχος της παρούσας εργασίας είναι η εξαγωγή τόσο συγκεντρωτικών, όσο και αξιολογών ποιοτικά αποτελεσμάτων, βασικό στάδιο αποτέλεσε η επικύρωση και αξιολόγηση των δεδομένων με στοιχεία από ηλεκτρονικές βάσεις δεδομένων. Η πρώτη αξιολόγηση έγινε με χρήση της MIRNASNP2 κατά την οποία ανακτήσαμε πληροφορίες δεδομένων έκφρασης SNP για διαφόρους τύπους καρκίνου (μόνο δεδομένα για καρκίνους είναι διαθέσιμα σε αυτή την βάση δεδομένων) και συγκρίναμε αυτά τα SNP με τα δεδομένα που πρόεκυψαν από τη δική μας έρευνα. Αξιοσημείωτο είναι ότι βρέθηκε ταύτιση των ζευγών κατά 92% .

Η δεύτερη αξιολόγηση έγινε με έλεγχο με τη χρήση του Diana tool κατά την οποία ελέχθησαν τα βιοχημικά μονοπάτια στα οποία εμπλέκονται τα συγκεκριμένα miRNA καθώς και οι ασθένειες που αυτά τα μονοπάτια επηρεάζουν. Τέλος έγινε

1. Εισαγωγή

σύγκριση των ασθενειών που πρόεκυψαν από την έρευνα στο Diana με τις ασθένειες που πρόεκυψαν από τη δική μας έρευνά και βρέθηκε μεγάλος βαθμός ταύτισης.

Επομένως, έχοντας συγκεντρώσει, αναλύσει, αξιολογήσει και επικυρώσει τα αποτελέσματα αυτής της εργασίας, επιτυγχάνεται και ο τελικός σκοπός της παρούσας πτυχιακής, ο οποίος είναι η συλλογή και ανακάλυψη των λίγων πιο σημαντικών και ουσιωδών συσχετίσεων SNPs σε microRNAs με ανθρώπινες ασθένειες, που θα μπορούν να χρησιμοποιηθούν από βιο-επιστήμονες για περαιτέρω βιολογικές αναλύσεις, ακόμα και για εμπορική χρήση .

1.2 microRNA

1.2.1 Ιστορική αναδρομή

Μέχρι τις αρχές του 20^{ου} αιώνα τα microRNAs κατατάσσονταν ως μη σημαντικά τμήματα RNA, με λειτουργική χρήση μόνο σε μη θηλαστικά. Το πρώτο microRNA που ανακαλύφθηκε ήταν αποτέλεσμα της έρευνας των Victor Ambros, Lee Rosalind C., Rhonda L. Feinbaum [1], σε συνδυασμό με τις δημοσιεύσεις των Gary Ruvkun, Wightman Bruce, Ilho Ha [2]. Και τα δύο εργαστήρια μελετούσαν την κατασταλτική δράση του γονιδίου lin-4, το οποίο όπως ήταν γνωστό ότι ελέγχει την ανάπτυξη του σκουληκιού *Caenorhabditis elegans*, πάνω στο γονίδιο lin-14. Κατά την διάρκεια του πειράματος των Ambros et al., όταν απομόνωσαν το lin-4 συνειδητοποίησαν πως αντί παραχθεί ένα κανονικό μόριο mRNA, το οποίο φυσιολογικά κωδικοποιεί μια πρωτεΐνη, παράχθηκε ένα πρόδρομο γονίδιο μήκους 61 νουκλεοτιδίων, το οποίο ωρίμασε σε ένα μη κωδικοποιημένο RNA μήκους 22 νουκλεοτιδίων. Παράλληλα η ομάδα των Ruvkun et al. ανακάλυψαν ότι η σύνθεση της πρωτεΐνης lin-14 είναι ρυθμιζόμενη μεταγραφικά και ότι τα επίπεδα lin-14 είναι αντιστρόφως ανάλογες με εκείνα του RNA lin-4. Ανάλυση της αλληλουχίας του RNA των 22 νουκλεοτιδίων αποκάλυψε πως περιείχε ακολουθίες σχεδόν συμπληρωματικές με διάφορα τμήματα του 3' άκρου του lin-14, οδηγώντας στο συμπέρασμα ότι το lin-4 ρυθμίζει το lin-14, και αποκαλύπτοντας την πρώτη αλληλεπίδραση miRNA και mRNA στόχου.

Για χρόνια θεωρούταν πως το μικρό RNA του lin-4 αποτελούσε ιδιομορφία του συγκεκριμένου οργανισμού, έως το 2000 όπου ανακαλύφθηκε ένα ακόμη μικρό RNA, στο *Caenorhabditis elegans*, το let-7, το οποίο καταπιέζει την έκφραση των lin-41, lin-14, lin-28, lin-42 και daf-12 κατά την ανάπτυξη. Η ταυτοποίηση της σχέσης των lin-4 και let-7 με τον ρυθμό ανάπτυξης πολλών σπονδυλωτών ειδών, συμπεριλαμβανομένων των ανθρώπων [4], οδήγησε σε υποψίες ότι τα μικρά RNA καθορίζουν την ρύθμιση της ανάπτυξης σε πλήθος ειδών. Ένα χρόνο αργότερα, το 2001, αυτές επιβεβαιώθηκαν, καθώς αναγνωρίστηκαν και κλωνοποιήθηκαν μόρια μικρών RNA από αρκετούς οργανισμούς, συμπεριλαμβανομένου του ανθρώπου, και οι νουκλεοτιδικές τους αλληλουχίες βρέθηκαν φυλογενετικά συντηρημένες [5][6][7]. Σε αυτό το σημείο, η ερευνητική κοινότητα ξεκίνησαν να χρησιμοποιούν τον όρο "microRNA" για να αναφερθούν σε αυτή την κατηγορία μικρών ρυθμιστικών RNAs και αποτελέσματα πλήθους μελετών αποκάλυψαν τις βασικές αρχές της βιογένεσης και λειτουργίας των microRNAs.

1. Εισαγωγή

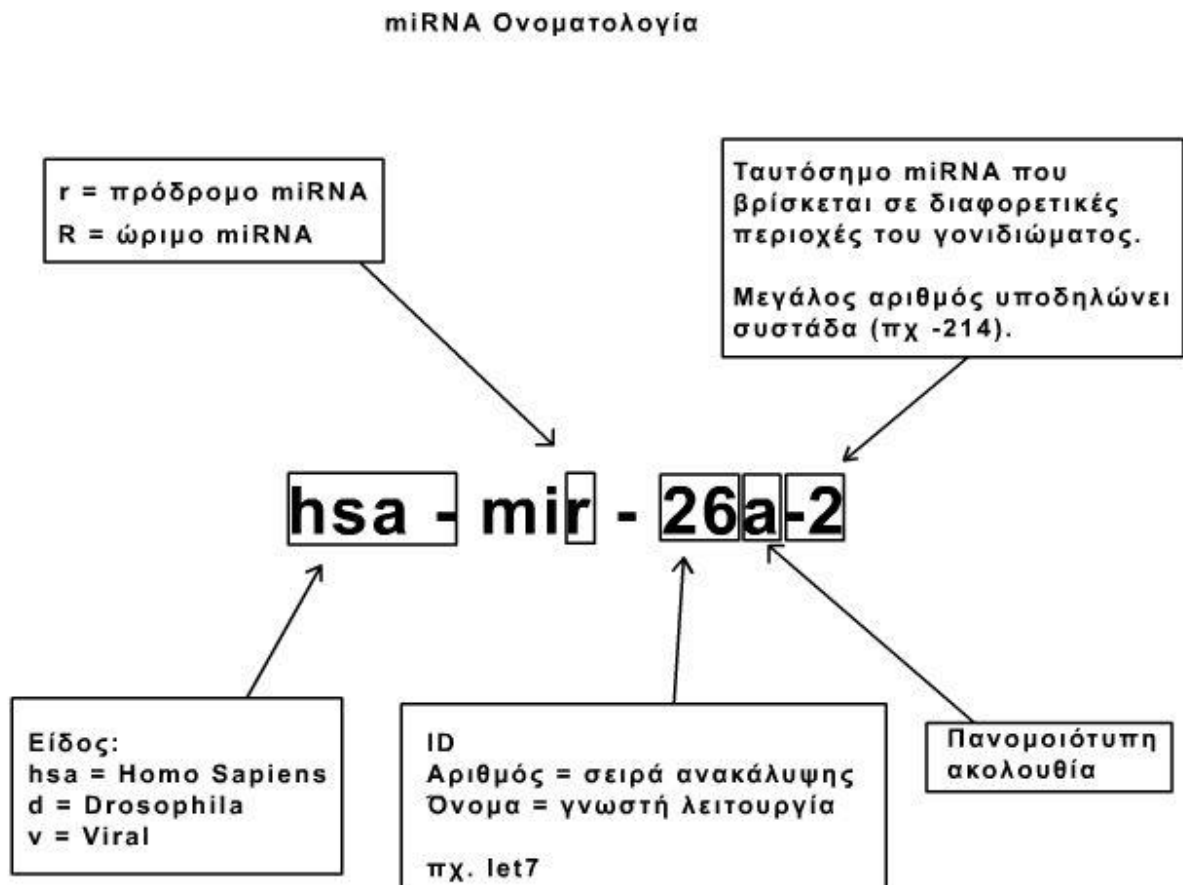
Η πρώτη ένδειξη για εμπλοκή των microRNAs σε ανθρώπινη νόσο παρουσιάστηκε το 2002 από την ομάδα των Calin et al., οι οποίοι έδειξαν μια συσχέτιση μεταξύ της υποέκφρασης των miR-15a και miR-16-1 και της εμφάνισης χρόνιας λεμφοκυτταρικής λευχαιμίας, υποδεικνύοντας τον ογκοκατασταλτικό ρόλο των miR-15a και miR-16-1 [8][9]. Πλέον, η ανάλυση μικροσυστοιχιών και οι προσεγγίσεις βαθιάς αλληλουχίας επέτρεψαν στους ερευνητές να συσχετίσουν την δυσλειτουργία των microRNAs στην εξέλιξη πολλών διαφορετικών ασθενειών, τόσο στα ζώα όσο και στους ανθρώπους.

1.2.2 Ονοματολογία των microRNAs

Ένα πρότυπο σύστημα ονοματολογίας για τον χαρακτηρισμό των microRNAs δημιουργήθηκε [10][11][12], σύμφωνα με το οποίο για να δοθεί όνομα σε ένα miRNA είναι απαραίτητη η πειραματική απόδειξη του πριν την δημοσίευση. Το σύστημα έχει τα ακόλουθα χαρακτηριστικά :

- Η γενική μορφή του ονόματος του microRNA είναι το πρόθεμα “miR” ακολουθούμενο από παύλα και ένα αριθμό.
- Ο αριθμός συνήθως υποδηλώνει την σειρά ανακάλυψης, π.χ. το miR-146 ανακαλύφθηκε πριν το miR-499.
- Το πρόθεμα “miR” χαρακτηρίζει ένα ώριμο microRNA, το “mir” ένα πρόδρομο microRNA (pre-miRNA και pri-miRNA), ενώ το πρόθεμα “MIR” απευθύνεται στ γονίδιο που το κωδικοποιεί.
- microRNAs με πανομοιότυπη ακολουθία, δηλαδή διαφέρουν σε 1-2 νουκλεοτίδια τότε χρησιμοποιείται ο ίδιος αριθμός και προστίθεται στο τέλος ένα επιπλέον γράμμα παρακάτω της αγγλικής αλφαβήτου π.χ. miR-146a και miR-146b.
- Όταν πρόδρομα microRNAs που βρίσκονται σε διαφορετικές περιοχές στο γονιδίωμα, αλλά οδηγούν στην ακριβής παραγωγή ίδιου ώριμου microRNA τότε συμβολίζονται προσθέτοντας παύλα και έναν αριθμό π.χ. . mir-196a-1 και mir-196a-2.
- Όταν δύο ώριμα microRNAs προέρχονται από αντίθετους βραχίονες του ίδιου πρώιμου miRNA και βρίσκονται σε περίπου παρόμοιες ποσότητες, σημειώνονται με ένα επίθημα -3p ή -5p, ανάλογα το άκρο που βρίσκονται στο τέλος του ονόματος.
- Ο οργανισμός στον οποίο ανήκει το microRNA δηλώνεται με την εισαγωγή ενός ακρώνυμου στην αρχή του ονόματος π.χ., για τον άνθρωπο hsa-miR-149 όπου το πρόθεμα hsa: Homo sapiens, ενώ για το ποντίκι mmu-miR-149 όπου mmu: Mus musculus.

1. Εισαγωγή



Εικόνα 1: Οδηγίες ονοματολογίας microRNA

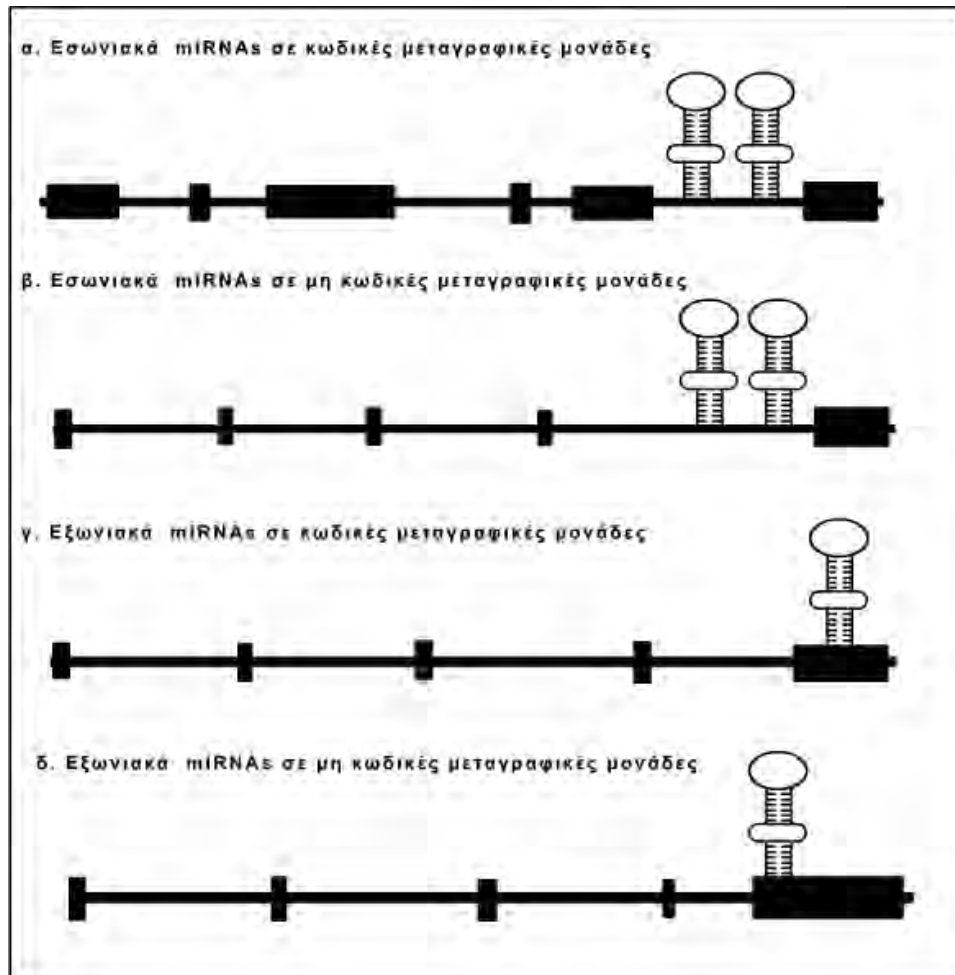
1.2.3 Βιογένεση των microRNAs

Ο όρος βιογένεση των microRNA αναφέρεται στη διαδικασία που απαιτείται ώστε να το αρχικό γονίδιο να μετατραπεί σε ώριμο μόριο microRNA, το οποίο έχει βιολογική δραστικότητα. Η διαδικασία αυτή περιλαμβάνει πλήθος σταδίων και ρυθμιστικών σημείων ελέγχου, κατά τα οποία τα microRNA αλλάζουν συνεχώς μορφή.

Η πλειοψηφία των πρώιμων microRNAs των θηλαστικών βρίσκονται σε εσώνια γονιδίων που κωδικοποιούν πρωτεΐνες ή σε μεταγραφικές μονάδες που εντοπίζονται μεταξύ γονιδίων και δεν κωδικοποιούν πρωτεΐνες, ενώ πιο σπάνια

1. Εισαγωγή

βρίσκονται σε εξώνια [5][6]. Με βάση την θέση τους στο γονιδίωμα τα γονίδια των microRNAs μπορούν να κατηγοριοποιηθούν ως εξής [13]: (α) εσωνιακά miRNAs σε κωδικές μεταγραφικές μονάδες, (β) εσωνιακά miRNAs σε μη κωδικές μεταγραφικές μονάδες, (γ) εξωνιακά miRNAs σε κωδικές μεταγραφικές μονάδες, (δ) εξωνιακά miRNAs σε μη κωδικές μεταγραφικές μονάδες. [Εικόνα 2]



Εικόνα 2: Πιθανές θέσεις εύρεσης microRNAs

Τα γονίδια των miRNAs που βρίσκονται σε εξώνια συνήθως ρυθμίζονται μαζί με τα γονίδια των ξενιστών τους [14][15][16]. Περίπου τα μισά γνωστά miRNAs όμως, έχουν βρεθεί να σχηματίζουν ομάδες και μεταγράφονται ως πολυκιστρονικά μετάγραφα [15]. Ενώ ορισμένα μπορούν να μεταγραφούν από τον δικό εκκινητή ως μονοκιστρονικά μετάγραφα [18].

Το πρώτο στάδιο της διαδικασίας της βιογένεσης αποτελεί η μεταγραφή του microRNA γονιδίου από το ένζυμο RNA πολυμεράση II [18], ενώ πιο σπάνια η μεταγραφή πραγματοποιείται από το ένζυμο RNA πολυμεράση III [19]. Η RNA πολυμεράση (II ή III), δεσμεύει μια πολύ συγκεκριμένη περιοχή του DNA και συγκεκριμένα στο σημείο εκείνο, το οποίο μετέπειτα θα γίνει το hair-pin loop του

1. Εισαγωγή

πρόδρομου miRNA ή precursor miRNA (pre- miRNA). Το προϊόν που παράγεται από την δράση των πολυμερασών είναι γνωστό ως πρώιμο miRNA ή primary miRNA (pri-miRNA), έχει μήκος πάνω από 100 νουκλεοτιδικές βάσεις και περιέχει τοπικές stem-loop δομές.

Η παραγωγή του pre-miRNA από το pri-miRNA συντελείται στον πυρήνα και αποδίδει ένα πρόδρομο μόριο μεγέθους περίπου 70 βάσεων. Η διαδικασία καταλύεται από το ένζυμο Drosha που ανήκει στην κατηγορία των ριβονουκλεασών τύπου III, σε συνδυασμό με συμπαραγόνα DGCR8, ο οποίο λειτουργεί ως ένας μοριακός κανόνας για την ακριβή θέση πέψης του miRNA, η οποία πρέπει να απέχει 11 βάσεις από την περιοχή που ξεκινάει η 3' περιοχή και 22 βάσεις από την περιοχή του loop [20].

Το pre-miRNA μεταφέρεται στο κυτταρόπλασμα ώστε να υποστεί την απαραίτητη ωρίμανση. Για την μεταφορά το pre-miRNA δημιουργεί ένα σύμπλοκο με τον νουκλεοκυτταροπλασματικό μεταφορικό παράγοντα Exportin-5 και RanGTP, ο οποίος αποτρέπει την πυρηνική αποικοδόμηση και διευκολύνει τη μετατόπιση στο κυτταρόπλασμα. Ένα κύριο προτεινόμενο μοντέλο λειτουργίας του παραπάνω συμπλόκου υποδηλώνει πως η εξαγωγή του pre-miRNA αρχίζει όταν η EXP5 αναγνωρίζει το μοτίβο μήκους 14 βάσεων double-stranded RNA (dsRNA) stem loop με μια προεξοχή δύο νουκλεοτιδίων που απομένει από το ένζυμο Drosha RNase III στο 3' άκρο της φουρκέτας του pre-miRNA. Το σύμπλεγμα εξάγεται από τον πυρήνα, όπου η υδρόλυση της GTP έχει ως αποτέλεσμα την απελευθέρωση του pre-miRNA [21][22][23][24][25].

Όταν πλέον το pre-miRNA βρεθεί στο κυτταρόπλασμα διασπάται από το ένζυμο Dicer, το οποίο ανήκει στην οικογένεια των RNase III ενδονουκλεασών. Το Dicer συνδέεται με το pre-miRNA στην προεξοχή μήκους 2 νουκλεοτιδίων στο 3' άκρο, η οποία είχε δημιουργηθεί προηγουμένως από το ένζυμο Drosha. Συνήθως οι θέσεις διάσπασης του Dicer βρίσκονται σε σταθερή απόσταση από το άκρο 3' των dsRNAs, σε απόσταση περίπου 21-25 νουκλεοτίδια και εξαρτάται από το είδος και τον τύπο του Dicer [26][27][28]. Στα θηλαστικά και τις μύγες, υπάρχει ένας επιπλέον μηχανισμός για τον προσδιορισμό της θέσης διάσπασης του pre-miRNA σύμφωνα με τον οποίο το Dicer δένεται στο 5' άκρο του pre-miRNA και το διασπά 22 νουκλεοτίδια μακριά από το άκρο αυτό [29]. Η σύνδεση της Dicer στο 5' άκρο πραγματοποιείται όταν το άκρο είναι θερμοδυναμικά ασταθές, αλλά όχι όταν το άκρο είναι ισχυρά ζευγαρωμένο, όπως μέσω ζευγών βάσης G-C. Το προϊόν που παράγεται από την δράση της Dicer είναι ένα μόριο περίπου 22 νουκλεοτιδικών ζευγών το οποίο αποτελεί και το ώριμο microRNA.

Το ώριμο μόριο microRNA που έχει δημιουργηθεί είναι ακόμα ατελές. Στην συνέχεια ενσωματώνεται σε ένα σύμπλοκο με πρωτεΐνες της οικογενείας Ago, οι οποίες είναι απαραίτητες για την επαγόμενη από mi-RNA σίγαση (miRNA-induced

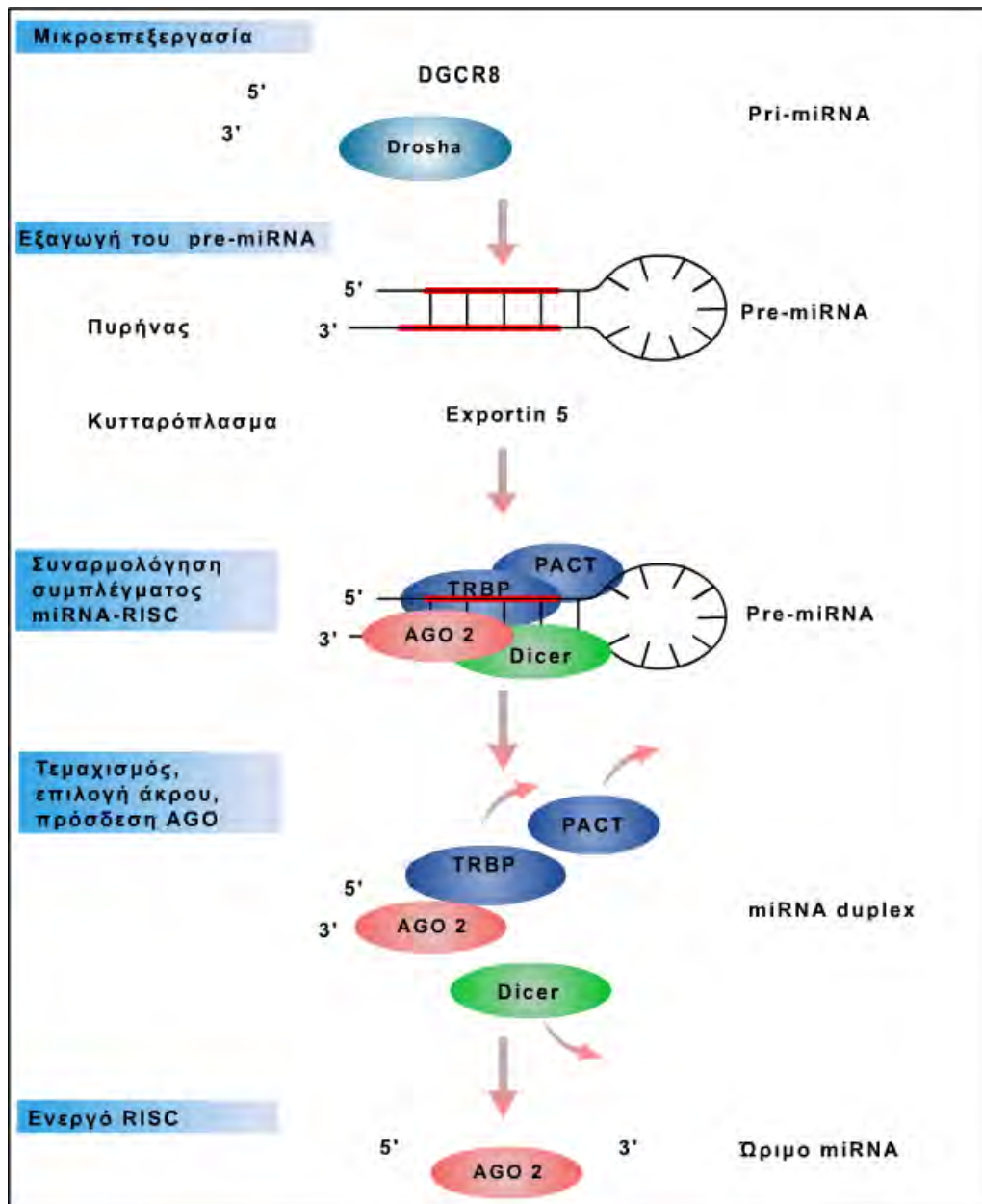
1. Εισαγωγή

silencing). Έχοντας δύο συγκεκριμένες περιοχές σχεδιασμένες να προσδένονται η μία με το 3' άκρο του ώριμου microRNA, και η άλλη με το 5' άκρο της αλυσίδας που θα επιτελέσει τον οδηγό, προσδένουν το ώριμο microRNA και το οδηγούν προς την αλληλεπίδραση του με τον mRNA στόχο. Όπως υπονοήθηκε ήδη μόνο ο ένας κλώνος του ώριμου microRNA που έχει δημιουργήσει το Dicer αποικοδομείται, ο κλώνος επιβάτη ή miRNA*. Ο άλλος κλώνος παραμένει προσδεμένος στην Ago πρωτεΐνη και καλείται οδηγός ή miRNA.

Το microRNA σε συνδυασμό με τις πρωτεΐνες Dicer, Ago και τις TRBP,PACT συμβάλουν στην δημιουργία του συμπλέγματος RISC (RNA induced silencing complex)[30]. Αρχικά δημιουργείται ένα σύμπλοκο φόρτωσης του RISC (RISC loading complex,RLC), του οποίου ο ακριβής τρόπος δράσης δεν είναι προσδιορισμένος, αλλά σύμφωνα με μελέτες υποστηρίζεται πως το σταθερό άκρο από το ώριμο microRNA, ύστερα από την απελευθέρωση του από την Dicer, προσδένεται σε ενεργές πρωτεΐνες του συμπλόκου RLC, ενώ το ασταθές με πρωτεΐνες τις οικογένειας Ago [31][32][33]. Με αυτόν τον τρόπο η μη επιλεγμένη αλυσίδα ξετυλίγεται και απομακρύνεται και ο κλώνος οδηγός, miRNA, οδηγεί το σύμπλοκο RISC στον mRNA στόχο, ο οποίος καταστέλλεται είτε μέσω αποικοδόμησης είτε μέσω καταστολή της μετάφρασης του. Αυτή είναι και η κύρια λειτουργία των microRNAs, οποία θα αναλυθεί στο κεφάλαιο που ακολουθεί.

Η εικόνα 3 περιγράφει όλη την διαδρομή που ακολουθεί το microRNA ώστε από πρώιμο μόριο να ωριμάσει και να αποκτήσει ενεργή λειτουργία.

1. Εισαγωγή



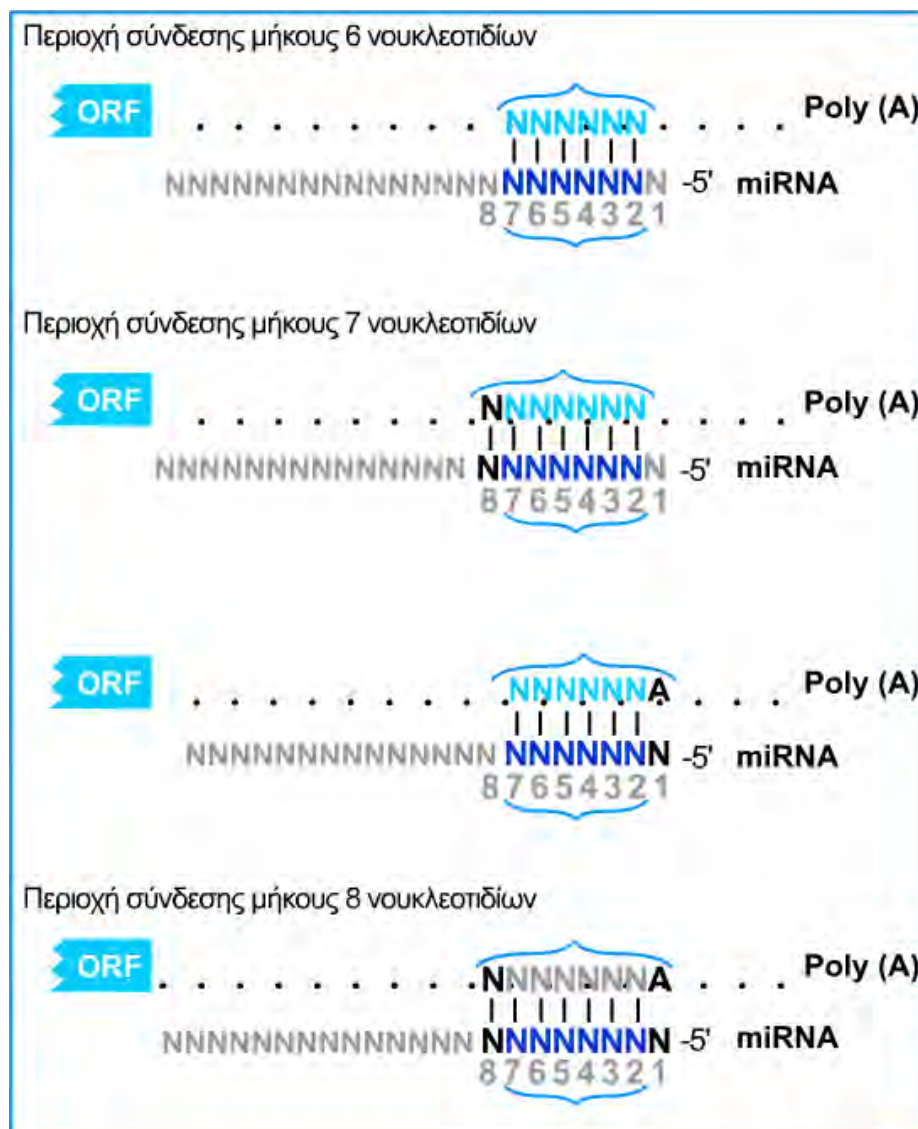
Εικόνα 3: Διαδικασία βιογέννησης του microRNA

1.2.4 Λειτουργία των microRNAs

Όπως αναφέρθηκε και στην προηγούμενη παράγραφο, τα miRNAs ρυθμίζουν τη γονιδιακή έκφραση είτε μέσω αποδόμησης των mRNAs στόχων, είτε μέσω παρεμπόδισης της μετάφρασης τους. Η μετα-μεταγραφική γονιδιακή σίγηση αρχίζει όταν το miRNA οδηγός, το οποίο είναι μέλος στο ενεργοποιημένο σύμπλοκο RISC, προσδέεται στο 3' άκρο του mRNA στόχου, μέσω του κανόνα συμπληρωματικότητας των βάσεων. Η αναγνώριση στόχου εξαρτάται σε μεγάλο

1. Εισαγωγή

βαθμό από τη συμπληρωματικότητα των βάσεων μεταξύ της περιοχής σπόρου (seed region SD) η οποία είναι τα 2-8 νουκλεοτίδια στο 5' άκρο του miRNA οδηγού [34][35][36][37]. Ο βαθμός και η φύση των συμπληρωματικών θέσεων μεταξύ του οδηγού και του στόχου φαίνεται να καθορίζουν τον μηχανισμό σίγασης των γονιδίων. Η αποικοδόμηση μέσω της ενδονουκλεάσης Ago2 γενικά ευνοείται από σχεδόν τέλεια και εκτεταμένη συμπληρωματικότητα βάσεων ενώ αναστολή μετάφρασης φαίνεται να απαιτεί πολλαπλές συμπληρωματικές θέσεις με μόνο μέτριο και περιορισμένο βαθμό συμπληρωματικότητας σε κάθε θέση [38]. Στα ζώα και στον άνθρωπο, το miRNA είναι γενικά 100% συμπληρωματικό στην seed region, αλλά όχι σε ολόκληρο το miRNA, γεγονός που οδηγεί σε ατελή υβρίδια RNA με χαρακτηριστικές εξογκώσεις [38][39][40]. Η δέσμευση στον σπόρο είναι σημαντική για τη θερμική σταθερότητα της αλληλεπίδρασης, η οποία συμβάλλει στην εξειδίκευση και δραστηριότητα του miRNA [41].



Εικόνα 4: Κατηγορίες περιοχών σύνδεσης miRNA

1. Εισαγωγή

Ένας μοναδικός οδηγός miRNA μπορεί να ρυθμίσει αρκετούς mRNA στόχους και αντίστροφα πολλαπλά miRNAs μπορούν να ρυθμίσουν σε συνεργασία έναν mRNA στόχο [38]. Ωστόσο, τα μέσα σίγασης παραμένουν ασαφή. Τα αποδεικτικά στοιχεία υποστηρίζουν δύο διαφορετικούς μηχανισμούς σίγασης, την αποδόμηση του mRNA στόχου και την καταστολή της μετάφρασης του, και οι δύο τελικά οδηγούν σε απορύθμιση της γονιδιακής έκφρασης [42] [43][44]. Μια σημαντική διαφορά μεταξύ των αποτελεσμάτων είναι η αναστρεψιμότητα, η αποσύνθεση του mRNA είναι μια μη αναστρέψιμη διαδικασία. Αντιθέτως, η αναστολή μετάφρασης είναι αναστρέψιμη επειδή το σταθερό mRNA μπορεί να μεταφραστεί μετά την εξάλειψη της καταστολής της μετάφρασης [44]. Παρακάτω αναλύονται και οι 2 μηχανισμοί σίγασης του mRNA.

1.2.4.1 Αποικοδόμηση mRNA

Έχει αποδειχθεί ότι όταν τα miRNAs έχουν ένα υψηλό βαθμό συμπληρωματικότητας αλληλουχιών, τότε οι διεργασίες αποικοδόμησης του mRNA στόχου διευκολύνονται μέσω της δραστηριότητας της Ago πρωτεΐνης. Το γεγονός ότι τα mRNAs μειώνονται υπό παρουσίας μιας πληθώρας miRNAs υποδηλώνει ότι τα miRNAs είναι υπεύθυνα για τις διεργασίες αποικοδόμησης του mRNA [44][45][46]. Πρόσφατες μελέτες έχουν δείξει ότι η διαδικασία αποικοδόμησης δεν καταλείπεται μόνο από την επίδραση της Ago στο mRNA, αλλά εμπλέκονται και άλλοι μηχανισμοί όπως η αποαδενυλίωση (deadenylation), η αποκόλληση (decapping) και η εξωνουκλεολυτική πέψη του mRNA [44][47][48]. Η αποικοδόμηση του mRNA με miRNA απαιτεί τις πρωτεΐνες Ago, GW182, και το μηχανισμό κυτταρικής αποκοπής και αποαδενυλίωσης. Η ακριβής διαδικασία επιλογής στόχων δεν έχει ακόμη καθοριστεί, ωστόσο, έχει αποδειχθεί ότι ο αριθμός, ο τύπος και η θέση των αναντιστοιχιών στο ζεύγος miRNA με mRNA στόχο διαδραματίζουν κρίσιμο ρόλο στην επιλογή των μηχανισμών καταστολής αποικοδόμησης ή μετάφρασης.

1.2.4.2 Καταστολή της μετάφρασης του mRNA στόχου

Πολλαπλές συμπληρωματικές θέσεις με ατελή συμπληρωματικότητα βάσεων δημιουργούν διογκώσεις στο διπλό μόριο RNA και αναστέλλουν τη δραστηριότητα του Ago2 [1][2][38] [49][50]. Αυτό δεν επηρεάζει την ικανότητα των πρωτεϊνών Argonaute να καταστέλλουν τη μετάφραση του mRNA στόχου [51][52]. Δεν υπάρχει ξεχωριστό μοντέλο για το πώς τα miRNAs αναστέλλουν τη μετάφραση, αλλά είναι σαφές ότι αυτό μπορεί να συμβεί με διάφορους τρόπους. Η μετάφραση λαμβάνει

1. Εισαγωγή

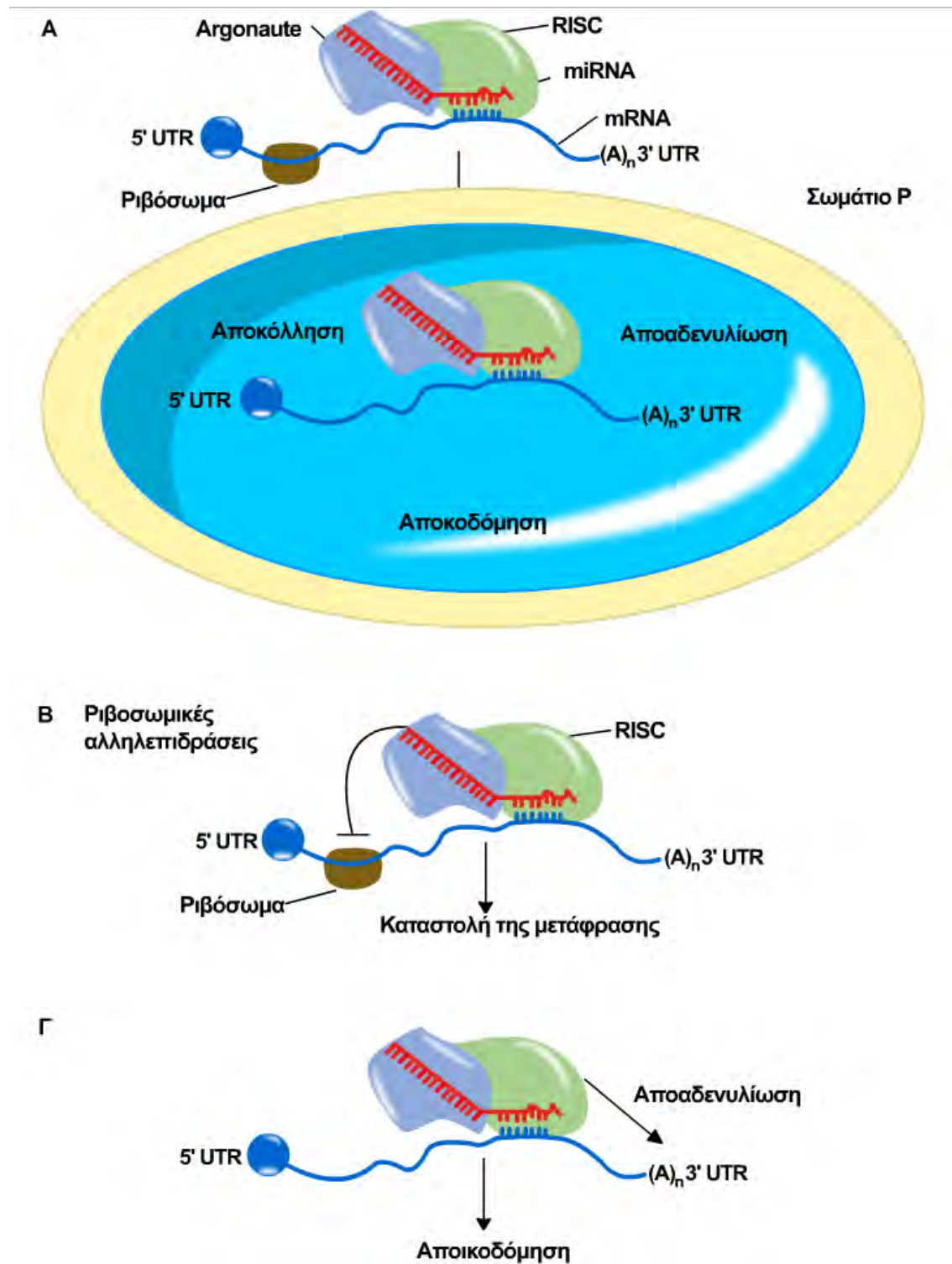
χώρα σε τρία στάδια: έναρξη, επιμήκυνση και τερματισμός, τα οποία απαιτούν συντονισμό πολλών πρωτεϊνικών παραγόντων. Πειράματα που εκτελέστηκαν με κύτταρα θηλαστικών παρείχαν στοιχεία που υποστηρίζουν την αναστολή της μετάλλαξης του miRNA τόσο στα στάδια έναρξης όσο και στην επιμήκυνση [50][53]. Ωστόσο, ο μηχανισμός επιλογής είναι ασαφής. Πρόσφατα δεδομένα υποδεικνύουν ότι ο υποκινητής που χρησιμοποιήθηκε για τη μεταγραφή του mRNA στόχου καθορίζει ποιοι μηχανισμοί καταστολής της μετάφρασης θα χρησιμοποιηθούν. Η μετάφραση μπορεί επίσης να ρυθμίζεται έμμεσα με χωρικό διαχωρισμό των συστατικών, έτσι ώστε το mRNA να απομακρύνεται από το μεταφραστικό μηχανισμό σε κυτταροπλασματικές εστίες γνωστές ως σωμάτια P (P-bodies ή Dcp) [54]. Επιπρόσθετα, το miRNA μπορεί να επιταχύνει την αποαδενυλίωση του mRNA στόχου και την απόσπαση ανεξάρτητα από τη δραστικότητα του τεμαχιστή, επηρεάζοντας, συνεπώς, την αποτελεσματικότητα της έναρξης της μετάφρασης και τη σταθερότητα της μεταγραφής. Το τελευταίο αποτέλεσμα θα οδηγήσει σε αποσύνθεση του mRNA στόχου μέσω του ίδιου εξωσώματος και στην ενζυματική αποικοδόμηση του Xrn1p ως εξαρτώμενη από τον τεμαχιστή σίγαση [44][45]. Τα ανεπίλυτα ζητήματα στην διαδικασία σίγασης του mRNA στόχου υποδεικνύουν τη συμμετοχή άγνωστων εναλλακτικών οδών αποσύνθεσης που μένουν ακόμα να ανακαλυφθούν.

1.2.4.3 P-Σωμάτια (P-Bodies)

Τα σωμάτια P είναι δυναμικές μικρές κυτταροπλασματικές πρωτεΐνες σφαιροειδούς περιοχής που βρίσκονται σε κύτταρα διαφόρων οργανισμών συμπεριλαμβανομένου του ανθρώπου[43]. Υπάρχουν πολλά παρόμοια κυτταροπλασματικά πεδία που βρίσκονται σε κύτταρα, συμπεριλαμβανομένων σωματιδίων που καθορίζουν το στρες, εξωσωμάτων και πολυκυστικών σωμάτων που περιέχουν μερικά από τα ίδια πρωτεϊνικά συστατικά. Ωστόσο, μερικοί δείκτες πρωτεϊνών φαίνεται να είναι ειδικοί για τα σώματα P όπως τα Dcp1, 4E-T, GE-1 / hedls, p54 / RCK και Xrn1 [55]. Τα σώματα P επηρεάζονται από μια ποικιλία κυτταρικών παραγόντων που περιλαμβάνουν το επίπεδο γλυκόζης, την ωσμωτική πίεση και τον πολλαπλασιασμό των κυττάρων [56]. Η συλλογική έρευνα δείχνει ότι τα σώματα P είναι η λειτουργική θέση της αναστρέψιμης καταστολής του mRNA και της αποσύνθεσης του mRNA, τα οποία μεσολαβούν από μια ποικιλία μηχανισμών που συμπεριλαμβάνουν τη διέγερση που προκαλείται από το miRNA [54][57][58]. Τα σώματα P περιέχουν mRNA μαζί με μια ποικιλία ενζύμων και παραγόντων που απαιτούνται για διεργασίες όπως απομάκρυνση mRNA, απελευθέρωση, αποικοδόμηση RNA και καταστολή της μετάφρασης. Περαιτέρω διερεύνηση της σχέσης μεταξύ της σίγασης του miRNA και των σωμάτων P αποκάλυψε ότι οι πρωτεΐνες τελεστών RISC, Ago 1-4 εντοπίζονται σε σώματα P [50][54][57]. Αυτό υποδηλώνει ότι εκτός από το ότι είναι

1. Εισαγωγή

η λειτουργική θέση για την κυτταρική κυκλοφορία του mRNA, τα σωματίδια P είναι επίσης η λειτουργική θέση της διέγερσης γονιδίων που προκαλείται από το miRNA. Ο ρόλος των P σωματιδίων στην σίγαση του miRNA διερευνάται ενεργά τα τελευταία χρόνια.



Εικόνα 5: Τρόποι ελέγχου μετα-μεταγραφικής ρύθμισης από τα miRNAs

1.3 Γενετικοί Πολυμορφισμοί και SNPs

1.3.1 Πολυμορφισμοί

Πολυμορφισμοί DNA είναι οι εναλλακτικές μορφές ενός χρωμοσωμικού τόπου, οι οποίες διαφέρουν ως προς την νουκλεοτιδική αλληλουχία τους[59]. Για την ποικιλομορφία τους συνήθως ευθύνεται το γεγονός έχουν διαφορετικό αριθμό αντιγράφων μίας διαδοχικά επαναλαμβανόμενης νουκλεοτιδικής αλληλουχίας μικρού μήκους. Είναι σημαντικό να διαχωριστούν από τις απλές μεταλλάξεις, καθώς στην περίπτωση του πολυμορφισμού η συχνότητα εμφάνισης τους είναι τουλάχιστον 1% σε έναν πληθυσμό ή από άποψη γονιδιώματος 1 ανά 350 ζεύγη νουκλεοτιδικών βάσεων[60]. Αυτή η συχνότητα εμφάνισης υποδηλώνει ότι ένας πολυμορφισμός εμφανίζεται φυσικά, είτε με ουδέτερο είτε με ευεργετικό αποτέλεσμα. Δεδομένου πως το ανθρώπινο γονιδίωμα έχει μέγεθος περίπου $3 \cdot 10^9$ ζεύγη βάσεων, με την συχνότητα που προαναφέρθηκε, υπολογίζεται πως υπάρχουν περίπου 9 εκατομμύρια πολυμορφικές θέσεις.

Οι πολυμορφισμοί θα μπορούσαν να χωριστούν σε δύο μεγάλες κατηγορίες, τους μονονουκλεοτιδικούς πολυμορφισμούς (Single Nucleotide Polymorphisms, SNPs) και τις ποικίλου αριθμού διαδοχικές επαναλήψεις (Variable Number Tandem Repeats, VNTR). Υποκατηγορία των VNTRs που συναντάται συχνά στον πληθυσμό είναι οι βραχείες διαδοχικές επαναλήψεις (Short Tandem Repeats, STR). Παρακάτω θα αναλυθούν περαιτέρω οι πολυμορφισμοί ενός νουκλεοτιδίου, οι οποίοι και συναντώνται κατά κόρον στα microRNAs τα οποία μελετώνται στην παρούσα εργασία.

Στοιχεία Διαδοχικών Επαναλήψεων

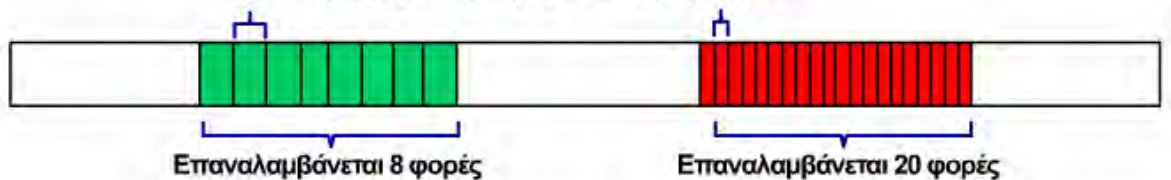
Ποικίλου αριθμού διαδοχικές επαναλήψεις (Variable Number Tandem Repeats, VNTR)

Μέγεθος μονάδας επανάληψης = 100 ζεύγη βάσεων



Βραχείες διαδοχικές επαναλήψεις (Short Tandem Repeats, STR)

Μέγεθος μονάδας επανάληψης = 2 - 6 ζεύγη βάσεων



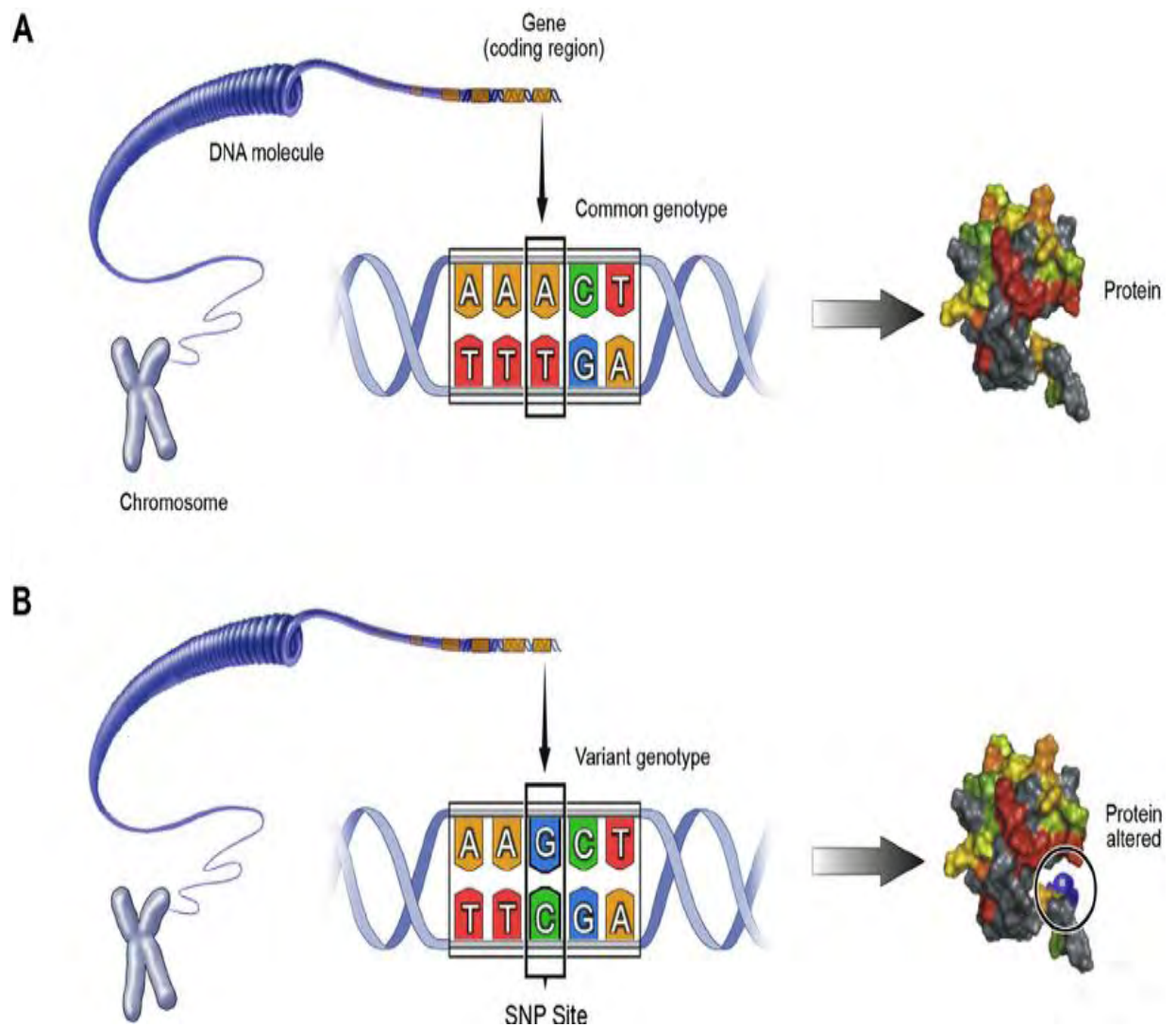
Εικόνα 6: Δύο από τους τρεις πιο συχνούς γενετικούς πολυμορφισμούς

1.3.2 Single Nucleotide Polymorphism SNP

Ένας μονονουκλεοτιδικός πολυμορφισμός είναι μία αλλαγή σε ένα ζεύγος βάσεων σε μία θέση του γονιδιώματος [59]. Οι SNPs αντιπροσωπεύουν την πιο συνηθισμένη μορφή γενετικής διακύμανσης στον άνθρωπο, αντιπροσωπεύοντας το 90-95% της ποικιλομορφίας του ανθρώπινου DNA.

Οι περισσότεροι μονονουκλεοτιδικοί πολυμορφισμοί υπάρχουν σε μη κωδικές περιοχές του γονιδιώματος, ενώ εκείνοι που βρίσκονται σε κωδικές καλούνται κωδικοί SNPs και συχνά συμβολίζονται cSNP (coding SNP). Σύμφωνα με μελέτες συναντώνται περίπου τέσσερις cSNP σε κάθε γονίδιο, οι μισοί μόνο από τους οποίους προκαλούν αλλαγή σε κωδικόνιο του mRNA, ενώ οι υπόλοιπες προκαλούν σιωπηλές μεταλλάξεις. Το εάν θα προκύψει κάποια αλλαγή στον φαινότυπο του οργανισμού εξαρτάται από την αμινοξική αλλαγή που θα προκληθεί στην πρωτεΐνη.

1. Εισαγωγή



Εικόνα 7: Αλλαγή που μπορεί να επιφέρει ένας SNP σε μία πρωτεΐνη.

Πηγή εικόνας : Camp, Kathryn M., and Elaine Trujillo. "Position of the Academy of Nutrition and Dietetics: nutritional genomics." *Journal of the Academy of Nutrition and Dietetics* 114.2 (2014): 299-312.

Καθώς περίπου οι μισοί από τους SNPs που προκαλούν αλλαγές σε πρωτεΐνες φαίνεται να συνδέονται με γενετικές ασθένειες στον άνθρωπο, έχει πλέον επικρατήσει η άποψη ότι δεν αποτελούν τυχαίες μεταλλάξεις, αλλά η ύπαρξη τους σχετίζεται άμεσα με παθολογικές για τον άνθρωπο καταστάσεις, είτε με αρνητική είτε με θετική επίδραση, επηρεάζοντας αλλά όχι αλλάζοντας τη δράση κάποιας σημαντικής πρωτεΐνης.

Εξαιτίας λοιπόν της προφανής αυτής σύνδεσης των SNPs με ασθένειες έχει μελετηθεί εκτεταμένα η χρήση τους ως εργαλεία διάγνωσης, θεραπείας αλλά και σε πληθυσμιακές μελέτες ταυτοποίησης γονιδίων και στη δημιουργία γενετικών χαρτών [61]. Τα βασικά πλεονεκτήματα της χρήσης τους είναι :

1. Εισαγωγή

- Η υψηλή συχνότητα με την οποία βρίσκονται τα SNPs στο γονιδίωμα, η οποία τους δίνει καθοριστική χρησιμότητα για στην ανακάλυψης γονιδίων χαρακτηριστικών σε ασθένειες. Επομένως μπορούν να χρησιμοποιηθούν ως δείκτες για εκτεταμένες και πυκνές μελέτες χαρτογράφησης γονιδίων σε θέσεις κλωνοποίησης ή ακόμα και ως άμεσος υπαίτιος για λειτουργικές μεταλλάξεις σχετιζόμενες με ασθένειες [62].
- Η θέση τους, καθώς τα SNPs βρίσκονται σε όλο το γονιδίωμα, όπως σε εξόνια, ιντρόνια, διαγονιδιακές περιοχές, σε προαγωγούς ή ενισχυτές και άλλα. Επομένως, είναι πιθανότερο να αποδώσουν ένα λειτουργικό ή σχετικό αλληλόμορφο από άλλα είδη πολυμορφισμού. Αυτό που είναι εξαιρετικά ενδιαφέρον από αυτή την άποψη είναι η φύση του αποτελέσματος που μπορεί να έχει μια απλή υποκατάσταση ζεύγους βάσεων σε ένα γνώρισμα ή μια ασθένεια. Έτσι, ένα SNP στην κωδική περιοχή μπορεί να επηρεάσει άμεσα μια σχετική πρωτεΐνη, ένα ιντρονικό SNP μπορεί να επηρεάσει τη συρραφή, ένα SNP σε έναν προαγωγό μπορεί να επηρεάσει την γονιδιακή έκφραση και ούτω καθεξής. Ο βαθμός στον οποίο κάθε είδος SNP επηρεάζει τη φαινοτυπική έκφραση είναι πιθανό να λάβει μεγάλη προσοχή καθώς όλο και περισσότερα SNP αναγνωρίζονται και μελετώνται [63].
- Επειδή τα νέα αλληλόμορφα SNPs προκύπτουν ως μεταλλάξεις σε διαφορετικούς τόπους και σε διαφορετικά χρονικά σημεία και επειδή εμφανίζονται με τόσο μεγάλη αφθονία στο γονιδίωμα, ομάδες γειτονικών SNPs μπορεί να έχουν αλληλόμορφα που παρουσιάζουν διακριτά μοτίβα ανισορροπίας σύνδεσης, δηλαδή η παρουσία ενός αλληλόμορφου σε ένα χρωμόσωμα μπορεί να υποδηλώνει μεγάλη πιθανότητα ότι ένα συγκεκριμένο αλληλόμορφο θα υπάρχει σε μια γειτονική περιοχή στο ίδιο χρωμόσωμα. Το φαινόμενο αυτό δημιουργεί μια απλοτυπική ποικιλομορφία που μπορεί να αξιοποιηθεί στη γενετική σύνδεση και στις άμεσες μελέτες σύνδεσης του πολυμορφισμού με ασθένειες.
- Τα SNPs είναι εύκολα στη χρήση τους σε γονοτυπικές αναλύσεις, αφού έχουν μια απλή δομή ως αλλαγές βάσεων συγκριτικά με άλλους πολυμορφισμούς, επομένως μέσω μικροσυστοιχιών και άλλων τεχνολογιών μπορούν γονοτυπηθούν ταχέως και με αποτελεσματικότητα εκατοντάδων ατόμων για εκατοντάδες SNPs [64].
- Επειδή τα SNPs τυπικά έχουν μόνο δύο αλληλόμορφα, θα έχουν συχνότητες αλληλόμορφων που θα εκφράζονται ως συνάρτηση της δυναμικής διαφορετικών πληθυσμών. Αυτό μπορεί να δημιουργήσει διαφορές συχνότητας αλληλόμορφων, επάγοντας την αξιοποίηση τους σε πολλές μελέτες που βασίζονται στον πληθυσμό, όπως η χαρτογράφηση των προσμείξεων [65].
- Τα SNPs γενικά παρουσιάζουν λιγότερες μεταλλάξεις από άλλες μορφές πολυμορφισμού, καθώς ένα νέο SNP συνήθως οφείλεται σε κάποιο σφάλμα

1. Εισαγωγή

κατά την αντιγραφή [59][60]. Αυτή η σταθερότητα προσφέρει έναν πιο αξιόπιστο τρόπο για να εκτίμησης των απλότυπων, τις συσχετίσεις θέσεων και τα φαινόμενα συνδιαχωρισμού, δεδομένου ότι οι συσχετισμοί δεν θα συγχέονται από αλληλόμορφα που έχουν μεταλλαχθεί σε διαφορετικές μορφές κατά τη διάρκεια της μετάδοσης αλληλόμορφων από γενιά σε γενιά.

- Τέλος δεδομένου ότι τα SNPs μπορούν να εμφανιστούν πολύ κοντά το ένα στο άλλο, η μελέτη των μοτίβων απλοτύπων, αποκαλύπτει θέσεις επαναλαμβανόμενης μετάλλαξης, μετατροπής γονιδίων ή ανασυνδυασμού. Οπότε μπορεί να αξιολογηθεί μια περιοχή γονιδιώματος για σύνδεση ή συσχετισμό με ένα συγκεκριμένο γνώρισμα ή ασθένεια.

1.3.3 SNPs και microRNAs

Οι μονονουκλεοτιδικοί πολυμορφισμοί σε microRNAs (miRSNPs) μπορεί να επηρεάσουν την έκφραση και τη λειτουργία του miRNA, το οποίο με τη σειρά του μπορεί να επηρεάσει την έκφραση αρκετών γονιδίων. Τα τελευταία χρόνια αρκετές μελέτες έδειξαν τη σύνδεση των miRSNPs με διάφορες μορφές καρκίνου και άλλες ασθένειες. Τα miRSNPs μπορούν να ταξινομηθούν σε τρεις κατηγορίες [66]: (α) SNPs σε γονίδια που εμπλέκονται στη βιογένεση και επεξεργασία του miRNA, (β) SNPs σε γονίδια miRNA και (γ) SNPs σε θέσεις δέσμευσης miRNA σε γονίδια-στόχους. Η συνάφεια τους με την εκάστοτε μελετώμενη ασθένεια μπορεί να εξεταστεί όσον αφορά τον κίνδυνο εμφάνισης, την πρόγνωση της εξέλιξης της νόσου και της επιβίωσης, καθώς και την ανταπόκριση στην θεραπεία.

1.3.3.1 SNPs σε γονίδια που εμπλέκονται στη βιογένεση και επεξεργασία του miRNA

Οι πολυμορφισμοί που επηρεάζουν την έκφραση ή τη λειτουργία των πρωτεϊνών που εμπλέκονται στη βιογένεση των miRNAs μπορεί να παρεμβαίνουν στα δίκτυα ρύθμισης miRNA στο κύτταρο. Αυτή η κατηγορία miRSNP θα είχε πιθανότατα την ευρύτερη επίδραση στη ρύθμιση με τη μεσολάβηση miRNA, καθώς θα μπορούσε να επηρεάσει την παγκόσμια βιογένεση miRNA και να απορυθμίσει ολόκληρο το microRNA του κυττάρου, επηρεάζοντας την έκφραση όλων των γονιδίων που ρυθμίζονται από miRNAs.

1. Εισαγωγή

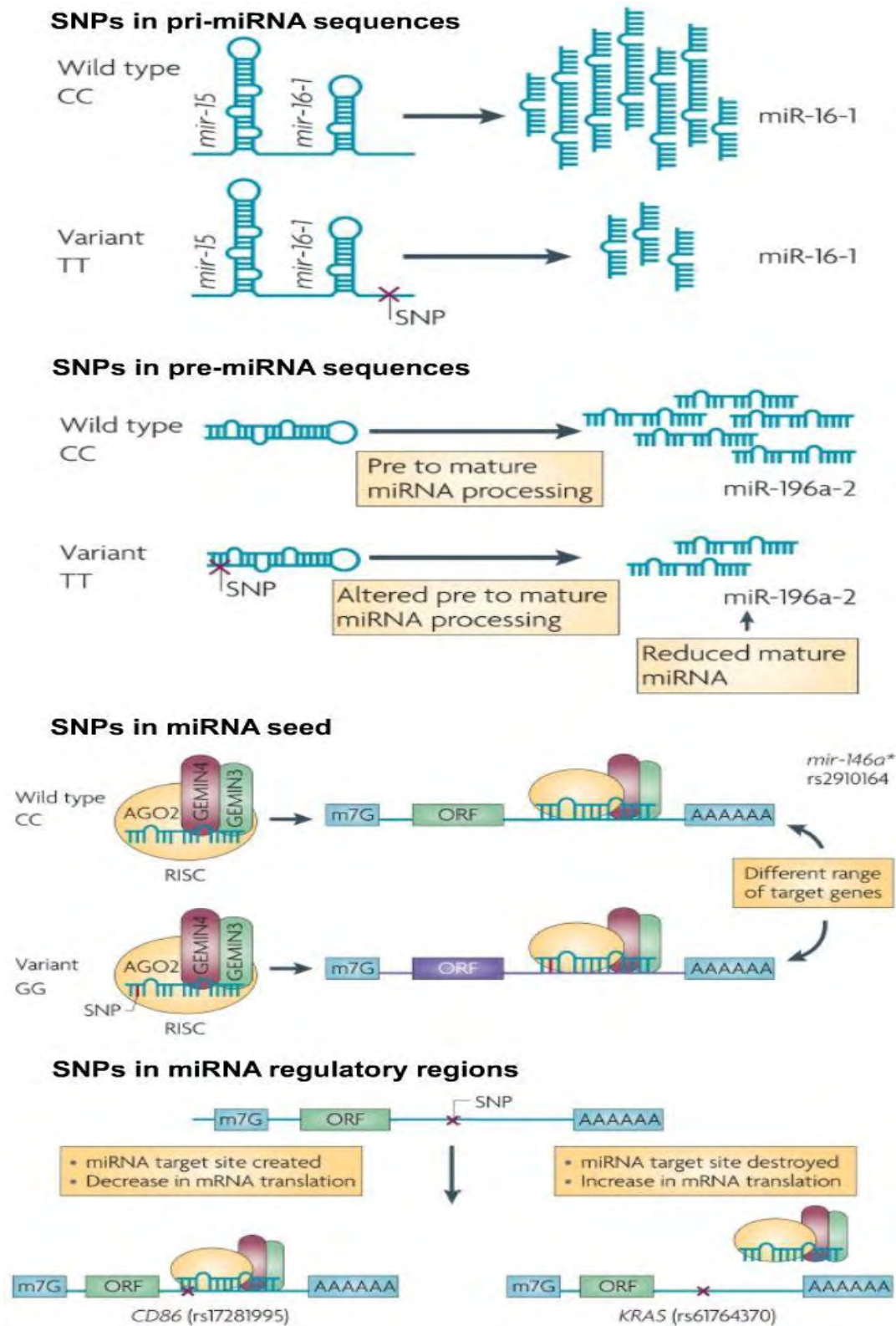
1.3.3.2 SNPs σε γονίδια miRNA

Ο πολυμορφισμός στα pri- και pre-miRNAs μπορεί να επηρεάσει την επεξεργασία τους, ενώ οι πολυμορφισμοί σε ώριμα miRNAs μπορεί να επηρεάσουν τη δέσμευσή τους σε mRNA στόχους. Επίσης SNPs σε υποκινητές γονιδίων miRNA ή γονιδίων ξενιστή miRNA μπορεί να έχουν αντίκτυπο στην έκφραση των miRNAs. Αυτή η κατηγορία miRSNP θα μπορούσε να έχει ευρύτατο αντίκτυπο για το κύτταρο, καθώς οποιοδήποτε miRNA πιθανόν να μπορεί να ρυθμίζει πολλά γονίδια που εμπλέκονται σε διάφορες διεργασίες.

1.3.3.3 SNPs σε θέσεις δέσμευσης miRNA σε γονίδια-στόχους

Οι πολυμορφισμοί στις θέσεις πρόσδεσης του miRNA στα γονίδια-στόχους ενδέχεται να μεταβάλλουν την ισχύ των αλληλεπιδράσεων του miRNA-mRNA, απορυθμίζοντας έτσι τα επίπεδα της πρωτεΐνης. Αυτή η κατηγορία miRSNPs πιθανότατα θα έχει το πιο περιορισμένο αντίκτυπο για το κύτταρο, επηρεάζοντας μόνο την έκφραση του γονιδίου που φιλοξενεί το SNP. Ωστόσο, η έμμεση επίδραση των SNPs στις θέσεις δέσμευσης του miRNA μπορεί να είναι πολύ ευρύτερη, καθώς θα επηρεάσει την ομάδα των miRNAs που είναι διαθέσιμα για άλλα αντίγραφα τα οποία επίσης ρυθμίζονται από ένα δεδομένο miRNA. Αυτό θα μπορούσε να οδηγήσει σε απορύθμιση των παγκόσμιων δικτύων miRNA.

1. Εισαγωγή



Εικόνα 8: Θέσεις εμφάνισης SNP σε microRNA

Πηγή εικόνας: Ryan, Bríd M., Ana I. Robles, and Curtis C. Harris. "Genetic variation in microRNA networks: the implications for cancer research." *Nature reviews. Cancer* 10.6 (2010): 389.

1.4 Μετα-ανάλυση και Συστηματική Ανασκόπηση

1.4.1 Μετα-ανάλυση

Η στατιστική ανάλυση η οποία συνδέει επιστημονικές μελέτες, που έγιναν στο παρελθόν, για την παροχή ακριβέστερης επίδρασης των αποτελεσμάτων, ορίζεται ως μετα-ανάλυση.

Η βασική της αρχή είναι ότι υπάρχει ένα κοινό αποτέλεσμα (κοινή αλήθεια) σε θεματικά όμοιες μελέτες, το οποίο έχει υπολογιστεί με ένα διάστημα σφάλματος. Ο στόχος είναι να χρησιμοποιηθούν προσεγγίσεις από τα στατιστικά στοιχεία για να προκύψει μια συγκεντρωτική εκτίμηση πλησιέστερη στο άγνωστο κοινό αποτέλεσμα με βάση τον τρόπο με τον οποίο γίνεται αντιληπτό αυτό το σφάλμα.

Πρόκειται για ένα ποσοτικό, επίσημο, επιδημιολογικό σχεδιασμό μελέτης που χρησιμοποιείται για τη συστηματική αξιολόγηση προηγούμενων ερευνητικών μελετών προκειμένου να εξαχθούν συμπεράσματα σχετικά με το εν λόγω σύνολο μελετών. Τα αποτελέσματα μιας μετα-ανάλυσης μπορεί να περιλαμβάνουν μια ακριβέστερη εκτίμηση της επίδρασης της θεραπείας ή του παράγοντα επικινδυνότητας για ασθένεια ή άλλων αποτελεσμάτων από οποιαδήποτε μεμονωμένη μελέτη που συνεισφέρει στην συγκεντρωτική ανάλυση. Η εξέταση της μεταβλητότητας ή της ετερογένειας στα αποτελέσματα της μελέτης είναι επίσης ένα κρίσιμο αποτέλεσμα [67].

Κατά την πραγματοποίηση μιας μετα-ανάλυσης, ο ερευνητής πρέπει να κάνει επιλογές που μπορούν να επηρεάσουν τα αποτελέσματα, συμπεριλαμβανομένης της λήψης αποφάσεων για την αναζήτηση μελετών, την επιλογή μελετών βάσει αντικειμενικών κριτηρίων, την αντιμετώπιση ελλιπών δεδομένων και την ανάλυση των δεδομένων.

1.4.1.1 Ιστορική Αναδρομή

Το 1904 ο Βρετανός στατιστικός Karl Pearson φαίνεται να ήταν ο πρώτος που εφάρμοσε μεθόδους για να συνδυάσει παρατηρήσεις από διαφορετικές κλινικές μελέτες [68]. Ζητήθηκε να αναλύσει δεδομένα που συγκρίνουν τη μόλυνση και τη θνησιμότητα μεταξύ στρατιωτών που είχαν προσφερθεί εθελοντικά για εμβολιασμό

1. Εισαγωγή

κατά του τυφοειδούς πυρετού σε διάφορα μέρη της Βρετανικής Αυτοκρατορίας με εκείνους άλλων στρατιωτών που δεν είχαν προσφερθεί εθελοντικά.

Μέχρι τα μέσα του 20ου αιώνα, ωστόσο, ο τεράστιος όγκος των ερευνητικών αναφορών ανάγκασε τους ερευνητές να εξετάσουν πώς να αναπτύξουν και να εφαρμόσουν μεθόδους για την εξαγωγή των αποτελεσμάτων. Το 1940, για παράδειγμα, η ποσοτική σύνθεση χρησιμοποιήθηκε σε μια ανάλυση των αποτελεσμάτων 60 ετών έρευνας από ψυχολόγους σχετικά με την υπερβατική αντίληψη.

Οι Αμερικανοί επιστήμονες κοινωνικών επιστημών και στατιστικοί, είχαν κατακλυστεί από έρευνες οπότε άρχισαν να αναπτύσσουν και να εφαρμόζουν μεθόδους για την ποσοτική σύνθεση των αποτελεσμάτων ξεχωριστών αλλά παρόμοιων μελετών. Το 1976, ένας από αυτούς, ο Gene Glass, χρησιμοποίησε τον όρο «μετα-ανάλυση» για να αναφερθεί «στη στατιστική ανάλυση μιας μεγάλης συλλογής αποτελεσμάτων ανάλυσης από μεμονωμένες μελέτες με σκοπό την ενσωμάτωση των ευρημάτων» [69].

Παρόλο που υπάρχουν περιστασιακά προηγούμενα παραδείγματα μετα-ανάλυσης που χρησιμοποιούνται από ιατρικούς ερευνητές, μετά τη δεκαετία του 1970 άρχισαν να εμφανίζονται σε μεγαλύτερο αριθμό. Ιδιαίτερα σημαντικό ρόλο έπαιξε η πρώτη τυχαίοποιημένη δοκιμή που διενεργήθηκε από τον Peter Elwood, τον Archie Cochrane και τους συναδέλφους τους για να αξιολογήσουν εάν η ασπιρίνη μείωσε την επανεμφάνιση εμφράγματος [70]. Τα αποτελέσματα υποδηλώνουν ευεργετική επίδραση, αλλά όχι στατιστικά σημαντική, έτσι όπως αναφέρθηκαν επιπρόσθετες μελέτες, οι Elwood και Cochrane συγκέντρωσαν και συνέθεσαν τα αποτελέσματά τους χρησιμοποιώντας μετα-ανάλυση. Αυτό δεν άφηνε την αμφιβολία ότι η ασπιρίνη θα μπορούσε να μειώσει τον κίνδυνο επανεμφάνισης εμφράγματος και τα αποτελέσματα δημοσιεύθηκαν το 1980 σε μία ανώνυμη συντακτική ομάδα του Lancet. Η δημοσίευση γράφτηκε από τον βρετανό ιατρικό στατιστικολόγο Richard Peto [71]. Με βάση τις προηγούμενες εργασίες, ο Peto και οι συνάδελφοί του συνέχισαν να παρέχουν ένα λεπτομερές παράδειγμα για να ενθαρρύνουν τους κλινικούς ιατρούς να επανεξετάζουν συστηματικά τυχαίοποιημένες δοκιμές και να συνδυάζουν εκτιμήσεις των επιδράσεων θεραπειών που θεωρούνται ίδιες, βάσει ενημερωμένης κλινικής κρίσης

1.4.1.2 Η μεθοδολογία της μετα-ανάλυσης

Κάθε έρευνα, σε βασικό επίπεδο, διεξάγεται ακολουθώντας τρία βήματα:

- διατύπωση της ερώτησης προς έρευνα

1. Εισαγωγή

- συλλογή και ανάλυση των δεδομένων
- αναφορά των αποτελεσμάτων

Όσον αφορά τη μετα-ανάλυση, σύμφωνα με τους Egger, Smith και Philips (1997) [72], για να έχουμε ένα επιτυχημένο αποτέλεσμα, πρέπει να λάβουμε υπ' όψιν μερικά επιπλέον βήματα ως μέρος της διαδικασίας.

Βήμα 1: Ορίζονται τα κριτήρια επιλεξιμότητας για τα δεδομένα που θα περιληφθούν. Τα κριτήρια θα καθορίσουν πόσο συμβατά θα πρέπει να είναι τα άρθρα που θα επιλεγούν μεταξύ τους για να υπάρχει αξιοκρατία στο αποτέλεσμα. Αυτό θα βασίζεται στην ποιότητα και τον τύπο των δοκιμών, στους ασθενείς, στα αποτελέσματα και στη διάρκεια παρακολούθησης. Ωστόσο, η σωστή εκτίμηση της ποιότητας των δοκιμών, εξαρτάται σε μεγάλο βαθμό από την κρίση του ερευνητή, και επομένως μια ανάλυση ευαισθησίας πρέπει να εκτελείται εκ των υστέρων.

Βήμα 2: Καθορίζεται στρατηγική για τον εντοπισμό των σχετικών μελετών. Είναι σημαντικό, σε συνδυασμό με τα κριτήρια επιλεξιμότητας, να περιγραφθεί ποια θα είναι η καλύτερη στρατηγική, σύμφωνα με την ερευνητική ερώτηση και τη σημασία της μελέτης, για να επιλεγούν οι μελέτες που θα συμπεριληφθούν στην ανάλυση. Ιδιαίτερα περίπτωση είναι η συμπερίληψη μη δημοσιευμένων μελετών εξαιτίας φαινόμενων μεροληψίας που συσχετίζονται συχνά με δημοσιευμένες μελέτες.

Βήμα 3: Δημιουργία μια τυποποιημένης φόρμας για τη συλλογή δεδομένων. Κατά την εξαγωγή των δεδομένων, η πιο αξιόπιστη μέθοδος είναι να χρησιμοποιηθούν δύο ανεξάρτητοι παρατηρητές. Για αυτό το στάδιο, υπάρχουν κλίμακες που έχουν σχεδιαστεί για να εκτιμούν την ποιότητα, για παράδειγμα, κλινικών δοκιμών. Επιπλέον, τα άτομα που θα είναι υπεύθυνα για την αξιολόγηση θα πρέπει να μην έχουν πρόσβαση σε παράγοντες που θα μπορούσαν να επηρεάσουν την αξιολόγησή τους, και συγκεκριμένα: τους συγγραφείς και τα ιδρύματά τους, τα ονόματα περιοδικών, πηγές χρηματοδότησης και λοιπές γνώμες.

Βήμα 4: Τυποποιούνται τα μεμονωμένα αποτελέσματα για τη σύγκριση μεταξύ των μελετών. Αφού συλλεχθούν τα δεδομένα, για να συγκριθεί τα αποτελέσματα, θα πρέπει να τα τυποποιηθεί σε κάτι ομοιογενές. Για συνεχή αποτελέσματα, θα πρέπει να βρεθεί η μέση διαφορά μεταξύ θεραπείας και ομάδας ελέγχου, σε μονάδες τυπικής απόκλισης. Για δυαδικά αποτελέσματα, θα πρέπει να ληφθεί υπόψη οι πιθανότητες ή οι σχετικοί κίνδυνοι.

Βήμα 5: Υπολογίζεται το συνολικό αποτέλεσμα συνδυάζοντας τα δεδομένα. Σύμφωνα με το προηγούμενο βήμα, φαίνεται από μεθοδολογική και στατιστική άποψη ότι οι απλοί αριθμητικοί μέσοι όροι δεν αποτελούν αξιόπιστο τρόπο σύγκρισης των αποτελεσμάτων. Για παράδειγμα, τα διαφορετικά μεγέθη δειγμάτων έχουν διαφορετική στατιστική ισχύ και, ως εκ τούτου, πρέπει να δίδονται

1. Εισαγωγή

διαφορετικά βάρη. Επομένως, πρέπει να χρησιμοποιείται ένας σταθμισμένος μέσος όρος των αποτελεσμάτων, στον οποίο οι μεγαλύτερες δοκιμές έχουν μεγαλύτερη επιρροή από τις μικρότερες. Οι στατιστικές τεχνικές για την αντιμετώπιση αυτού του προβλήματος μπορούν να ταξινομηθούν σε δύο μοντέλα: μοντέλο "σταθερών αποτελεσμάτων" και μοντέλο "τυχαίων αποτελεσμάτων". Η επιλογή μεταξύ του ενός ή του άλλου θα εξαρτηθεί από τον τρόπο αντιμετώπισης της μεταβλητότητας των αποτελεσμάτων μεταξύ των μελετών που έχουν επιλεγεί. Σύμφωνα με το πρώτο μοντέλο η τυχαία διακύμανση είναι η μόνη αιτία για αυτή τη μεταβλητότητα, δηλαδή το μέγεθος των μελετών δεν έχει σχέση με τον τύπο των αποτελεσμάτων που δίνουν. Το δεύτερο μοντέλο "τυχαίων αποτελεσμάτων" θεωρεί ως πηγές διαφοροποίησης διαφορετικές υποκείμενες επιδράσεις σε κάθε μελέτη, οδηγώντας σε μεγαλύτερα διαστήματα εμπιστοσύνης από το μοντέλο σταθερών αποτελεσμάτων.

Βήμα 6: Τέλος η ανάλυση και η ερμηνεία των αποτελεσμάτων. Υπάρχουν διάφορα προγράμματα στατιστικής διαθέσιμα για τον υπολογισμό των μεγεθών αποτελεσμάτων (effect size) για μετα-αναλύσεις. Τα μεγέθη αποτελεσμάτων δηλώνονται μαζί με εύρος διαστήματος εμπιστοσύνης π.χ. (CI) 95% και παρουσιάζονται τόσο σε ποσοτική μορφή όσο και σε γραφική. Παράδειγμα γραφικής απεικόνισης αποτελεσμάτων είναι τα δενδρογράμματα παρουσιάζουν οπτικά κάθε δοκιμή ως οριζόντιο σχήμα διαμαντιού με το μέσο του να αντιπροσωπεύει το μέγεθος αποτελέσματος. Αυτά τα διαμάντια εμφανίζονται σε γράφημα με κεντρική γραμμή που αντιπροσωπεύει το σημείο μηδέν. Συχνά η αριστερή πλευρά του γραφήματος (<μηδέν) αντιπροσωπεύει την πλευρά που ευνοεί τη θεραπεία, ενώ η δεξιά πλευρά (> μηδέν) αντιπροσωπεύει την πλευρά που ευνοεί την συνθήκη ελέγχου. Στο κάτω μέρος του γραφήματος υπάρχει ένα συνολικό μέγεθος αποτελέσματος ή διαμάντι που αντιπροσωπεύει όλες τις μεμονωμένες μελέτες που συγκεντρώνονται μαζί. Στην ιδανική περίπτωση, όλο αυτό το διαμάντι (το μέγεθος του αποτελέσματος και τα δύο άκρα του διαστήματος εμπιστοσύνης) πέφτει κάτω από το μηδέν, υποδεικνύοντας ότι η παρέμβαση ευνοείται στον έλεγχο. Επιπλέον, τα περισσότερα προγράμματα υπολογίζουν μια τιμή ετερογένειας για να υποδείξουν αν οι μεμονωμένες μελέτες είναι αρκετά παρόμοιες για να συγκριθούν. Σε αυτή την περίπτωση, είναι προτιμότερο να υπάρχουν μη σημαντικά αποτελέσματα για ετερογένεια. Είναι, βέβαια, δυνατή η συγκέντρωση μελετών ακόμη κι όταν υπάρχει σημαντική ετερογένεια, παρόλα αυτά τα αποτελέσματα αυτά θα πρέπει να ερμηνεύονται με προσοχή ή θα πρέπει να διερευνηθούν οι λόγοι για την ύπαρξη ετερογένειας.

1.4.1.3 Μεμονωμένα ή συγκεντρωτικά δεδομένα

1. Εισαγωγή

Η πλειονότητα των μετα-αναλύσεων βασίζεται σε μια σειρά μελετών για την παραγωγή ενός σημείου εκτίμησης ενός αποτελέσματος και των μετρήσεων της ακρίβειας αυτής της εκτίμησης. Ωστόσο, έχουν αναπτυχθεί μέθοδοι για τις μετα-αναλύσεις που πρέπει να διεξαχθούν με δεδομένα που λαμβάνονται από τις αρχικές δοκιμές [72][73]. Αυτή η προσέγγιση μπορεί να θεωρηθεί ως "χρυσό πρότυπο" στη μετα-ανάλυση επειδή προσφέρει πλεονεκτήματα σε σχέση με τις αναλύσεις χρησιμοποιώντας συγκεντρωτικά δεδομένα, συμπεριλαμβανομένης της μεγαλύτερης ικανότητας επικύρωσης της ποιότητας των δεδομένων και διεξαγωγής κατάλληλης στατιστικής ανάλυσης. Περαιτέρω, είναι ευκολότερο να διερευνηθούν οι διαφορές στην επίδραση σε όλες τις υποομάδες στον πληθυσμό της μελέτης παρά σε συγκεντρωτικά δεδομένα. Η χρήση τυποποιημένων πληροφοριών σε ατομικό επίπεδο μπορεί να βοηθήσει στην αποφυγή των προβλημάτων που συναντώνται σε μετα-αναλύσεις προγνωστικών παραγόντων [74][75]. Είναι ο καλύτερος τρόπος για να αποκτήσουμε μια πιο σφαιρική εικόνα της φυσικής ιστορίας και των προγνωστικών του κινδύνου για σημαντικά αποτελέσματα. Η προσέγγιση αυτή βασίζεται στη συνεργασία μεταξύ των ερευνητών που πραγματοποίησαν τις σχετικές μελέτες [76][77]. Οι ερευνητές που γνωρίζουν τη δυνατότητα να συνεισφέρουν ή να διεξάγουν αυτές τις μελέτες θα παράσχουν και θα αποκτήσουν πρόσθετα οφέλη μέσω της προσεκτικής συντήρησης των πρωτότυπων βάσεων δεδομένων και της διάθεσής διεξαγωγής μελλοντικών μελετών.

1.4.1.4 Αναζήτηση βιβλιογραφίας

Μια υγιής μετα-ανάλυση χαρακτηρίζεται από μια εμπεριστατωμένη και πειθαρχημένη αναζήτηση βιβλιογραφίας. Τα θεμέλια για μία τέτοια διερεύνηση είναι ένας σαφής ορισμός των υποθέσεων που πρέπει να διερευνηθούν. Σύμφωνα με τις οδηγίες του PRISMA, πρέπει να παρέχεται ρητή δήλωση των ερωτήσεων με αναφορά σε συμμετέχοντες, παρεμβάσεις, συγκρίσεις, αποτελέσματα και σχεδιασμό μελέτης (PICOS – Participants, Interventions, Comparisons, Outcomes, Study Design) [78][79]. Είναι σημαντικό να λαμβάνονται όλες οι σχετικές μελέτες, επειδή η απώλεια των μελετών μπορεί να οδηγήσει σε προκατάληψη στη μελέτη. Τυπικά, τα δημοσιευμένα έγγραφα και περιλήψεις αναγνωρίζονται από μια ηλεκτρονική βιβλιογραφική αναζήτηση ηλεκτρονικών βάσεων δεδομένων όπως το PubMed (www.ncbi.nlm.nih.gov/entrez/query.fcgi), το ScienceDirect (www.sciencedirect.com), το Scirus (www.scirus.com/srsapp), το ISI Web of Knowledge (www.isiwebofknowledge.com), το Google Scholar (<http://scholar.google.com>) και το CENTRAL (Κεντρικό Μητρώο Ελεγχόμενων Δοκιμών Cochrane) και άλλες. Οι οδηγίες του PRISMA συνιστούν να παρουσιαστεί μια πλήρης ηλεκτρονική στρατηγική αναζήτησης για τουλάχιστον μία σημαντική βάση δεδομένων [80]. Οι αναζητήσεις βάσεων δεδομένων θα πρέπει να αυξηθούν με χειρωνακτικές αναζητήσεις πόρων βιβλιοθήκης για σχετικά έγγραφα, βιβλία,

1. Εισαγωγή

περιλήψεις και εργασίες συνεδρίων. Ο διασταυρούμενος έλεγχος των αναφορών, οι αναφορές σε έγγραφα ανασκόπησης και η επικοινωνία με επιστήμονες που δουλεύουν στον σχετικό τομέα είναι σημαντικές μέθοδοι που χρησιμοποιούνται για την παροχή ολοκληρωμένης αναζήτησης. Προφανώς δεν είναι εφικτό να βρεθεί απολύτως κάθε σχετική μελέτη για ένα θέμα, καθώς πολλές μελέτες ενδέχεται να μην δημοσιεύονται και αυτές που δημοσιεύονται, ενδέχεται να μην είναι καταταγμένες σε βάσεις δεδομένων που μπορούν να αναζητηθούν από υπολογιστή. Χρήσιμες πηγές για αδημοσίευτες δοκιμές είναι τα μητρώα κλινικών δοκιμών, όπως η ιστοσελίδα της ClinicalTrials.gov της Εθνικής Βιβλιοθήκης Ιατρικής. Οι ανασκοπήσεις θα πρέπει να είναι όσο το δυνατόν πιο αποτελεσματικές. Δηλαδή να έχουμε στη διάθεσή μας όσο το δυνατόν περισσότερες μελέτες, με ελάχιστη προκατάληψη.

1.4.1.5 Στατιστική ανάλυση

Τα πιο συνηθισμένα μέτρα επίδρασης που χρησιμοποιούνται για διχοτομημένα δεδομένα είναι ο σχετικός κίνδυνος (risk ratio ή relative risk) και ο λόγος συμπληρωματικών πιθανοτήτων (odds ratio). Η κυρίαρχη μέθοδος που χρησιμοποιείται για συνεχή δεδομένα είναι η εκτίμηση της τυποποιημένης μέσης διαφοράς (Standardized Mean Difference). Οι μέθοδοι που χρησιμοποιούνται στη μετα-ανάλυση ευρημάτων από πειράματα είναι σχετικά ειδικές για τη μετα-ανάλυση και περιλαμβάνουν ανάλυση ετερογένειας, ανάλυση ευαισθησίας και αξιολόγηση της μεροληψίας δημοσίευσης.

Όλες οι χρησιμοποιούμενες μέθοδοι πρέπει να επιτρέπουν τη στάθμιση των μελετών. Η έννοια της στάθμισης εκφράζουν την αξία των στοιχείων μιας συγκεκριμένης μελέτης. Συνήθως, οι μελέτες σταθμίζονται σύμφωνα με το αντίστροφο της διακύμανσής τους. Συνεπώς είναι σημαντικό να κατανοείται πως οι μικρότερες μελέτες, συνήθως, συνεισφέρουν λιγότερο στις εκτιμήσεις του συνολικού αποτελέσματος. Ωστόσο, καλά διεξαγόμενες μελέτες με αυστηρό έλεγχο της διακύμανσης των μετρήσεων και των πηγών πληροφοριών, συνεισφέρουν περισσότερο στις εκτιμήσεις της συνολικής επίδρασης από μια μελέτη ίδιου μεγέθους λιγότερο καλά διεξαγόμενη.

Μια από τις σημαντικότερες αποφάσεις που πρέπει να λαμβάνονται κατά τη διεξαγωγή μιας μετα-ανάλυσης είναι εάν θα χρησιμοποιηθεί ένα μοντέλο σταθερών αποτελεσμάτων ή τυχαίων αποτελεσμάτων. Το μοντέλο σταθερών αποτελεσμάτων παρέχει ένα σταθμισμένο μέσο όρο μιας σειράς εκτιμήσεων μελέτης. Το αντίστροφο της διακύμανσης των εκτιμήσεων χρησιμοποιείται συνήθως ως το βάρος μελέτης, επομένως, μεγαλύτερες μελέτες τείνουν να συμβάλλουν περισσότερο από τις μικρότερες μελέτες στον σταθμισμένο μέσο όρο. Συνεπώς,

1. Εισαγωγή

όταν μελέτες στο πλαίσιο μιας μετα-ανάλυσης επισκιάζονται από μια πολύ μεγάλη μελέτη, τα ευρήματα από μικρότερες μελέτες πρακτικώς αγνοούνται. Το μοντέλο σταθερών αποτελεσμάτων υποθέτει ότι όλες οι συμπεριλαμβανόμενες μελέτες διερευνούν τον ίδιο πληθυσμό, χρησιμοποιούν τους ίδιους ορισμούς μεταβλητών και αποτελεσμάτων. Αυτή η υπόθεση είναι συνήθως μη ρεαλιστική καθώς η έρευνα είναι συχνά επιρρεπής σε διάφορες πηγές ετερογένειας, για παράδειγμα, αποτελέσματα θεραπείας μπορεί να διαφέρουν ανάλογα με την περιοχή, τα επίπεδα δοσολογίας, τις συνθήκες μελέτης.

Το άλλο μοντέλο που χρησιμοποιείται για τη σύνθεση ετερογενούς έρευνας είναι το μοντέλο τυχαίων αποτελεσμάτων της μετα-ανάλυσης. Αυτός είναι απλά ο σταθμισμένος μέσος όρος των μεγεθών αποτελεσμάτων μιας ομάδας μελετών. Το βάρος που εφαρμόζεται επιτυγχάνεται αρχικά μέσω την αντίστροφη στάθμιση διακύμανσης και την συνέχεια μέσω αποσύνδεσης αυτής της αντίστροφης στάθμισης διακύμανσης εφαρμόζοντας μια συνιστώσα μεταβλητότητας τυχαίων επιδράσεων (REVC – Random Effect Variance Component) που προέρχεται από την έκταση της μεταβλητότητας των μεγεθών των επιπτώσεων των υποκειμένων μελετών. Αυτό σημαίνει ότι όσο μεγαλύτερη είναι αυτή η μεταβλητότητα των μεγεθών των αποτελεσμάτων (ετερογένεια), τόσο λιγότερο μπορούν να αποδοθούν βάρη και αυτό μπορεί να φτάσει σε ένα σημείο όταν το αποτέλεσμα της μετα-ανάλυσης των τυχαίων αποτελεσμάτων γίνεται απλώς το μη σταθμισμένο μέσο μέγεθος αποτελεσμάτων στις μελέτες. Από την άλλη, όταν όλα τα μεγέθη των αποτελεσμάτων είναι παρόμοια (ή η μεταβλητότητα δεν υπερβαίνει το σφάλμα δειγματοληψίας), δεν εφαρμόζεται REVC και η μετα-ανάλυση των τυχαίων αποτελεσμάτων αποδίδει απλά μια μετα-ανάλυση σταθερού αποτελέσματος.

1.4.1.6 Ετερογένεια

Το όφελος από τη διεξαγωγή της μετα-ανάλυσης είναι να εξεταστούν, αν υπάρχουν, πηγές ετερογένειας μεταξύ των μελετών. Η εκτίμηση της ετερογένειας της μετα-ανάλυσης είναι ένα κρίσιμο ζήτημα επειδή η παρουσία έναντι της απουσίας πραγματικής ετερογένειας (μεταβλητότητα μεταξύ των μελετών) μπορεί να επηρεάσει το στατιστικό μοντέλο που ο μετα-αναλυτής αποφασίζει να εφαρμόσει στη μετα-αναλυτική βάση δεδομένων. Έτσι, όταν τα αποτελέσματα των μελετών διαφέρουν μόνο από το σφάλμα δειγματοληψίας (ομοιογενής περίπτωση), μπορεί να εφαρμοστεί ένα μοντέλο σταθερών αποτελεσμάτων για να επιτευχθεί ένα μέσο μέγεθος αποτελέσματος. Αντίθετα, εάν τα αποτελέσματα της μελέτης διαφέρουν περισσότερο από το σφάλμα δειγματοληψίας (ετερογενής περίπτωση), τότε ο μετα-αναλυτής μπορεί να αναλάβει ένα μοντέλο τυχαίων αποτελεσμάτων, προκειμένου να λάβει υπόψη τόσο τη μεταβλητότητα εντός όσο και μεταξύ των μελετών ή

1. Εισαγωγή

μπορεί να αποφασίσει να αναζητήσει μεταβλητές του συντονιστή από ένα μοντέλο σταθερών αποτελεσμάτων.

Ο συνηθισμένος τρόπος αξιολόγησης του αν υπάρχει πραγματική ετερογένεια σε μια μετα-ανάλυση είναι η χρήση του τεστ Q, ενός στατιστικού τεστ που καθορίστηκε από τον Cochran (1954) [81]. Η δοκιμή Q υπολογίζεται με το άθροισμα των τετραγωνικών αποκλίσεων της εκτίμησης του αποτελέσματος κάθε μελέτης από την εκτίμηση συνολικού αποτελέσματος, σταθμίζοντας τη συμβολή κάθε μελέτης με την αντίστροφη διακύμανσή της. Υπό την υπόθεση της ομοιογένειας μεταξύ των μεγεθών των επιδράσεων, η στατιστική Q ακολουθεί μια κατανομή chi-square (χ^2) με $k - 1$ βαθμούς ελευθερίας, όπου k είναι ο αριθμός των μελετών. Η μη απόρριψη της υπόθεσης ομοιογένειας οδηγεί συνήθως τον μετα-αναλυτή να υιοθετήσει ένα μοντέλο σταθερών αποτελεσμάτων, διότι υποτίθεται ότι τα εκτιμώμενα μεγέθη των αποτελεσμάτων διαφέρουν μόνο με σφάλμα δειγματοληψίας. Αντίθετα, η απόρριψη της παραδοχής ομοιογένειας μπορεί να οδηγήσει στην εφαρμογή ενός μοντέλου τυχαίων αποτελεσμάτων που περιλαμβάνει τόσο την μεταβλητότητα εντός όσο και μεταξύ των μελετών. Ένα μειονέκτημα της στατιστικής Q είναι ότι έχει μικρή ισχύ για να ανιχνεύσει πραγματική ετερογένεια μεταξύ των μελετών όταν η μετα-ανάλυση περιλαμβάνει ένα μικρό αριθμό μελετών και υπερβολική ισχύ για να ανιχνεύσει αμελητέα μεταβλητότητα με μεγάλο αριθμό μελετών. Έτσι, ένα μη σημαντικό αποτέλεσμα για το τεστ Q με ένα μικρό αριθμό μελετών μπορεί να οδηγήσει έναν αναλυτή να εφαρμόσει, λανθασμένα, ένα μοντέλο σταθερών αποτελεσμάτων όταν υπάρχει πραγματική ετερογένεια μεταξύ των μελετών και αντίστροφα. Από την άλλη πλευρά, η στατιστική Q δεν μας ενημερώνει για την έκταση της πραγματικής ετερογένειας, μόνο για τη στατιστική σημασία της.

Μια άλλη στρατηγική για την ποσοτικοποίηση της πραγματικής ετερογένειας σε μια μετα-ανάλυση αποτελείται από την εκτίμηση της μεταξύ των μελετών διακύμανσης, τ^2 . Εφαρμόζοντας ένα μοντέλο τυχαίων αποτελεσμάτων, η διακύμανση μεταξύ των μελετών εκφράζει το πόσο διαφορετικά μεγέθη των πραγματικών επιπτώσεων του πληθυσμού στις μεμονωμένες μελέτες μιας μετα-ανάλυσης διαφέρουν. Δεδομένου ότι το τ^2 εξαρτάται από το μοντέλο αποτελεσμάτων που χρησιμοποιείται σε μια μετα-ανάλυση, δεν είναι δυνατόν να συγκριθούν οι τιμές τ^2 που υπολογίζονται από μετα-αναλύσεις που έχουν χρησιμοποιήσει διαφορετικούς δείκτες μεγέθους αποτελέσματος (π.χ. τυποποιημένες μέσες διαφορές, συντελεστές συσχέτισης, λόγος συμπληρωματικών πιθανοτήτων κ.λπ.).

Προκειμένου να ξεπεραστούν οι αδυναμίες του τεστ Q και του τ^2 , οι Higgins και Thompson πρότειναν τρεις δείκτες για την εκτίμηση της ετερογένειας σε μια μετα-ανάλυση: τους H^2 , R^2 και I^2 . Επειδή είναι αλληλένδετοι, εδώ εστιάζουμε στον δείκτη I^2 , λόγω της εύκολης ερμηνείας του. Ο δείκτης I^2 μετράει την έκταση της πραγματικής ετερογένειας που διαχωρίζει τη διαφορά μεταξύ του αποτελέσματος

1. Εισαγωγή

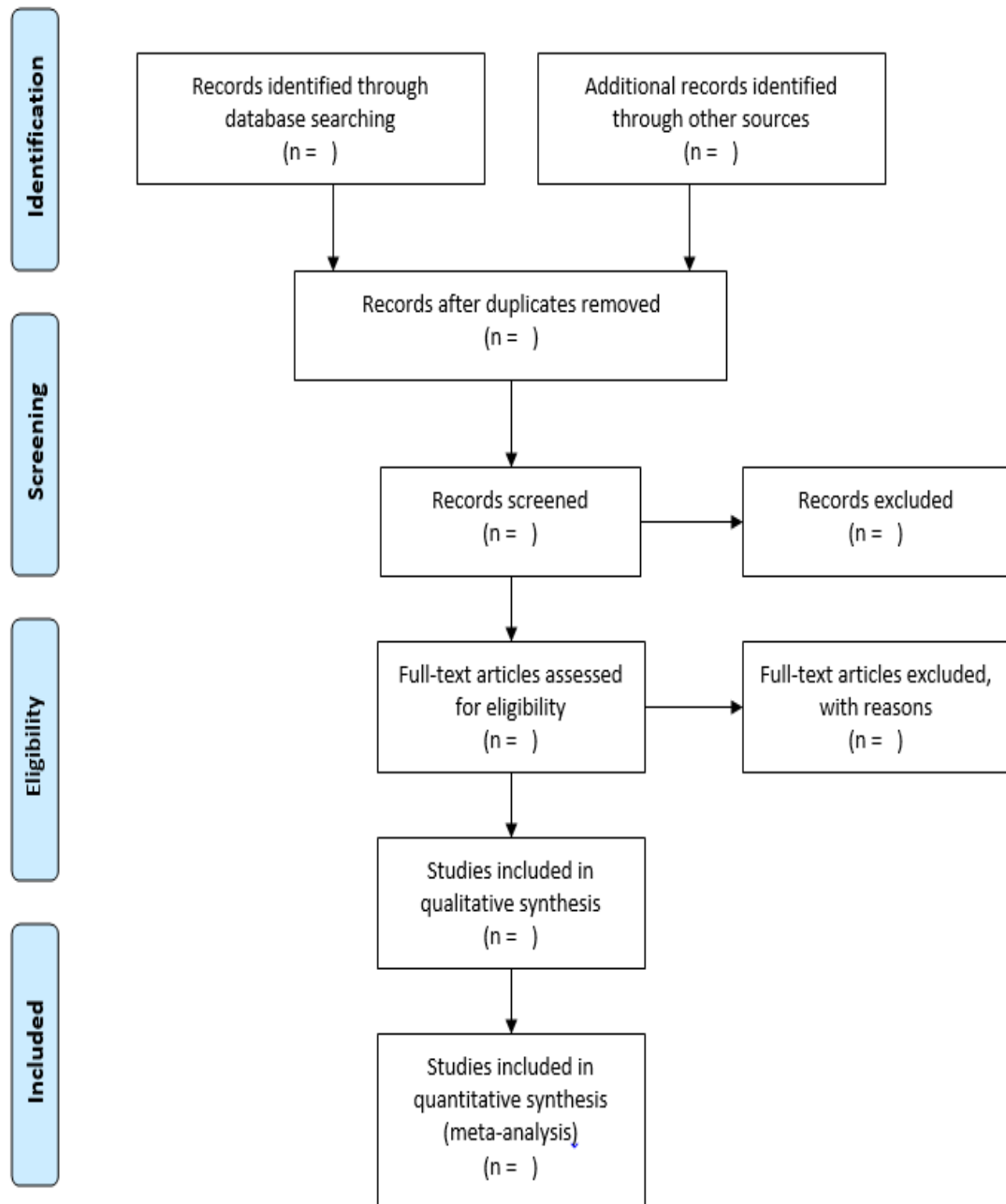
του Q τεστ και των βαθμών ελευθερίας ($k - 1$) από την ίδια την τιμή Q και πολλαπλασιάζεται με το 100 [82]. Ο δείκτης I^2 μπορεί να ερμηνευτεί ως το ποσοστό της συνολικής μεταβλητότητας σε ένα σύνολο μεγεθών αποτελεσμάτων λόγω της πραγματικής ετερογένειας, δηλαδή της μεταβλητότητας μεταξύ των μελετών. Για παράδειγμα, μια μετα-ανάλυση με $I^2 = 0$ σημαίνει ότι όλες οι μεταβλητότητες στις εκτιμήσεις μεγέθους αποτελέσματος οφείλονται σε σφάλμα δειγματοληψίας εντός των μελετών. Από την άλλη πλευρά, μια μετα-ανάλυση με $I^2 = 50$ σημαίνει ότι το ήμισυ της συνολικής μεταβλητότητας μεταξύ των μεγεθών των αποτελεσμάτων δεν προκαλείται από σφάλμα δειγματοληψίας, αλλά από την πραγματική ετερογένεια μεταξύ των μελετών. Οι Higgins και Thompson πρότειναν μια προσωρινή ταξινόμηση των τιμών I^2 με σκοπό να συμβάλουν στην ερμηνεία του μεγέθους του. Έτσι, ποσοστά περίπου 25% ($I^2 = 25$), 50% ($I^2 = 50$) και 75% ($I^2 = 75$) θα σήμαινε χαμηλή, μεσαία και υψηλή ετερογένεια, αντίστοιχα. Ο δείκτης I^2 και η διακύμανση μεταξύ των μελετών, τ^2 , σχετίζονται άμεσα: όσο υψηλότερο είναι το τ^2 , τόσο υψηλότερος είναι ο δείκτης I^2 . Ωστόσο ένα πλεονέκτημα του δείκτη I^2 σε σχέση με το τ^2 είναι ότι οι δείκτες I^2 που λαμβάνονται από μετα-αναλύσεις με διαφορετικούς αριθμούς μελετών και διαφορετικές μετρήσεις αποτελεσμάτων είναι άμεσα συγκρίσιμες.

Μαζί με αυτήν την περιγραφική ερμηνεία του δείκτη I^2 , οι Higgins και Thompson όρισαν ένα διάστημα εμπιστοσύνης το οποίο μπορεί να χρησιμοποιηθεί με τον ίδιο τρόπο που χρησιμοποιείται το Q τεστ, για να εκτιμηθεί η ετερογένεια της μετα-ανάλυσης. Επομένως, αν το διάστημα εμπιστοσύνης γύρω από το I^2 περιέχει την τιμή 0%, τότε ο μετα-αναλυτής μπορεί να κρατήσει την υπόθεση της ομοιογένειας. Εάν, αντίθετα, το διάστημα εμπιστοσύνης δεν περιλαμβάνει την τιμή 0%, τότε υπάρχουν στοιχεία για την ύπαρξη πραγματικής ετερογένειας.

1.4.1.7 Παρουσίαση των αποτελεσμάτων

Ένα χρήσιμο διάγραμμα, που παρουσιάζεται στις οδηγίες του PRISMA, είναι το διάγραμμα ροής τεσσάρων φάσεων.

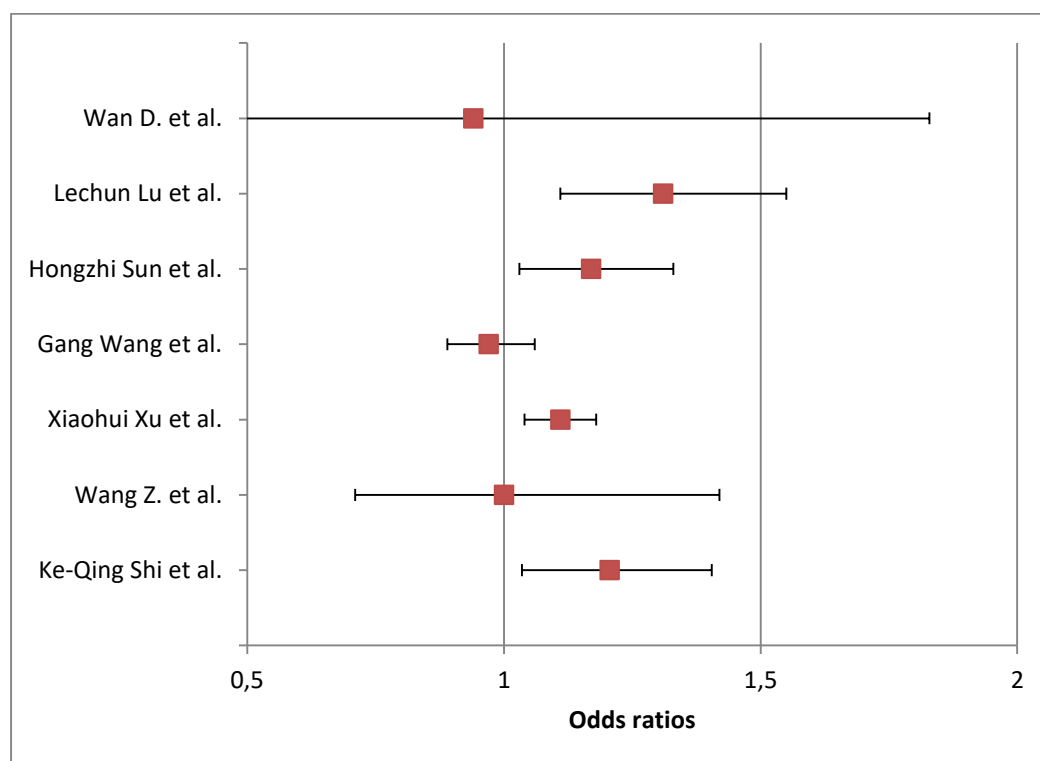
1. Εισαγωγή



Διάγραμμα 1: Διάγραμμα ροής σύμφωνα με το πρότυπο PRISMA

Αυτό το διάγραμμα ροής απεικονίζει τη ροή πληροφοριών μέσω των διαφόρων φάσεων μιας συστηματικής ανασκόπησης ή μετα-ανάλυσης. Καθορίζει τον αριθμό των εγγραφών που προσδιορίστηκαν, περιλήφθηκαν και εξαιρέθηκαν και τους λόγους για τους αποκλεισμούς. Τα αποτελέσματα των μετα-αναλύσεων παρουσιάζονται συχνά σε ένα δενδρόγραμμα, όπου κάθε μελέτη παρουσιάζεται με το μέγεθος της επίδρασης και το αντίστοιχο διάστημα εμπιστοσύνης 95%.

1. Εισαγωγή



Διάγραμμα 2: Forest plot από τμήμα μετα-αναλύσεων που συμπεριλαμβάνεται στην εργασία δημιουργημένο σε περιβάλλον Microsoft excel 2010

Το συγκεντρωτικό αποτέλεσμα και το διάστημα εμπιστοσύνης 95% εμφανίζονται στο κάτω μέρος στην ίδια γραμμή με το "Συνολικά". Στο δεξιό πλαίσιο της παραπάνω εικόνας παρουσιάζεται γραφικά η σωρευτική μετα-ανάλυση, όπου τα δεδομένα εισάγονται διαδοχικά, συνήθως με τη σειρά της χρονολογικής τους εμφάνισης [83]. Αυτή η σωρευτική μετα-ανάλυση μπορεί να αναγνωρίσει αναδρομικά το σημείο στο χρόνο, όταν ένα αποτέλεσμα θεραπείας έφθασε για πρώτη φορά σε συμβατικά επίπεδα σπουδαιότητας.

Η σωρευτική μετα-ανάλυση είναι ένας αποδοτικός τρόπος για να εξεταστούν οι τάσεις στην εξέλιξη του μεγέθους της συνοπτικής επίδρασης και να αξιολογηθεί ο αντίκτυπος μιας συγκεκριμένης μελέτης στα συνολικά συμπεράσματα [83].

1.4.2 Συστηματική ανασκόπηση

Μια συστηματική ανασκόπηση επιδιώκει τη συστηματική αναζήτηση, αξιολόγηση και σύνθεση των προσεκτικά σχεδιασμένων διαθέσιμων μελετών ως προς ένα συγκεκριμένο θέμα και αντιμετωπίζει ερωτήματα μεγάλου εύρους για την

1. Εισαγωγή

παραγωγή της καλύτερης σύνθεσης στοιχείων. Συχνά η συστηματική ανασκόπηση περιέχει μετα-αναλύσεις ως καλύτερο μέσω αξιολόγησης και εξαγωγής συμπερασμάτων.

Αυτές οι ανασκοπήσεις είναι περίπλοκες και εξαρτώνται σε μεγάλο βαθμό από τις διαθέσιμες κλινικές δοκιμές, τον τρόπο διεξαγωγής τους, δηλαδή την ποιότητα των δοκιμών, και τα αποτελέσματα που μετρήθηκαν. Οι συγγραφείς των ανασκοπήσεων συσσωρεύουν αριθμητικά δεδομένα σχετικά με τα αποτελέσματα του θέματος που μελετάται μέσω μιας διαδικασίας που ονομάζεται μετα-ανάλυση. Στη συνέχεια, οι συγγραφείς αξιολογούν τα στοιχεία για τυχόν οφέλη ή βλάβες που εγκλείονται στις έρευνες. Με αυτόν τον τρόπο, οι συστηματικές ανασκοπήσεις είναι σε θέση να συνοψίσουν την υπάρχουσα κλινική έρευνα σε ένα θέμα.

1.4.2.1 Στάδια συστηματικής ανασκόπησης

Οι συγγραφείς ακολουθούν πολύ μεθοδικά, ένα σχέδιο που ονομάζεται πρωτόκολλο. Το πρωτόκολλο περιγράφει τα βήματα που θα γίνουν κατά την προετοιμασία μιας αναθεώρησης. Τα πρωτόκολλα Cochrane τα πιο διαδεδομένα πρωτόκολλα στον κλάδο των ιατρικών και γενικότερα του κλάδου των επιστημών ζωής ερευνών. Σύμφωνα με το πρωτόκολλο ορίζεται:

- Ο τρόπος με τον οποίο βρίσκονται οι υπάρχουσες μελέτες
- Πώς αξιολογούνται οι σχετικές μελέτες όσον αφορά τη χρησιμότητά τους στην απάντηση στην ερώτηση επανεξέτασης
- Πώς τα αποτελέσματα των ξεχωριστών μελετών συγκεντρώνονται για να δώσουν ένα συνολικό μέτρο αποτελεσματικότητας

Τα βήματα που περιγράφονται παρακάτω βασίζονται σε αυτά που απαιτούνται για τη βιβλιοθήκη Cochrane. Παρέχουν μια σχολαστική διαδικασία μέσω της οποίας οι ερευνητές μπορούν να συνθέσουν δεδομένα από μια σειρά μελετών.

Βήμα 1: Οι ερευνητές πρέπει πρώτα να αποφασίσουν για ποια ερευνητική ερώτηση χρειάζονται απάντηση, με έμφαση όμως στο γεγονός πως ότι το ζήτημα πρέπει να είναι πολύ συγκεκριμένο.

Βήμα 2: Επιλογή των μελετών που θα συμπεριληφθούν στην ανασκόπηση. Αυτό θα αποφασιστεί εν μέρει από το ερευνητικό ερώτημα, αλλά περαιτέρω "κριτήρια επιλεξιμότητας" θα καθορίσουν εκ των προτέρων ποιες μελέτες θα συμπεριλάβει ή θα αποκλείσει η μελέτη.

1. Εισαγωγή

Βήμα 3: Αναζήτηση για τις μελέτες σε συγκεκριμένες πηγές, βάσεις αναζήτησης, με χρήση κατάλληλων όρων κλειδιών. Η έρευνα θα πρέπει επίσης να περιλαμβάνει προσπάθειες αποκάλυψης μη δημοσιευμένων μελετών.

Βήμα 4: Τα δεδομένα λαμβάνονται από μελέτες που πληρούν τα προκαθορισμένα κριτήρια επιλεξιμότητας. Τα δεδομένα μπορεί να προέρχονται από διάφορες μορφές (άρθρα, περιοδικά, βιβλία κ.α.).

Βήμα 5: Αξιολόγηση του κίνδυνου μεροληψίας στις μελέτες που περιλαμβάνονται, ώστε να εξασφαλιστεί ότι όλες οι μελέτες που εξετάζονται είναι σχετικές και αξιόπιστες.

Βήμα 6: Ανάλυση των δεδομένων και πραγματοποίηση μετα-αναλύσεων αν κριθεί απαραίτητο. Αυτή είναι η βασική διαδικασία μιας συστηματικής αναθεώρησης και το κύριο βήμα προς τη σύνταξη συμπερασμάτων.

Βήμα 7: Αντιμετώπιση οποιασδήποτε μεροληψίας δημοσίευσης, που μπορεί να οδηγήσει σε ψευδή αναφορά των πραγματικών αποτελεσμάτων της μελέτης.

Βήμα 8: Παρουσίαση των τελικών αποτελεσμάτων της ανασκόπησης επεξηγηματικά και με χρήση γραφημάτων και πινάκων.

1.4.2.2 Πλεονεκτήματα της Συστηματικής Ανασκόπησης

Οι παραδοσιακές ανασκοπήσεις της βιβλιογραφίας περιορίζονται πολύ συχνά στη βιβλιογραφία που είναι ήδη γνωστή στους συγγραφείς ή στη βιβλιογραφία που διαπιστώνεται με τη διενέργεια λίγων μόνο βιαστικών αναζητήσεων. Αυτό σημαίνει ότι οι ίδιες μελέτες αναφέρονται συχνά και αυτό εισάγει μια επίμονη μεροληψία σε βιβλιογραφικές αναθεωρήσεις. Οι συστηματικές ανασκοπήσεις συμβάλλουν στη μείωση της προκατάληψης του ερευνητή.

Μέσω της υιοθέτησης ευρέων στρατηγικών αναζήτησης, προκαθορισμένων αλυσίδων αναζήτησης και ομοιόμορφων κριτηρίων ένταξης και αποκλεισμού, οι συστηματικές αξιολογήσεις καθιστούν αποτελεσματικά τους ερευνητές να αναζητούν μελέτες πέρα από τις δικές τους περιοχές και δίκτυα. Ταυτόχρονα, η προσεκτική αποδόμηση του ερευνητικού ζητήματος από την αρχή, από την άποψη του πληθυσμού, της παρέμβασης, του μελετητή και του αποτελέσματος, διασφαλίζει ότι η διαδικασία αναθεώρησης παραμένει στενά εστιασμένη. Αυτό βελτιώνει την πιθανότητα δημιουργίας μιας σαφέστερης, πιο αντικειμενικής απάντησης στο ερευνητικό ερώτημα.

Επιπλέον, οι ανασκοπήσεις της παραδοσιακής βιβλιογραφίας στη διεθνή έρευνα για την ανάπτυξη συχνά επικεντρώνονται αποκλειστικά στα αποτελέσματα άλλων

1. Εισαγωγή

μελετών, χωρίς να λαμβάνεται υπόψη ο σχεδιασμός της μελέτης, τα δεδομένα και οι αναλυτικές μέθοδοι που χρησιμοποιούνται. Αντίθετα, οι συστηματικές αναθεωρήσεις επικεντρώνονται περισσότερο σε στοιχεία, επιπτώσεις, εγκυρότητα και αιτιότητα. Εξάγοντας πληροφορίες σχετικά με τον σχεδιασμό της έρευνας, τη στρατηγική δειγματοληψίας και μεθόδους συλλογής δεδομένων.

Ακόμα οι συστηματικές ανασκοπήσεις είναι αποτελεσματικές για τη μέτρηση της αξιοπιστίας των αποδείξεων. Η ταξινόμηση της ποιότητας και των χαρακτηριστικών των μελετών αντίκτυπου εναντίον τυποποιημένων κριτηρίων επιτρέπει επίσης τη δημιουργία συγκρίσεων και μετα-αναλύσεων μεταξύ των διαφόρων μελετών, οι οποίες είναι πολύτιμες για τη χάραξη πολιτικών βάσει τεκμηριωμένων και εμπειρικών στοιχείων.

Τέλος, η χρήση ενός σαφούς πρωτοκόλλου συστηματικής ανασκόπησης είναι αποτελεσματική όχι μόνο για την καθοδήγηση των ερευνητών καθ' όλη τη διάρκεια της διαδικασίας, αλλά και για τη βελτίωση της μεθοδολογικής διαφάνειας της ανασκόπησης [85]. Η αξιολόγηση του πρωτοκόλλου και της διαδικασίας εξασφαλίζει περαιτέρω μείωση της μεροληψίας του ερευνητή. Επιπλέον, οι συστηματικές ανασκοπήσεις είναι σε θέση να παράγουν μια σχετικά αντικειμενική βάση, σύμφωνα με την οποία μπορούν να αξιολογηθούν μελλοντικές έρευνες.

1.4.2.3 Μειονεκτήματα της Συστηματικής Ανασκόπησης

Παρόλο που η διεξαγωγή μια συστηματικής ανασκόπησης, όπως παρουσιάζεται παραπάνω, θα θεωρούταν η ιδανική επιλογή για το συνδυασμό των πληροφοριών σε ένα ακριβές αποτέλεσμα περικλείει όμως, όπως σε οποιαδήποτε άλλη στατιστική ή μεθοδολογική προσέγγιση της έρευνας μειονεκτήματα.

Αρχικά οι διαθέσιμες δημοσιευμένες εργασίες δεν έχουν ισορροπία ως προς τα αποτελέσματα που καταλήγουν καθώς οι πιθανότητες δημοσίευσης των αποτελεσμάτων είναι μεγαλύτερες αν τα αποτελέσματα ήταν σημαντικά, δηλαδή κατέληγαν σε επιθυμητό αποτέλεσμα. Σαν συνέπεια πολλές μελέτες δεν δημοσιεύονται λόγω μη θετικού αποτελέσματος, δημιουργώντας την επιλεκτική υποβολή αντί της επιλεκτικής αποδοχής

Ένα μειονέκτημα που είναι δύσκολο να καταπολεμηθεί δημιουργείται από τον ίδιο τον τρόπο λειτουργίας των βάσεων δεδομένων. Παρά το γεγονός ότι είναι μια πολύ καλή πηγή επιστημονικής πληροφόρησης, οι πιο συχνά προσβάσιμες βάσεις δεδομένων δημοσιεύουν κυρίως αναφορές από ανεπτυγμένες χώρες και τείνουν να παραμελούν την υπάρχουσα βιβλιογραφία από λιγότερο ανεπτυγμένες χώρες.

1. Εισαγωγή

Σύμφωνα με όσα αναφέρθηκαν παραπάνω όσον αφορά την ύπαρξη καλών κριτηρίων συμπερίληψης των εγγράφων για ανάλυση, πολλές φορές οι κατάλογοι αναφοράς των χειρογράφων περιλαμβάνουν συχνά μελέτες που υποστηρίζουν έντονα τα δημοσιευμένα αποτελέσματα [84].

Επίσης πολλές φορές ο μετα-αναλυτής στην προσπάθεια συγκέντρωσης άρθρων συναντά πλήθος δημοσιεύσεων από την ίδια ομάδα ερευνητών με βάση όμως μία μόνο έρευνα, καθιστώντας δύσκολο να διακρίνει κανείς εάν δύο έγγραφα είναι αρχικές έρευνες ή αντίγραφα από τις ίδιες δοκιμές [84].

Τέλος ακόμα και αν ο αναλυτής συγκεντρώσει μεγάλο πλήθος δεδομένων και από πολλές πηγές, μέσα σε αυτό το υλικό μπορεί να υπάρξουν έρευνες από τις οποίες αδυνατεί να αποκτήσει τα πλήρη στοιχεία, διότι οι συγγραφείς δεν τα καθιστούν διαθέσιμα. Κάτι τέτοιο μπορεί να οδηγήσει σε αλλαγή της κατεύθυνσης των αποτελεσμάτων της έρευνας.

2. Μεθοδολογία

2.1 Εργαλεία

Σε αυτήν τη ενότητα θα παρουσιαστούν συνοπτικά τα διάφορα υπολογιστικά εργαλεία που χρησιμοποιήθηκαν στην εκπλήρωση της εργασίας.



Η PubMed [86] είναι μία δωρεάν βάση δεδομένων που αναπτύσσεται και συντηρείται από το Εθνικό Κέντρο Βιοτεχνολογικών Πληροφοριών (NCBI), στην Εθνική Βιβλιοθήκη Ιατρικής των ΗΠΑ (NLM), που βρίσκεται στα Εθνικά Ινστιτούτα Υγείας (NIH). Περιλαμβάνει πάνω από 26 εκατομμύρια αναφορές για βιοϊατρική βιβλιογραφία από το MEDLINE, περιοδικά για επιστήμες υγείας, και άλλα online περιοδικά. Επιπλέον παρέχει πρόσβαση σε δημοσιεύσεις που πραγματοποιήθηκαν πριν από την ηλεκτρονική έκδοση του Index Medicus το 1951, αλλά και πολύ πρόσφατες που δεν έχουν καταχωρηθεί ακόμα στο MEDLINE, στο Medical Subject Headings (MeSH) και το Index Medicus. Τέλος έχει αναφορές σε μια συλλογή από διαθέσιμα βιβλία πλήρους κειμένου και άλλα υποσύνολα των αρχείων της Εθνικής Βιβλιοθήκης Ιατρικής των ΗΠΑ [87].



Το Microsoft Excel είναι πρόγραμμα λογιστικών φύλλων που αναπτύχθηκε από τη Microsoft, αποτελεί μέρος του Microsoft Office και χρησιμοποιείται για την αποθήκευση, την οργάνωση και τον χειρισμό δεδομένων. Τα υπολογιστικά φύλλα παρουσιάζουν πίνακες τιμών διατεταγμένων σε σειρές και στήλες που μπορούν να τροποποιηθούν μαθηματικά χρησιμοποιώντας βασικές και πολύπλοκες αριθμητικές πράξεις και συνθήκες. Επιπλέον παρέχει την δυνατότητα δημιουργία γραφημάτων και συγκεντρωτικών πινάκων, καθώς και μια γλώσσα προγραμματισμού μακροεντολών με όνομα Visual Basic for Applications.

2. Μεθοδολογία

Adobe Flash



Το Adobe Flash είναι μια πλατφόρμα λογισμικού πολυμέσων που χρησιμοποιείται για την παραγωγή γραφικών, κινούμενων εικόνων, εφαρμογών και ηλεκτρονικών παιχνιδιών. Δημιουργήθηκε από την Adobe System και το πλήθος εκδόσεων που παρέχει καλύπτει τις ανάγκες ποικίλων ειδικοτήτων χρηστών. Οι εικόνες στην παρούσα εργασία (το 95%)δημιουργήθηκε με το συγκεκριμένο πρόγραμμα.



Cytoscape

Το Cytoscape είναι ένα πρόγραμμα λογισμικού ανοιχτού κώδικα για την ενσωμάτωση δικτύων βιομοριακής αλληλεπίδρασης με δεδομένα έκφρασης υψηλής απόδοσης σε ένα ενοποιημένο θεματικό πλαίσιο. Παρόλο που μπορεί να εφαρμοστεί σε οποιοδήποτε σύστημα μοριακών στοιχείων και αλληλεπιδράσεων, το Cytoscape είναι πιο ισχυρό όταν χρησιμοποιείται σε συνδυασμό με μεγάλες βάσεις δεδομένων με σχέσεις πρωτεϊνών-πρωτεϊνών, πρωτεϊνών-DNA και γενετικών αλληλεπιδράσεων ανθρώπων και άλλων οργανισμών. Το λογισμικό του παρέχει βασικές λειτουργίες για να διαταχθεί το δίκτυο, να γίνουν αναζητήσεις μέσα σε αυτό, να πραγματοποιηθεί οπτική ενσωμάτωση του δικτύου με προφίλ έκφρασης, φαινοτύπους και άλλες μοριακές καταστάσεις και επιπλέον να συνδέσει το δίκτυο με βάσεις δεδομένων λειτουργικών σχολιασμών [88].



miRNASNP

miRNASNP είναι μία δωρεάν βάση δεδομένων, η οποία περιέχει SNPs σχετιζόμενα με microRNAs, είτε SNPs σε pre-miRNAs ανθρώπου και άλλων ειδών, είτε κέρδος και απώλεια από SNPs σε seed regions του miRNA ή στο 3'UTR άκρο των mRNA στόχων. Αυτοί οι πολυμορφισμοί μπορεί να επηρεάσουν τη δέσμευση του mRNA στόχου ή και την βιογένεση του miRNA. Παρείχε άφθονα δεδομένα σχετικά με την έκφραση του miRNA, τους επικυρωμένους στόχους και τις σχετικές φαινοτυπικές παραλλαγές, ενώ στην τελευταία διαθέσιμη έκδοση παρέχονται δυνατότητες όπως η εύρεση δεδομένων σχετικά με το επίπεδο έκφρασης αλλά και τη συσχέτιση της έκφρασης των miRNAs και των γονιδίων στόχων σε διαφορετικούς ιστούς, την σύνδεση των SNPs με τα αποτελέσματα ευρέων μελετών σύνδεσης γονιδιώματος,

2. Μεθοδολογία

την ενσωμάτωση πειραματικά επικυρωμένων miRNA: mRNA αλληλεπιδράσεων, και τέλος την προσθήκη πολλαπλών φίλτρων για την ιεράρχηση των λειτουργικών SNP. Επιπλέον υπάρχουν διαθέσιμα τρία online εργαλεία για την ανάλυση SNPs σε microRNAs που επιλέγει ο χρήστης [89].



Η βάση δεδομένων μονονουκλεοτιδικών πολυμορφισμών (Single Nucleotide Polymorphism Database, dbSNP) είναι ένα δωρεάν δημόσιο αρχείο που περιέχει γενετικούς πολυμορφισμούς στο γονιδίωμα διαφόρων ειδών. Δημιουργήθηκε και υποστηρίζεται από το Εθνικό Κέντρο Βιοτεχνολογικών Πληροφοριών (NCBI) σε συνεργασία με το Εθνικό Ινστιτούτο Ανθρώπινου Γονιδιώματος (NHGRI). Αποτελεί επί του παρόντος την μεγαλύτερη βάση δεδομένων για νουκλεοτιδικούς πολυμορφισμούς, καθώς περιέχει πάνω από 12 εκατομμύρια πολυμορφισμούς, πάνω από 1 δισεκατομμύριο μεμονωμένους γονότυπους από το HarMap και συνολικά πάνω από 200GB δεδομένων που συνεχώς αυξάνονται.

Diana tools **DIANA TOOLS**

Στόχος των εργαλείων DIANA είναι η παροχή αλγορίθμων, βάσεων δεδομένων και λογισμικού για την ερμηνεία και την αρχειοθέτηση των δεδομένων σε ένα συστηματικό πλαίσιο που κυμαίνεται από την ανάλυση της ρύθμισης της έκφρασης από τα δεδομένα βαθιάς αλληλουχίας, τον σχολιασμό ρυθμιστικών στοιχείων και στόχων των miRNAs, αλλά και την κατανόηση του ρόλου των μη κωδικών miRNAs σε διάφορες ασθένειες και βιολογικά μονοπάτια. Τα εργαλεία περιέχουν αλγόριθμους πρόβλεψης mRNA στόχων, βάσεις δεδομένων πειραματικώς επαληθευμένων στόχων miRNA σε κώδικα και μη RNAs, λογισμικό ικανό να αναγνωρίζει δυνητικά μεταβαλλόμενες μοριακές οδούς με την έκφραση ενός ή πολλαπλών miRNAs και άλλες επιλογές επεξεργασίας δεδομένων από τον χρήστη [90].



KEGG

Η KEGG (Kyoto Encyclopedia of Genes and Genomes) είναι μία πηγή πληροφοριών από ένα σύνολο βάσεων δεδομένων, η οποία δημιουργήθηκε και αναπτύσσεται από την εργαστηριακή ομάδα του καθηγητή Kanehisa [91]. Περιέχει πληροφορίες

2. Μεθοδολογία

για γονιδιακές, χημικές και συστημικές λειτουργίες και τον τρόπο με τον οποίο αυτές αλληλεπιδρούν με τα γονίδια, ώστε να προκύψουν οι βασικές λειτουργίες του βιολογικού συστήματος. Επιπλέον περιέχει πληροφορίες για ένα μεγάλο εύρος νοσημάτων και φαρμάκων, καθώς και για τις αλληλεπιδράσεις μορίων, γονιδίων, πρωτεϊνών και μονοπατιών σε αυτά.

2.2 Βήματα

2.2.1 Επιλογή Άρθρων και Μετα-αναλύσεων

Στην παρούσα εργασία αναλύθηκαν στοιχεία από μετα-αναλύσεις με θέμα την σχέση μονονουκλεοτιδικών πολυμορφισμών σε microRNAs με ασθένειες στον άνθρωπο. Αρχικά αναζητήθηκαν στην βάση δεδομένων Pubmed δημοσιευμένη βιβλιογραφία μετα-αναλύσεων σε πολυμορφισμούς, με τους όρους (microRNA OR MiR OR mirna) and meta-analysis . Τα άρθρα που προέκυψαν ως απάντηση περιείχαν πλήθος μη σχετικών αναζητήσεων. Για την εκκαθάριση και επιλογή μόνο των απαραίτητων κειμένων ακολουθήθηκαν τα βήματα που ορίζονται από το πρότυπο PRISMA. Το τελικό σύνολο δεδομένων αποτελούταν από 120 άρθρα που το καθένα είχε είτε μία είτε περισσότερες μετα-αναλύσεις.

Προσεκτική επιλογή έπρεπε να γίνει και στο ποιων μετα-αναλύσεων τα δεδομένα θα καταγραφούν, καθώς όπως προαναφέρθηκε αρκετά άρθρα πραγματοποιούσαν πολλαπλές μετα-αναλύσεις είτε για διαφορετική φυλή, είτε για περισσότερα του ενός microRNA ή πολυμορφισμό, είτε τέλος για πλήθος ασθενειών. Σε αυτή την περίπτωση εάν η εκάστοτε μετα-ανάλυση πληρούσε χαρακτηριστικά όπως ο αριθμός των case-control μελετών που περιλάμβανε, ο συνολικός αριθμός των συμμετεχόντων, τα στατιστικά δεδομένα που εξήγαγε και η αιτιολόγηση των αποτελεσμάτων τότε στο φύλλο καταγραφής των δεδομένων μπορεί να εμπεριέχεται παραπάνω της μίας μετα-ανάλυσης από το ίδιο άρθρο. Δεν καταγραφόταν μία μετα-ανάλυση, ακόμα και αν πληρούσε τις παραπάνω προϋποθέσεις, , αν υπήρχε σημαντικός βαθμός επικάλυψης της μίας μετα-ανάλυσης με την άλλη μέσα στο ίδιο άρθρο, δηλαδή οι έρευνες που εμπεριέχονται στην μία καταγραφή να αποτελούν και το μεγαλύτερο ποσοστό (περίπου της τάξεως του 80% και άνω) κάποιας άλλης καταγραφής ως προς το ίδιο πολυμορφισμό και ασθένεια.

2.2.2 Καταγραφή Δεδομένων

Το επόμενο βήμα ήταν η καταγραφή στοιχείων από τα 120 επιλεγμένα άρθρα, τα οποία ήταν τόσο ποιοτικού όσο και ποσοτικού χαρακτήρα. Πιο αναλυτικά καταγράφηκαν για κάθε άρθρο:

- i. Το PubMed id, δηλαδή ο χαρακτηριστικός αριθμός για κάθε δημοσιευμένη μελέτη, ο οποίος είναι μοναδικός.

2. Μεθοδολογία

- ii. Το όνομα του ή των συγγραφέων κάθε άρθρου με την σειρά που αναγραφόταν στην δημοσίευση.
- iii. Η φυλή των συμμετεχόντων στις μετα-αναλύσεις. Οι τρεις κατηγορίες οι οποίες συναντήθηκαν κατά την περαιώση της καταγραφής ήταν η ασιατική (Asian), η καυκάσια (Caucasian) και η μεικτή (Mixed) η οποία μπορεί να περιέχει και πληθυσμούς από αφρικανική φυλή
- iv. Το όνομα του microRNA που εξετάζεται στην κάθε μετα-ανάλυση. Για λόγους συντομίας τα microRNAs συμβολίζονται miR-number, άλλα όλα ανήκουν σε ανθρώπινο γονιδίωμα και η πλειοψηφία τους είναι pre-miRNA.
- v. Το όνομα του πολυμορφισμού στο συγκεκριμένο microRNA με την μορφή rs number. Το πρόθεμα rs προέρχεται από την αγγλική references SNP (RefSNP, rs) και ακολουθείται από μία χαρακτηριστική ακολουθία αριθμών που είναι μοναδική για κάθε μονονουκλεοτιδικό πολυμορφισμό.
- vi. Την ασθένεια της οποίας η συσχέτιση με τον πολυμορφισμό μελετάται. Σε αρκετές περιπτώσεις είναι κατηγορία ασθενειών όπως για παράδειγμα ο καρκίνος ή τα αυτοάνοσα νοσήματα.
- vii. Τον αριθμό των αρχικών πειραματικών ερευνών που περιλαμβάνει κάθε μετανάλυση. Πολλά άρθρα διεξήγαγαν μετα-αναλύσεις με την ύπαρξη μόνο δύο ερευνών. Στην παρούσα εργασία έχει οριστεί κατώτερο όριο των τριών ερευνών ώστε να συμπεριληφθεί μια μετανάλυση στα καταγεγραμμένα δεδομένα.
- viii. Ο αριθμός των ασθενών που συμμετέχουν στο σύνολο των cases-controls μελετών και
- ix. Ο αριθμός αντίστοιχα των υγείων που συμμετέχουν. Όπως και στο πλήθος των ερευνών, έτσι και για τον αριθμό ασθενών και υγείων ελέγχθηκε να μην είναι πολύ χαμηλός ώστε να υπάρχει αντιπροσωπευτικότητα στο δείγμα.
- x. Η ύπαρξη ή η απουσία των στοιχείων (συγγραφείς, χρονολογία δημοσίευσης κ.α.) των ερευνών που περιλάμβανε η κάθε μετανάλυση ώστε να είναι εφικτός ο έλεγχος των δεδομένων.
- xi. Το odds ratio του αποτελέσματος της μετα-ανάλυσης για κάθε έναν από τους 5 γονότυπους επικρατής (dominant) , υπολειπόμενος (recessive) , ομόζυγος (homozygous) , ετερόζυγος (heterozygous) και αλληλόμορφου (allelic) όταν αυτός εξεταζόταν.
- xii. Το διάστημα εμπιστοσύνης (Confidence Interval) για 95% πιθανότητα, αντίστοιχα για καθέναν από τους 5 γονότυπους όταν εξεταζόταν.
- xiii. Την τιμή p για την συσχέτιση (p-value for association) και
- xiv. Την τιμή p για την ετερογένεια (p-value for heterogeneity) για όσους γονότυπους υπήρχε.
- xv. Τον αριθμό Z (Z-score) και
- xvi. Την τιμή του I^2 όπου αυτά είχαν εξαχθεί από τους μελετητές, και αντίστοιχα για όσους από τους 5 γονότυπους.

2. Μεθοδολογία

- xvii. Αν για την συγκεκριμένη μετα-ανάλυση είχε διεξαχθεί bias test , για την αποφυγή λαθών προκαταλήψεων.
- xviii. Αν το αποτέλεσμα της μετα-ανάλυσης αποδείκνυε συσχέτιση της ασθένειας με το συγκεκριμένο πολυμορφισμό, τόσο από τα συμπεράσματα των συγγραφέων όσο και από τις διαθέσιμες καταγεγραμμένες τιμές που το υποδείκνυαν.
- xix. Και τέλος, στην περίπτωση ύπαρξη συσχέτισης, σε ποιους γονότυπους αυτή εμφανίζεται, πάλι βάσει των λεγομένων των συγγραφέων αλλά και παρατήρησης των στατιστικών τιμών.

2.2.3 Επεξεργασία καταγραφών

Ορισμένα από τα χαρακτηριστικά που καταχωρήθηκαν απαιτούσαν επεξεργασία, ώστε να αυξηθεί είτε η ευκρίνεια τους, είτε η αξιοπιστία τους. Αρχικά όλες οι ασθένειες μετονομάστηκαν συμπεριλαμβάνοντας στον τίτλο τους τον κωδικό από το σύστημα ICD -10 (International Classification of Diseases). Η δέκατη αναθεώρηση της Διεθνούς Στατιστικής Ταξινόμησης των Νόσων και των Σχετικών Προβλημάτων Υγείας (ICD -10) είναι ένα σύστημα κατηγοριών στο οποίο καταγράφονται οι νοσολογικές οντότητες σύμφωνα με καθιερωμένα κριτήρια. Η στατιστική ταξινόμηση των νόσων περιλαμβάνει ολόκληρο το φάσμα των νοσολογικών καταστάσεων μέσα σε έναν ικανό αριθμό κατηγοριών. Με αυτό τον τρόπο προσφέρεται μία κοινή βάση αναφοράς και ανεξαρτητοποιείται η μελέτη τόσο από την υποκειμενικότητα της λεκτικής ονομασίας όσο και από την εθνικότητα του αναγνώστη.

Όπως έχει ήδη αναφερθεί, έχει καθοριστεί ότι στην περίπτωση ύπαρξης πολλαπλών μετα-αναλύσεων σε ένα άρθρο, τότε εκείνες που συμπεριέλαβαν μόνο 2 cases-controls έρευνες δεν θα καταγράφονταν. Επειδή όμως υπήρχαν είτε μετα-αναλύσεις που γίνονταν αποκλείστηκα σε 2 έρευνες και τα ευρήματά τους ήταν σημαντικά από στατιστικής απόψεως , είτε αυτές οι 2 ήταν μεγάλου πληθυσμού, τότε έχουν καταγραφεί και εκείνων τα δεδομένα.

2.2.4 Υπολογισμός στατιστικών τιμών

Ένα μεγάλο ποσοστό από τους μετα-αναλυτές δεν δημοσιεύει η' δεν πραγματοποιεί όλα τα στατιστικά τεστ που στην παρούσα μελέτη κρίνονται ως απαραίτητα. Για τον λόγο αυτόν τα ήδη καταγεγραμμένα δεδομένα επεξεργάστηκαν, υπολογίστηκαν τιμές και αναθεωρήθηκαν αποτελέσματα.

2. Μεθοδολογία

Αρχικά όλες οι τιμές των odds ratio μετατράπηκαν σε τιμές μεγαλύτερες του 1 χρησιμοποιώντας την συνθήκη:

$$IF \begin{cases} Or_{old} < 1 & \text{then } or_{new} = or_{old} \\ Or_{old} > 1 & \text{then } or_{new} = \frac{1}{or_{old}} \end{cases} \quad [1]$$

όπου or_{old} : Odds ratio από τις μετα-αναλύσεις,

or_{new} : Odds ratio σύμφωνα με την συνθήκη

Ελέχθησαν όλα τα διαστήματα εμπιστοσύνης των μετα-αναλύσεων, με έμφαση σε εκείνες που εμφάνιζαν συσχέτιση μεταξύ SNPs και ασθενειών, και χαρακτηρίστηκαν ως μην σχετικές όσες περιείχαν το 1 στο διάστημα τους.

Λόγω του μεγάλου ποσοστού έλλειψης τιμών Z, ενός χαρακτηριστικού σημαντικού, υπολογίστηκε, για όλες τις μετα-αναλύσεις και όλους τους γονότυπους, όπου υπήρχαν τα απαραίτητα δεδομένα, σύμφωνα με τους τύπους:

$$CI = \log(or_{new}) \pm 1,96 * SE \quad [2]$$

όπου CI: Confidence Interval από τις μετα-αναλύσεις,

or_{new} : Odds ratio γνωστό από την συνθήκη [1]

SE: Standard Error όπου είναι ο άγνωστος,

+ : όταν ως CI τοποθετείται το άνω όριο του διαστήματος

- : όταν ως CI τοποθετείται το κάτω όριο του διαστήματος

1,96: η τιμή του z_c (critical value, c) για διάστημα εμπιστοσύνης 95%

Στη συνέχεια εφαρμόζεται ο τύπος:

$$Z = \frac{\log(or_{new})}{SE} \quad [3]$$

Όπου Z: τιμή για την συσχέτιση (Z score of association)

or_{new} : Odds ratio γνωστό από την συνθήκη [1]

SE: Standard Error γνωστό από την συνθήκη [2]

2. Μεθοδολογία

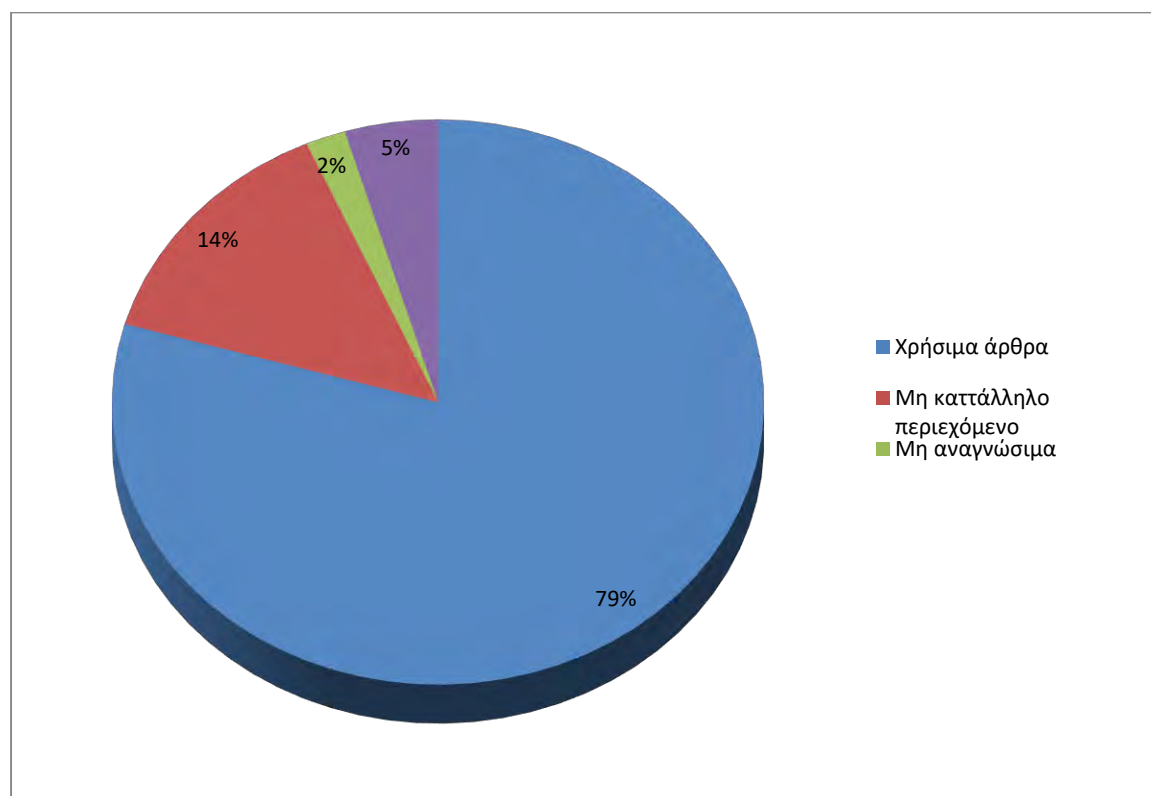
Οι καταγραφές των οποίων η τιμή Z ήταν ή πλησίαζε το 0 χαρακτηρίστηκαν ως μη συσχετιζόμενες.

Αφού λοιπόν τα δεδομένα υπέστησαν όσες τροποποιήσεις και προσθήκες θεωρήθηκε απαραίτητο παρακάτω ακολουθεί η ανάλυση των καταγραφών αυτών.

3. Αποτελέσματα

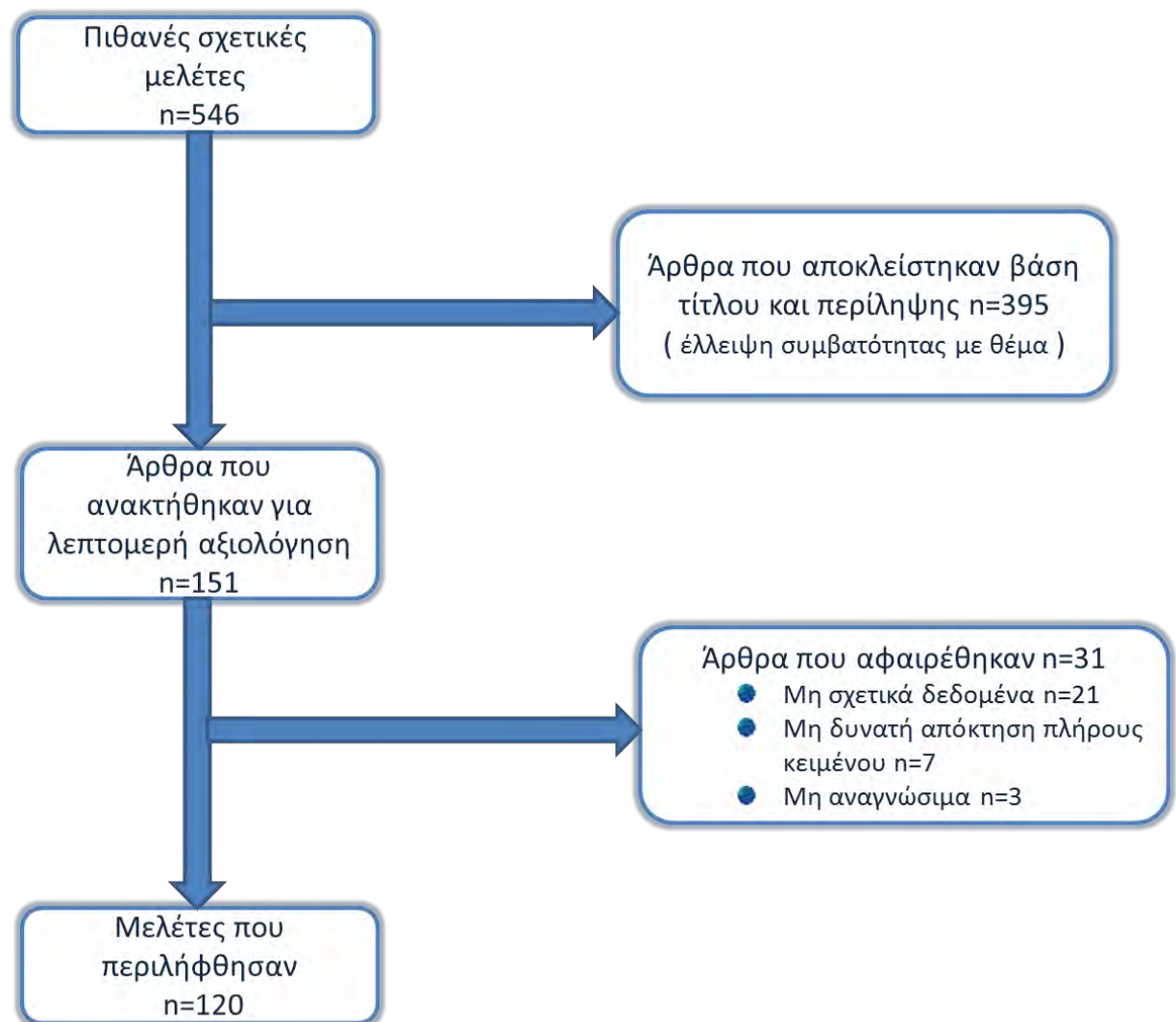
3.1 Περιγραφική Στατιστική

Συνολικά από την αναζήτηση στην PubMed Βρέθηκαν 546 άρθρα, τα οποία αναγνώστηκαν και διαχωρίστηκαν. Από αυτά τα 151 που αφορούσαν πολυμορφισμούς σε miRNAs. Τα 151 άρθρα μελετήθηκαν σε εκτενέστερο βαθμό με στόχο να απομείνουν μόνο εκείνα τα οποία καλύπτουν πλήρως το θέμα και τα δεδομένα τους είναι επαρκή για την διεξαγωγή της μελέτης. Αποτέλεσμα ήταν η εξάλειψη 21 άρθρων ως μη κατάλληλα για την παρούσα έρευνα καθώς είτε αφορούσαν πολυμορφισμούς σε γονίδια και πρωτεΐνες, είτε δεν πραγματοποιούνταν συσχέτιση με κάποια ασθένεια, είτε δεν ήταν μετανάλυση. Επιπλέον 7 άρθρα αφαιρέθηκαν καθώς δεν ήταν δυνατή η απόκτηση του πλήρους κειμένου τους, ενώ 3 ακόμα ήταν σε μη αναγνώσιμη γλώσσα (Ελληνικά, Αγγλικά) αλλά στην μητρική γλώσσα των συγγραφέων (Κινέζικα). Τελικά επιλέχθηκαν 120 μετα-αναλύσεις να αποτελέσουν την πηγή δεδομένων. [Γράφημα 1] [Διάγραμμα 3]



Γράφημα 1: Ποσοστιαία αναλογία κατηγοριοποίησης άρθρων

3. Αποτελέσματα

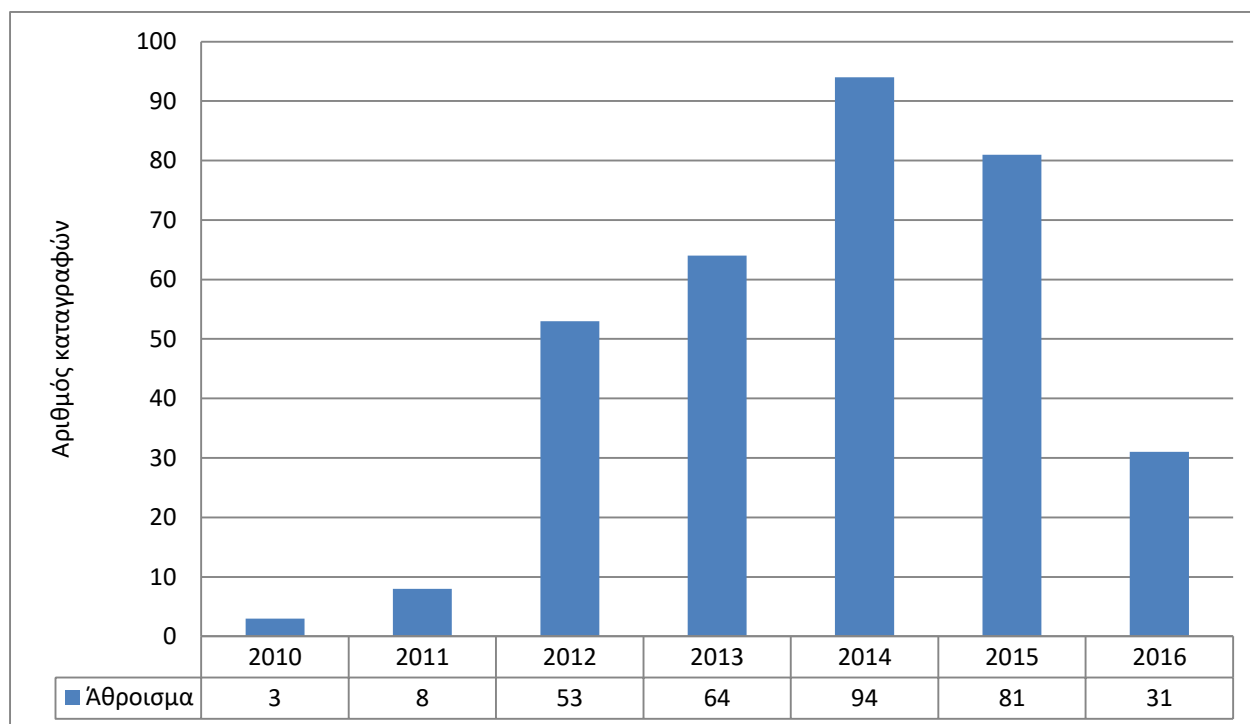


Διάγραμμα 3: Διάγραμμα ροής επιλογής μελετών

Από τα 120 άρθρα, τα οποία περιλήφθηκαν στην μελέτη, προέκυψαν 334 καταγραφές/μετα-αναλύσεις. Η διαφορά στον αριθμό άρθρων - καταγραφών υπάρχει λόγω του φαινομένου που προαναφέρθηκε, της ύπαρξης δηλαδή άνω της μίας μετα-ανάλυσης στο ίδιο άρθρο.

3. Αποτελέσματα

Το χρονολογικό εύρος των μελετών κυμαίνεται από το 2010 έως το 2016. Η αιτία για την μη ύπαρξη παλιότερων σχετικών μελετών στην βάση δεδομένων είναι το γεγονός πως η ύπαρξη και κατανόηση των microRNA συνέβη στις αρχές του 21^{ου} αιώνα. Οπότε η διαπραγμάτευση ενός ακόμα πιο ειδικού θέματος, όπως είναι οι μονονουκλεοτιδικοί πολυμορφισμοί στα microRNAs, και η ύπαρξη αρκετών εργαστηριακών μελετών για την πραγματοποίηση μετα-ανάλυσης άνοιξε μετά την πρώτη δεκαετία. Το κατώφλι του έτους 2016 υπάρχει καθώς τέλος του συγκεκριμένου έτους άρχισε η παρούσα εργασία. Όπως είναι ορατό στο γράφημα που ακολουθεί τα πρώτα δύο έτη, 2010-2011 η δημοσιευμένες μετα-αναλύσεις για τον παρών θέμα είναι ελάχιστες, αποτελώντας μόλις το 1% των καταγραφών. Από το 2012 και μετά αρχίζει σταδιακά να αυξάνεται ο αριθμός με κορύφωση το έτος 2014 όπου δημοσιεύτηκαν το 28% των μετα-αναλύσεων που έχουν καταγραφή εδώ. Την επόμενη χρονιά το ποσοστό πέφτει ενώ οι καταγραφές από το 2016 είναι μόλις το 9%. Το μικρό ποσοστό των δημοσιεύσεων το 2016 μπορεί να οφείλεται στο πλήθος μελετών από το 2016 που δεν είχε ακόμα καταχωρηθεί στην βάση δεδομένων της PubMed. [Γράφημα 2]

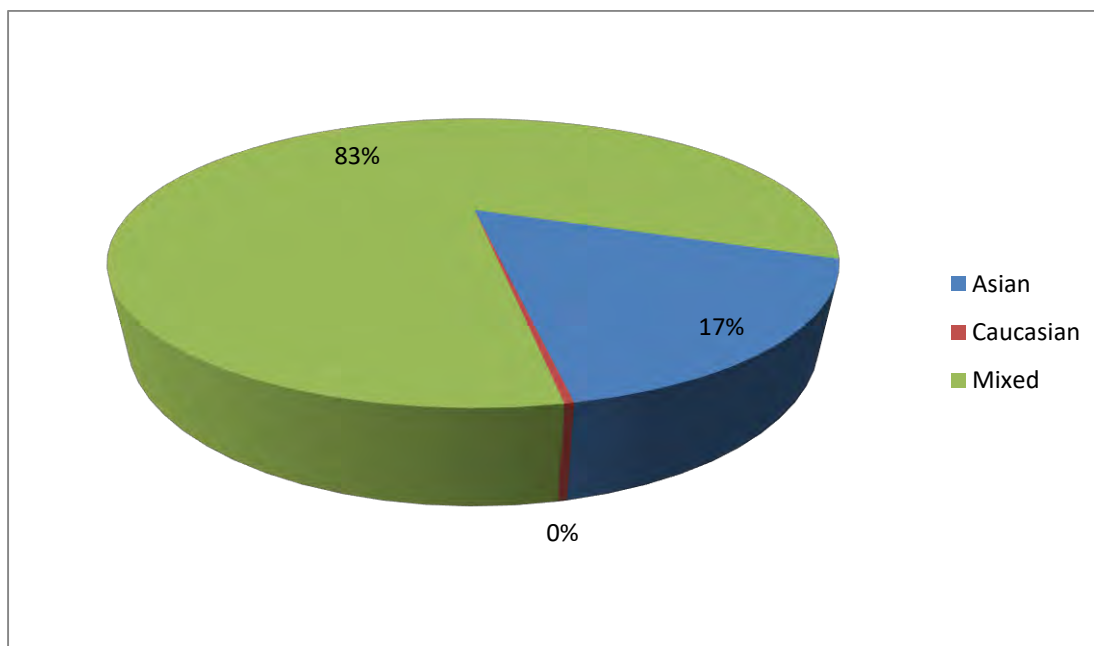


Γράφημα 2: Χρονική κατανομή δημοσίευσης μετα-αναλύσεων

Αν εξεταστούν οι καταγραφές από άποψη εθνικότητας των συμμετεχόντων, τότε η πλειοψηφία περιλάμβανε άτομα από διάφορες φυλές (Mixed), αγγίζοντας τον αριθμό των 278 μετα-αναλύσεων. Μία μία σημαντική μερίδα ερευνών

3. Αποτελέσματα

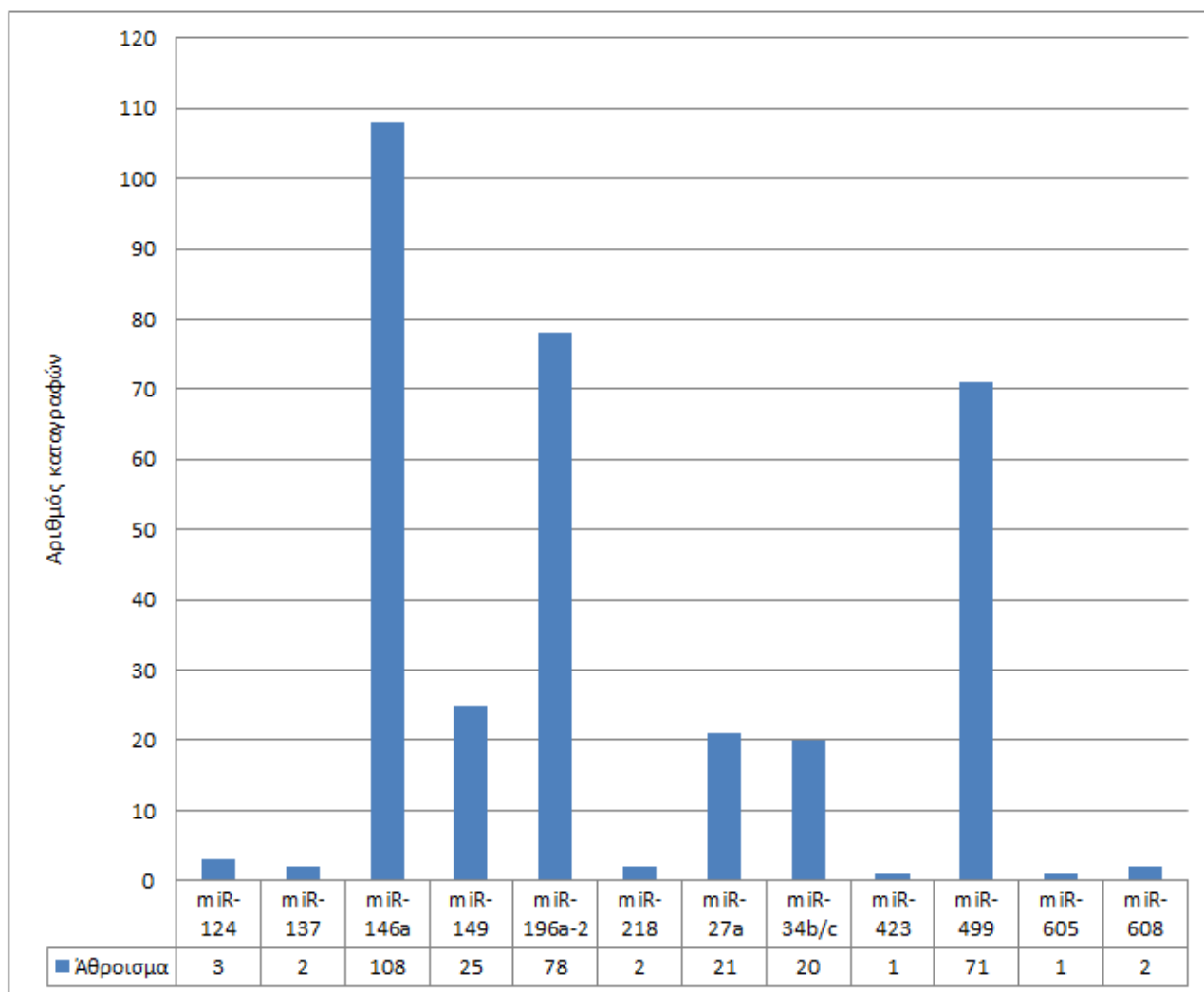
πραγματοποιήθηκε σε ασθενείς και υγιείς ασιατικής φυλής (Asian), ενώ μόνο μία μετα-ανάλυση περιλάμβανε έρευνες σε αμιγής καυκάσιους πληθυσμούς (Caucasian). Για την πολλαπλάσια ύπαρξη ερευνών με ασιατικής προέλευσης πληθυσμούς έναντι άλλων, ρόλο διαδραματίζει και το γεγονός πως η πλειοψηφία των μετα-αναλυτών, αλλά και των φορέων που υποστήριζαν τις μελέτες είναι από χώρες της Ασίας. [Γράφημα 3]



Γράφημα 3: Ποσοστιαία αναλογία εθνικότητας συμμετεχόντων σε καταγραφές

Όσο αφορά τα microRNAs που εξετάστηκαν στις μετα-αναλύσεις ήταν σύνολο 12, από τα οποία ορισμένα μελετούνταν με μεγαλύτερη συχνότητα, ενώ άλλα σε ελάχιστες μόνο μετα-αναλύσεις. Πιο συγκεκριμένα η πλειοψηφία των μετα-αναλύσεων ασχολήθηκε με 3 microRNAs, το miR-146a σε ποσοστό 32%, το miR-196a-2 σε ποσοστό 23% και το miR-499 σε ποσοστό 21%. Σε ένα μικρότερο αλλά ικανοποιητικό βαθμό μελετήθηκαν τα microRNAs miR-149, miR-27a, miR-34b/c, σε ποσοστά 7%, 6% και 6% αντίστοιχα ως προς το σύνολο το μετα-αναλύσεων. Τα υπόλοιπα microRNAs, miR-124, miR-137, miR-218, miR-423, miR-605, miR-608, αποτέλεσαν μόνο 1% των καταγραφών έκαστος. [Γράφημα 4]

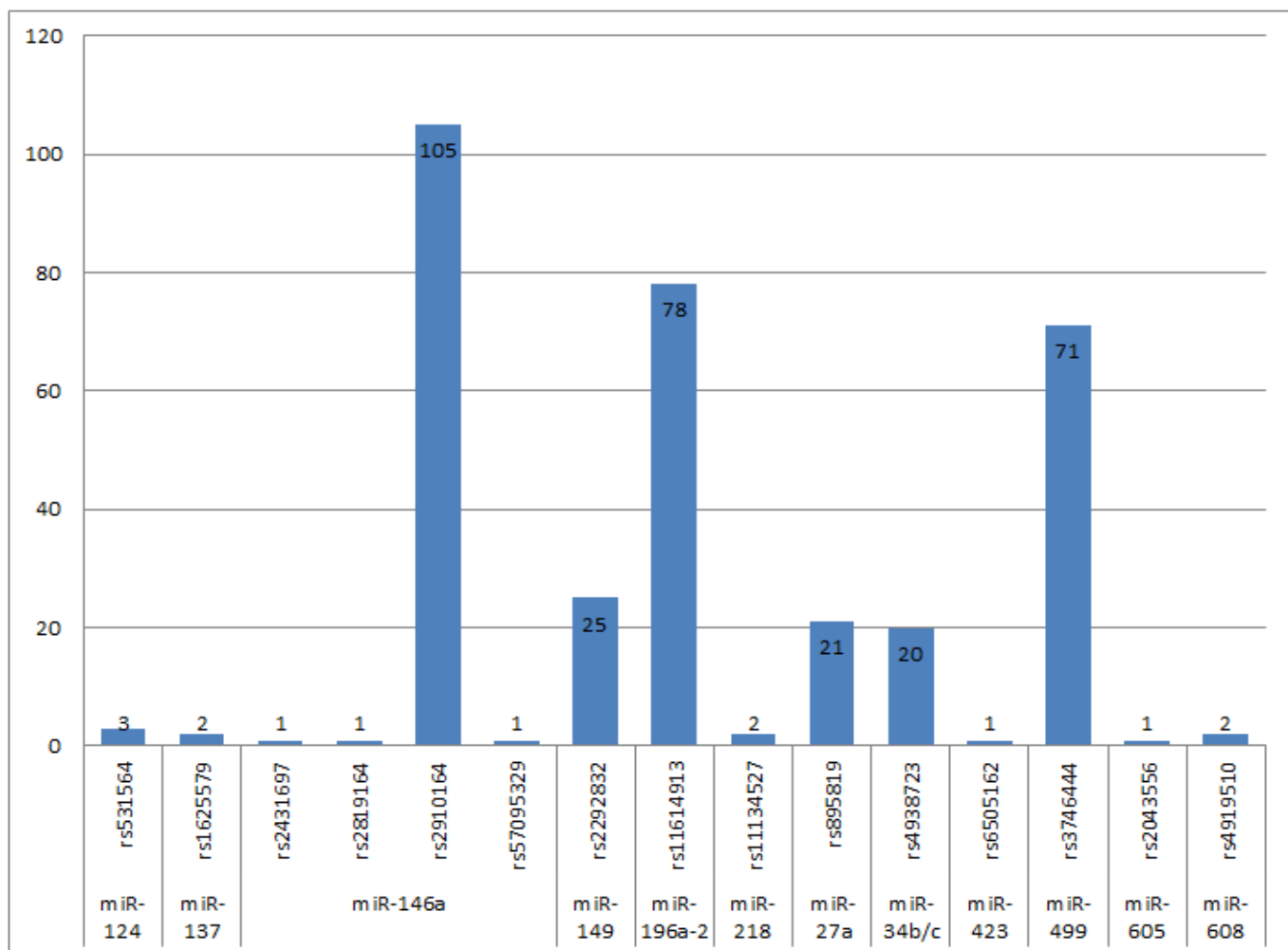
3. Αποτελέσματα



Γράφημα 4: Κατανομή των καταγραφών ως προς τα microRNAs

Σε αντίστοιχη λογική με τα microRNAs, είναι κατανεμημένοι και οι πολυμορφισμοί SNPs αυτών. Συνολικά υπάρχουν 15 SNPs, από τα οποία τα 3 συναντώνται σε ένα microRNA το miR-146a. Πάλι η πλειοψηφία των καταγραφών προέρχονταν από τρεις πολυμορφισμούς, στα τρία microRNAs που εξετάστηκαν περισσότερο. Ο SNP rs2910164 στο miR-146a σε ποσοστό 31% εκ των συνολικών μετα-αναλύσεων. Ο rs11614913 στο miR-196a-2 σε ποσοστό 23%. Ο rs3746444 στο miR-499 σε ποσοστό 21%. Ακλουθούν σε ποσοστό 6-7% έκαστος οι πολυμορφισμοί rs2292832, rs895819, rs4938723 στα microRNAs miR-149, miR-27a, miR-34b/c. Τέλος οι SNPs που το ποσοστό του καθένα δεν ξεπερνάει το 1% είναι rs531564, rs1625579, rs11134527, rs6505162, rs2043556, rs4919510 στα microRNAs , miR-124, miR-137, miR-218, miR-423, miR-605, miR-608 αντιστοίχως και τρεις μονονουκλεοτιδικό πολυμορφισμοί στο miR-146a, οι rs2431697, rs2819164 και rs57095329. [Γράφημα 5]

3. Αποτελέσματα

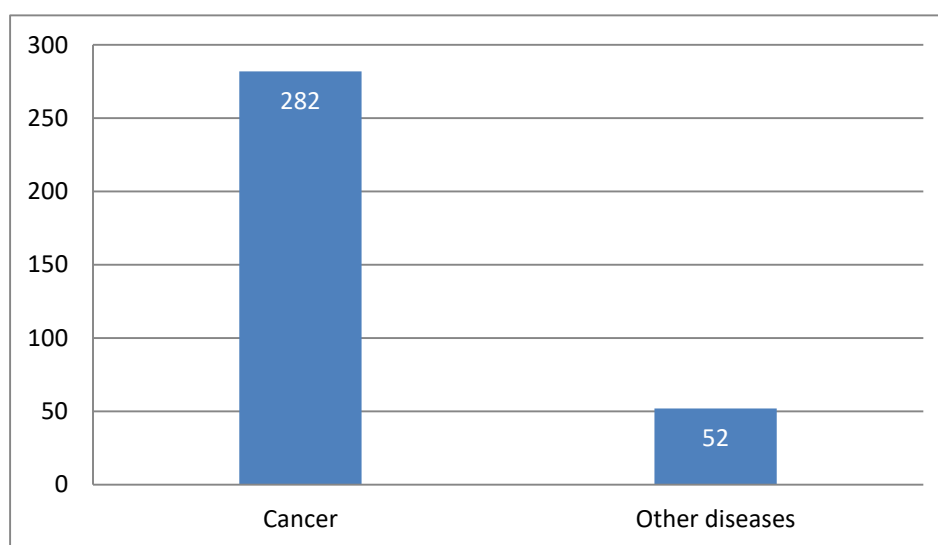


Γράφημα 5: Κατανομή των καταγραφών ως προς τους SNPs

Όσο αφορά τις ασθένειες, συνολικά καταγράφηκαν 34 ασθένειες και κατηγορίες ασθενειών, με πιο εξεταζόμενη τον καρκίνο, ως γενική κατηγορία, σε ποσοστό 20%. Σε σειρά συχνότητας εμφάνισης στις μετα-αναλύσεις που καταγράφηκαν ακολουθούν τα κακοήθη νεοπλάσματα του μαστού με ποσοστό 13%, ο καρκίνος του παχέος εντέρου με ποσοστό 11%, το ηπατοκυτταρικό καρκίνωμα με 10%, τα κακοήθη νεοπλάσματα του στομάχου με 9%, με 6% ο καρκίνος του πνεύμονα, 4% ο καρκίνος του πεπτικού συστήματος, με ποσοστό 3% έκαστος το ισχαιμικό εγκεφαλικό και ο καρκίνος στο συκώτι, η ρευματοειδή αρθρίτιδα, ο συστηματικός ερυθηματώδης λύκος, ο καρκίνος του οισοφάγου και διάφορα αυτοάνοσα νοσήματα καταλαμβάνουν ένα ποσοστό της τάξεως του 2% το καθένα από αυτά. Επιπλέον με ποσοστό μικρότερο της μονάδας έχουν καταγραφή μετα-αναλύσεις από πάνω από 10 ασθένειες όπως η σχιζοφρένεια, η πολλαπλή σκλήρυνση, η

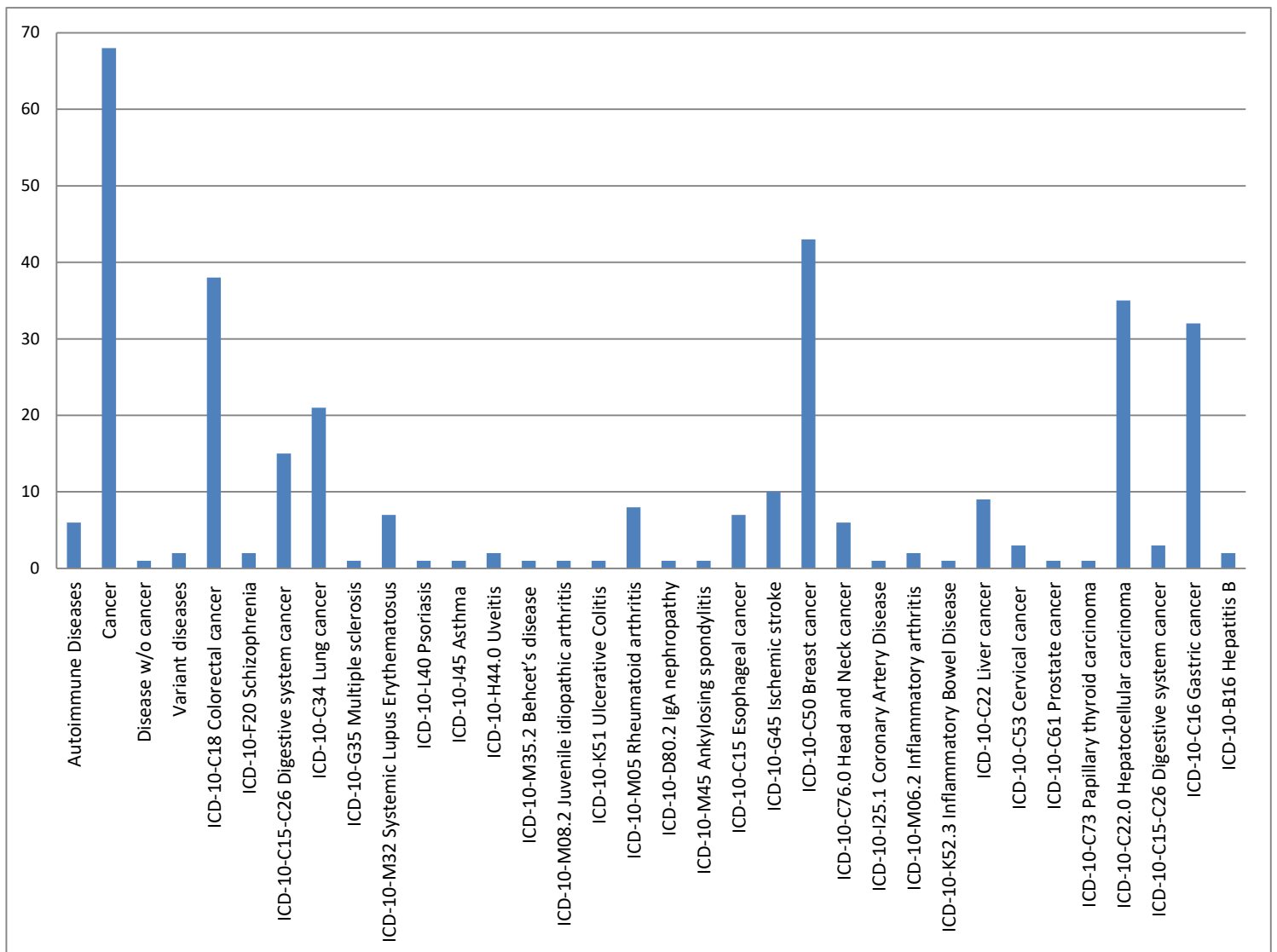
3. Αποτελέσματα

ψωρίαση, το άσθμα, ο ουβεΐτης, η νόσος Behcet, η νεανική αρθρίτιδα, η ελκώδης κολίτιδα, η ανεπάρκεια ανοσφαιρίνης A, η αγκυλοποιητική σπονδυλίτιδα, η νόσος στεφανιαίας αρτηρίας, η ρευματοειδής αρθρίτιδα, η λοιμώδης εντερίτιδα και κολίτιδα, ο καρκίνος τραχήλου της μήτρας, ο καρκίνος του προστάτη, καρκινώματα θυρεοειδούς αδένων, ο καρκίνος πεπτικού συστήματος, η ηπατίτιδα B και διάφορες άλλες ασθένειες εκτός καρκίνου. Ακόμα, αν χωριστούν οι ασθένειες σε 2 κατηγορίες, διάφοροι καρκίνοι και υπόλοιπες ασθένειες, η αναλογία είναι συντριπτική αφού το 84% των μετα-αναλύσεων διαπραγματεύεται καρκινώματα. [Γράφημα6][Γράφημα 7]



Γράφημα 6: Κατανομή των καταγραφών ως προς μετα-αναλύσεις σε καρκίνους και μη

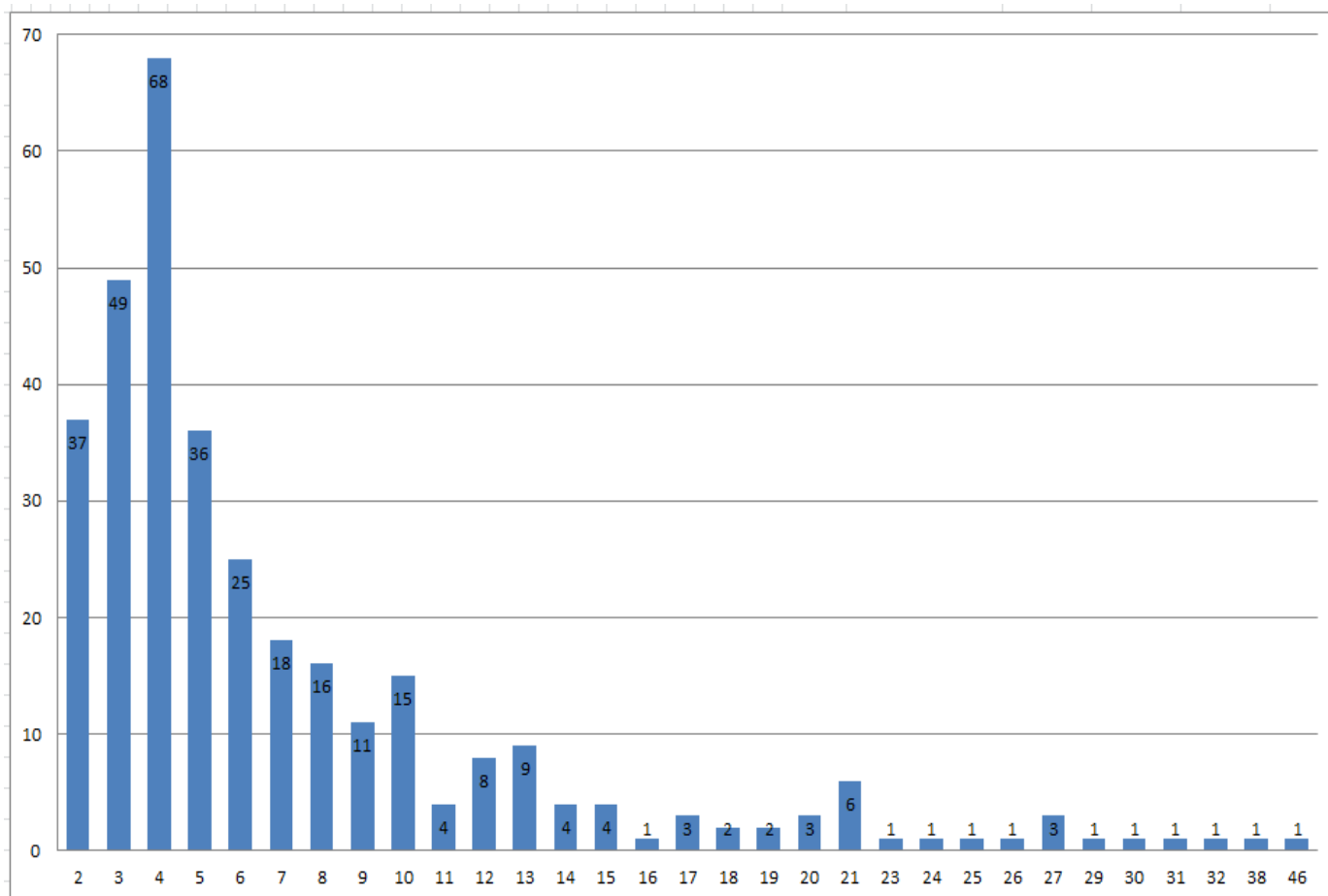
3. Αποτελέσματα



Γράφημα 7: Κατανομή των καταγραφών ως προς τις ασθένειες

Μελετώντας το πλήθος των κλινικών ερευνών, που περιλαμβάνει κάθε μετα-ανάλυση, κυμαίνεται σε εύρος από 2 έρευνες έως τον μέγιστο αριθμό των 46 ερευνών. Η πλειοψηφία των μετα-αναλύσεων περιλαμβάνουν από 2 έως 5 πρωτότυπες έρευνες. Πιο αναλυτικά ένα 20% των μετα-αναλύσεων έχουν πραγματοποιηθεί από τέσσερις case-control μελέτες, 15% από τρεις, 11% από δύο ή πέντε, 7% από έξι, 4%-5% από μετα-αναλύσεις με επτά, οχτώ ή δέκα έρευνες, 2% με 3% από εννιά, δώδεκα, δεκατρείς ή είκοσι μία πρωτότυπες έρευνες, ενώ το 1% περίπου καταλαμβάνει το καθένα από τα μη αναφερθέντα πιθανά νούμερα ερευνών στο εύρος 2 με 46. [Γράφημα 8]

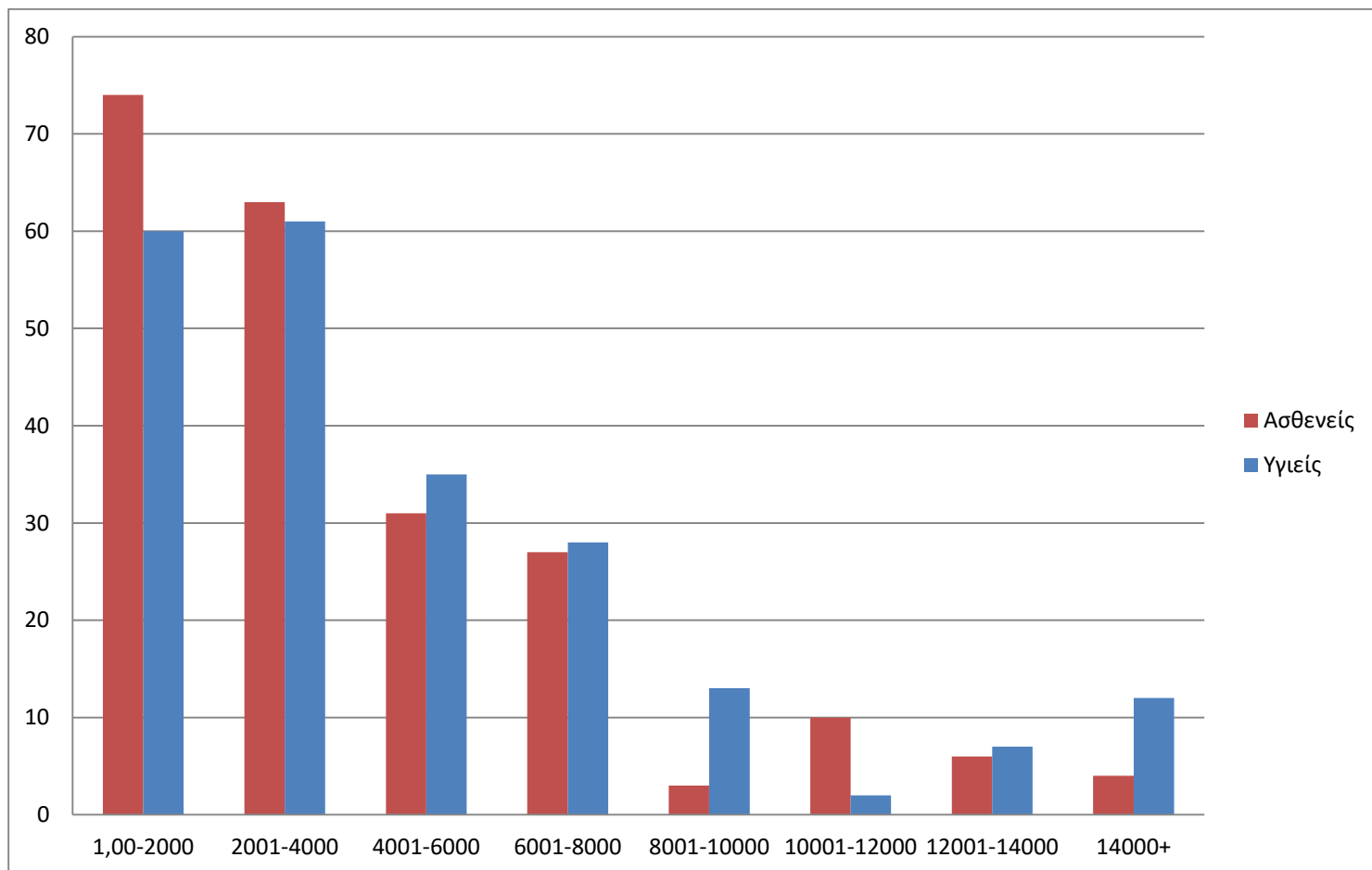
3. Αποτελέσματα



Γράφημα 8: Κατανομή αριθμού case-control studies σε κάθε μετανάλυση που καταγράφηκε

Συνολικά οι μετα-αναλύσεις που συμπεριλήφθηκαν περιλάμβαναν 836.441 ασθενείς και 979.600 υγιείς. Σε κάθε μετα-ανάλυση ο αριθμός των ασθενών και ο αριθμός των υγιών που συμμετείχαν κυμαινόταν για τους ασθενείς από 360 έως 25888, ενώ για τους υγιείς από 595 έως 25143. Επειδή το εύρος είναι μεγάλο και ένα απλό ραβδόγραμμα δεν θα μπορούσε να απεικονίσει ικανοποιητικά την κατανομή των ασθενών και υγιών στο πλήθος των καταγραφών κατασκευάστηκε ιστόγραμμα. Ένα ιστόγραμμα είναι μια γραφική παράσταση που απεικονίζει την διανομή συχνοτήτων ενός συνόλου συνεχών δεδομένων με στόχο την βαθύτερη κατανόηση της κατανομής τους. [Γράφημα 9]

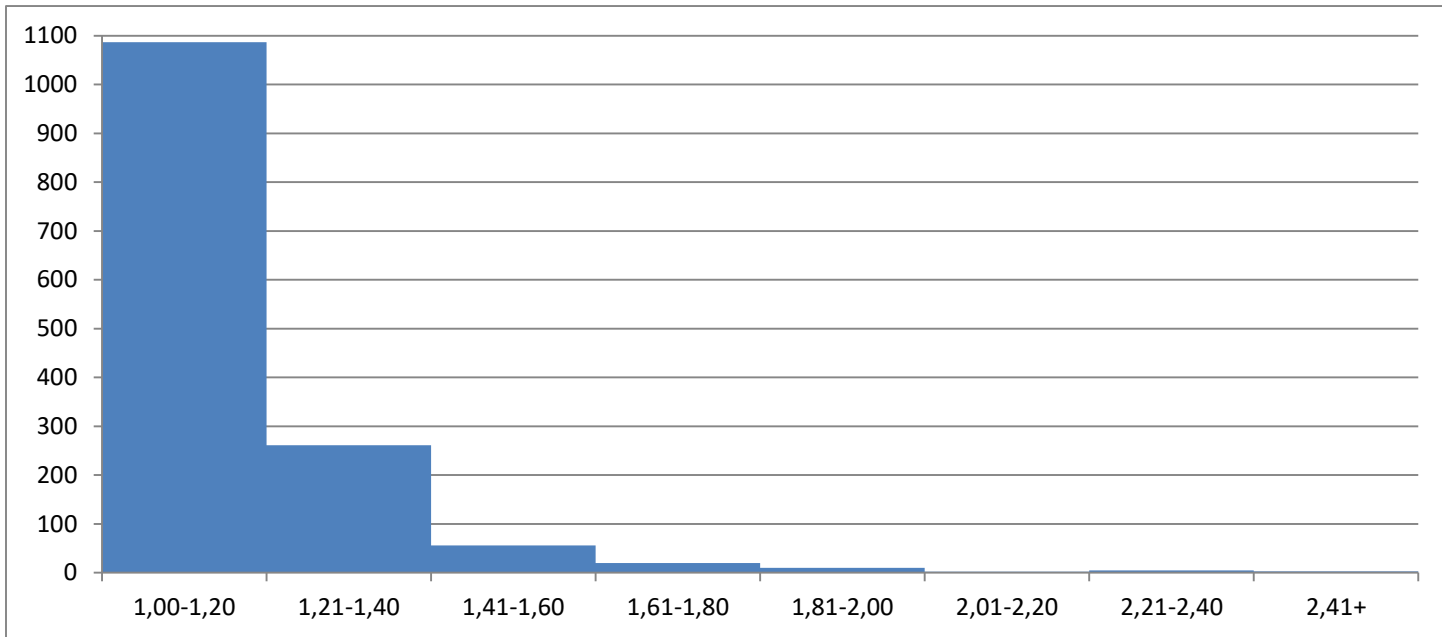
3. Αποτελέσματα



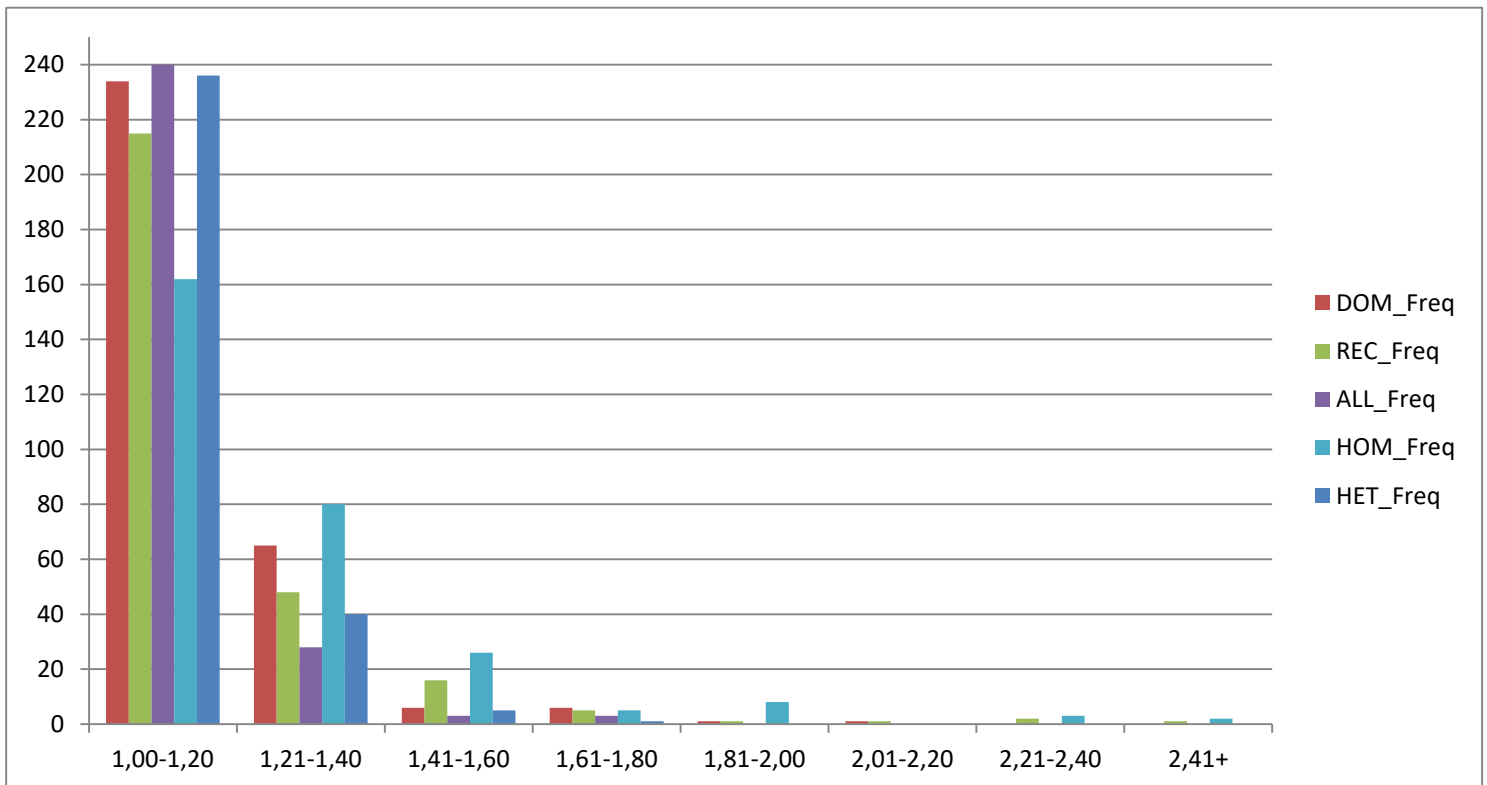
Γράφημα 9: Ιστόγραμμα συχνοτήτων αριθμού ασθενών και υγιών σε μετα-αναλύσεις

Επιπλέον ιστόγραμμα δημιουργήθηκε για την παρατήρηση των τιμών odds ratio από όλες τις καταγραφές και όλους τους γονοτύπους. Σημαντική παρατήρηση είναι ότι η κατανομή της τείνει προς την κανονική. Επιπλέον διάγραμμα συχνοτήτων δημιουργήθηκε με τα odds ratio κάθε γονοτύπου. Η ίδια παρατήρηση ως προς την κατανομή τους ισχύει και εδώ. Στην κανονική κατανομή παρατηρούμε συμμετρικότητα ως προς ένα σημείο, και η γραφική παράσταση του ιστογράμματος θυμίζει καμπάνα, γι' αυτό συχνά αναφέρατε και έτσι (bell curve). Είναι η πιο συχνά παρατηρούμενη κατανομή στην φύση και σε αποτελέσματα βιολογικών, ιατρικών και άλλων τομέων ερευνών. Στην παρούσα απεικόνιση εμφανίζεται μόνο το μισό, θετικό μέρος. [Γράφημα10] [Γράφημα 11]

3. Αποτελέσματα



Γράφημα 10: Ιστόγραμμα συχνοτήτων odds ratio όλων των καταγραφών



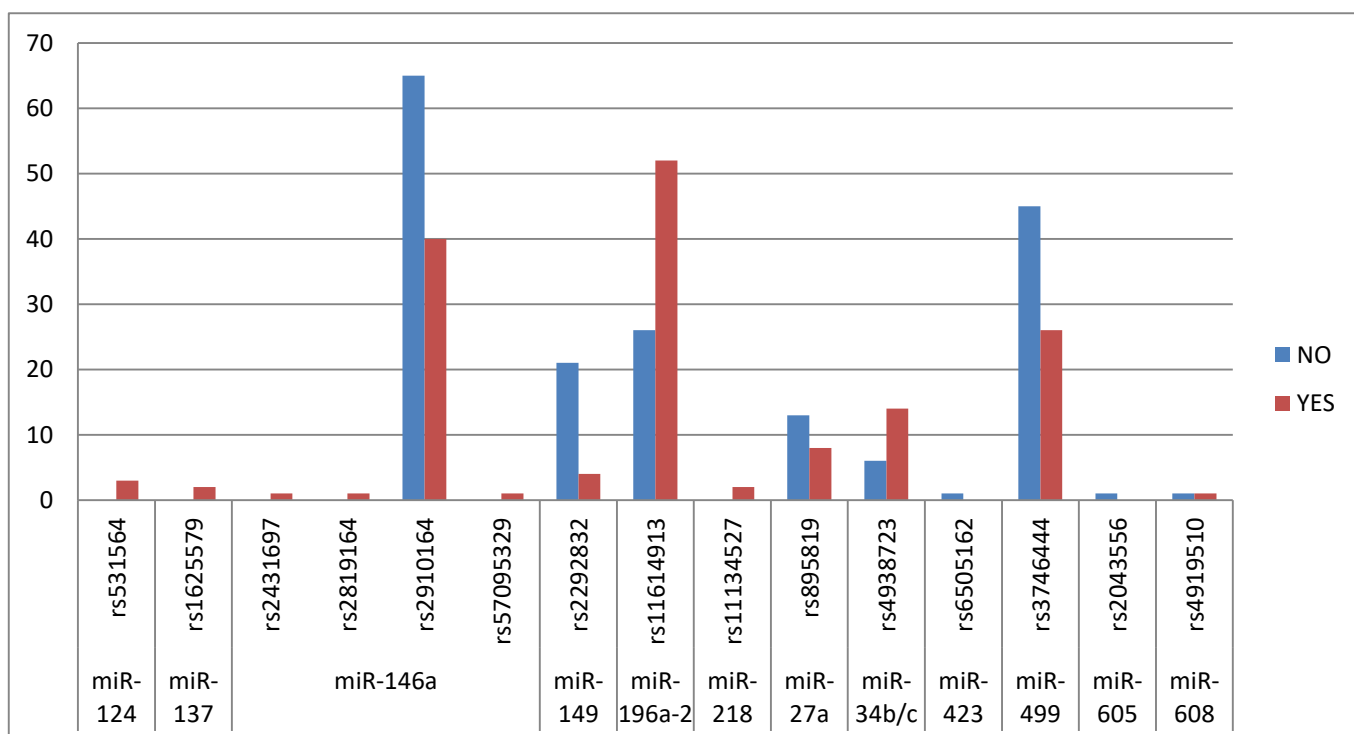
Γράφημα 11: Ιστόγραμμα συχνοτήτων odds ratio σε κάθε γονότυπο

3. Αποτελέσματα

Κάποια επιπλέον στοιχεία που είναι χρήσιμο να αναφερθούν σε αυτό το στάδιο είναι πως:

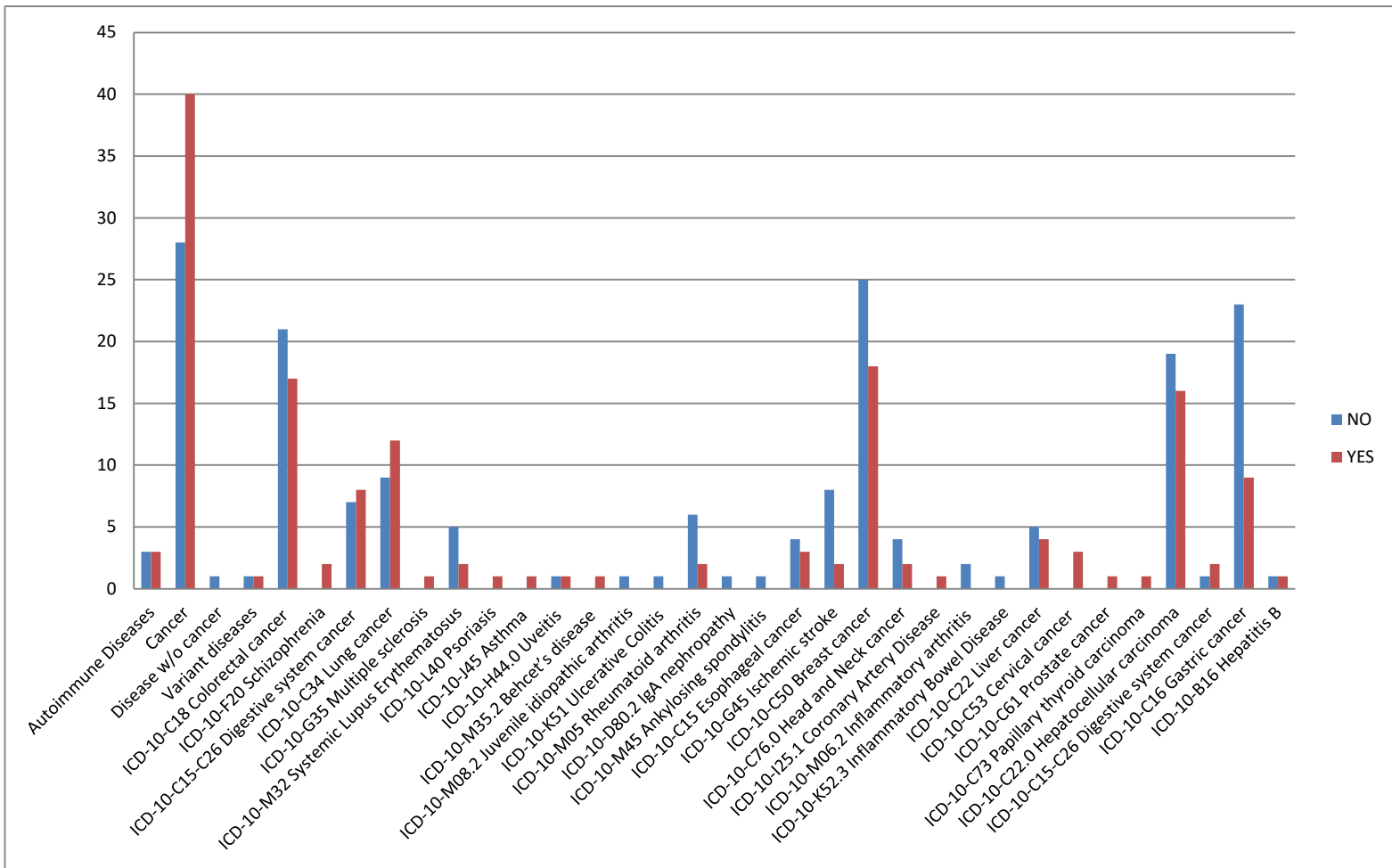
- Μόνο το 2% των καταγραφών δεν είχαν διαθέσιμα τα στοιχεία (Data) των ερευνών που συμμετείχαν στην μετα-ανάλυση τους.
- Το 94% των μετα-αναλύσεων μελέτησε την συσχέτιση του επικρατές γονότυπου, 87% τον υπολειπόμενο γονότυπο, 86% τον ομόζυγο, 85% τον ετερόζυγο ενώ λιγότερο απασχόλησε τους μελετητές, κατά ένα ποσοστό 82% ο έλεγχος για αλληλόμορφο.
- Το 100% των μετα-αναλύσεων που έχουν συμπεριληφθεί έχει πραγματοποιήσει bias test, για την εξάλειψη τυχών μεροληψιών.
- Η τιμή p _ association δεν έχει υπολογιστεί στο 29% των καταγραφών, ενώ το Z _ association παραλείπεται στην μεγάλη αναλογία του 85%.

Τελικά, από το σύνολο των μετα-αναλύσεων και όλων των παραπάνω συνθηκών και στατιστικών μέτρων που εφαρμόστηκαν, προέκυψε ότι το 54% των καταγραφών δεν απέδιδαν συσχέτιση μεταξύ ενός SNP σε microRNA και μίας ασθένειας, ενώ το 46% κατέληγε σε ύπαρξη θετικής σχέσης. [Γράφημα 12][Γράφημα 13]



Γράφημα 12: Κατανομή εγγραφών ως προς τους SNPs που συσχετίζονται θετικά και μη

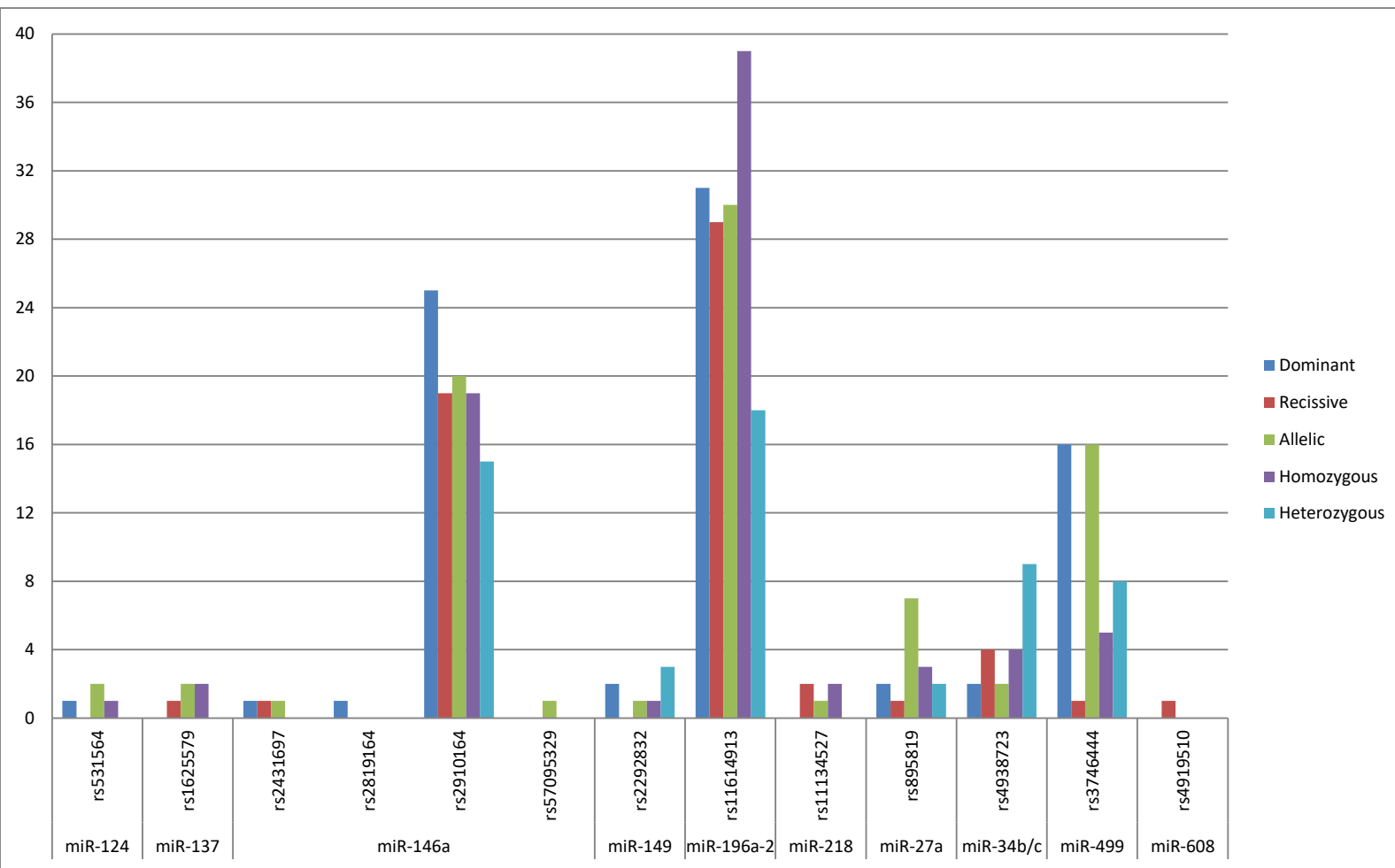
3. Αποτελέσματα



Γράφημα 13: Κατανομή εγγραφών ως προς τις ασθένειες που συσχετίζονται θετικά και μη

Επιπλέον για τους διάφορους ελέγχους μετα-ανάλυσης βρέθηκε πώς πιο συχνά γίνεται έλεγχος για το αλληλόμορφο σε ποσοστό 54%. Ακολουθεί με 52% ο έλεγχος για τον επικρατή χαρακτήρα, με 49% ο έλεγχος των ομόζυγων, με ένα ποσοστό 38% ο έλεγχος για τον υπολειπόμενο χαρακτήρα και τέλος ο έλεγχος ετερόζυγων με 35%. [Γράφημα 14]

3. Αποτελέσματα



Γράφημα 14: Κατανομή εγγραφών ως προς τις ασθένειες που συσχετίζονται και μη ανά γονότυπο

3.2. Έλεγχοι γενετικής συσχέτισης miR-SNPs με ασθένειες

Από την ολοκλήρωση των προηγούμενων βημάτων καταλήξαμε στην ύπαρξη συσχέτισης μεταξύ 13 μονονουκλεοτιδικών πολυμορφισμών σε microRNAs και 26 ασθενειών στον άνθρωπο. Ο πίνακας που ακολουθεί παρουσιάζει τα ζεύγη συσχέτισεων SNP με ασθένεια για το σύνολο το ευρημάτων [Πίνακας 1]. Όπως είναι εμφανές δύο SNPs εμφανίζουν συσχέτιση με μεγαλύτερο πλήθος ασθενειών, αλλά και έχουν αποτελέσει κα το θέμα του μεγαλύτερου μέρους των μετα-αναλύσεων που κατέληξαν σε ύπαρξη θετικής σχέσης. Αυτά είναι το rs2819164 στο miR-146a, το οποίο συσχετίζεται με την εμφάνιση 16 ασθενειών και τον καρκίνο ως γενική κατηγορία, σε ένα σύνολο 40 αναλύσεων, και το rs11614913 στο miR-196a-2, το οποίο συνδέεται με 9 ασθένειες και με τον καρκίνο γενικά, σε 52 μετα-αναλύσεις. Σημαντικά αποτελέσματα αποδίδουν και τα rs3746444, rs4938723, rs895819 στα microRNAs miR-499a, miR-34b/c και miR-27a αντίστοιχα. Ο πρώτος συνδέεται με 5 ασθένειες, τον καρκίνο ως ομάδα και ένα σύνολο ποικίλων ασθενειών, σε 26 έρευνες. Ο δεύτερος με 4 ασθένειες και τον καρκίνο σε 14 μελέτες, ενώ ο τρίτος με 3 και τον καρκίνο, σε 8 μετα-αναλύσεις. Τα υπόλοιπα ζεύγη μελετούνται σε αισθητά μικρότερο βαθμό.

3. Αποτελέσματα

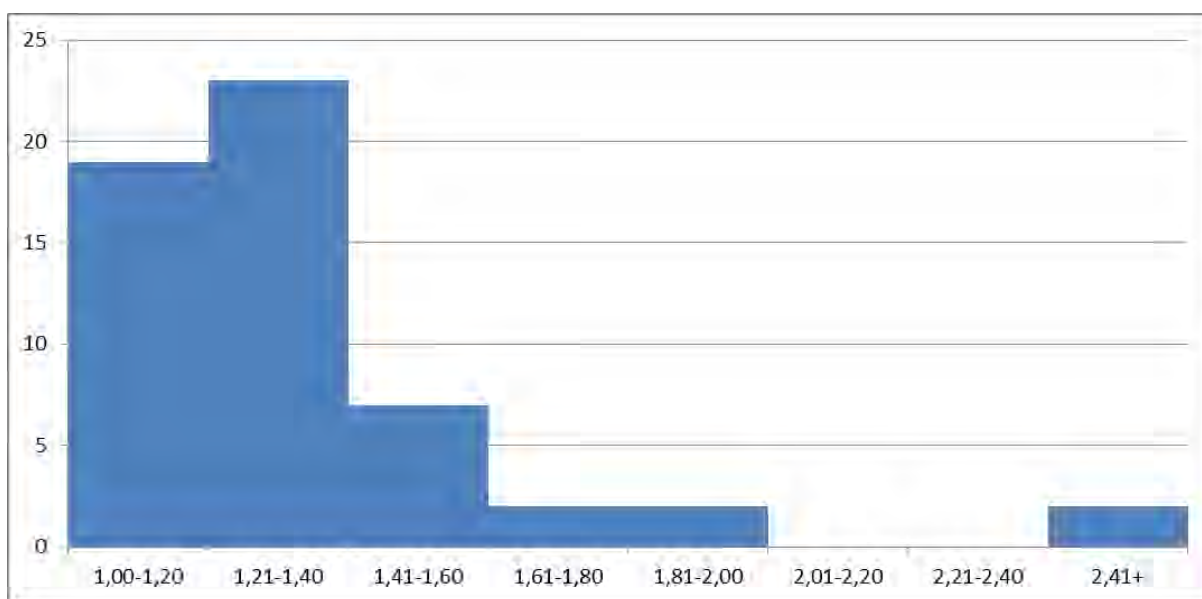
miRNA	Polymorphism	Disease	Meta-analysis
miR-124	rs531564	Cancer	2
		Esophageal cancer	1
miR-137	rs1625579	Schizophrenia	2
miR-146a	rs2431697	Systemic Lupus Erythematosus	1
	rs2910164	Behcet's disease	1
	rs57095329	Systemic Lupus Erythematosus	1
	rs2819164	Hepatocellular carcinoma	1
		Autoimmune Diseases	1
		Cancer	4
		Gastric cancer	8
		Digestive system cancer	4
		Liver cancer	1
		Hepatocellular carcinoma	8
		Lung cancer	2
		Cervical cancer	3
		Prostate cancer	1
		Papillary thyroid carcinoma	1
		Head and Neck cancer	1
		Multiple sclerosis	1
	Uveitis	1	
	Coronary Artery Disease	1	
	Asthma	1	
	Psoriasis	1	
miR-149	rs2292832	Cancer	1
		Gastric cancer	1
		Colorectal cancer	1
		Ischemic stroke	1
miR-196a-2	rs11614913	Cancer	11
		Hepatitis B	1
		Digestive system cancer	3
		Colorectal cancer	10
		Liver cancer	2
		Hepatocellular carcinoma	5
		Lung cancer	10
		Breast cancer	8
		Head and Neck cancer	1
Ischemic stroke	1		
miR-218	rs11134527	Cancer	2
miR-27a	rs895819	Cancer	1
		Digestive system cancer	1
		Colorectal cancer	3
		Breast cancer	3

3. Αποτελέσματα

miR-34b/c	rs4938723	Cancer	3
		Esophageal cancer	1
		Digestive system cancer	2
		Colorectal cancer	3
		Hepatocellular carcinoma	2
miR-499a	rs3746444	Autoimmune Diseases	2
		Cancer	12
		Esophageal cancer	1
		Liver cancer	1
		Breast cancer	7
		Rheumatoid arthritis	2
		Variant diseases	1
miR-608	rs4919510	Cancer	1

Πίνακας 1: Συγκεντρωτικός πίνακας αποτελεσμάτων συσχετίσεων SNPs σε microRNAs με ασθένειες. Στην τελευταία στήλη αναγράφεται ο αριθμός των μετα-αναλύσεων που απέδειξαν θετική σχέση για κάθε ζεύγος SNP/ασθένεια.

Για κάθε ζεύγος επιλέχθηκε μία καταγραφή η οποία κρίθηκε πιο αντιπροσωπευτική βάση κάλυψης όλων των σχετικών άρθρων, ημερομηνίας δημοσίευσης και πλήθος δείγματος. Από αυτήν την καταγραφή επιλέχθηκε η μέγιστη τιμή Odds ratio, σε όποιον έλεγχο γονοτύπου αυτή βρισκόταν, ως χαρακτηριστικό του βαθμού συσχέτισης. [Γράφημα 15]



Γράφημα 15: Ιστόγραμμα συχνότητας επιλεγμένων odds ratio σε συσχετιζόμενα ζεύγη SNP:ασθένεια








3.3 Δημιουργία, Ανάλυση και Αξιολόγηση δικτύου γενετικής συσχέτισης

Επειδή η απεικόνιση των συσχετίσεων με την χρήση κλασικών γραφημάτων δεν απέδιδε πλήρως τα αποτελέσματα, δημιουργήθηκε με την χρήση του προγράμματος Cytoscape, δίκτυο με τις σχέσεις των 11 μονονουκλεοτιδικών πολυμορφισμών σε microRNAs και των 23 ασθενειών. Στο δίκτυο δεν συμπεριλήφθηκαν οι κατηγορίες ασθενειών “καρκίνος”, “αυτοάνοσα νοσήματα” και “διάφορες ασθένειες”, καθώς περιέπλεκαν άσκοπα το δίκτυο. Το SNP:disease δίκτυο που δημιουργήθηκε, λοιπόν, είναι ένα διμερές γράφημα στο οποίο ένα SNP συνδέεται με μία ασθένεια, αν έχει αποδειχθεί από την ανασκόπηση που πραγματοποιήθηκε ότι υπάρχει μεταξύ τους θετική συσχέτιση.

Είναι εύκολα ορατό το γεγονός πως συγκεκριμένα SNP παρουσιάζουν συσχέτιση με πλήθος ασθενειών, ενώ άλλα με ελάχιστες μόνο νόσους. Πιο αναλυτικά με περισσότερους κόμβους ασθενειών, 15 σε αριθμό, συνδέεται ο rs2910164. Με 9 ασθένειες συνδέεται ο πολυμορφισμός rs11614913, ενώ ακολουθούν με 4 ακμές έκαστος οι SNPs rs3746444 και rs4938723. Ακόμα με 3 ασθένειες συνδέονται οι rs2292832 και rs895819. Η πλειοψηφία των μονονουκλεοτιδικών πολυμορφισμών έχει μόνο μία ακμή, άρα σύνδεση με μόνο μία ασθένεια.

Μελετώντας τον διαμοιρασμό των ασθενειών στο δίκτυο, κάθε μία συνδέεται κατά ελάχιστο με 1 πολυμορφισμό και μέγιστο με 4. Ο καρκίνος του πεπτικού συστήματος και ο καρκίνος του παχέος εντέρου είναι οι παθήσεις που έχουν το μέγιστο αριθμό 4 ακμών. Ασθένειες που συνδέονται με 3 SNPs είναι τα κακοήγη νεοπλασμάτα του μαστού, ο καρκίνος του οισοφάγου και ο καρκίνος στο συκώτι. Οι υπόλοιπες ασθένειες έχουν μόλις 1 ή 2 ακμές.

Περιγράφοντας το δίκτυο μορφολογικά παρουσιάζονται:

- Με σχήμα ρόμβου και κόκκινο χρώμα  το όνομα του SNP πολυμορφισμού
- Με κύκλο  το όνομα της ασθένειας και ανάλογα με το πόσες ακμές έχει τα χρώματα που λαμβάνει είναι:
 -  όταν συνδέεται με μία ακμή
 -  όταν συνδέεται με δύο ακμές
 -  όταν συνδέεται με τρεις ακμές
 -  όταν συνδέεται με τέσσερις ακμές
- Στις ακμές, ευθείες της μορφής , αναγράφονται οι τιμές του Odds ratio, που έχει επιλεχθεί ως αντιπροσωπευτικό.

3. Αποτελέσματα

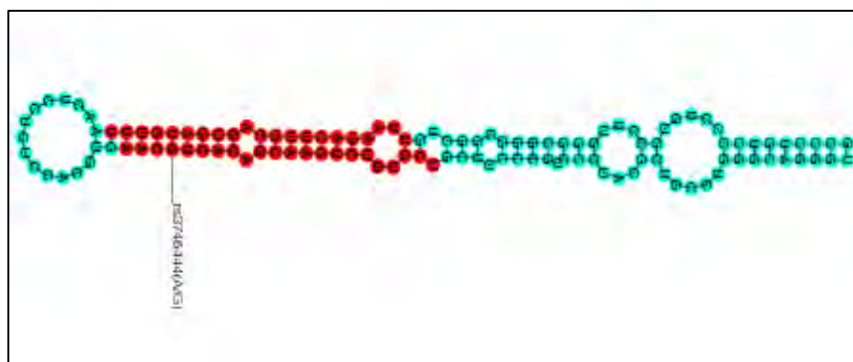
3.3.1 Αξιολόγηση με δεδομένα από MIRNASNP2

Για την καλύτερη κατανόηση και περαιτέρω επεξεργασία των ευρημάτων, εισάγαμε στην βάση δεδομένων miRNASNP κάθε ένα από τους SNPs που συσχετίστηκαν με ασθένειες. Από τους 13 πολυμορφισμούς μόνο οι 8 ήταν καταχωρημένοι στα δεδομένα της βάσης για τους οποίους και ανακτήσαμε βασικά τους στοιχεία.

Από τους 8 αυτούς SNPs οι 6 βρίσκονται σε pre-miRNA και συγκεκριμένα:

- οι 2 σε seed regions
 - rs3746444

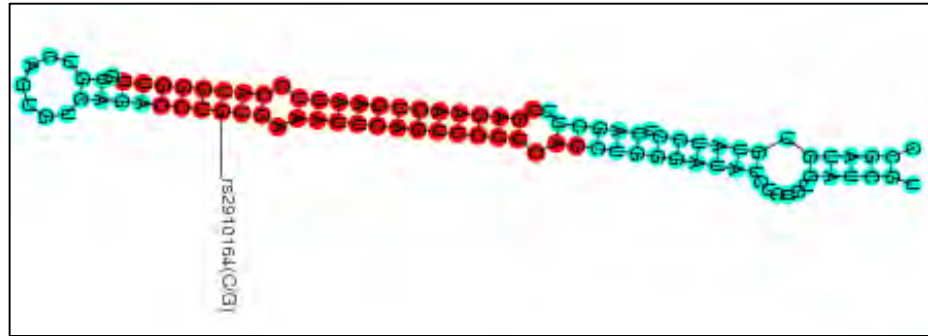
Chromosome	Chr20
Position	33578251
SNP Location relative to pre-miRNA	73
Reference allele	A/G



- rs2910164

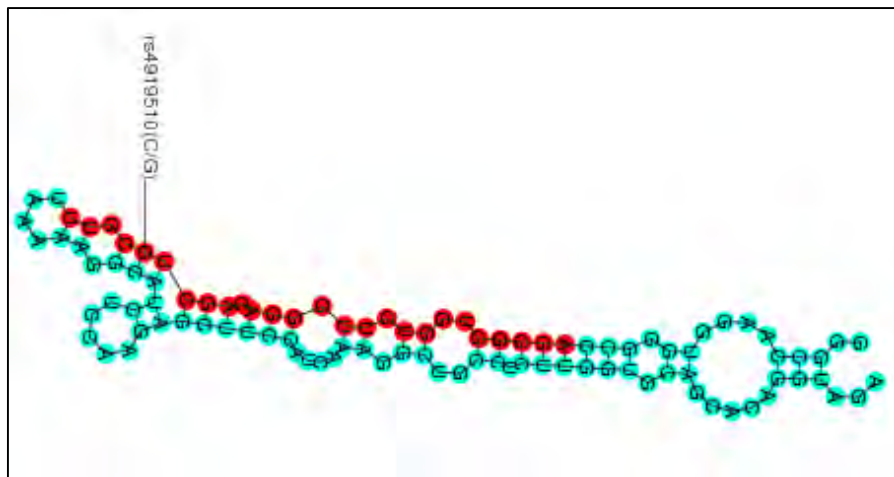
Chromosome	Chr15
Position	159912418
SNP Location relative to pre-miRNA	60
Reference allele	C/G

3. Αποτελέσματα



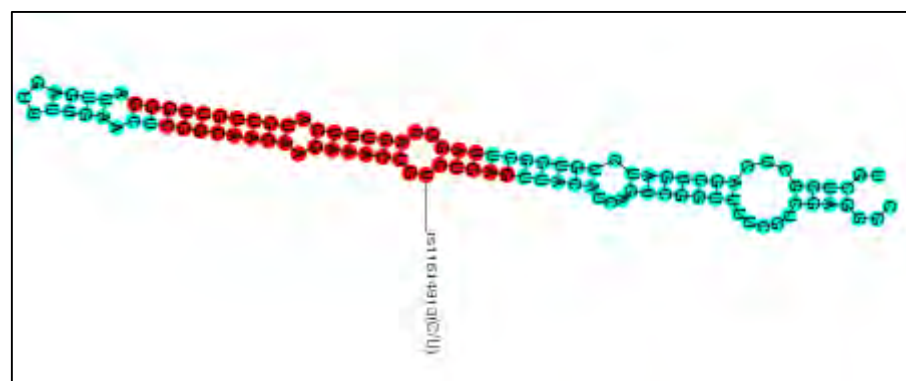
- οι 2 στο ώριμο pre-miRNA
 - rs4919510

Chromosome	Chr10
Position	102734778
SNP Location relative to pre-miRNA	37
Reference allele	C/G



- rs11614913

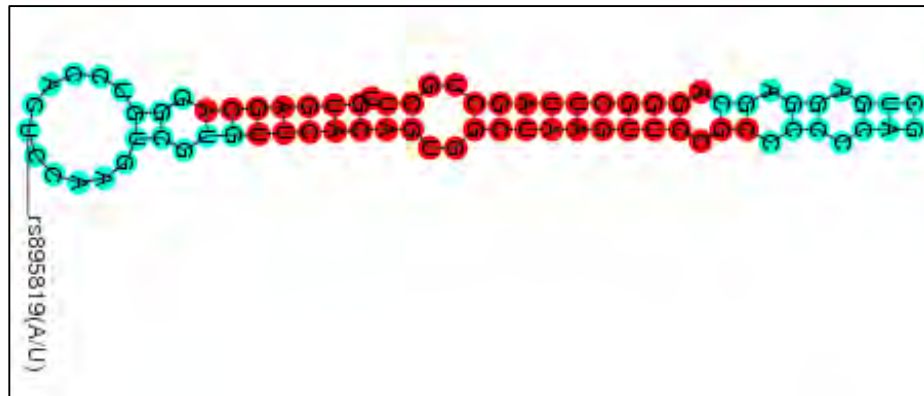
Chromosome	Chr12
Position	54385599
SNP Location relative to pre-miRNA	78
Reference allele	C/T



3. Αποτελέσματα

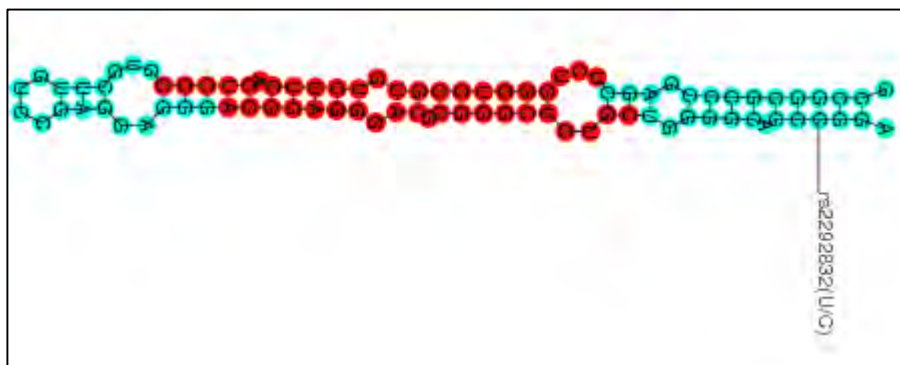
- οι 2 στο pre-miRNA γενικά
 - rs895819

Chromosome	Chr19
Position	13947292
SNP Location relative to pre-miRNA	40
Reference allele	T/A



- rs2292832

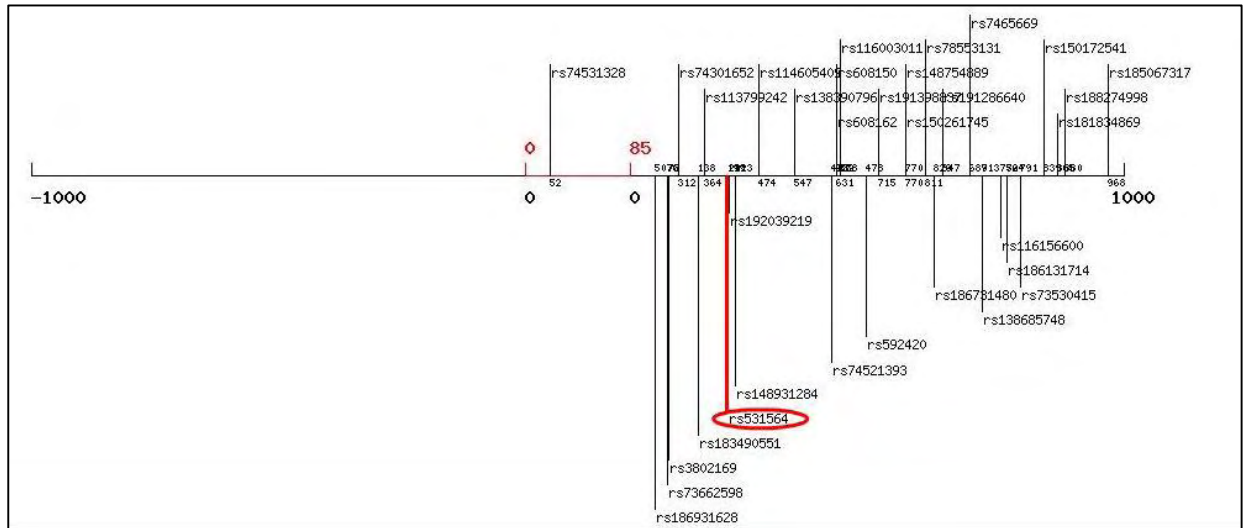
Chromosome	Chr2
Position	241395503
SNP Location relative to pre-miRNA	86
Reference allele	T/C



3. Αποτελέσματα

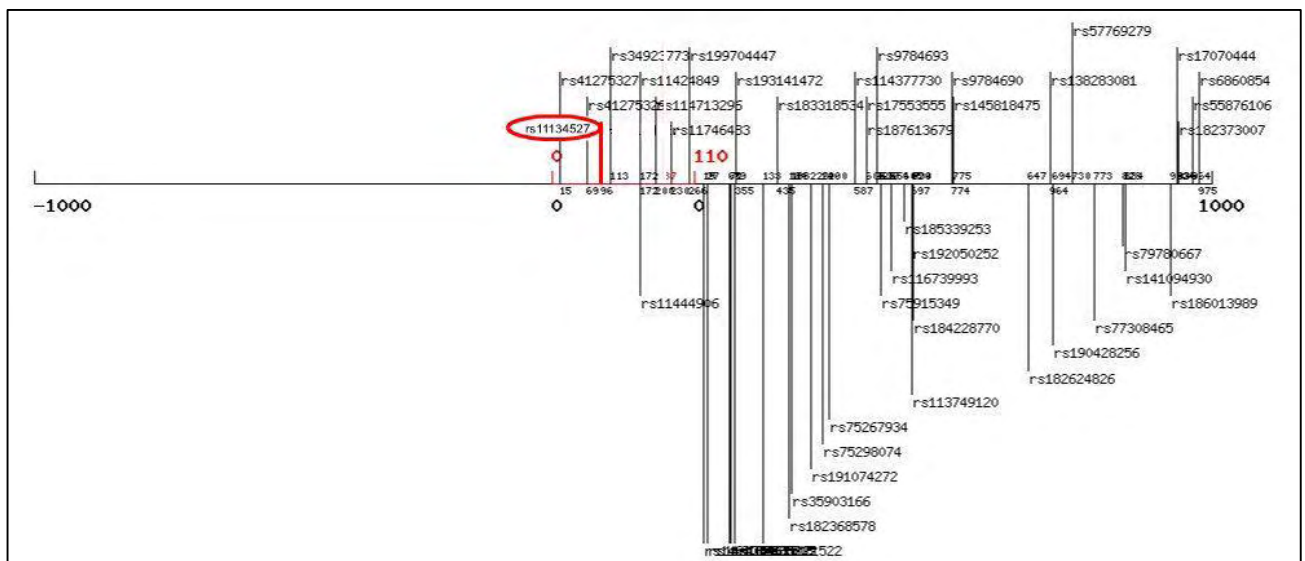
- Ακόμα 2 πολυμορφισμοί συναντούνται στις πλευρικές περιοχές (flanking regions) του pre-miRNA.
 - rs11134527

Chromosome	Chr5
Position	168195356
SNP Location relative to 5' pre-miRNA	96
Reference allele	G/A



- rs531564

Chromosome	Chr8
Position	9760699
SNP Location relative to 5' pre-miRNA	199
Reference allele	G/C



3. Αποτελέσματα

Για όσα SNPs δεν μπορέσαμε να βρούμε δεδομένα στην miRNASNP, αναζητήσαμε στην βάση δεδομένων dbSNP και καταγράψαμε κάποια βασικά στοιχεία για το καθένα από τα 5.

SNP	miRNA	Chromosome	Position	Reference Allele
rs1625579	miR-137	1	98502934	G/T
rs11134527	miR-218	5	16876835	G/A
rs2431697	miR-146a	5	160452971	T/C
rs2819164	miR-146a	1	64186383	C/T
rs57095329	miR-146a	5	160467840	A/G

Ένα επιπλέον δεδομένο, που προσφέρει η βάση δεδομένων miRNASNP2, είναι η ένδειξη αλλαγής στην έκφραση του microRNA σε σύνδεση με 28 μορφές καρκίνου, σύμφωνα με την μέθοδο RPKM. Αξιοποιώντας αυτή την δυνατότητα, για όσους πολυμορφισμούς είχε αυτή την επιλογή η miRNASNP2, συγκρίναμε την ένδειξη την ένδειξη απορύθμισης έκφρασης με δικά μας αποτελέσματα. Πιο αναλυτικά για κάθε ζευγάρι SNP:ασθένεια, για το οποίο εμείς έχουμε καταλήξει ότι υπάρχει θετική συσχέτιση, ελέγξαμε αν και η βάση δεδομένων παρουσιάζει αλλαγές στην έκφραση του SNP στην συγκεκριμένη μορφή καρκίνου. Καταλήξαμε σε ένα ποσοστό επιβεβαίωσης της τάξεως του 92%. Τα δεδομένα αυτά συνοψίστηκαν και παρουσιάζονται στον Πίνακα 2. Επιπρόσθετα υπάρχουν και 21 ζευγάρια SNP:ασθένεια, για παθήσεις εκτός καρκίνου, για τις οποίες η βάση δεδομένων δεν έχει στοιχεία.

3. Αποτελέσματα

miRNA	Polymorphism	Associated diseases according present study	Deregulated expression of SNP according MIRNASNP2 and association in present study
miR-146a	rs2910164	Autoimmune Diseases	N/A
		Gastric cancer	√
		Digestive system cancer	√
		Liver cancer	√
		Hepatocellular carcinoma	√
		Lung cancer	√
		Cervical cancer	√
		Prostate cancer	×
		Papillary thyroid carcinoma	√
		Head and Neck cancer	√
		Multiple sclerosis	N/A
		Uveitis	N/A
		Coronary Artery Disease	N/A
		Asthma	N/A
		Psoriasis	N/A
Behcet's disease	N/A		
miR-149	rs2292832	Gastric cancer	√
		Colorectal cancer	√
		Ischemic stroke	N/A
miR-196a-2	rs11614913	Hepatitis B	N/A
		Digestive system cancer	√
		Colorectal cancer	√
		Liver cancer	√
		Hepatocellular carcinoma	√
		Lung cancer	√
		Breast cancer	√
		Head and Neck cancer	√
		Ischemic stroke	N/A
miR-27a	rs895819	Digestive system cancer	√
		Colorectal cancer	√
		Breast cancer	√
miR-499a	rs3746444	Autoimmune Diseases	N/A
		Cancer	√
		Esophageal cancer	×
		Liver cancer	√
		Breast cancer	√
Rheumatoid arthritis	N/A		

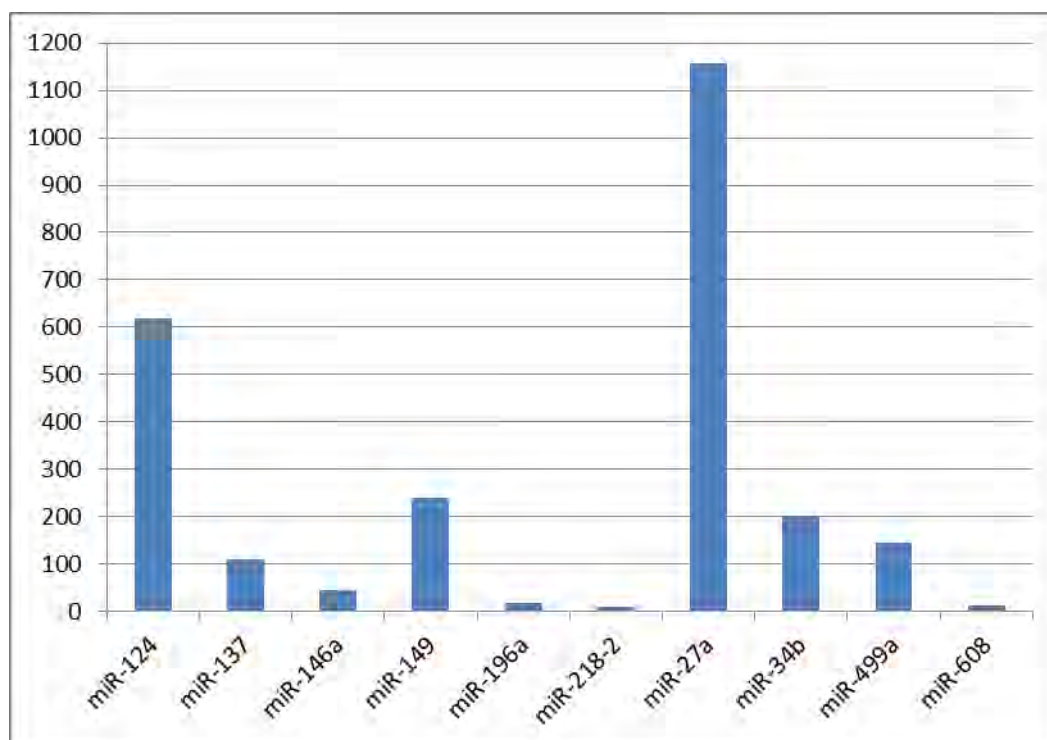
Πίνακας 2: Πίνακας σύγκρισης αποτελεσμάτων παρούσας μελέτης με δεδομένα έκφρασης από MIRNASNP2. Η ένδειξη N/A (not available) αποδόθηκε σε ζευγάρια, για τα οποία λείπουν δεδομένα από την βάση.

3. Αποτελέσματα

3.3.2 Αποτελέσματα από ανάλυση με Diana tools

Για κάθε ένα από τα microRNA, στα οποία βρίσκονται τα συσχετιζόμενα SNPs, έγινε αναζήτηση τους στο πακέτο εργαλείων Diana, ώστε να βρεθεί η σχέση τους τόσο με γονίδια, όσο και με βιολογικά μονοπάτια.

Όσο αφορά τους στόχους στους οποίους συνδέονται τα microRNAs που μας απασχόλησαν είναι σύνολο 2033 γονίδια και δημιουργούν, με τα 10 microRNAs, 2556 ζεύγη ενδιαφέροντος. Οι συνδέσεις αυτές δεν είναι ισότιμα κατανομημένες, καθώς το 45% των γονιδίων συνδέονται με το miR-27a, στο οποίο βρίσκεται και ένας από τους πέντε πιο συσχετιζόμενους με ασθένεια πολυμορφισμούς, σύμφωνα με τα αποτελέσματα μας. Αντίθετα, παρόλο που το miR-124 συνδέεται με το αξιοσημείωτο ποσοστό των 24% εκ των όλων των συνδέσεων, στην μελέτη βρέθηκε σύνδεση μόνο σε 2 μετα-αναλύσεις. Επιπλέον αναντιστοιχία σε αποτελέσματα συναντάται και για τα miR-196a και miR-146a, τα οποία φαίνεται να συνδέονται με γονίδια σε ένα ποσοστό της τάξεως 1-2%, ενώ αποτελούν, σύμφωνα με την παρούσα μελέτη, τα πιο εξεταζόμενα και συσχετιζόμενα microRNAs όσο αφορά τους SNPs που φέρουν. Σε σχετική συμφωνία με τα ευρήματα των μετα-αναλύσεων βρίσκονται τα ποσοστά των miR-149, miR-34b και miR-499a, τα οποία είναι 9%, 8% και 6% αντίστοιχα. Τέλος τα microRNAs miR-137, miR-218-2, miR-608 απαρτίζουν το 2-4% των ευρημάτων έκαστος. [Γράφημα 16]

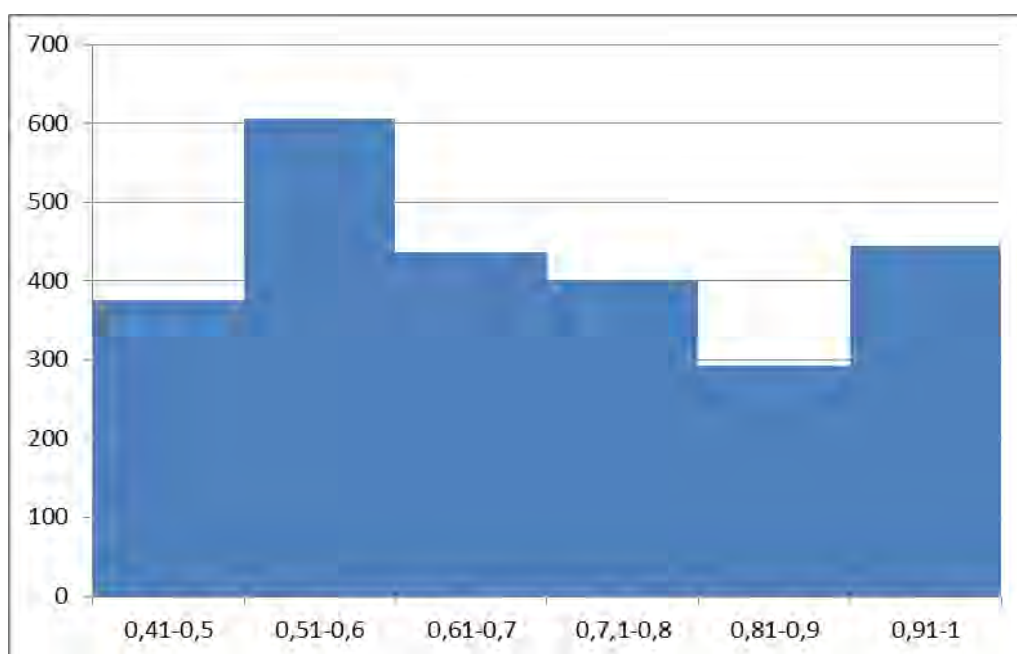


Γράφημα 16: Αριθμός γονιδίων που σχετίζονται με κάθε microRNA

3. Αποτελέσματα

Ομοιόμορφα είναι διαμοιρασμένες οι συνδέσεις ως προς τα γονίδια, αφού κάθε γονίδιο συνδέεται με 1 έως 3 microRNAs.

Για κάθε ζεύγος microRNA με γονίδιο υπάρχει ένας αριθμός που χαρακτηρίζει το βαθμό της πρόβλεψης, όπως αυτός προέκυψε από τα δεδομένα του Diana. Οι βαθμοί κυμαίνονται από το 0,45-1. Ο βαθμός 0,45 είναι το κατώφλι που θέσαμε, ώστε να θεωρηθεί ότι είναι σημαντική η πρόβλεψη σύνδεσης. Παρακάτω παρουσιάζεται το ιστόγραμμα συχνοτήτων των τιμών αυτών. Όπως παρουσιάζεται, η συχνότητα κορυφώνεται στις τιμές 0,6-1, ενώ δεύτερη τοπική κορυφή εμφανίζεται στις τιμές 0,9-1, όπου είναι και ο μέγιστος βαθμός πρόβλεψης. Όλα τα microRNAs έχουν τουλάχιστον σύνδεση με ένα γονίδιο με σκορ πρόβλεψης πάνω από 0,9. Η ύπαρξη μεγάλου πλήθους βαθμών πρόβλεψης στην κλίμακα 0,91-1 αυξάνει την εγκυρότητα της σύνδεσης του miRNA με το γονίδιο. [Γράφημα 17]



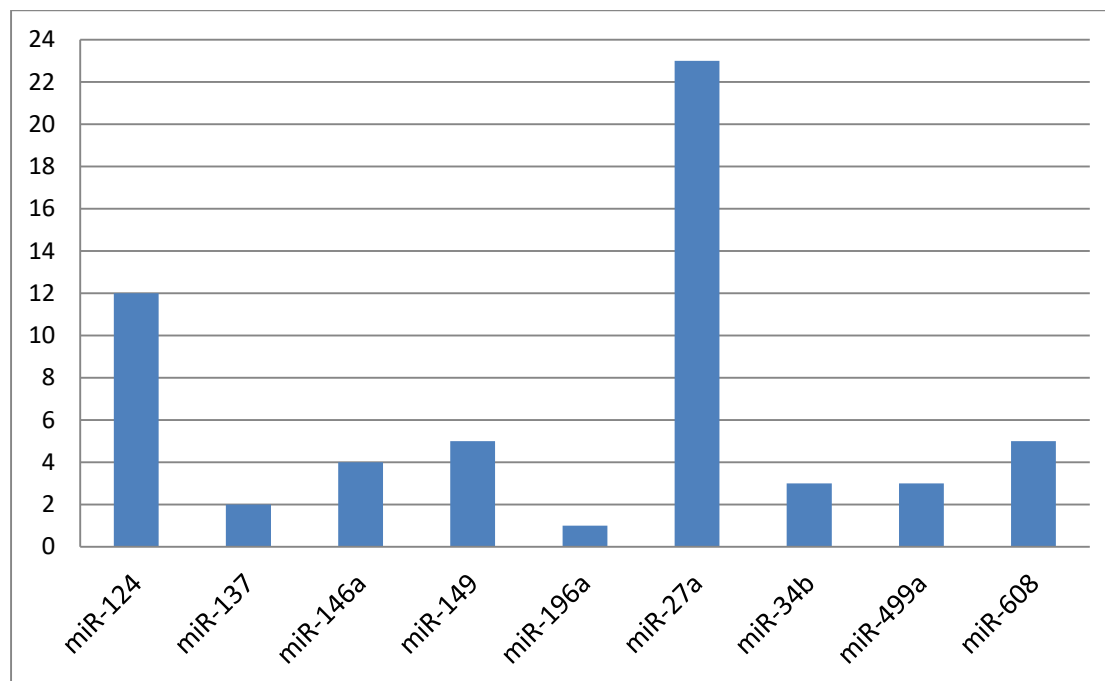
Γράφημα 17: Ιστόγραμμα συχνοτήτων Prediction score για σχέση microRNAs και γονιδίων

Αναζητώντας, στην συνέχεια, την σχέση των microRNAs με βιολογικά μονοπάτια, καταλήξαμε στην εύρεση 58 συνδέσεων miRNA:Pathways, που περιλάμβαναν όλα τα microRNAs στα οποία έχουμε βρει πολυμορφισμούς συσχετιζόμενους με μία ασθένεια, εκτός του miR-218 για το οποίο δεν υπάρχει κάποιο καταχωρημένο μονοπάτι.

Η πλειοψηφία των μονοπατιών σχετιζόταν με 2 microRNAs, το miR-27a και το miR-124, σε μια αναλογία του 40% και 21% αντίστοιχα. Σαν αποτέλεσμα είναι σχετικά αναμενόμενο καθώς είναι και τα δύο microRNAs που έχει προβλεφθεί πως έχουν και το μεγαλύτερο πλήθος συνδέσεων με γονίδια στόχους, ανάμεσα σε αυτά που

3. Αποτελέσματα

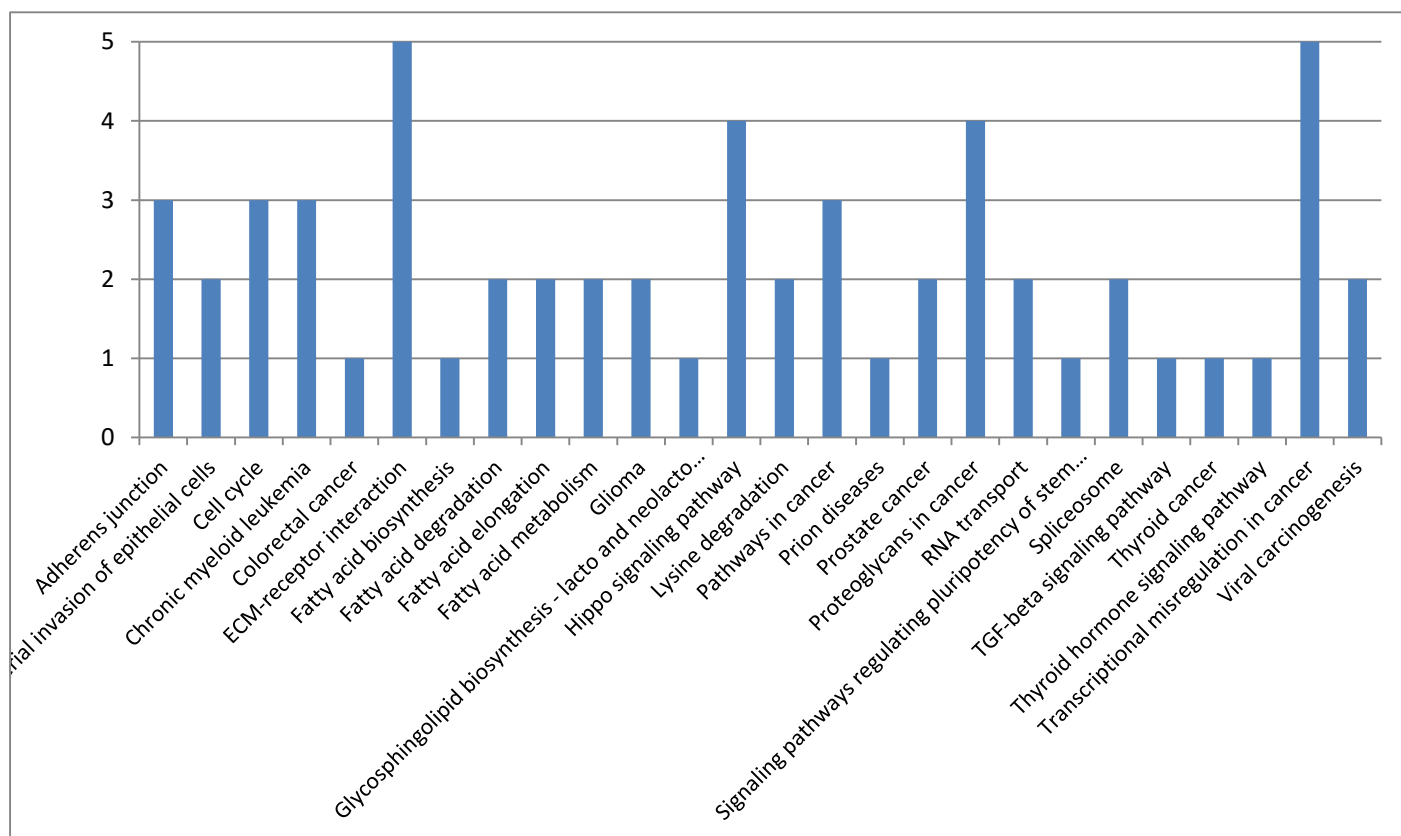
αναζητήσαμε. Σε ποσοστό περίπου 8% ακολουθούν τα miR-149, miR-608, miR-146a, από τα οποία τα δύο (miR-149, miR-146a) ανήκουν στα πιο συσχετιζόμενα με ασθένειες βάση αυτής της μελέτης. Τα ποσοστά των υπόλοιπων microRNAs κυμαίνονται στο 2% με 5% το καθένα. [Γράφημα 18]



Γράφημα 18: Αριθμός μονοπατιών που σχετίζονται με κάθε microRNA

Τα μονοπάτια, στα οποία εμπλέκονται περισσότερα miRNAs, σύμφωνα με το Diana, είναι ο υποδοχέας εξωκυτταρικής μήτρας και η μεταγραφική απορύθμιση στον καρκίνο. Τα δύο αυτά μονοπάτια συνδέονται με 5 microRNAs το καθένα, ενώ ακολουθούν σε ποσοστά, με 4 συνδέσεις, το μονοπάτι σηματοδότησης του υποθαλάμου και οι πρωτεογλυκάνες στον καρκίνο. Τα υπόλοιπα μονοπάτια συνδέονται με 1 έως 3 miRNAs, ενώ σύνολο είναι 26 βιολογικά μονοπάτια εμπλεκόμενα με τα 9 microRNAs που αναζητήθηκαν. [Γράφημα 19]

3. Αποτελέσματα



Γράφημα 19: Αριθμός microRNAs που σχετίζονται με κάθε μονοπάτι

3.3.3 Αξιολόγηση με δεδομένα από Kegg

Για κάθε ζεύγος miRNA:pathway έγινε περαιτέρω ανάλυση και βρέθηκαν από την βάση δεδομένων Kegg οι ασθένειες, οι οποίες εμπλέκονται τα συγκεκριμένα μονοπάτια [92]. Κάθε μονοπάτι εμπλεκόταν με τουλάχιστον 1 ασθένεια, ενώ τα μονοπάτια με τις πιο πολλές αλληλεπιδράσεις με παθήσεις ήταν ο υποδοχέας εξωκυτταρικής μήτρας, με 22 ασθένειες, και η μεταγραφική απορύθμιση στον καρκίνο, με 20. Για 4 μονοπάτια η βάση δεν παρείχε για συνδέσεις με ασθένειες. Τα δύο από αυτά τα μονοπάτια αφορούσαν τον καρκίνο και λειτουργίες που εμπλέκονται σε αυτόν. Ακόμα ένα από αυτά τα μονοπάτια είναι η εισβολή βακτηρίων στα επιθηλιακά κύτταρα, η οποία ανήκει στις μεταδοτικές ασθένειες, για τις οποίες η παρούσα εργασία δεν έχει βρει κάποια σύνδεση με miR-SNP. Επιπλέον, το μονοπάτι της πέψης των λιπαρών οξέων, το οποίο συμμετέχει στις διαδικασίες μεταβολισμού των λιπιδίων.

Στο σύνολο των ασθενειών, για τις οποίες η KEGG έδειχνε συσχέτιση με το κάθε miRNA:pathway, έγινε έλεγχος αν ανήκουν σε αυτές οι νόσοι που εμείς έχουμε βρει θετική συσχέτιση. Καταφέραμε και επαληθεύσαμε 16 συσχετίσεις miRNA:pathway:disease, 14 από τις οποίες είναι με καρκίνους και 2 με αυτοάνοσα νοσήματα. Επιπρόσθετα 36 ασθένειες, τις οποίες εμείς έχουμε βρει να σχετίζονται

3. Αποτελέσματα

με miR-SNPs, η KEGG δεν έδειχνε σχέση. Τα παραπάνω στοιχεία συνοψίστηκαν και απεικονίζονται στον πίνακα 3.

miRNA	Diana Kegg Pathways (number of associated diseases according Kegg database)	Associated diseases according present study	Common
miR-124	ECM-receptor interaction (22) Proteoglycans in cancer (N/A) Fatty acid elongation (5) Pathways in cancer (N/A) Glioma (1) Chronic myeloid leukemia (1) Adherens junction (13) Bacterial invasion of epithelial cells (N/A) Fatty acid degradation (8) Cell cycle (14) Fatty acid metabolism (1) Prostate cancer (1)	Cancer Esophageal cancer	√ √
miR-137	Hippo signaling pathway (1)	Schizophrenia	×
	ECM-receptor interaction (22)	Systemic Lupus Erythematosus	×
miR-146a	Hippo signaling pathway (1) Transcriptional misregulation in cancer (20) Glycosphingolipid biosynthesis - lacto and neolacto series (3) Lysine degradation (12)	Hepatocellular carcinoma Autoimmune Diseases Cancer Gastric cancer Digestive system cancer Liver cancer Hepatocellular carcinoma Lung cancer Cervical cancer Prostate cancer Papillary thyroid carcinoma Head and Neck cancer Multiple sclerosis Uveitis Coronary Artery Disease Asthma Psoriasis Behcet's disease	√ √ √ × × × × × × × √ √ × × × × × × × ×
miR-149	ECM-receptor interaction (22) Transcriptional misregulation in cancer (20) RNA transport (5) Spliceosome (7) Cell cycle (14)	Cancer	√

3. Αποτελέσματα

miR-196a-2		Gastric cancer	×
		Colorectal cancer	×
		Ischemic stroke	×
		Cancer	√
		Hepatitis B	×
	ECM-receptor interaction (22)	Digestive system cancer	×
		Colorectal cancer	×
		Liver cancer	×
		Hepatocellular carcinoma	×
		Lung cancer	×
		Breast cancer	×
		Head and Neck cancer	×
		Ischemic stroke	×
miR-27a	ECM-receptor interaction (22)		
	Proteoglycans in cancer (N/A)		
	Hippo signaling pathway (1)		
	Transcriptional misregulation in cancer (20)		
	Fatty acid elongation (5)		
	Pathways in cancer		
	Glioma (1)		
	Viral carcinogenesis (9)		
	Prion diseases (1)		
	Chronic myeloid leukemia (1)		
	Adherens junction (13)		
	Bacterial invasion of epithelial cells (N/A)	Cancer	√
	Thyroid hormone signaling pathway (1)		
	Lysine degradation (12)		
	Fatty acid degradation (8)		
	Signaling pathways regulating pluripotency of stem cells (1)		
	Fatty acid biosynthesis (N/A)		
	TGF-beta signaling pathway (29)		
	Cell cycle (14)		
	Thyroid cancer (1)		
	Fatty acid metabolism (1)		
	Colorectal cancer (5)		
	Prostate cancer (1)		
miR-34b/c		Digestive system cancer	√
	Proteoglycans in cancer (N/A)	Colorectal cancer	√
	Transcriptional misregulation in cancer (20)	Breast cancer	√
	Pathways in cancer (N/A)	Cancer	√
		Esophageal cancer	×



3. Αποτελέσματα

		Digestive system cancer Colorectal cancer Hepatocellular carcinoma	× × ×
miR-499a	Hippo signaling pathway (1) RNA transport (5) Spliceosome (7)	Autoimmune Diseases Cancer Esophageal cancer Liver cancer Breast cancer Rheumatoid arthritis	√ × × × × ×
miR-608		Cancer	√
	Proteoglycans in cancer (N/A) Transcriptional misregulation in cancer (20) Viral carcinogenesis (9) Chronic myeloid leukemia (1) Adherens junction (13)		

Πίνακας 3: Πίνακας σύγκρισης αποτελεσμάτων παρούσας μελέτης με δεδομένα από βάσεις δεδομένων Diana και Kegg. Η ένδειξη N/A (not available) αποδόθηκε σε ζευγάρια, για τα οποία λείπουν δεδομένα από την βάση.

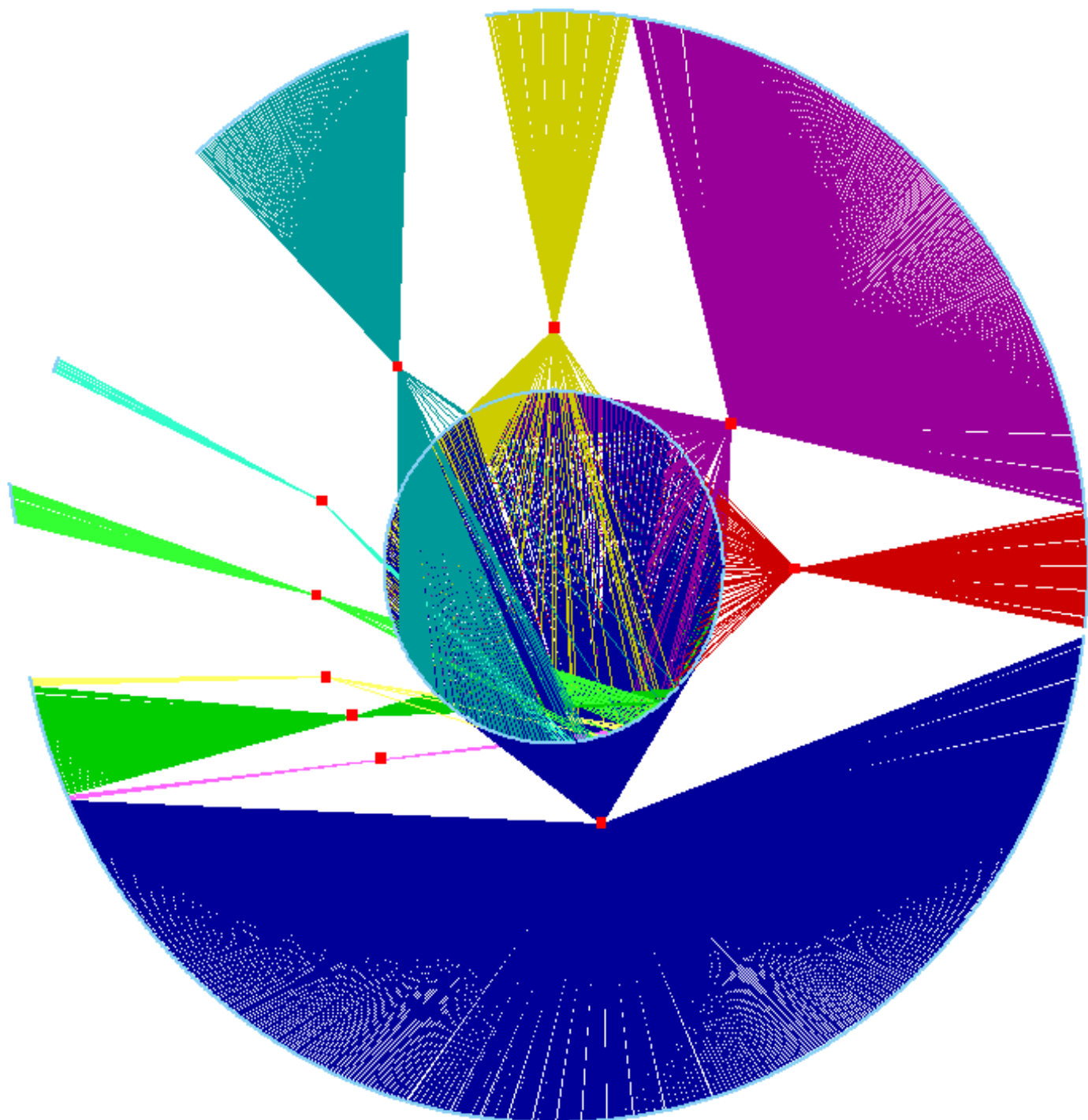
3.3.4 Δημιουργία δικτύων συσχέτισης miRNA:genes και miRNA:pathways

Τα στοιχεία που προέκυψαν από το πακέτο εργαλείων Diana, λόγω του μεγάλου αριθμού συνδέσεων που παρουσίαζαν, μετατράπηκαν σε δίκτυο συσχέτισης με την χρήση του προγράμματος Cytoscape.

Αρχικά το δίκτυο miRNA:genes αποτελεί ένα πυκνό δίκτυο με 2044 κόμβους και 2556 ακμές. Τα microRNAs συμβολίζονται με κόκκινους ρόμβους  , ενώ τα γονίδια με γαλάζια κύκλους  . Για την καλύτερη κατανόηση των συνδέσεων οι ακμές έχουν διαφορετικό χρώμα για κάθε miRNA από το οποίο ξεκινάνε. Παρόλο που η οπτικοποίηση ενός τόσο μεγάλου δικτύου με σαφή μορφή, σε συμβατή εικόνα, είναι δύσκολη, τα βασικά χαρακτηριστικά του δικτύου είναι ορατά όπως:

- το μεγαλύτερο μέρος των γονιδίων συνδέονται με το miR-27a (μπλε ακμές), ενώ ακολουθεί σε αριθμό συνδέσεων το miR-124 (μωβ ακμές).
- Με ελάχιστα γονίδια συνδέονται τα miR-196a (γαλάζιο), miR-608 (κίτρινο), miR-218 (ροζ).
- τα περισσότερα γονίδια συνδέονται με 1-3 μόνο microRNAs
- υπάρχει δυσαναλογία στο καταμερισμό των συνδέσεων





3. Αποτελέσματα



Δίκτυο 2: Δίκτυο συσχέτισης miRNA:genes

3. Αποτελέσματα

Το δίκτυο που δημιουργήθηκε για την συσχέτιση των microRNAs με τα μονοπάτια είναι αρκετά πιο απλό. Αποτελείται από 38 κόμβους και 58 ακμές. Τα σύμβολα που χρησιμοποιούνται στην απεικόνιση του είναι:

- Με σχήμα κόκκινου ρόμβου  το όνομα του microRNA
- Με κύκλο  το όνομα του μονοπατιού και ανάλογα με το πόσες ακμές έχει σκουραίνουν τα χρώματα ανάλογα από  για 1 ακμή έως  για 5.

Όπως και στο Δίκτυο 2, είναι εμφανές ότι τα περισσότερα μονοπάτια συνδέονται με ένα microRNA, το miR-27a, το οποίο έχει 16 ακμές. Με πολλά μονοπάτια, 12 σε αριθμό, συνδέεται και το miR-124. Τα υπόλοιπα microRNAs έχουν από 1 ακμή έως το μέγιστο των 5 ακμών.

4. Συμπεράσματα

Η βιολογία είναι ένας επιστημονικός κλάδος συνεχώς εξελισσόμενος. Τα ίδια τα κύρια συστατικά που αποτέλεσαν το θέμα της πτυχιακής, τόσο τα microRNAs όσο και τα SNPs αυτών, είναι ακόμα πρόσφατες ανακαλύψεις στο πεδίο της επιγενετικής ρύθμισης ασθενειών και η μελέτη τους υποδεικνύει συνεχώς νέες λειτουργίες που διαδραματίζουν. Η συνεχής προσπάθεια της ερευνητικής κοινότητας να βελτιώσει την ποιότητα ζωής των ανθρώπων, συμβαδίζει με την κατανόηση των παραγόντων που επηρεάζουν ή και ρυθμίζουν τις ανθρώπινες ασθένειες. Είναι επομένως αναμενόμενη η ύπαρξη πλήθους κλινικών ερευνών που διερευνούν τις πρόσφατες βιολογικές ανακαλύψεις ως προς την σχέση τους με ανθρώπινα νοσήματα. Σε μια προσπάθεια συγκέντρωσης και αξιολόγησης αυτών των ερευνών πολλές μετα-αναλύσεις έχουν διεξαχθεί.

Στην παρούσα εργασία πραγματοποιήθηκε συστηματική ανασκόπηση της διαθέσιμης βιβλιογραφίας αναφορικά με μονονουκλεοτιδικούς πολυμορφισμούς σε microRNAs και την σύνδεση τους με ανθρώπινες ασθένειες, όπως αυτή εξετάζεται σε μετα-αναλύσεις. Τα συγκεντρωτικά αποτελέσματα που προέκυψαν από την ανάλυση υπέδειξαν σημαντική συσχέτιση μεταξύ 13 SNPs σε 10 microRNAs με ένα σύνολο 26 ασθενειών, σχηματίζοντας 55 ζευγάρια miR-SNP:ασθένεια.

Τα ζεύγη αυτά απεικονίστηκαν σε ένα ευανάγνωστο δίκτυο συσχετίσεων, με την χρήση του εργαλείου Cytoscape. Το δίκτυο αυτό συνοψίζει τις συσχετίσεις miR-SNP με ασθένειες.

Από το σύνολο των 55 ζευγαριών miR-SNP:ασθένεια, τα 39 συνδέαν τα SNPs με διάφορους τύπους καρκίνου. Δεδομένα που σχετίζονται με αυτά τα ζευγάρια και επικυρώνουν τα ευρήματά μας, προσέφερε η online βάση δεδομένων MIRNASNP2. Η συγκεκριμένη βάση παρέχει πληροφορίες σχετικά με την διαφορική έκφραση αυτών των miR-SNPs σε διάφορες ασθένειες. Στο 92% των ζευγαριών τα οποία συγκρίθηκαν με δεδομένα της MIRNASNP2 υπήρξε επαλήθευση ευρημάτων.

Ιδιαίτερη έμφαση αξίζει να δοθεί στα 21 ζεύγη miR-SNP:ασθένειας, για των οποίων τα δεδομένα δεν υπήρχε πληροφορία. Μικρή σε αριθμό για τα ίδια ζεύγη είναι και η διαθέσιμη δημοσιευμένη βιβλιογραφία. Το δίκτυο λοιπόν που κατασκευάστηκε στην παρούσα πτυχιακή εργασία, μπορεί να λειτουργήσει ως ένας δείκτης για την περαιτέρω μελέτη των σχέσεων αυτών σε μελλοντικές έρευνες.

Επιπροσθέτως ανάλυση των microRNAs, στα οποία βρήκαμε τα SNPs που σχετίζονται με ασθένειες, με το πακέτο εργαλείων Diana, έδειξε την σχέση τους με

4. Συμπεράσματα

ένα μεγάλο αριθμό γονιδίων, 2557 ζεύγη miRNA:genes, και με 26 βιολογικά μονοπάτια, δίνοντας 59 ζεύγη miRNA:pathways.

Σύγκριση με δεδομένα για τα μονοπάτια από την βάση δεδομένων KEGG και των αποτελεσμάτων της παρούσας εργασίας επιβεβαίωσαν την σχέση 16 miRNA:pathways:diseases. Αυτό το αποτέλεσμα αποτελεί τόσο επικύρωση των ευρημάτων μας, όσο και δείκτη για την συσχέτιση των SNPs που βρίσκονται στα συγκεκριμένα miRNAs με τα αντίστοιχα μονοπάτια.

Τέλος τα δεδομένα, όπως προέκυψαν από την επεξεργασία στο Diana, απεικονίστηκαν μέσω του Cytoscape, σε δίκτυα με τις σχέσεις των 10 microRNAs, στα οποία βρίσκονται οι συσχετιζόμενοι SNPs, και (α) γονίδια στόχους και (β) βιολογικά μονοπάτια.

Η παρούσα πτυχιακή εργασία μπορεί να χρησιμοποιηθεί από ερευνητές κατά την έναρξη της ερευνάς τους, για επιλογή των πολυμορφισμών και ασθενειών που θα εξετάσουν, και γενικότερα για κάθε μελετητή που τον απασχολεί ο ρόλος που διαδραματίζουν τα SNPs σε microRNAs στην διάγνωση, εξέλιξη και θεραπεία των ανθρώπινων νόσων.

5. Βιβλιογραφία

- [1] Lee, Rosalind C., Rhonda L. Feinbaum, and Victor Ambros. "The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*." *Cell* 75.5 (1993): 843-854.
- [2] Wightman, Bruce, Ilho Ha, and Gary Ruvkun. "Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*." *Cell* 75.5 (1993): 855-862.
- [3] Reinhart, Brenda J., et al. "The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*." *nature* 403.6772 (2000): 901.
- [4] Pasquinelli, A. E., et al. "Srinivasan a, Fishman M, Finnerty J, Corbo J, Levine M, Leahy P, Davidson E, Ruvkun G. Conservation of the sequence and temporal expression of *let-7* heterochronic regulatory RNA." *Nature* 408 (2000): 86-9.
- [5] Lagos-Quintana, Mariana, et al. "Identification of novel genes coding for small expressed RNAs." *Science* 294.5543 (2001): 853-858.
- [6] Lau, Nelson C., et al. "An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*." *Science* 294.5543 (2001): 858-862.
- [7] Lee, Rosalind C., and Victor Ambros. "An extensive class of small RNAs in *Caenorhabditis elegans*." *Science* 294.5543 (2001): 862-864.
- [8] Calin, George Adrian, et al. "Frequent deletions and down-regulation of micro-RNA genes *miR15* and *miR16* at 13q14 in chronic lymphocytic leukemia." *Proceedings of the National Academy of Sciences* 99.24 (2002): 15524-15529.
- [9] Calin, George Adrian, et al. "MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias." *Proceedings of the National Academy of Sciences of the United States of America* 101.32 (2004): 11755-11760.
- [10] Ambros, Victor, et al. "A uniform system for microRNA annotation." *Rna* 9.3 (2003): 277-279.
- [11] Griffiths-Jones, Sam, et al. "miRBase: microRNA sequences, targets and gene nomenclature." *Nucleic acids research* 34.suppl_1 (2006): D140-D144.
- [12] Wright, Mathew W., and Elspeth A. Bruford. "Naming 'junk': human non-protein coding RNA (ncRNA) gene nomenclature." *Human genomics* 5.2 (2011): 90.
- [13] Wahid, Fazli, et al. "MicroRNAs: synthesis, mechanism, function, and recent clinical trials." *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* 1803.11 (2010): 1231-1243.
- [14] Rodriguez, Antony, et al. "Identification of mammalian microRNA host genes and transcription units." *Genome research* 14.10a (2004): 1902-1910.
- [15] Baskerville, Scott, and David P. Bartel. "Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes." *Rna* 11.3 (2005): 241-247.
- [16] Kim, Young-Kook, and V. Narry Kim. "Processing of intronic microRNAs." *The EMBO journal* 26.3 (2007): 775-783.

5. Βιβλιογραφία

- [17]Nelson, Peter, et al. "The microRNA world: small is mighty." *Trends in biochemical sciences* 28.10 (2003): 534-540.
- [18]Cai, Xuezhong, Curt H. Hagedorn, and Bryan R. Cullen. "Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs." *Rna* 10.12 (2004): 1957-1966.
- [19]Borchert, Glen M., William Lanier, and Beverly L. Davidson. "RNA polymerase III transcribes human microRNAs." *Nature structural & molecular biology* 13.12 (2006): 1097.
- [20]Lee, Y., et al. "dmark, S. Kim, VN Kim." *The nuclear RNase III Drosha initiates microRNA processing.* *Nature* 425.6956 (2003): 415-419.
- [21]Bohnsack, Markus T., Kevin Czaplinski, and DIRK GÖRLICH. "Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs." *Rna* 10.2 (2004): 185-191.
- [22]Okada, Chimari, et al. "A high-resolution structure of the pre-microRNA nuclear export machinery." *Science* 326.5957 (2009): 1275-1279.
- [23]Yi, Rui, et al. "Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs." *Genes & development* 17.24 (2003): 3011-3016.
- [24]Zeng, Yan, and Bryan R. Cullen. "Structural requirements for pre-microRNA binding and nuclear export by Exportin 5." *Nucleic acids research* 32.16 (2004): 4776-4785.
- [25]Lund, Elsebet, et al. "Nuclear export of microRNA precursors." *Science* 303.5654 (2004): 95-98.
- [26]Lund, E., and J. E. Dahlberg. "Substrate selectivity of exportin 5 and Dicer in the biogenesis of microRNAs." *Cold Spring Harbor symposia on quantitative biology.* Vol. 71. Cold Spring Harbor Laboratory Press, 2006.
- [27]Ji, Xinhua. "The mechanism of RNase III action: how dicer dices." *RNA interference* (2008): 99-116.
- [28]MacRae, Ian J., et al. "Structural basis for double-stranded RNA processing by Dicer." *Science* 311.5758 (2006): 195-198.
- [29]Park, Jong-Eun, et al. "Dicer recognizes the 5' end of RNA for efficient and accurate processing." *Nature* 475.7355 (2011): 201.
- [30]Chendrimada, Thimmaiah P., et al. "TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing." *Nature* 436.7051 (2005): 740.
- [31]Maniataki, Elisavet, and Zissimos Mourelatos. "A human, ATP-independent, RISC assembly machine fueled by pre-miRNA." *Genes & development* 19.24 (2005): 2979-2990.
- [32]MacRae, Ian J., et al. "In vitro reconstitution of the human RISC-loading complex." *Proceedings of the National Academy of Sciences* 105.2 (2008): 512-517.
- [33]Gregory, Richard I., et al. "Human RISC couples microRNA biogenesis and posttranscriptional gene silencing." *Cell* 123.4 (2005): 631-640.
- [34]Lewis, Benjamin P., et al. "Prediction of mammalian microRNA targets." *Cell* 115.7 (2003): 787-798.
- [35]Doench, John G., and Phillip A. Sharp. "Specificity of microRNA target selection in translational repression." *Genes & development* 18.5 (2004): 504-511.
- [36]Krek, Azra, et al. "Combinatorial microRNA target predictions." *Nature genetics* 37.5 (2005): 495-500.

5. Βιβλιογραφία

- [37] Brennecke, Julius, et al. "Principles of microRNA–target recognition." *PLoS biology* 3.3 (2005): e85.
- [38] Bartel, David P. "MicroRNAs: genomics, biogenesis, mechanism, and function." *cell* 116.2 (2004): 281-297.
- [39] Zeng, Yan, Rui Yi, and Bryan R. Cullen. "MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms." *Proceedings of the National Academy of Sciences* 100.17 (2003): 9779-9784.
- [40] Lewis, Benjamin P., et al. "Prediction of mammalian microRNA targets." *Cell* 115.7 (2003): 787-798.
- [41] Doench, John G., and Phillip A. Sharp. "Specificity of microRNA target selection in translational repression." *Genes & development* 18.5 (2004): 504-511.
- [42] Doench, John G., and Phillip A. Sharp. "Specificity of microRNA target selection in translational repression." *Genes & development* 18.5 (2004): 504-511.
- [43] Sheth, Ujwal, and Roy Parker. "Decapping and decay of messenger RNA occur in cytoplasmic processing bodies." *Science* 300.5620 (2003): 805-808.
- [44] Wu, Ligang, Jihua Fan, and Joel G. Belasco. "MicroRNAs direct rapid deadenylation of mRNA." *Proceedings of the National Academy of Sciences of the United States of America* 103.11 (2006): 4034-4039.
- [45] Bagga, Shveta, et al. "Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation." *Cell* 122.4 (2005): 553-563.
- [46] Lim, Lee P., et al. "Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs." *Nature* 433.7027 (2005): 769.
- [47] Behm-Ansmant, Isabelle, et al. "mRNA degradation by miRNAs and GW182 requires both CCR4: NOT deadenylase and DCP1: DCP2 decapping complexes." *Genes & development* 20.14 (2006): 1885-1898.
- [48] Giraldez, Antonio J., et al. "Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs." *science* 312.5770 (2006): 75-79.
- [49] Valencia-Sanchez, Marco Antonio, et al. "Control of translation and mRNA degradation by miRNAs and siRNAs." *Genes & development* 20.5 (2006): 515-524.
- [50] Pillai, Ramesh S., et al. "Inhibition of translational initiation by Let-7 MicroRNA in human cells." *Science* 309.5740 (2005): 1573-1576.
- [51] Meister, Gunter, et al. "Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs." *Molecular cell* 15.2 (2004): 185-197.
- [52] Liu, Jidong, et al. "Argonaute2 is the catalytic engine of mammalian RNAi." *Science* 305.5689 (2004): 1437-1441.
- [53] Petersen, Christian P., et al. "Short RNAs repress translation after initiation in mammalian cells." *Molecular cell* 21.4 (2006): 533-542.
- [54] Liu, Jidong, et al. "MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies." *Nature cell biology* 7.7 (2005): 719.
- [55] Kedersha, Nancy, and Paul Anderson. "Mammalian stress granules and processing bodies." *Methods in enzymology* 431 (2007): 61-81.
- [56] TEIXEIRA, DANIELA, et al. "Processing bodies require RNA for assembly and contain nontranslating mRNAs." *Rna* 11.4 (2005): 371-382.
- [57] Liu, Jidong, et al. "A role for the P-body component, GW182, in microRNA function." *Nature cell biology* 7.12 (2005): 1261.

5. Βιβλιογραφία

- [58]Parker, Roy, and Ujwal Sheth. "P bodies and the control of mRNA translation and degradation." *Molecular cell* 25.5 (2007): 635-646.
- [59]Russell, Peter J. *Igenetics Coursecompass Student Access Kit: A Mendelian Approach*. Benjamin-Cummings, 2005.
- [60]Brookes AJ. The essence of SNPs. *Gene*. 1999;234:177–86.
- [61]Collins, Francis S., Mark S. Guyer, and Aravinda Chakravarti. "Variations on a theme: cataloging human DNA sequence variation." *Science* 278.5343 (1997): 1580-1581.
- [62]Kruglyak, Leonid. "The use of a genetic map of biallelic markers in linkage studies." *Nature genetics* 17.1 (1997): 21-24.
- [63]Schork, Nicholas J., Daniele Fallin, and Jerry S. Lanchbury. "Single nucleotide polymorphisms and the future of genetic epidemiology." *Clinical genetics* 58.4 (2000): 250-264.
- [64]Ramsay, Graham. "DNA chips: state-of-the art." *Nature biotechnology* 16.1 (1998): 40-44.
- [65]McKeigue, Paul M. "Mapping genes that underlie ethnic differences in disease risk: methods for detecting linkage in admixed populations, by conditioning on parental admixture." *The American Journal of Human Genetics* 63.1 (1998): 241-251.
- [66]Sun, Guihua, et al. "SNPs in human miRNA genes affect biogenesis and function." *Rna* 15.9 (2009): 1640-1651.
- [67]Haidich, Anna-Bettina. "Meta-analysis in medical research." *Hippokratia*14.Suppl 1 (2010): 29.
- [68]Simpson, R. J. S., and Karl Pearson. "Report on certain enteric fever inoculation statistics." *The British Medical Journal* (1904): 1243-1246.
- [69]Glass, Gene V. "Primary, secondary, and meta-analysis of research." *Educational researcher* 5.10 (1976): 3-8.
- [70]Elwood, Peter Creighton, et al. "A randomized controlled trial of acetyl salicylic acid in the secondary prevention of mortality from myocardial infarction." *British Medical Journal* 1.5905 (1974): 436.
- [71]Peto, Richard. "Cancer, cholesterol, carotene, and tocopherol." *The Lancet*318.8237 (1981): 97-98.
- [72]Egger, Matthias, George Davey Smith, and Andrew N. Phillips. "Meta-analysis: principles and procedures." *BMJ: British Medical Journal* 315.7121 (1997): 1533.
- [73]Simmonds, Mark C., et al. "Meta-analysis of individual patient data from randomized trials: a review of methods used in practice." *Clinical Trials* 2.3 (2005): 209-217.
- [74]Stewart, Lesley A. "Practical methodology of meta-analyses (overviews) using updated individual patient data." *Statistics in medicine* 14.19 (1995): 2057-2079.
- [75]Altman, Douglas G. "Systematic reviews in health care: Systematic reviews of evaluations of prognostic variables." *BMJ: British Medical Journal* 323.7306 (2001): 224.
- [76]Simon, R., and David G. Altman. "Statistical aspects of prognostic factor studies in oncology." *British journal of cancer* 69.6 (1994): 979.
- [77]Ioannidis, John PA, et al. "Mortality in systemic sclerosis: an international meta-analysis of individual patient data." *The American journal of medicine* 118.1 (2005): 2-10.

5. Βιβλιογραφία

- [78]Oxman, Andrew D., Michael J. Clarke, and Lesley A. Stewart. "From science to practice: meta-analyses using individual patient data are needed." *JAMA* 274.10 (1995): 845-846.
- [79]Moher, David, et al. "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement." *PLoS medicine* 6.7 (2009): e1000097.
- [80]Liberati, Alessandro, et al. "The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration." *PLoS medicine* 6.7 (2009): e1000100.
- [81]Cochran, William G. "The combination of estimates from different experiments." *Biometrics* 10.1 (1954): 101-129.
- [82]Higgins, Julian, and Simon G. Thompson. "Quantifying heterogeneity in a meta-analysis." *Statistics in medicine* 21.11 (2002): 1539-1558.
- [83]Lau, Joseph, Christopher H. Schmid, and Thomas C. Chalmers. "Cumulative meta-analysis of clinical trials builds evidence for exemplary medical care." *Journal of clinical epidemiology* 48.1 (1995): 45-57.
- [84]Egger, Matthias, and G. Davey Smith. "Bias in location and selection of studies." *BMJ: British Medical Journal* 316.7124 (1998): 61.
- [85]Gough, David, and Diana Elbourne. "Systematic research synthesis to inform policy, practice and democratic debate." *Social policy and society* 1.3 (2002): 225-236.
- [86]<https://www.ncbi.nlm.nih.gov/pubmed>
- [87]<https://www.nlm.nih.gov/>
- [88]Shannon, Paul, et al. "Cytoscape: a software environment for integrated models of biomolecular interaction networks." *Genome research* 13.11 (2003): 2498-2504.
- [89]Gong, Jing, et al. "An update of miRNASNP database for better SNP selection by GWAS data, miRNA expression and online tools." *Database* 2015 (2015).
- [90]<http://diana.imis.athena-innovation.gr/DianaTools/index.php?r=site/index>
- [91]Kanehisa, Minoru, and Susumu Goto. "KEGG: kyoto encyclopedia of genes and genomes." *Nucleic acids research* 28.1 (2000): 27-30.
- [92]Kanehisa, Minoru, et al. "KEGG: new perspectives on genomes, pathways, diseases and drugs." *Nucleic acids research* 45.D1 (2017): D353-D361.

