



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ ΒΙΟΪΑΤΡΙΚΗ

**ΕΡΓΑΣΤΗΡΙΟ ΜΟΡΙΑΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΙΚΗΣ ΒΙΟΛΟΓΙΑΣ ΚΑΙ
ΓΕΝΕΤΙΚΗΣ (ΕΜΥΒΓ)**

Μεθοδολογίες υπολογιστικής ανάλυσης και μετα-ανάλυσης δεδομένων γονιδιακής έκφρασης

Διδακτορική Διατριβή

Κωνσταντίνα Βέννου

Λαμία, Απρίλιος 2020

«Η παρούσα διδακτορική διατριβή συγχρηματοδοτείται από την Ελλάδα και την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) μέσω του Επιχειρησιακού Προγράμματος «Ανάπτυξη Ανθρώπινου Δυναμικού, Εκπαίδευση και Διά Βίου Μάθηση», στο πλαίσιο της Πράξης «Ενίσχυση του ανθρώπινου ερευνητικού δυναμικού μέσω της υλοποίησης διδακτορικής έρευνας» (MIS-5000432), που υλοποιεί το Ίδρυμα Κρατικών Υποτροφιών (ΙΚΥ)»



**Επιχειρησιακό Πρόγραμμα
Ανάπτυξη Ανθρώπινου Δυναμικού,
Εκπαίδευση και Διά Βίου Μάθηση**



Science is a way of thinking much more than it is a body of knowledge

Carl Sagan

Τριμελής συμβουλευτική επιτροπή

1) Παντελής Μπάγκος (Επιβλέπων καθηγητής)

Καθηγητής, Τμήμα Πληροφορικής με Εφαρμογές στη Βιοϊατρική, Πανεπιστήμιο Θεσσαλίας

2) Βασίλειος Πλαγιανάκος,

Καθηγητής, Τμήμα Πληροφορικής με Εφαρμογές στη Βιοϊατρική, Πανεπιστήμιο Θεσσαλίας

3) Μαρία Αδάμ

Αναπληρώτρια Καθηγήτρια, Τμήμα Πληροφορικής με Εφαρμογές στη Βιοϊατρική,
Πανεπιστήμιο Θεσσαλίας

Επταμελής εξεταστική επιτροπή

1) Παντελής Μπάγκος

Καθηγητής, Τμήμα Πληροφορικής με Εφαρμογές στη Βιοϊατρική, Πανεπιστήμιο Θεσσαλίας

2) Βασίλειος Πλαγιανάκος,

Καθηγητής, Τμήμα Πληροφορικής με Εφαρμογές στη Βιοϊατρική, Πανεπιστήμιο Θεσσαλίας

3) Μαρία Αδάμ

Αναπληρώτρια Καθηγήτρια, Τμήμα Πληροφορικής με Εφαρμογές στη Βιοϊατρική,
Πανεπιστήμιο Θεσσαλίας

4) Άρτεμις Χατζηγεωργίου

Καθηγήτρια, Τμήμα Πληροφορικής με Εφαρμογές στη Βιοϊατρική, Πανεπιστήμιο Θεσσαλίας

5) Γεωργία Μπράλιου

Επίκουρη Καθηγήτρια, Τμήμα Πληροφορικής με Εφαρμογές στη Βιοϊατρική, Πανεπιστήμιο
Θεσσαλίας

6) Κωνσταντίνος Παπαλουκάς

Αναπληρωτής Καθηγητής, Τμήμα Βιολογικών Εφαρμογών και Τεχνολογιών, Πανεπιστήμιο
Ιωαννίνων

7) Γεώργιος Νικολόπουλος

Επίκουρος Καθηγητής, Ιατρική Σχολή, Πανεπιστήμιο Κύπρου

Ευχαριστίες

Η παρούσα διδακτορική διατριβή εκπονήθηκε από τον Μάρτιο 2016 έως τον Μάρτιο 2020 στο Τμήμα Πληροφορικής με Εφαρμογές στη Βιοϊατρική του Πανεπιστημίου Θεσσαλίας.

Αρχικά θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου Καθηγητή κ. Παντελή Μπάγκο, για την άριστη η συνεργασία μας τα τελευταία δέκα χρόνια η οποία είχε ως εναρκτήριο την εκπόνηση της πτυχιακής μου εργασίας ως προπτυχιακή φοιτήτρια στο Τμήμα Βιολογικών Εφαρμογών και Τεχνολογιών του Πανεπιστημίου Ιωαννίνων το 2010, συνεχίστηκε με τη μεταπτυχιακή διπλωματική μου εργασία το 2016 στο διατμηματικό μεταπτυχιακό της Υπολογιστικής Βιολογίας και Ιατρικής του Πανεπιστημίου Θεσσαλίας έως και σήμερα που μου έκανε την τιμή να είναι ο επιβλέπων καθηγητής μου στη διδακτορική διατριβή μου. Φυσικά ένα ευχαριστώ είναι λίγο για την συνεχή καθοδήγηση, στήριξη, βοήθεια καθώς και για τον πολύτιμο χρόνο που ήταν πάντα διατεθειμένος να μου αφιερώσει, παρά το υπερβολικά φορτωμένο πρόγραμμά του.

Ακόμη, θα ήθελα να ευχαριστήσω τον Καθηγητή του Τμήματος Πληροφορικής με Εφαρμογές στη Βιοϊατρική κ. Βασίλειο Πλαγιανάκο και την Αναπληρώτρια Καθηγήτρια του Τμήματος Πληροφορικής με Εφαρμογές στη Βιοϊατρική, κα. Μαρία Αδάμ, οι οποίοι παρά το φορτωμένο πρόγραμμά τους δέχτηκαν να συνεπιβλέψουν τη διατριβή μου και φυσικά για την άριστη συνεργασία που είχαμε όλο αυτό το διάστημα.

Ευχαριστώ πολύ τα μέλη της εξεταστικής επιτροπής, την Καθηγήτρια κα Αρτέμιδα Χατζηγεωργίου, την Επίκουρη Καθηγήτρια κα Γεωργία Μπράλιου, τον Αναπληρωτή Καθηγητή και παλιό μου καθηγητή κ. Κωνσταντίνο Παπαλουκά και τον Επίκουρο Καθηγητή κ. Γεώργιο Νικολόπουλο για τις υποδείξεις και τον πολύτιμο χρόνο που αφιέρωσαν κατά την κριτική ανάγνωση και την υποστήριξη της παρούσας διατριβής.

Ιδιαίτερες ευχαριστίες θα ήθελα να εκφράσω σε όλα τα μέλη του Εργαστηρίου Μοριακής και Υπολογιστικής Βιολογίας και Γενετικής του Πανεπιστημίου Θεσσαλίας για την πολύτιμη βοήθειά τους και για την άριστη συνεργασία μας. Αρχικά την Επίκουρη Καθηγήτρια κα Γεωργία Μπράλιου για την συμμετοχή της στην επταμελή εξεταστική επιτροπή της παρούσας διατριβής, αλλά κυρίως για την άριστη συνεργασία μας, γιατί ήταν πάντα πρόθυμη να βοηθήσει σε κάθε δυσκολία και κάθε απορία. Τη μεταδιδάκτορα κα Παναγιώτα Κοντού για την άριστη συνεργασία μας όπως επίσης και για το χρόνο που έβρισκε ώστε να με βοηθήσει σε οποιοδήποτε πρόβλημα αντιμετώπιζα. Ακόμη θα ήθελα να ευχαριστήσω την διδάκτορα κα Γεωργία Καπούλα, τη

μεταδιδάκτορα κα Μαργαρίτα Θεοδωροπούλου, τον υποψήφιο διδάκτορα κ. Ahmed Roumia, και τον υποψήφιο διδάκτορα κ. Ιωάννη Ταμπόση για την πολύ καλή συνεργασία που είχαμε όλο αυτό το διάστημα. Επίσης, ευχαριστώ πολύ την εντεταλμένη διδασκαλίας κα Ευφροσύνη-Άλκηστη Παρασκευοπούλου-Κόλλια για την βοήθειά της και σωστά ελληνικά της. Τους ευχαριστώ όλους για τις διάφορες γνώσεις που αποκόμισα από τον καθένα και για το πολύ ωραίο κλίμα που επικρατούσε τόσο στο εργαστήριο όσο και στο γραφείο κάνοντας πιο εύκολες τις πολλές ώρες εργασίας.

Να ευχαριστήσω ακόμη τον Επίκουρο Καθηγητή του Humanitas University της Ιταλίας, κ. Στέφανο Μπονόβα και τον ερευνητή στο Istituto Clinico Humanitas, IRCCS, κ. Daniele Piovani για την άριστη συνεργασία μας.

Φυσικά, θα ήθελα να ευχαριστήσω την οικογένειά μου. Τους γονείς μου, την αδερφή μου και τα υπέροχα ανιψάκια μου για την αμέριστη αγάπη τους. Μαμά, Μπαμπά, ευχαριστώ πολύ για όλη την στήριξη που μου δίνετε και που είστε πάντα δίπλα μου. Εύη σε ευχαριστώ που πάντα πιστεύεις σε έμένα, ακόμη και όταν εγώ δε το κάνω, και που πάντα με ωθείς προς κάτι καλύτερο. Και τέλος τον Θεοδωρή για την αγάπη, την υπομονή και την στήριξή του.

Κλείνοντας, θα ήθελα να αναφέρω ότι η παρούσα διδακτορική διατριβή συγχρηματοδοτήθηκε από την Ελλάδα και την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) μέσω του Επιχειρησιακού Προγράμματος «Ανάπτυξη Ανθρώπινου Δυναμικού, Εκπαίδευση και Διά Βίου Μάθηση», στο πλαίσιο της Πράξης «Ενίσχυση του ανθρώπινου ερευνητικού δυναμικού μέσω της υλοποίησης διδακτορικής έρευνας» (MIS-5000432), που υλοποιεί το Ίδρυμα Κρατικών Υποτροφιών (ΙΚΥ)».

Κωνσταντίνα Βέννου

Λαμία, Απρίλιος 2020

Δημοσιεύσεις σε επιστημονικά περιοδικά

1. Vennou KE, Piovani D, Kontou PI, Bonovas S, Bagos PG. **Multiple outcome meta-analysis of gene-expression data in Inflammatory Bowel Disease.** *Genomics*, 2019 Oct 18.
2. Vennou KE, Kontou PI, Braliou GG, Bagos PG. **Meta-analysis of gene expression profiles in preeclampsia.** *Pregnancy Hypertension*, 2019 Dec 20.
3. Vennou KE, Piovani D, Kontou PI, Bonovas S, Bagos PG. **Multiple outcome method for meta-analysis of gene-expression data.** *MethodsX*, 2020 Feb 21.

Εργασίες σε συνέδρια

Διεθνή συνέδρια

1. K.E. Vennou, P.I. Kontou and P.G. Bagos. **An approach for multiple outcome meta-analysis of microarray data.** 17th European Conference on Computational Biology, 8-12 October 2018, Athens, Greece.
2. G.V. Kapoula, P.I. Kontou, K.E. Vennou, G.K. Nikolopoulos, P.G. Bagos. **Endothelial Nitric Oxide Synthase (eNOs) 4a/b polymorphism and the risk of IgA nephropathy: A meta-analysis.** 7th International Conference of the Cyprus Society of Human Genetics, 7-8 December 2018, Nicosia, Cyprus.

Πανελλήνια συνέδρια

1. K.E. Vennou, P.I. Kontou and P.G. Bagos. **Meta-analysis of Gene Expression Profiles in Preeclampsia.** 29th Annual Conference of the Hellenic Society for Biological Sciences, 25-27 May 2017, Lamia, Greece.
2. K. E. Vennou, P.I. Kontou and P.G. Bagos. **Meta-analysis of Gene Expression Profiles in Preeclampsia, An update.** 12th Conference of the Hellenic Society for Computational Biology and Bioinformatics (HSCBB17), 11-13 October 2017, Athens, Greece.
3. K.E. Vennou, P.I. Kontou and P.G. Bagos. **Multiple outcomes meta-analysis of gene-expression data.** 31st Conference of the Greek Statistical Institute, 4-6 May 2018, Lamia, Greece.

4. K.E. Vennou, D. Piovani, P.I. Kontou, S. Bonovas, P.G. Bagos. **Multiple outcome meta-analysis of gene expression data in Inflammatory Bowel Disease.** 14th Conference of the Hellenic Society for Computational Biology and Bioinformatics (HSCBB19), 8-10 December 2019 Patras, Greece.

5. A.S. Mexias, K.E. Vennou, I.A. Tamposis, P.I. Kontou, P.G. Bagos. **Gene symbol updater (GISU): a tool for converting probe to gene identifiers in microarray datasets.** 14th Conference of the Hellenic Society for Computational Biology and Bioinformatics (HSCBB19), 8-10 December 2019 Patras, Greece.

Πρόλογος

Η παρούσα διδακτορική διατριβή ασχολήθηκε με την εμπειριστατωμένη μελέτη και εφαρμογή των μεθόδων και στατιστικών τεχνικών της μετα-ανάλυσης, σε δεδομένα υψηλής απόδοσης, και πιο συγκεκριμένα σε δεδομένα γονιδιακής έκφρασης από πειράματα μικροσυστοιχιών. Η μετα-ανάλυση είναι ένα πολύτιμο εργαλείο για τη σύνθεση των στοιχείων σε ένα ευρύ φάσμα μελετών, συμπεριλαμβανομένων των πειραμάτων υψηλής απόδοσης, όπως μελέτες συσχέτισης ολόκληρου του γονιδιώματος (GWAS) και μελέτες γονιδιακής έκφρασης.

Ένα από τα θέματα με τα οποία ασχολήθηκε η παρούσα διατριβή ήταν η διερεύνηση βέλτιστων πρακτικών στη μεθοδολογία μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών. Πραγματοποιήθηκαν διάφοροι έλεγχοι, σε σχετικά μικρά καθώς και σε μεγαλύτερα δοκιμαστικά τυχαία σύνολα δεδομένων (99 και 462 ατόμων αντίστοιχα), χρησιμοποιώντας την παλιά αλλά κλασσική μέθοδο της διόρθωσης του Hedges και διερευνώντας την εμφάνιση καλύτερων ή όχι αποτελεσμάτων με χρήση της μεθόδου αναδειγματοληψίας Bootstrap (είτε με 200 είτε με 500 επαναλήψεις). Τα αποτελέσματα των ελέγχων έδειξαν ότι η χρήση της διόρθωσης του Hedges είναι σχεδόν αναγκαία σε μικρά σύνολα δεδομένων. Ακόμη διερευνήθηκε εάν όντως ισχύει ότι η διόρθωση με τη μέθοδο αναδειγματοληψίας Bootstrap (200 επαναλήψεων) δίνει καλύτερα αποτελέσματα σε σχέση μόνο με την διόρθωση του Hedges, κάτι που εξακριβώθηκε και στην παρούσα μελέτη, ωστόσο χρειάζεται πολύ περισσότερο χρόνο. Επιπλέον παρατηρήθηκε ότι η χρήση της μεθόδου αναδειγματοληψίας Bootstrap των 200 επαναλήψεων επιστρέφει σχεδόν ίδιας στατιστικής ισχύος και χρησιμότητας αποτελέσματα με την πολύ πιο χρονοβόρα μεθοδολογία αναδειγματοληψίας Bootstrap των 500 επαναλήψεων (το οποίο θεωρείται στην βιβλιογραφία ότι επιστέφει πολύ αξιόλογα αποτελέσματα).

Στα πλαίσια διερεύνησης διαφόρων πολυπαραγοντικών ασθενειών πραγματοποιήθηκε μετα-ανάλυση γονιδιακών δεδομένων από πειράματα μικροσυστοιχιών σχετικά με την ασθένεια της προεκλαμψίας, ώστε να ανακαλυφθούν διαφορικά εκφραζόμενα γονίδια με σκοπό να βρεθούν και να ταυτοποιηθούν μηχανισμοί οι οποίοι θα μπορούσαν να ευθύνονται για την εμφάνιση της νόσου καθώς και για τους διαφορετικούς υποτύπους της. Με βάση αυτή την κατεύθυνση διερευνήθηκαν οι βάσεις δεδομένων Pubmed και GEO για την ύπαρξη δεδομένων γονιδιακής έκφρασης σε δείγματα πλακούντα από μελέτες ασθενών-μαρτύρων. Πραγματοποιήθηκε μετα-ανάλυση με την οποία βρέθηκαν 629 στατιστικώς σημαντικά διαφορικά εκφραζόμενα γονίδια να συσχετίζονται με την προεκλαμψία σε επίπεδο σημαντικότητας 1%, τα οποία αναλύθηκαν

περαιτέρω με εργαλεία Βιοπληροφορικής και προέκυψαν τα δίκτυα αλληλεπίδρασης και τα βιοχημικά μονοπάτια στα οποία συμμετέχουν. Η ανάλυση του δικτύου έδειξε λίγα γονίδια, πολύ διασυνδεδεμένα, τα οποία εμπλέκονται στην φυσική ανοσία και σε μονοπάτια μεταγωγής σήματος, ενδεικτικά χαρακτηριστικά μιας πολύπλευρης νόσου με αιτιολογική ετερογένεια. Η ανάλυση των βιοχημικών μονοπατιών αποκάλυψε ότι αυτά τα διαφορετικά εκφραζόμενα γονίδια συμμετέχουν κυρίως στον μεταβολισμό υδατανθράκων, αμινοξέων και πυριμιδίνης, στο σύστημα του κερκαδικού ρολογιού και σε μονοπάτια μεταγωγής σήματος.

Όπως αναφέρθηκε η μετα-ανάλυση είναι ισχυρό εργαλείο στη σύνθεση ερευνητικών μελετών. Υπάρχουν ωστόσο καταστάσεις στις οποίες έχουμε είτε πολλαπλά αποτελέσματα είτε πολλαπλές θεραπείες, σε αυτές είναι αναγκαίο ένα πολυπαραγοντικό μοντέλο μετα-ανάλυσης, το οποίο θα εκτελεί κοινή μοντελοποίηση των διαφόρων ποσοτήτων ενδιαφέροντος όπου θα προσφέρει σημαντικά πλεονεκτήματα, όπως την αύξηση της στατιστικής ισχύος και τη δυνατότητα διεξαγωγής συνολικών ελέγχων. Για τους λόγους αυτούς προσαρμόσαμε τη μεθοδολογία της πολυμεταβλητής μετα-ανάλυσης ώστε να εφαρμόζεται και σε δεδομένα γονιδιακής έκφρασης, σε ένα εύκολο και αξιόπιστο λογισμικό, με τη χρήση του στατιστικού πακέτου Stata. Η υλοποίηση της μεθόδου είναι διαθέσιμη στο ευρύ κοινό στην ιστοσελίδα www.compgen.org/tools/multivariate-microarrays.

Τέλος χρησιμοποιώντας το προαναφερθέν λογισμικό πραγματοποιήθηκε πολυμεταβλητή μετα-ανάλυση σε δεδομένα γονιδιακών δεδομένων από πειράματα μικροσυστοιχιών σχετικά με την νόσο του Crohn και την ελκώδη κολίτιδα, οι οποίες είναι οι κύριες μορφές της Ιδιοπαθούς Φλεγμονώδους Νόσου του Εντέρου. Οι δύο ασθένειες διαθέτουν κοινά συμπτώματα αλλά διαφέρουν στη θέση και την έκταση της φλεγμονής καθώς και στις επιλοκές τους. 249 διαφορετικά εκφραζόμενα γονίδια βρέθηκαν, από τη μετα-ανάλυση, στατιστικώς σημαντικά σχετιζόμενα με τη νόσο του Crohn και 38 με την ελκώδη κολίτιδα σε επίπεδο σημαντικότητας 1%. 20 από τα διαφορετικά εκφραζόμενα γονίδια ήταν κοινά και στις δύο ασθένειες. Ένας συνολικός έλεγχος έδειξε 260 διαφορετικά εκφραζόμενα γονίδια σχετιζόμενα με την Ιδιοπαθή Φλεγμονώδη Νόσο του Εντέρου, εκ των οποίων 53 δεν θεωρήθηκαν σημαντικά σε καμία από τις δύο ασθένειες. Ακόμη ανακαλύφθηκαν στατιστικώς σημαντικά βιοχημικά μονοπάτια, με χρήση βιοπληροφορικών εργαλείων, τα οποία ενδεχομένως να εμπλέκονται στην παθογένεση της Ιδιοπαθούς Φλεγμονώδους Νόσου του Εντέρου. Χαρακτηριστικά είναι το μονοπάτι των μεταγραφικών παραγόντων των JAK/STAT και το σηματοδοτικό μονοπάτι της ιντερφερόνης γ . Επιπροσθέτως

γονίδια τα οποία συμμετέχουν στην κυτταρική διαίρεση, στην απόπτωση και στην καρκινογέννεση, βρέθηκαν στατιστικώς σημαντικά. Μεταξύ άλλων τα γονίδια BCAT2 και GZMB είναι δύο ενδιαφέροντα νέα διαφορετικά εκφραζόμενα γονίδια τα οποία αξίζουν επικείμενη μελέτη σε πειραματικά μοντέλα.

Abstract

This PhD thesis dealt with meta-analysis methodologies and its statistical techniques and applications in high-throughput experiments, and in particular on gene expression data from microarray experiments. Meta-analysis is a valuable tool for the synthesis of evidence across a wide range of study types including high-throughput experiments such as genome-wide association studies (GWAS) and gene expression studies.

One of the issues that this thesis dealt with, was the investigation of optimal meta-analysis techniques. Various tests were performed on a relatively small and a larger random testing dataset (99 and 462 individuals respectively) using the old but classic method of Hedges correction and investigating whether better results can be obtained using the resampling method Bootstrap (of 200 or 500 repetitions). The results showed that the use of the Hedges correction is almost always necessary in small datasets. It was also investigated whether Bootstrap (200 repetitions) correction is giving better results, as being told, than using the Hedges correction only. Which was confirmed in the present study, however, it takes much longer. Point was also made that using 200 Bootstrap repetitions gives almost the same results in statistical power and utility with the time consuming 500 Bootstrap repetitions (which is considered more reliable in the literature).

In the context of investigating various multifactorial diseases meta-analysis of gene expression data from microarray experiments related to preeclampsia was performed; to identify differentially expressed genes, to help elucidate the identification of the disease etiological mechanisms and even help deciphering differentiating mechanisms underlying the different subtypes of preeclampsia. According to this the databases Pubmed and GEO were searched for gene expression data on placental samples from case-control studies. Meta-analysis was performed, and 629 statistically significant differentially expressed genes were identified to be associated with preeclampsia at an FDR 1%. The identified differentially expressed genes were further analyzed with bioinformatics tools and we retrieve interaction networks and overrepresented biochemical pathways, in which they are taking part in. Network analysis showed few, highly interconnected, genes involved in innate immunity and signal transduction pathways which is indicative of a multifaceted disease with etiological heterogeneity. Over-representation analysis revealed that these genes participate mainly in carbohydrates, amino acids and pyrimidine metabolism, circadian clock system and signal transduction pathways.

As noted, meta-analysis is a powerful tool for the synthesis of research studies. There are situations though, in which we have multiple outcomes or multiple treatments, in which multivariate meta-analysis framework which performs a joint modeling of the different quantities of interest could offer important advantages, such as increasing statistical power and allowing performing global tests. For these reasons, we adapted the multivariate meta-analysis methodology so that it could be applied in gene expression data, and we create an easy and reliable software, using the statistical package Stata. The implementation of the method is available at www.compgen.org/tools/multivariate-microarrays.

Last but not least using the aforementioned software we performed a multivariate meta-analysis of microarray data in Crohn's disease and ulcerative colitis, which are the main forms of Inflammatory Bowel Disease. They share similar symptoms but differ in the location and extent of inflammation and their complications. 249 differentially expressed genes were identified to be associated with Crohn's disease and 38 with ulcerative colitis at an FDR of 1%. 20 of the differentially expressed genes were common to both diseases. A global test identified 260 differentially expressed genes associated with inflammatory bowel disease, 53 of which were not found in any of the disorders. With the use of bioinformatics methods important molecular pathways were identified, which could be implicated in the pathogenesis of inflammatory bowel disease, such as the JAK/STAT and interferon-gamma signaling pathways. Additionally genes involved in cell adhesion, apoptosis and carcinogenesis were found to be statistically important. Among others BCAT1 and GZMB are interesting novel differentially expressed genes that deserve further investigation in experimental models.

Περιεχόμενα

Κεφάλαιο 1 ^ο : Εισαγωγή	1
1.1 Το Γενετικό Υλικό	3
1.2 Το «Κεντρικό Δόγμα» της Μοριακής Βιολογίας.....	5
1.3 Πρωτεΐνες.....	8
1.4 Γονιδιακή Έκφραση	9
1.5 Μικροσυστοιχίες	11
1.5.1 Κατασκευή των μικροσυστοιχιών	12
1.5.2 Η Πειραματική διαδικασία των μικροσυστοιχιών.....	14
1.5.3 Βάσεις δεδομένων μικροσυστοιχιών	18
1.6 Μετα-ανάλυση	19
1.6.1 Πλεονεκτήματα μετα-ανάλυσης	19
1.6.2 Μεθοδολογία μετα-ανάλυσης.....	20
1.6.3 Έλεγχος ετερογένειας	21
1.6.4 Πολυμεταβλητή Μετα-ανάλυση (Multivariate meta-analysis).....	21
1.6.4.1 Πλεονεκτήματα της πολυμεταβλητής μετα-ανάλυσης	22
1.6.5 Πλειοτροπία	22
1.6.6 Μετα-ανάλυση σε δεδομένα μικροσυστοιχιών.....	23
1.6.7 Μέθοδοι διόρθωσης πολλαπλών συγκρίσεων	24
1.6.8 Μέθοδοι αναδειγματοληψίας (Resampling methods).....	25
1.7 Προεκλαμψία	25
1.8 Ιδιοπαθής Φλεγμονώδης Νόσος του Εντέρου.....	26
1.8.1 Νόσος του Crohn	26
1.8.2 Ελκώδης Κολίτιδα	27
1.9 Σκοπός της Παρούσας Διατριβής	30

Κεφάλαιο 2 ^ο : Υλικά και Μέθοδοι.....	33
2.1 Ανάλυση δεδομένων μικροσυστοιχιών.....	35
2.1.1 Συλλογή Δεδομένων	35
2.1.2 Ανάκτηση και καταγραφή δεδομένων γονιδιακής έκφρασης από τις μελέτες.....	36
2.1.3 Μεθοδολογία ανάλυσης δεδομένων μικροσυστοιχιών.....	36
2.1.4 Ονοματολογία γονιδίων	37
2.2 Μεθοδολογία μετα-ανάλυσης.....	39
2.2.1 Στατιστικό μοντέλο.....	39
2.2.1.1 Μοντέλο σταθερών επιδράσεων (fixed effects model, FEM)	40
2.2.1.2 Μοντέλο τυχαίων επιδράσεων (random effects model, REM).....	40
2.2.2 Έλεγχος ετερογένειας	41
2.3 Μεθοδολογία μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών	43
2.3.1 Μέθοδος βασισμένη στο t-test.....	43
2.3.2 Μέθοδος συνδυασμού των p-values	44
2.3.3 Μέθοδος υπολογισμού του γινομένου των βαθμών κατάταξης (rank product)	45
2.3.4 Στατιστικά πακέτα μετα-ανάλυσης.....	46
2.3.4.1 R.....	46
2.3.4.2 Stata.....	47
2.4 Μεθοδολογία πολυμεταβλητής μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών.....	47
2.4.1 Συνολικοί έλεγχοι (global tests).....	49
2.5 Μέθοδοι διόρθωσης πολλαπλών συγκρίσεων	50
2.6 Μέθοδοι αναδειγματοληψίας.....	51
2.7 IDR και IRR.....	52
2.8 Εργαλεία εμπλουτισμού.....	53
2.8.1 PANTHER.....	54

2.8.2 STRING	54
2.8.3 MGI.....	55
Κεφάλαιο 3 ^ο : Αποτελέσματα	57
3.1 Διερεύνηση της μεθοδολογίας της μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών.....	59
3.1.1 Σύγκριση εφαρμογής και μη εφαρμογής διόρθωσης Hedges	59
3.1.2 Σύγκριση εφαρμογής διόρθωσης Hedges χωρίς και με 200 επαναλήψεις Bootstrap..	61
3.1.3 Έλεγχος αριθμών επαναλήψεων Bootstrap με εφαρμογή διόρθωσης Hedges	63
3.2 Μετα-ανάλυση δεδομένων γονιδιακής έκφρασης στην προεκλαμψία	65
3.2.1 Αποτελέσματα.....	65
3.3 Πολυμεταβλητή μετα-ανάλυση δεδομένων γονιδιακής έκφρασης της ιδιοπαθούς φλεγμονώδους νόσου του εντέρου.....	85
3.3.1 Αποτελέσματα.....	85
3.4 Δημοσιεύσεις σε επιστημονικά περιοδικά	93
Κεφάλαιο 4 ^ο : Συζήτηση - Συμπεράσματα	95
4.1 Συμπεράσματα της διερεύνηση της μεθοδολογίας της μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών	97
4.2 Συμπεράσματα της μετα-ανάλυσης δεδομένων γονιδιακής έκφρασης στην προεκλαμψία	98
4.3 Συμπεράσματα της μεθοδολογίας της πολυμεταβλητής μετα-ανάλυσης δεδομένων γονιδιακής έκφρασης	101
4.4 Συμπεράσματα της πολυμεταβλητής μετα-ανάλυσης σε δεδομένα γονιδιακής έκφρασης για την Ιδιοπαθή Φλεγμονώδη Νόσο του Εντέρου.....	102
Βιβλιογραφία	107
Παράρτημα.....	129

Εικόνες

Εικόνα 1 Η δομή του DNA.....	4
Εικόνα 2 Σχηματική απεικόνιση του «Κεντρικού Δόγματος» της Μοριακής Βιολογίας	5
Εικόνα 3 Γονιδιακή έκφραση	9
Εικόνα 4 Το πειραματικό πρωτόκολλο των μικροσυστοιχιών.	13
Εικόνα 5 Συνδυασμένη εικόνα μικροσυστοιχίας.....	16
Εικόνα 6 Ένας απλός χάρτης γονότυπου-φαινοτύπου που δείχνει το φαινόμενο της πλειοτροπίας.	23
Εικόνα 7 Περιοχές φλεγμονής της νόσου του Crohn και της ελκώδους κολίτιδας αντίστοιχα. ..	28
Εικόνα 8 Ένα τμήμα του πίνακα γονιδιακής έκφρασης, από μια μελέτη μικροσυστοιχίας. Εμφανίζονται πολλαπλοί ανιχνευτές του ίδιου γονιδίου με διαφορετικές τιμές έκφρασης.	38
Εικόνα 9 Ένα τμήμα του πίνακα γονιδιακής έκφρασης, από μια μελέτη μικροσυστοιχίας. Η τιμή έκφρασης κάθε γονιδίου υπολογίστηκε από τον μέσο όρο των πολλαπλών ανιχνευτών.....	38
Εικόνα 10 Το μοντέλο σταθερών επιδράσεων.....	40
Εικόνα 11 Το μοντέλο τυχαίων επιδράσεων.	41
Εικόνα 12 Διάγραμμα Venn των στατιστικώς σημαντικών γονιδίων τα οποία προέκυψαν από τη μετα-ανάλυση με και χωρίς διόρθωση Hedges σε ένα μικρό σύνολο δεδομένων (99 άτομα).	60
Εικόνα 13 Διάγραμμα Venn των στατιστικώς σημαντικών γονιδίων τα οποία προέκυψαν από τη μετα-ανάλυση με και χωρίς διόρθωση Hedges σε ένα μεγάλο σετ δεδομένων (426 άτομα).	60
Εικόνα 14 Διάγραμμα Venn των στατιστικώς σημαντικών γονιδίων τα οποία προέκυψαν από τη μετα-ανάλυση με και χωρίς bootstrap 200 επαναλήψεων σε ένα μικρό σύνολο δεδομένων (99 άτομα).	61
Εικόνα 15 Διάγραμμα Venn των στατιστικώς σημαντικών γονιδίων τα οποία προέκυψαν από τη μετα-ανάλυση με και χωρίς bootstrap 200 επαναλήψεων σε ένα μεγάλο σύνολο δεδομένων (426 άτομα).	61
Εικόνα 16 Διάγραμμα του αρνητικού λογάριθμου των p -values όλων των γονιδίων σε ένα μικρό σύνολο δεδομένων (99 άτομα) σε σχέση με την χρήση ή όχι της μεθόδου Bootstrap με 200 επαναλήψεις.....	62

Εικόνα 17 Διάγραμμα του αρνητικού λογάριθμου των p -values όλων των γονιδίων σε ένα μεγάλο σύνολο δεδομένων (426 άτομα) σε σχέση με την χρήση ή όχι της μεθόδου Bootstrap με 200 επαναλήψεις.....	62
Εικόνα 18 Διάγραμμα Venn των στατιστικώς σημαντικών γονιδίων τα οποία προέκυψαν από τη μετα-ανάλυση με διαφορετικό αριθμό επαναλήψεων. Συγκεκριμένα 200 εναντίον 500 επαναλήψεων.	63
Εικόνα 19 Διάγραμμα του αρνητικού λογάριθμου των p -values όλων των γονιδίων σε σχέση με τον αριθμό των επαναλήψεων της μεθόδου Bootstrap. Συγκεκριμένα 200 εναντίον 500 επαναλήψεων.	64
Εικόνα 20 Διάγραμμα ροής της μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών για την ασθένεια της προεκλαμψίας.....	65
Εικόνα 21 Διάγραμμα Venn σύγκρισης των ΔΕΓ που προέκυψαν από τις μελέτες μεμονωμένα και από τη μετα-ανάλυση.....	71
Εικόνα 22 Δίκτυο συσχέτισης γονιδίων/πρωτεϊνών των 525 ΔΕΓ (από τα 629) για την προεκλαμψία, σε FDR 1%, με τιμή ορίου αλληλεπίδρασης 0,9.....	73
Εικόνα 23 Δίκτυο συσχέτισης γονιδίων/πρωτεϊνών των τοπ 50 (41 αναγνωρίστηκαν) ΔΕΓ τα οποία αναγνωρίστηκαν για την συνολική προεκλαμψία, με τιμή ορίου αλληλεπίδρασης 0,4.....	74
Εικόνα 24 Διάγραμμα Venn της σύγκρισης των ΔΕΓ που προέκυψαν από τη μετα-ανάλυση της συνολικής προεκλαμψίας και των υποτύπων της.	74
Εικόνα 25 Δίκτυο συσχέτισης γονιδίων/πρωτεϊνών των 1.071 ΔΕΓ (από τα 1.342) για την προεκλαμψία, σε FDR 5%, με τιμή ορίου αλληλεπίδρασης 0,9.....	79
Εικόνα 26 Διάγραμμα ροής της πολυμεταβλητής μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών για τις ασθένειες της ΙΦΝΕ.....	85
Εικόνα 27 Διάγραμμα Venn ταξινόμησης των ΔΕΓ που προέκυψαν από τη μετα-ανάλυση.....	87
Εικόνα 28 Δίκτυο συσχέτισης γονιδίων/πρωτεϊνών των 232 (από τα 260) ΔΕΓ τα οποία αναγνωρίστηκαν από τον συνολικό έλεγχο, με τιμή ορίου αλληλεπίδρασης 0,7.....	89
Εικόνα 29 Δίκτυο συσχέτισης γονιδίων/πρωτεϊνών των 224 (από τα 249) ΔΕΓ για την νόσο του Crohn, με τιμή ορίου αλληλεπίδρασης 0,7.	90
Εικόνα 30 Δίκτυο συσχέτισης γονιδίων/πρωτεϊνών των 29 (από τα 38) ΔΕΓ για την ελκώδη κολίτιδα, με τιμή ορίου αλληλεπίδρασης 0,7.	91

Εικόνα 31 Δίκτυο συσχέτισης γονιδίων/πρωτεϊνών των 49 (από τα 53) ΔΕΓ τα οποία εντοπίστηκαν μόνο από τον συνολικό έλεγχο, με τιμή ορίου αλληλεπίδρασης 0,7.....	91
Εικόνα 32 Διάγραμμα Venn σύγκρισης των στατιστικώς σημαντικών μονοπατιών της PANTHER μεταξύ των ασθενειών και του συνολικού ελέγχου.....	93

Πίνακες

Πίνακας 1 Ο Γενετικός Κώδικας	8
Πίνακας 2 Ευρήματα που διαφοροποιούν τη νόσο του Crohn από την ελκώδη κολίτιδα	29
Πίνακας 3 Χαρακτηριστικά των μελετών που συμπεριλήφθηκαν στη μετα-ανάλυση για την ασθένεια της προεκλαμψίας.....	66
Πίνακας 4 Τα βιοχημικά μονοπάτια που συμμετέχουν τα στατιστικώς σημαντικά ΔΕΓ όπως προέκυψαν από την PANTHER, για την συνολική προεκλαμψία	76
Πίνακας 5 Τα βιοχημικά μονοπάτια που συμμετέχουν τα στατιστικώς σημαντικά ΔΕΓ όπως προέκυψαν από την PANTHER, για την ήπια προεκλαμψία	77
Πίνακας 6 Τα βιοχημικά μονοπάτια που συμμετέχουν τα στατιστικώς σημαντικά ΔΕΓ όπως προέκυψαν από την PANTHER, για τη σοβαρή προεκλαμψία.....	78
Πίνακας 7 Τα βιοχημικά μονοπάτια που συμμετέχουν τα στατιστικώς σημαντικά ΔΕΓ όπως προέκυψαν από την PANTHER, για την προεκλαμψία σε FDR 5%	80
Πίνακας 8 Σύγκριση μεταξύ προηγούμενων μετα-αναλύσεων και της συγκεκριμένης μελέτης .	83
Πίνακας 9 Χαρακτηριστικά των μελετών που χρησιμοποιήθηκαν στην πολυμεταβλητή μετα-ανάλυση	86
Πίνακας 10 Στατιστικώς σημαντικά PANTHER βιοχημικά μονοπάτια στα οποία συμμετέχουν τα ΔΕΓ της μετα-ανάλυσης	92

Κεφάλαιο 1^ο:
Εισαγωγή

1.1 Το Γενετικό Υλικό

Η επιβίωση όλων των ζωντανών οργανισμών που υπήρχαν και υπάρχουν σε αυτό τον πλανήτη βασίζεται στην ικανότητα των κυττάρων να αποθηκεύουν, να ανακαλούν και να μεταφράζουν γενετικές πληροφορίες απαραίτητες για την κάθε κατάσταση. Τα ερωτήματα γύρω από το ποιος, πού και γιατί είναι πολύ παλιά, ωστόσο απαντήθηκαν σχετικά πρόσφατα. Η ύπαρξη του δεοξυριβονουκλεϊκού οξέος (Deoxyribonucleic Acid, DNA) στον πυρήνα των κυττάρων πιστοποιήθηκε το 1869 από τον Miescher, στα μέσα όμως του 20^{ου} αιώνα οι ερευνητές ξεκίνησαν να υποθέτουν ότι μπορεί να αποθηκεύει την γενετική πληροφορία. Αυτή η ιδιότητά του επιβεβαιώθηκε το 1944 από τους Avery, MacLeod, και McCarty (Avery, Macleod et al. 1944) και το 1952 από τους Hershey και Chase οι οποίοι έδειξαν ότι οι βακτηριοφάγοι (ιοί) μολύνουν τα βακτήρια εισάγοντάς τους το γονικό DNA (Hershey and Chase 1952). Το 1953 οι Watson και Crick, με τη βοήθεια δεδομένων της Franklin και την επιβεβαίωση του Wilkins, πρότειναν τη δομή του DNA με το μοντέλο της διπλής έλικας (Watson and Crick 1953).

Το DNA είναι ένα μακρομόριο που απαρτίζεται από νουκλεοτίδια. Τα νουκλεοτίδια αποτελούνται από μια πεντόζη, συγκεκριμένα μια δεοξυριβόζη στην οποία συνδέεται μια φωσφορική ομάδα και μια αζωτούχος βάση. Στο μόνο που διαφέρουν τα νουκλεοτίδια μεταξύ τους είναι στην αζωτούχο βάση, η οποία μπορεί να είναι είτε Αδενίνη (A), είτε Κυτοσίνη (C), είτε Γουανίνη (G), είτε Θυμίνη (T). Τα νουκλεοτίδια συνδέονται μεταξύ τους με ομοιοπολικούς δεσμούς, μεταξύ του υδροξυλίου του 3'-άνθρακα της πεντόζης του πρώτου νουκλεοτιδίου και της φωσφορικής ομάδας του 5'-άνθρακα του επόμενου νουκλεοτιδίου, σχηματίζοντας μια αλυσίδα εναλλασσόμενων μονάδων σακχάρου-φωσφορικής ομάδας-σακχάρου και ούτω καθεξής με κατεύθυνση 5'-3'. Η κατεύθυνση αυτή υπάρχει καθώς στο πρώτο νουκλεοτίδιο της αλυσίδας είναι ελεύθερη η φωσφορική ομάδα του 5'-άνθρακα της πεντόζης και στο τελευταίο νουκλεοτίδιο της αλυσίδας είναι ελεύθερο το υδροξύλιο του 3'-άνθρακα. Όπως προαναφέραμε το DNA έχει μια δομή διπλής έλικας, όπου οι δύο αντιπαράλληλες (η πολικότητα του ενός κλώνου είναι αντίθετου προσανατολισμού από την πολικότητα του άλλου) αλυσίδες συγκρατούνται με δεσμούς υδρογόνου (Εικόνα 1). Οι δεσμοί υδρογόνου σχηματίζονται μεταξύ των αζωτούχων βάσεων των δύο αντιπαράλληλων αλυσίδων· συγκεκριμένα σχηματίζονται 2 δεσμοί υδρογόνου μεταξύ της A και της T και 3 δεσμοί μεταξύ της C και της G (Alberts, Bray et al. 2004).



Εικόνα 1 Η δομή του DNA

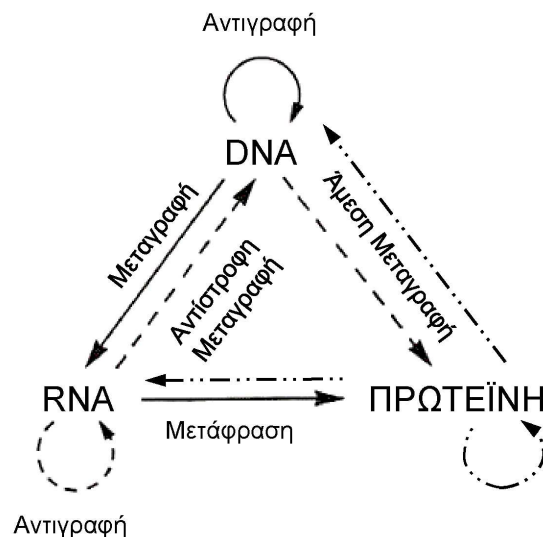
Ένα εξίσου σημαντικό μακρομόριο δεν είναι άλλο από το ριβονουκλεϊκό οξύ (Ribonucleic Acid, RNA), το οποίο ανακαλύφθηκε μαζί με το DNA το 1896. Το RNA απαρτίζεται και αυτό από νουκλεοτίδια, τα οποία αποτελούνται από μια πεντόζη, συγκεκριμένα την ριβόζη, μια φωσφορική ομάδα και μια αζωτούχο βάση. Οι αζωτούχες βάσεις του RNA είναι και πάλι τέσσερις και είναι η Αδενίνη (A), η Κυτοσίνη (C), η Γουανίνη (G), όπως και στο DNA, αλλά αντί της Θυμίνης (T) (το RNA) έχει την Ουρακίλη (U). Η Ουρακίλη σχηματίζει και αυτή με την Αδενίνη δύο δεσμούς υδρογόνου (Alberts, Bray et al. 2004). Το RNA είναι μικρότερο σε μέγεθος από το DNA και κατά κύριο λόγο είναι μονόκλωνο. Έχει γίνει γνωστό ότι υπάρχουν πάνω από 30 διαφορετικοί τύποι του (Antisense RNA, Long noncoding RNA, MicroRNA, Retrotransposon κ.α.) σε διάφορους οργανισμούς με πιο γνωστούς τους παρακάτω τέσσερις κύριους τύπους.

- Το αγγελιοφόρο RNA (messenger RNA, mRNA) το οποίο μεταφέρει τη γενετική πληροφορία για τη σύνθεση των πρωτεϊνών, καθώς είναι το μετάγραφο που προκύπτει κατά τη μεταγραφή του DNA,
- Το μεταφορικό RNA (transfer RNA, tRNA) το οποίο μεταφέρει, με βάση τον κανόνα της συμπληρωματικότητας, τα κατάλληλα αμινοξέα κατά τη μετάφραση του mRNA,
- Το ριβοσωμικό RNA (ribosomal RNA, rRNA), το οποίο σχηματίζει το κέντρο των ριβοσωμάτων, στα οποία γίνεται η μετάφραση του mRNA σε πρωτεΐνη και

- Το μικροπυρηνικό RNA (small-nuclear RNA, snRNA) το οποίο συμμετέχει μαζί με κάποιες πρωτεΐνες στην διαδικασία ωρίμανσης του mRNA στους ευκαρυωτικούς οργανισμούς.

1.2 Το «Κεντρικό Δόγμα» της Μοριακής Βιολογίας

Το 1958 ο Crick δημοσίευσε το «Κεντρικό Δόγμα» της Μοριακής Βιολογίας (Crick 1970), το οποίο ήταν η απεικόνιση της ροής της γενετικής πληροφορίας μέσα στα βιολογικά συστήματα. Η άποψη ήταν ότι η πληροφορία υπάρχει στο DNA μεταφέρεται στο RNA και καταλήγει στις πρωτεΐνες. Η πληροφορία δεν είναι άλλο από την ακριβή ακολουθία των νουκλεοτιδίων. Βέβαια ο Crick είχε αναφέρει ότι υπάρχουν εννέα διαφορετικές κατευθύνσεις που μπορούν να προκύψουν από αυτή την απεικόνιση, από τις οποίες τρεις είναι οι γενικές, τρεις είναι οι ειδικές και τρεις οι άγνωστες (Cobb 2017). Η σχηματική αναπαράσταση των διαφορετικών κατευθύνσεων φαίνεται στην Εικόνα 2.



Εικόνα 2 Σχηματική απεικόνιση του «Κεντρικού Δόγματος» της Μοριακής Βιολογίας

Οι γενικές είναι (Εικόνα 2 με συνεχή γραμμή):

- Η αντιγραφή, δηλαδή η διεργασία του αυτοδιπλασιασμού του DNA. Συμβαίνει με ημισυντηρητικό τρόπο και από αυτήν προκύπτουν δύο θυγατρικά κύτταρα ταυτόσημα ως προς τις γενετικές τους πληροφορίες, τα οποία περιέχουν μια παλιά και μία νεοσυντιθέμενη αλυσίδα. Στην αντιγραφή συμμετέχουν αρκετά ένζυμα, ορισμένα από τα οποία είναι η DNA ελικάση, η DNA πολυμεράση, η DNA δεσμάση, τα επιδιορθωτικά ένζυμα κ.α. (Alberts, Bray et al. 2004),

- Η μεταγραφή, η διεργασία κατά την οποία τμήματα του DNA, τα γονίδια, μεταγράφονται σε RNA με βάση τον κανόνα συμπληρωματικότητας [A απέναντι από U (στην περίπτωση του RNA)/T (στην περίπτωση του DNA) και G απέναντι από C και το αντίστροφο]. Η μεταγραφή για να εκτελεστεί χρειάζεται ειδικές αλληλουχίες DNA πριν από το ενεργό γονίδιο (εκκινητές), ειδικά ένζυμα (RNA πολυμεράσες) και ειδικές πρωτεΐνες (μεταγραφικοί παράγοντες, ρυθμιστικές πρωτεΐνες). Με τη διαδικασία της μεταγραφής παράγεται ένα μονόκλωνο μόριο mRNA, το οποίο περιέχει τη γενετική πληροφορία του γονιδίου (Alberts, Bray et al. 2004),
- Η τρίτη γενική διεργασία δεν είναι άλλη από τη μετάφραση, κατά την οποία η αλληλουχία των νουκλεοτιδίων του mRNA με τη βοήθεια των ριβοσωμάτων (ένα σύμπλοκο παρασκευής πρωτεϊνών) και των tRNAs μεταφράζεται σε αλληλουχία αμινοξέων με τελικό αποτέλεσμα την παραγωγή της πρωτεΐνης (Alberts, Bray et al. 2004).
Οι ειδικές είναι (Εικόνα 2 με διακεκομμένη γραμμή):
- Η αντίστροφη μεταγραφή, είναι μια διεργασία κατά την οποία δημιουργείται DNA με βάση την αλληλουχία του RNA και φυσικά ακολουθώντας τον κανόνα της συμπληρωματικότητας. Αρχικά ανακαλύφθηκε ότι αυτή η διεργασία πραγματοποιείται όταν ένας ρετροϊός (είδος RNA ιού) μολύνει ένα ευκαρυωτικό κύτταρο και αντιγράφει το γονιδιώμα του (Baltimore 1970, Temin and Mizutani 1970). Στη συνέχεια βρέθηκε να πραγματοποιείται και στους ευκαρυώτες και ότι χρησιμοποιείται στα κινητά μεταφερόμενα στοιχεία γενετικής ποικιλότητας, τα ρετροτρανσποζόνια (Dombroski, Feng et al. 1994) και κατά την αντιγραφή των τελομερών (χρωμοσωμικά άκρα) (Weinrich, Pruzan et al. 1997). Τέλος στα βακτήρια αντίστροφη μεταγραφή πραγματοποιείται κατά τη σύνθεση εξωχρωμοσωματικών χιμαιρικών DNA/RNA, μονόκλωνων πολλαπλών αντιγράφων DNA (multicopy single-stranded DNA, msDNA) (Inouye, Herzer et al. 1990, Poole, Andrews et al. 2001),
- Η αντιγραφή του RNA βρέθηκε να συμβαίνει κυρίως στους RNA ιούς κατά την δημιουργία νέων ιών (Inouye, Herzer et al. 1990),
- Η απευθείας μεταγραφή από DNA σε πρωτεΐνη έχει καταδειχθεί σε ένα σύστημα χωρίς κύτταρα (in vitro), ωστόσο, δεν είναι σαφές εάν αυτός ο μηχανισμός μετάφρασης θα μπορούσε να αντιστοιχιστεί ειδικά στον γενετικό κώδικα (McCarthy and Holland 1965, Uzawa, Yamagishi et al. 2002).

Οι άγνωστες είναι (Εικόνα 2 με διακεκομμένη εστιγμένη γραμμή):

- Από πρωτεΐνη σε DNA,
- Από πρωτεΐνη σε RNA,
- Αντιγραφή της πρωτεΐνης.

Από τις εννέα διεργασίες που αναφέραμε, στη συγκεκριμένη εργασία μας ενδιέφεραν περισσότερο η μεταγραφή και η μετάφραση. Όπως προαναφέραμε κατά τη διαδικασία της μεταγραφής παράγεται ένα μονόκλωνο μόριο mRNA, το οποίο περιέχει τη γενετική πληροφορία του γονιδίου. Εάν το κύτταρο, το οποίο παράγει το μετάγραφο είναι ευκαρυωτικό τότε παράγεται αρχικά ένα πρόδρομο μόριο mRNA, το οποίο δέχεται τροποποιήσεις μέσα στον πυρήνα του κυττάρου από ένα σύμπλοκο, το σωματίδιο συρραφής (το οποίο αποτελείται από πρωτεΐνες και RNA). Κατά αυτή την τροποποίηση οι κωδικοποιητικές αλληλουχίες του γονιδίου (εξώνια), συρράπτονται μεταξύ τους καθώς έχουν αφαιρεθεί οι μη κωδικοποιητικές αλληλουχίες (ιντρόνια) οι οποίες τις διακόπτουν. Φυσικά πριν από αυτή τη διαδικασία και την έξοδο του ώριμου mRNA στο κυτταρόπλασμα έχουν προηγηθεί ο σχηματισμός της καλύπτρας του RNA (συμβαίνει μια τροποποίηση στο 5' άκρου) και η πολυαδενυλίωση (συμβαίνει στο 3' άκρο και του προσδίδει μια εξειδικευμένη δομή).

Στη συνέχεια από το ώριμο mRNA, με τη βοήθεια των ριβοσωμάτων (ένα σύμπλοκο παρασκευής πρωτεϊνών) και των tRNAs, μεταφράζεται η αλληλουχία των νουκλεοτιδίων σε αλληλουχία αμινοξέων με τελικό αποτέλεσμα την παραγωγή της πρωτεΐνης. Όπως είναι γνωστό η αλληλουχία των νουκλεοτιδίων του mRNA διαβάζεται ανά τριάδες, γνωστές ως κωδικόνια, και κάθε κωδικόνιο αντιστοιχεί σε ένα αμινοξύ. Αυτές οι αντιστοιχίσεις συνολικά αναφέρονται ως ο γενετικός κώδικας (Πίνακας 1).

Ο Γενετικός κώδικας είναι οικουμενικός (είναι ίδιος σχεδόν σε όλους τους οργανισμούς), είναι κώδικας τριπλέτας όπως προαναφέραμε (τρία νουκλεοτίδια αντιστοιχούν σε ένα αμινοξύ), είναι συνεχής και μη επικαλυπτόμενος (δεν υπάρχουν νουκλεοτίδια τα οποία είτε να μην ανήκουν σε κανένα κωδικόνιο είτε να ανήκουν σε δυο), είναι εκφυλισμένος (περισσότερα του ενός κωδικόνια μπορούν να αντιστοιχούν στο ίδιο αμινοξύ, εξαίρεση αποτελούν τα αμινοξέα Τρυπτοφάνη και Μεθειονίνη) και τέλος υπάρχει ένα κωδικόνιο έναρξης (AUG το οποίο κωδικοποιεί το αμινοξύ Μεθειονίνη) καθώς και τρία κωδικόνια λήξης τα οποία δεν κωδικοποιούν κανένα αμινοξύ, αλλά σηματοδοτούν τη λήξη της μετάφρασης (UGA, UAA, UAG) (Alberts, Bray et al. 2004).

Πίνακας 1 Ο Γενετικός Κώδικας

Δεύτερο Γράμμα									
Πρώτο Γράμμα	U		C		A		G		Τρίτο Γράμμα
	U	UUU	Φαινυλαλαλίνη (F)	UCU		UAU	Τυροσίνη (Y)	UGU	
	UUC		UCC		UAC		UGC		C
	UUA		UCA	Σερίνη (S)	UAA		UGA	Λήξη	A
	UUG	Λευκίνη (L)	UCG		UAG	Λήξη	UGG	Τρυπτοφάνη (W)	G
C	CUU		CCU		CAU	Ιστιδίνη (H)	CGU		U
	CUC	Λευκίνη (L)	CCC	Προλίνη (P)	CAC		CGC	Αργινίνη (R)	C
	CUA		CCA		CAA	Γλουταμίνη (Q)	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU		ACU		AAU	Ασπαραγίνη (N)	AGU	Σερίνη (S)	U
	AUA	Ισολευκίνη (I)	ACC		AAC		AGC		C
	AUC		ACA	Θρεονίνη (T)	AAA		AGA		A
	AUG	Μεθειονίνη (M)	ACG		AAG	Λυσίνη (K)	AGG	Αργινίνη (R)	G
G	GUU		GCU		GAU	Ασπαρτικό οξύ (D)	GGU		U
	GUA	Βαλίνη (V)	GCC	Αλανίνη (A)	GAC		GGC	Γλυκίνη (G)	C
	GUC		GCA		GAA	Γλουταμινικό οξύ (E)	GGA		A
	GUG		GCG		GAG		GGG		G

1.3 Πρωτεΐνες

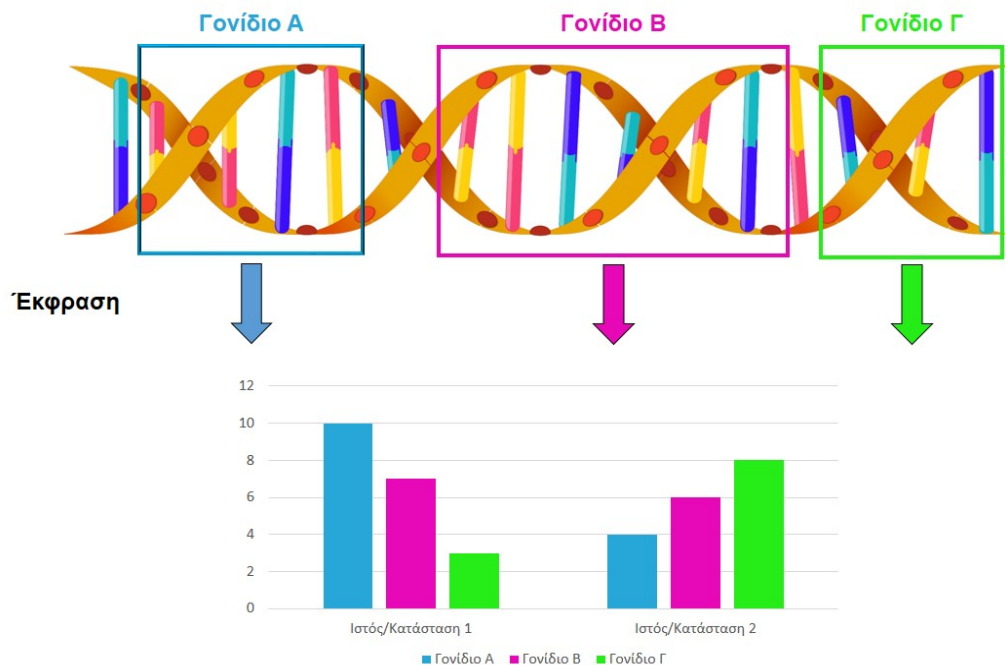
Οι πρωτεΐνες είναι μόρια μεγάλης βιολογικής σημασίας. Συνήθως ρυθμίζουν συγκεκριμένες βιολογικές διεργασίες, συμμετέχουν στη μεταφορά ιόντων και μορίων μέσα στα κύτταρα, καταλύουν αντιδράσεις (δρώντας κυρίως ως ένζυμα), συμμετέχουν στην προστασία του οργανισμού από διάφορους εισβολείς, αποτελούν θρεπτικά συστατικά, συμβάλλουν στην υποστήριξη συγκεκριμένων δομών κ.α.

Δομικοί λίθοι των πρωτεϊνών αποτελούν τα αμινοξέα. Τα αμινοξέα όπως προαναφέραμε είναι μικρά οργανικά μόρια, υπάρχουν 20 διαφορετικά και όταν συνδυαστούν μεταξύ τους

σχηματίζουν μεγάλους μήκους αλυσίδες, τα πολυπεπίδια. Τα πολυπεπίδια για να είναι λειτουργικά πρέπει να έχουν φυσικά την κατάλληλη αλληλουχία αλλά και την κατάλληλη τρισδιάστατη δομή. Η κατάλληλη αλληλουχία προκύπτει, όπως προαναφέραμε, από τη μεταγραφή της αλληλουχίας του DNA και πιο συγκεκριμένα ενός γονιδίου. Φυσικά μια πρωτεΐνη μπορεί να αποτελείται από περισσότερα του ενός πολυπεπίδια, και άρα να προέρχεται από περισσότερα του ενός γονίδια (Tropp 2012).

1.4 Γονιδιακή Έκφραση

Η γονιδιακή έκφραση είναι η διαδικασία κατά την οποία χρησιμοποιούνται οι πληροφορίες ενός γονιδίου για τη σύνθεση ενός λειτουργικού γονιδιακού προϊόντος και αναφέρεται κυρίως στην ποσότητα αυτού του παραγόμενου προϊόντος σε μια συγκεκριμένη χρονική στιγμή μέσα στο κύτταρο. Τα προϊόντα αυτά μπορεί να είναι είτε πρωτεΐνες είτε διάφορα είδη RNAs (tRNAs, miRNAs, snRNAs κ.α.). Είναι γνωστό ότι σε ένα συγκεκριμένο κύτταρο ή ιστό, εκφράζεται μόνο ένα υποσύνολο γονιδίων από ολόκληρο το γονιδίωμα (Εικόνα 3). Η ταυτότητα και η διαφορετική ένταση των εκφραζόμενων γονιδίων καθορίζει τον χαρακτήρα, τη λειτουργία (Hayes and Meyerson 2005), τα διαφορετικά αναπτυξιακά στάδια, είτε την εκδήλωση κάποιας ασθένειας (Luk, Kapushesky et al. 2010) ενός συγκεκριμένου κυτταρικού τύπου.



Εικόνα 3 Γονιδιακή έκφραση

Η ανάλυση της γονιδιακής έκφρασης είναι πρωταρχικής σημασίας στη σύγχρονη Μοριακή Βιολογία. Υπάρχουν διάφορες τεχνικές, με τα θετικά τους και τα αρνητικά τους, για την ανάλυση και ποσοτικοποίηση της γενετικής έκφρασης. Χαρακτηριστικά παραδείγματα είναι οι τεχνικές:

- Της αλυσιδωτής αντίδρασης πολυμεράσης διαφορικής έκφρασης (Differential Display Polymerase Chain Reaction, PCR) (Liang and Pardee 1992), η οποία είναι από τις παλαιότερες τεχνικές μεταφραστικής έκφρασης, με βασικά μειονεκτήματα το ότι μόνο ένα τμήμα των μεταγραφών μπορεί να αναλυθεί σε μια μεμονωμένη αντίδραση και το ότι είναι επιρρεπής σε ψευδώς θετικά αποτελέσματα,
- Της σειριακής ανάλυσης της γονιδιακής έκφρασης (serial analysis of gene expression, SAGE) (Velculescu, Zhang et al. 1995), η οποία είναι τεχνολογία ανοιχτού τύπου, έχει απεριόριστη δυναμική εμβέλεια και ενσωματωμένο έλεγχο ποιότητας, αλλά δεν είναι πολύ χρήσιμη σε γονιδιώματα μη πλήρους αλληλούχησης καθώς είναι δύσκολο να συσχετιστούν τμήματα των 15 ζευγών βάσεων, με τα γονίδια τους,
- Της PCR πραγματικού χρόνου (Real Time, RT) (Holden and Wang 2008), η οποία είναι ευαίσθητη τεχνική με εξαιρετικό ποσοτικό προσδιορισμό υψηλής αναπαραγωγιμότητας, μπορεί να ανιχνεύσει μέχρι και 5 μόρια και θεωρείται ο «χρυσός κανόνας» για τον ποσοτικό προσδιορισμό του mRNA, όμως δεν είναι τεχνική υψηλής απόδοσης (χρειάζεται 10 λεπτά για 100 γονίδια) και είναι επιρρεπής στο να υπάρξει επιμόλυνση RNA, οδηγώντας έτσι σε ψευδώς θετικά αποτελέσματα,
- Του στυπώματος τύπου Northern (Alberts, Wilson et al. 2008), η οποία είναι μια ποσοτική μέθοδος μέτρησης της αφθονίας των μεταγράφων, πολύ χρησιμοποιούμενη και κατανοητή τεχνολογία και διαθέτει άπειρη δυναμική εμβέλεια, έχει, όμως, προβλήματα ελέγχου ποιότητας και χρησιμοποιεί ραδιενεργό σήμανση,
- Των μικροσυστοιχιών, η οποία είναι τεχνική υψηλής απόδοσης μέτρησης της αφθονίας των μεταγράφων, είναι εύκολη στη χρήση, γρήγορη, μπορεί να μελετήσει πολύ μεγάλο αριθμό (σχεδόν το σύνολο) των μεταγράφων και δεν χρησιμοποιεί ραδιενέργεια, αλλά χρειάζεται προσοχή καθώς η ποιότητα και ο έλεγχός της ποικίλλουν και η ανάλυση καθώς και η ερμηνεία των αποτελεσμάτων συχνά είναι δύσκολη,
- Της αλληλούχησης RNA (RNA sequencing, RNA-Seq) (Wang, Gerstein et al. 2009), η οποία είναι από τις πιο πρόσφατες τεχνολογίες, παρέχει με μεγάλη ακρίβεια τη μέτρηση των επιπέδων των μεταγράφων και των ισομορφών τους και δεν απαιτείται υβριδισμός για

να αναλύσει την αλληλουχία, αλλά απαιτεί εγκαταστάσεις υπολογιστών υψηλής ισχύος, εμφανίζει προβλήματα σε οργανισμούς με αρκετές επαναλαμβανόμενες αλληλουχίες οι οποίοι δεν διαθέτουν γνωστό γονιδίωμα και δυσκολεύεται κατά την ανάλυση ενναλακτικού ματίσματος γονιδίων.

1.5 Μικροσυστοιχίες

Οι μικροσυστοιχίες (συγκεκριμένα οι μικροσυστοιχίες DNA, DNA chips) αναπτύχθηκαν αρχικά για την ανάλυση αλληλουχιών, γύρω στα τέλη του 1980, έγιναν όμως ευρέως γνωστές όταν επισημάνθηκε και μια άλλη εφαρμογή τους, η μέτρηση των επιπέδων της γονιδιακής έκφρασης. Στις αρχές του 1990 ξεκίνησε η παραγωγή συσκευασμένων συστοιχιών ολιγονουκλεοτιδικών ανιχνευτών (probes). Οι ανιχνευτές χρησίμευαν στον υβριδισμό στοχευμένων αλληλουχιών DNA (Choudhuri 2004). Υβριδισμός (hybridization) είναι η διαδικασία κατά την οποία μια μονόκλωνη νουκλεοτιδική αλυσίδα (κυρίως DNA), σε συγκεκριμένες συνθήκες περιβάλλοντος (θερμοκρασία, pH), σχηματίζει δεσμούς υδρογόνου με βάση τον κανόνα συμπληρωματικότητας, με μια άλλη μονόκλωνη νουκλεοτιδική αλυσίδα (είτε DNA είτε RNA) και σχηματίζουν διπλή έλικα. Τα νουκλεϊκά μόρια ακόμη και αν έχουν διαφορετική προέλευση, αν είναι συμπληρωματικά θα υβριδοποιηθούν. Η έννοια της υβριδοποίησης αποτελεί τη βάση σε πολλές φυσικές διεργασίες μέσα στα κύτταρα, και ακολούθως τη βάση σε πληθώρα πειραμάτων Μοριακής Βιολογίας.

Η τεχνική την οποία χρησιμοποιούν οι μικροσυστοιχίες δεν ήταν εντελώς νέα, αλλά βασίστηκε και διεύρυνε δύο γνωστές τεχνολογίες. Την τεχνολογία ενζυμικής ανοσοδοκιμασίας ELISA (Δοκιμασία Ενζυμο-Συζευγμένης Ανοσοπροσρόφησης, Enzyme-Linked Immunosorbent Assay) και τις τεχνικές ανίχνευσης ειδικής αλληλουχίας νουκλεϊκού οξέος με μορφή κηλίδων (Choudhuri 2004).

Οι μικροσυστοιχίες παρέχουν τη δυνατότητα μέτρησης γεγονότων τα οποία συμβαίνουν στο γονιδίωμα. Γεγονός μπορεί να είναι η μεταγραφή ενός γονιδίου, η μεταβολή του αριθμού των αντιγράφων μίας γενετικής θέσης (locus), η παρουσία ή η απουσία μίας μετάλλαξης, η σύνδεση μιας πρωτεΐνης σε ένα τμήμα του DNA, μια αλλαγή στη μεθυλίωση του DNA, και γενικά οποιαδήποτε δραστηριότητα ή κατάσταση η οποία σχετίζεται είτε με το DNA είτε με το RNA. Ως γονιδιωματική ανάγνωση, οι μικροσυστοιχίες προσδιορίζουν εάν τα γεγονότα αυτά έχουν συμβεί (Seidel 2008).

Αν και στην αρχή του 1990 οι μικροσυστοιχίες ήταν συνώνυμες με τις μικροσυστοιχίες DNA, υπάρχουν ωστόσο αρκετοί τύποι τους. Χαρακτηριστικά υπάρχουν:

- Μικροσυστοιχίες DNA (DNA microarrays, με διάφορους υπότυπους π.χ. single cell microarray),
- MMChips (για τον έλεγχο microRNAs),
- Μικροσυστοιχίες πρωτεϊνών (Protein microarrays),
- Μικροσυστοιχίες πεπτιδίων (Peptide microarrays),
- Μικροσυστοιχίες ιστών (Tissue microarrays),
- Κυτταρικές Μικροσυστοιχίες είτε μικροσυστοιχίες επιμόλυνσης (Cellular microarrays/transfection microarrays),
- Μικροσυστοιχίες χημικών ενώσεων (Chemical compound microarrays),
- Μικροσυστοιχίες αντισωμάτων (Antibody microarrays),
- Συστοιχίες υδατανθράκων (Glycan arrays, carbohydrate arrays),
- Μικροσυστοιχίες φαινοτύπων (Phenotype microarrays),
- Μικροσυστοιχίες λύσης πρωτεϊνών αντίστροφης φάσης (Reverse phase protein lysate microarrays, RPMA),
- Interferometric reflectance imaging sensor (IRIS).

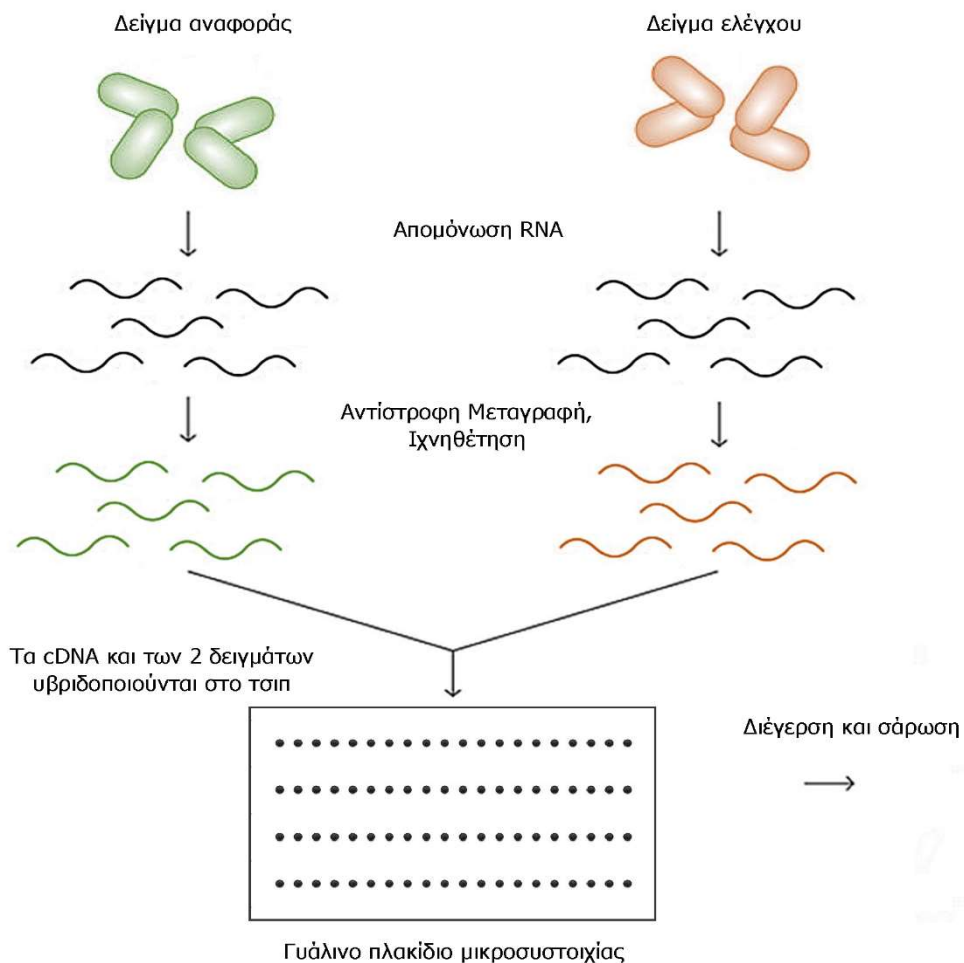
1.5.1 Κατασκευή των μικροσυστοιχιών

Η δομή της μικροσυστοιχίας DNA απαρτίζεται από ένα γυάλινο πλακίδιο (τσιπ, chip) επάνω στο οποίο υπάρχουν συγκεκριμένα τμήματα αλληλουχιών DNA συμπληρωματικά με συγκεκριμένα γονίδια. Αυτά τα τμήματα είναι οι ανιχνευτές (probes) και είναι σταθεροποιημένοι σε συγκεκριμένες θέσεις, σε μια κουκίδα (spot) του τσιπ μικροσυστοιχιών. Σε ένα τσιπ συνήθως υπάρχουν χιλιάδες κουκίδες σε μια σειρά και κάθε μια κουκίδα μπορεί να περιέχει τμήμα αλληλουχίας που να αντιστοιχεί σε ένα μοναδικό γονίδιο. Με χρήση κατάλληλων συνθηκών, οι αλληλουχίες που βρίσκονται στο τσιπ υβριδοποιούνται με αντίγραφα νουκλεϊκών οξέων από βιολογικά δείγματα, τα οποία προηγουμένως έχουν ιχνηθετηθεί με φθορίζουσες ουσίες, τους ονομαζόμενους στόχους (targets) (Brazma and Vilo 2000, Babu 2004, Lee, Hsu et al. 2004).

Οι στόχοι διαφέρουν ανάλογα με το πρωτόκολλο που ακολουθείται. Σε πρωτόκολλο ανάλυσης γονιδιακής έκφρασης πραγματοποιείται σύγκριση των επιπέδων έκφρασης ενός συνόλου γονιδίων από κάποιο κύτταρο συγκεκριμένων συνθηκών και ιδιοτήτων (κατάσταση Α)

σε αντιδιαστολή με τα επίπεδα έκφρασης ίδιου συνόλου γονιδίων προερχόμενα από κάποιο κύτταρο αναφοράς, κύτταρο υπό φυσιολογικές συνθήκες, (κατάσταση B) (Babu 2004). Για τον λόγο αυτό συνήθως χρησιμοποιείται το mRNA των κυττάρων, από το οποίο κατασκευάζεται το συμπληρωματικό του DNA (complementary DNA, cDNA), το οποίο ιχνηθετείται με διαφορετικό χρώμα φθορισμού για την κάθε κατάσταση και, όπως προαναφέρθηκε, υβριδοποιείται με το τσιπ μικροσυστοιχιών.

Ακολούθως, οι κουκίδες στο υβριδοποιημένο πλέον τσιπ μικροσυστοιχιών διεγείρονται από ένα λέιζερ και σαρώνονται στα κατάλληλα, για την ανίχνευση των φθορίζουσών ουσιών, μήκη κύματος. Το μέγεθος του φθορισμού το οποίο εκπέμπεται μετατρέπεται σε μια ποσοτική τελική έξοδο η οποία αντιπροσωπεύει το σχετικό επίπεδο έκφρασης του γονιδίου αυτού, δηλαδή την ποσότητα των αντιγράφων του κάθε mRNA που υπάρχει στο υπό μελέτη δείγμα (Εικόνα 4) (Babu 2004). Οι μικροσυστοιχίες, όπως γίνεται αντιληπτό, πραγματοποιούν μέτρηση των επιπέδων της γονιδιακής έκφρασης χιλιάδων γονιδίων ταυτόχρονα σε ένα πείραμα υβριδοποίησης.



Εικόνα 4 Το πειραματικό πρωτόκολλο των μικροσυστοιχιών.

1^ο βήμα απομονώνεται το βιολογικό υλικό (mRNA) και από τις 2 καταστάσεις, 2^ο βήμα κατασκευάζονται τα αντίστοιχα cDNA που ιχνηθετούνται, 3^ο βήμα τα cDNA υβριδοποιούνται στο τσιπ και 4^ο βήμα το τσιπ διεγείρεται και σαρώνεται ώστε να δώσει την προς επεξεργασία και μελέτη εικόνα.

1.5.2 Η Πειραματική διαδικασία των μικροσυστοιχιών

Κατά την πραγματοποίηση ενός πειράματος μικροσυστοιχιών ακολουθούνται συγκεκριμένα στάδια, τα οποία πρέπει να εκτελεστούν ορθά και με επιτυχία ώστε να διασφαλιστεί ότι προκύπτουν αξιόπιστα και ακριβή συμπεράσματα από τα δεδομένα.

1^ο Στάδιο: Το βιολογικό ερώτημα

Το βιολογικό ερώτημα είναι ο ακρογωνιαίος λίθος του πειράματος. Το βιολογικό ερώτημα πρέπει να διατυπώνεται προσεκτικά και με ακρίβεια πριν ξεκινήσει η διαδικασία του πειράματος μικροσυστοιχιών και αναμένεται να απαντηθεί στο τέλος. Το ερώτημα εξαρτάται από τον τομέα ενδιαφέροντος της μελέτης (Ramasamy, Mondry et al. 2008). Σημαντικός παράγοντας κατά την διατύπωση του ερωτήματος και κατά τη διεξαγωγή του πειράματος είναι το βιολογικό υλικό που θα χρησιμοποιηθεί στο πείραμα, δηλαδή η διαθεσιμότητά του, η σταθερότητά του (για παράδειγμα το mRNA είναι πιο ασταθές από το DNA, αλλά πιο εύκολο στη συλλογή του), το μέγεθος της ποσότητάς του, διότι θα επηρεάσει τα αποτελέσματα του πειράματος.

2^ο Στάδιο: Επιλογή κατάλληλου τύπου τσιπ μικροσυστοιχίας

Εφόσον έχει τεθεί ένα συγκεκριμένο βιολογικό ερώτημα, πρέπει να επιλεγεί ο τύπος του τσιπ των μικροσυστοιχιών που ταιριάζει να χρησιμοποιηθεί στο πείραμα. Υπάρχουν δύο βασικοί τύποι μικροσυστοιχιών σε ευρεία χρήση στα περισσότερα εργαστήρια, οι μικροσυστοιχίες σύνθεσης επί τόπου (in situ) και οι μικροσυστοιχίες εκτύπωσης (spotted microarrays) στις οποίες οι ανιχνευτές εκτυπώνονται, μετά την παρασκευή τους, στην επιφάνεια του τσιπ. Οι μικροσυστοιχίες εκτύπωσης χωρίζονται σε δύο κατηγορίες, σε μικροσυστοιχίες ολιγονουκλεοτιδίων και σε μικροσυστοιχίες cDNA (Seidel 2008).

3^ο Στάδιο: Παρασκευή του δείγματος

Το πρώτο βήμα σε ένα πείραμα μικροσυστοιχιών είναι η απομόνωση του βιολογικού υλικού, το οποίο θα χρησιμοποιηθεί ως στόχος. Κατά την ανάλυση γονιδιακού προφίλ έκφρασης χρειάζεται απομόνωση RNA από κύτταρα ή ιστούς. Η γονιδιακή έκφραση μετριέται με βάση το mRNA, άρα είτε ιχνηθετείται μόνο το mRNA από το διάλυμα του ολικού RNA, είτε απομονώνεται συγκεκριμένα το mRNA και ιχνηθετείται. Η ιχνηθέτηση του δείγματος μπορεί να γίνει είτε με αντίστροφη μεταγραφή, είτε με χρήση RNA πολυμεράσης, είτε με τη διαδικασία Eberwine και άλλες.

Η πιο απλή και κλασσική μέθοδος είναι η ιχνηθέτηση χρησιμοποιώντας ιχνηθετημένα, με φθορίζουσα ουσία, νουκλεοτίδια κατά τη διαδικασία της αντίστροφης μεταγραφής του mRNA

προς cDNA. Σε πειράματα μικροσυστοιχιών στα οποία χρησιμοποιούνται εκτυπωμένα τσιπ, και των δυο κατηγοριών, χρησιμοποιούνται πάντα δύο δείγματα, ένα δείγμα ελέγχου και ένα δείγμα αναφοράς. Σε διχρωματικές μικροσυστοιχίες τα δείγματα mRNA αναφοράς και ελέγχου (για παράδειγμα ιστός υγιούς ατόμου και ιστός ασθενούς ατόμου, σε περίπτωση μελέτης κάποιας ασθένειας) μεταγράφονται αντίστροφα με χρήση ιχνηθετημένων νουκλεοτιδίων, με διαφορετική φθορίζουσα ουσία, σε cDNA. Συνήθως χρησιμοποιούνται οι φθορίζουσες ουσίες κυανίνη3 (Cy3), η οποία προσδίδει πράσινο χρώμα φθορισμού και η κυανίνη5 (Cy5), η οποία προσδίδει κόκκινο. Η Cy3 συνήθως χρησιμοποιείται κατά την ιχνηθέτηση του δείγματος αναφοράς και η Cy5 κατά την ιχνηθέτηση του δείγματος ελέγχου (Freeman, Robertson et al. 2000). Τα νεοσυντιθέμενα μόρια cDNA μεταφέρονται στα τσιπ των μικροσυστοιχιών ώστε να υβριδοποιηθούν με τους ανιχνευτές.

4ο Στάδιο: Υβριδοποίηση

Το αμέσως επόμενο στάδιο του πειράματος μικροσυστοιχιών είναι το στάδιο της υβριδοποίησης, κατά το οποίο τα μονόκλιωνα μόρια των στόχων δημιουργούν δεσμούς υδρογόνου με τα συμπληρωματικά νουκλεοτίδια των ανιχνευτών. Για να πραγματοποιηθεί η διαδικασία τα μόρια των στόχων, πρέπει να βρίσκονται σε κατάλληλο ρυθμιστικό διάλυμα, να τοποθετηθούν επάνω στη διάταξη της μικροσυστοιχίας και να επωásουν σε συγκεκριμένη θερμοκρασία για ένα συγκεκριμένο χρονικό διάστημα. Οι συνθήκες πρέπει να είναι ευνοϊκές ως προς τις ειδικές υβριδοποιήσεις (specific hybridization) και δυσμενείς ως προς τις μη ειδικές (non specific hybridization), για παράδειγμα μεταξύ των στόχων και του υποστρώματος. Η θερμοκρασία υβριδοποίησης, το pH, η συγκέντρωση των αλάτων, η σύσταση της αλληλουχίας, το μήκος των στόχων, το μήκος των ανιχνευτών, η δευτεροταγής δομή και άλλα είναι παράγοντες οι οποίοι επηρεάζουν την αποτελεσματικότητα της υβριδοποίησης και τη συνοχή των διμερών (Freeman, Robertson, and Vrana 2000). Διαδοχικές πλύσεις των επιφανειών των μικροσυστοιχιών ακολουθούν τη διαδικασία της υβριδοποίησης, ώστε να αποβληθεί το διάλυμα της υβριδοποίησης και να απομακρυνθούν τυχόν μη ειδικές υβριδοποιήσεις, που ενδέχεται να έχουν δημιουργηθεί. Τοιουτοτρόπως το τσιπ είναι έτοιμο για το επόμενο στάδιο τη σάρωσή του από ειδικό μηχάνημα.

5ο Στάδιο: Σάρωση του τσιπ μικροσυστοιχίας

Σε αυτό το στάδιο, της σάρωσης, λαμβάνουμε τις μετρήσεις του εκάστοτε δείγματος μελέτης (αναφοράς και ελέγχου). Αρχικά οι υβριδοποιημένες κουκίδες διεγείρονται με ένα λέιζερ και στη συνέχεια με τη χρήση ενός σαρωτή (scanner), σαρώνεται η επιφάνεια του τσιπ, σε κατάλληλα

μήκη κύματος για την ανίχνευση των κόκκινων και των πράσινων χρωμάτων, και μετέπειτα μετατρέπεται η ένταση του σήματος σε ψηφιακή εικόνα. Η ποσότητα φθορισμού που εκπέμπεται αντιστοιχεί στην ποσότητα των δεσμευμένων νουκλεϊκών οξέων (Babu 2004). Στην ανασυνδυασμένη ψηφιακή εικόνα, όπως φαίνεται και στην Εικόνα 5, κάθε κουκίδα του τσιπ εμφανίζεται με ένα χρώμα, το οποίο εξαρτάται από την ποσότητα του δείγματος ελέγχου και του δείγματος αναφοράς. Χαρακτηριστικά:

- Κουκίδα κίτρινου είτε λευκού χρώματος, υποδεικνύει ότι οι ποσότητες του δείγματος αναφοράς και του δείγματος ελέγχου είναι υψηλές και ίσες.
- Κουκίδα κόκκινου χρώματος, υποδεικνύει ότι η ποσότητα του δείγματος που εκπέμπει στο κόκκινο μήκος κύματος (το δείγμα ελέγχου) είναι υψηλή και η ποσότητα του δείγματος που εκπέμπει στο πράσινο μήκος κύματος (το δείγμα αναφοράς) είναι χαμηλή.
- Κουκίδα πράσινου χρώματος, υποδεικνύει ότι η ποσότητα του δείγματος που εκπέμπει στο πράσινο μήκος κύματος (το δείγμα αναφοράς) είναι υψηλή και η ποσότητα του δείγματος που εκπέμπει στο κόκκινο μήκος κύματος (το δείγμα ελέγχου) είναι χαμηλή.
- Κουκίδα μαύρου χρώματος, υποδεικνύει ότι είτε οι ποσότητες και των δύο δειγμάτων (αναφοράς και ελέγχου) είναι χαμηλές και ίσες, είτε κανένα από τα δύο δείγματα δεν έχει υβριδοποιηθεί (Baird 2012).



Εικόνα 5 Συνδυασμένη εικόνα μικροσυστοιχίας.

Πείραμα μικροσυστοιχιών με χρήση δύο δειγμάτων ελέγχου και αναφοράς ιχνηθετημένα με κόκκινο και πράσινο χρώμα αντίστοιχα

6^ο Στάδιο: Κανονικοποίηση (normalization)

Η κανονικοποίηση των αποτελεσμάτων των μικροσυστοιχιών στοχεύει στην ελαχιστοποίηση των συστηματικών σφαλμάτων, κυρίως στα εκτιμώμενα επίπεδα έκφρασης των γονιδίων, ώστε να είναι ευκολότερα αντιληπτές οι βιολογικές διαφορές των δεδομένων. Ακόμη με τη χρήση κανονικοποιημένων αποτελεσμάτων δίνεται η δυνατότητα σύγκρισης των επιπέδων έκφρασης σε δεδομένα από διαφορετικά τσιπ μικροσυστοιχιών. Με την κανονικοποίηση μπορούν να ρυθμιστούν διαφορές κατά την ιχνηθέτηση, την ανίχνευση της αποδοτικότητας των φθορίζοντων ιχνηθετών και τις διαφορετικές ποσότητες των αρχικών βιολογικών δειγμάτων τα οποία χρησιμοποιήθηκαν στο το πείραμα των μικροσυστοιχιών. Τέτοιου είδους προβλήματα ενδέχεται να προκαλέσουν μετατόπιση της μέσης αναλογίας των φθορίζουσών ουσιών, με αποτέλεσμα την ανάγκη αναβαθμονόμησης των εντάσεων για τη σωστή ανάλυση των αποτελεσμάτων.

Οι κυριότερες τεχνικές κανονικοποίησης δεδομένων γονιδιακής έκφρασης είναι η κανονικοποίηση ολικής έντασης, η γραμμική οπισθοδρόμηση τοπικών βαρών (LOcally WEighted Scatterplot Smoothing regression, Lowess), η μέση λογαριθμική κεντροποίηση (Mean log centring), η γραμμική παλινδρόμηση (linear regression) και η στατιστική των λόγων (Chen's ratio statistics).

Η πρώτη κανονικοποίηση παίρνει ως δεδομένο πως η ποσότητα του αρχικού mRNA είναι η ίδια και για τα δύο δείγματα των διαφορετικών καταστάσεων. Δέχεται ότι κάποια γονίδια εάν είναι υπερεκφραζόμενα στο δείγμα, κατ' αντιστοιχία θα υπάρχουν υποεκφραζόμενα γονίδια σε ίσο περίπου αριθμό. Η συνολική ένταση φθορισμού για όλα τα στοιχεία στο τσιπ θεωρείται ίδια και για τις δύο φθορίζουσες ουσίες καθώς θεωρείται ότι υβριδοποιείται ίσος αριθμός RNA και από τα δύο δείγματα. Με βάση αυτή την προσέγγιση υπολογίζεται ένας παράγοντας κανονικοποίησης, ο οποίος τροποποιεί τις τιμές έκφρασης των γονιδίων και ρυθμίζει τη μέση αναλογία ώστε να είναι ίση με τη μονάδα (Quackenbush 2001). Το δεύτερο είδος κανονικοποίησης προτείνεται ως μέθοδος η οποία καταφέρνει να εξαλείψει αποτελέσματα από τις λογαριθμικές αναλογίες, οι οποίες οφείλονται στην ένταση σήματος (Hoheisel 2006).

7^ο Στάδιο: Στατιστική ανάλυση

Τα δεδομένα τα οποία προκύπτουν από το τσιπ των μικροσυστοιχιών είναι ένας πίνακας γονιδιακής έκφρασης, όπου οι γραμμές αναπαριστούν τα γονίδια και οι στήλες τα δείγματα. Οι τιμές κάθε θέσης του πίνακα αντιστοιχούν στις τιμές έκφρασης του συγκεκριμένου γονιδίου στις συγκεκριμένες πειραματικές συνθήκες. Οι τιμές αυτές προκύπτουν, συνήθως, από τον λογάριθμο

με βάση το δύο του λόγου της έντασης του ενός χρωματικού καναλιού (π.χ. κόκκινο) προς την ένταση του άλλου (π.χ. πράσινο).

Τα αποτελέσματα που λαμβάνονται έστω και από έναν σχετικά μικρό αριθμό δειγμάτων πειράματος μικροσυστοιχιών είναι αρκετά πολύπλοκα για να αναλυθούν, για τον λόγο αυτό χρησιμοποιούνται διάφορες στατιστικές μεθοδολογίες, ώστε να βρεθούν πιθανές σχέσεις μεταξύ των δεδομένων. Οι μέθοδοι αυτοί περιλαμβάνουν:

- Την αλλαγή διπλώματος (fold-change, FC) (Ramasamy, Mondry et al. 2008).
- Έλεγχο βασισμένο στον έλεγχο t (t -test), στον οποίο πολλές φορές χρησιμοποιούνται και διάφορες μέθοδοι διόρθωσης της τιμής p (p -value), όπως η διόρθωση Bonferroni (Dudoit, Yang et al. 2000), η διόρθωση Sidak (Šidák 1967), η διόρθωση Bonferroni Step-Down ή Holm (Holm 1979), η διόρθωση Holland (Holland and Copenhaver 1987) και η διόρθωση ποσοστού ψευδούς ανακάλυψης (False Discovery Rate, FDR) (Benjamini and Hochberg 1995).
- Την αναδειγματοληψία [όπως το Bootstrap (Efron 1982, Efron and Tibshirani 1993) και το permutation (Tsai, Chen et al. 2003)].
- Μπεϋζιανές μεθόδους (Bayesian) (Fox and Dimmic 2006) και
- Κανονικοποιημένους ελέγχους t είτε ελέγχους t με ποινές (Baldi and Long 2001), που είναι γνωστοί και ως εμπειρικές μέθοδοι Bayes, μέθοδοι οι οποίες μεταβάλλουν τον παρονομαστή του ελέγχου t με σκοπό να αυξήσουν τη διακύμανση.

1.5.3 Βάσεις δεδομένων μικροσυστοιχιών

Αν και η τεχνολογία των μικροσυστοιχιών ξεκίνησε ως μια αρκετά ακριβή μεθοδολογία, πλέον εξαιτίας νέων οικονομικότερων τσιπ μικροσυστοιχιών, πραγματοποιούνται πολλά πειράματα ανάλυσης γονιδιακής έκφρασης, το οποίο συνεπάγεται έναν μεγάλο όγκο δεδομένων. Τα δεδομένα αυτά χρειάζεται κάπου να αποθηκευθούν ώστε στη συνέχεια να αναλυθούν. Για τον λόγο αυτό δημιουργήθηκαν διάφορες βάσεις δεδομένων, οι οποίες περιέχουν χιλιάδες πειράματα μικροσυστοιχιών.

Οι βάσεις δεδομένων μικροσυστοιχιών είναι ένα χρήσιμο εργαλείο τόσο για την καταχώρηση πειραματικών αποτελεσμάτων όσο και για την ανάκτηση των πειραματικών δεδομένων. Κάθε πείραμα μικροσυστοιχίας που υπάρχει σε κάθε βάση παρέχει πληροφορίες σχετικά με το είδος των δεδομένων, τα άτομα που συμμετείχαν στο πείραμα, τα γονίδια που

μελετώνται, την πλατφόρμα μικροσυστοιχιών που χρησιμοποιήθηκε και διάφορες άλλες πληροφορίες σχετικά με τα δείγματα του πειράματος.

Υπήρχαν κατά καιρούς πολλές διαδικτυακές βάσεις δεδομένων μικροσυστοιχιών, ωστόσο τώρα πια ενεργές και ευρέως χρησιμοποιούμενες είναι οι:

- Gene Expression Omnibus (GEO) <https://www.ncbi.nlm.nih.gov/geo/> (Barrett and Edgar 2006)
- Archive of Functional Genomics Data (ArrayExpress) <https://www.ebi.ac.uk/arrayexpress/> (Athar, Fullgrabe et al. 2019)
- ONCOMINE <https://www.oncomine.org/resource/login.html> (Rhodes, Yu et al. 2004)
- Center for Information Biology Gene Expression database (CIBEX) <https://bio.tools/cibex> (Ikeo, Ishi-i et al. 2003)
- The Cancer Genome Atlas (TCGA) <https://cancergenome.nih.gov/>
- Cancer Program Datasets <http://portals.broadinstitute.org/cgi-bin/cancer/datasets.cgi>.

1.6 Μετα-ανάλυση

Η συνεχής συσσώρευση βιοϊατρικών δεδομένων έχει καταστήσει επιτακτική την ανάγκη συνδυασμού αυτών των δεδομένων σε ένα ενοποιημένο πλαίσιο ανάλυσης. Η μετα-ανάλυση έρχεται να δώσει τη λύση σε αυτό. Η μετα-ανάλυση εφαρμόστηκε για πρώτη φορά στην Ψυχολογία στα τέλη της δεκαετίας του '70 και είναι μια στατιστική μεθοδολογία στην οποία συνδυάζεται ένα σύνολο αρχικών μελετών, οι οποίες μελετούν το ίδιο ερώτημα, και διερευνάται η πιθανή ποικιλομορφία τους (Greenland 1998). Είναι ένα πολύ ισχυρό εργαλείο στην αξιολόγηση των επιδράσεων υποψήφιων γονιδίων σε σύνθετους φαινότυπους. Η μετα-ανάλυση στο χώρο της ιατρικής έρευνας εφαρμόστηκε πρώτη φορά στις τυχαιοποιημένες κλινικές δοκιμές (Sacks, Berrier et al. 1987) και πλέον θεωρείται πολύτιμο εργαλείο για τον συνδυασμό μελετών γενετικής συσχέτισης, διαγνωστικών δοκιμασιών, επιδημιολογικών μελετών και μελετών φαρμακογονιδιοματικής (Trikalinos, Salanti et al. 2008).

1.6.1 Πλεονεκτήματα μετα-ανάλυσης

Η μετα-ανάλυση παρέχει μια αντικειμενική σύνθεση πολλών σχετικών μελετών/πειραμάτων για ένα συγκεκριμένο πρόβλημα έτσι ώστε να οδηγηθεί η έρευνα σε ένα μοναδικό συνολικό συμπέρασμα.

Επιπρόσθετα επιτρέπει την αξιολόγηση της μεροληψίας των υπάρχουσών μελετών, υπολογίζοντας τον αριθμό των μελετών οι οποίες οδήγησαν στο συμπέρασμα ύπαρξης μη σημαντικής επίδρασης. Οι μελέτες αυτές είχαν πραγματοποιηθεί πριν ο ερευνητής καταλήξει σε ένα αντιφατικό γενικό συμπέρασμα (ότι υπάρχει κάποια σημαντική επίδραση) το οποίο θα μπορούσε να ανατρέψει την πρότερη πεποίθηση.

Επίσης σημαντικό είναι ότι επιτρέπει, πριν τη συγκέντρωση και τη σύνθεση των αποτελεσμάτων, την αντικειμενική ερμηνεία της αξίας των μεθόδων οι οποίες είχαν χρησιμοποιηθεί σε περιορισμένες μελέτες, ώστε να δοθούν κατάλληλα σταθμικά βάρη στις στατιστικές συναρτήσεις ελέγχου των συμπεριλαμβανομένων μελετών. Καθώς επίσης επιτρέπει να δοθεί διαφορετική βαρύτητα στις μελέτες με βάση το μέγεθος του δείγματός τους.

Παρέχει δυνατότητες συνδυασμού των ελέγχων έτσι ώστε να αξιοποιηθούν αποτελέσματα μελετών οι οποίες είτε ανέφεραν μόνο τη στατιστική συνάρτηση t (π.χ. τύπος του Winner, Winner's formula) είτε ανέφεραν μόνο τις p -values, ανεξάρτητα από τη στατιστική συνάρτηση που είχε χρησιμοποιηθεί (π.χ. τύπος του Fisher, Fisher's formula).

Τέλος, δίνει τη δυνατότητα υπολογισμού του μεγέθους επίδρασης, δηλαδή την ποσότητα κατά την οποία η πραγματική εναλλακτική υπόθεση (δηλαδή το αποτέλεσμα της επίδρασης) διαφέρει από την ελεγχθείσα μηδενική υπόθεση, ώστε να καθοριστεί πρακτικά η σημασία μιας διαπίστωσης η οποία είναι στατιστικώς σημαντική (Πανάρετος and Ξεκαλάκη 2000).

1.6.2 Μεθοδολογία μετα-ανάλυσης

Αρχικό και σημαντικό βήμα κατά τη μετα-ανάλυση είναι η εκτενής αναζήτηση στη βιβλιογραφία, για το προς μελέτη αντικείμενο, ώστε να βρεθούν όλες οι διαθέσιμες μελέτες οι οποίες αναφέρονται στο θέμα που εξετάζεται (Normand 1999). Μόλις συγκεντρωθεί ένας ικανοποιητικός αριθμός μελετών, συνήθως περισσότερες από τρεις, καταγράφονται τα δεδομένα τα οποία θα χρησιμοποιηθούν στη μετα-ανάλυση. Η αξιοπιστία της μετα-ανάλυσης επηρεάζεται, όπως είναι αναμενόμενο, από την αξιοπιστία των συμπεριλαμβανόμενων στοιχείων, με αποτέλεσμα να απαιτείται ιδιαίτερη προσοχή κατά την ανάκτηση και επεξεργασία τους.

Βασικός παράγοντας κατά την πραγματοποίηση της μετα-ανάλυσης είναι το μέγεθος επίδρασης. Μέγεθος επίδρασης μπορεί να είναι είτε η διαφορά των μέσων τιμών, είτε ο συντελεστής συσχέτισης, είτε ο σχετικός λόγος πιθανοτήτων (Odds Ratio, OR), ο οποίος αποδίδει την πιθανότητα ένα γεγονός να συμβεί προς την πιθανότητα να μη συμβεί. Τα μοντέλα τα οποία

μπορούν να χρησιμοποιηθούν για το συνδυασμό των μεμονωμένων εκτιμήσεων του μεγέθους επίδρασης είναι το μοντέλο σταθερών επιδράσεων (Fixed effects model) και το μοντέλο τυχαίων επιδράσεων (Random effects model).

1.6.3 Έλεγχος ετερογένειας

Στο μοντέλο των τυχαίων επιδράσεων η ετερογένεια μεταξύ των μελετών λαμβάνεται υπόψη σε αντίθεση με το μοντέλο των σταθερών επιδράσεων. Ελλοχεύει ωστόσο ο κίνδυνος της ύπαρξης ετερογένειας στη μελέτη, καθώς τα δείγματα ενδέχεται να προέρχονται από διαφορετικούς πληθυσμούς. Η ανομοιογένεια μεταξύ των δειγμάτων μπορεί να επηρεάσει τα αποτελέσματα της μετα-ανάλυσης και να προκύψουν λάθος συμπεράσματα.

1.6.4 Πολυμεταβλητή Μετα-ανάλυση (Multivariate meta-analysis)

Γενικά όταν μιλάμε για μετα-ανάλυση συνήθως αναφερόμαστε στη μονομεταβλητή μετα-ανάλυση (univariate meta-analysis), όπου ερευνούμε δύο καταστάσεις, υπάρχουν όμως περιπτώσεις στις οποίες έχουμε πολλαπλές καταστάσεις (π.χ. πολλαπλές εκβάσεις ή πολλαπλές θεραπείες) και που επίσης μας ενδιαφέρουν οι ίδιες συγκρίσεις όπως και στη μονομεταβλητή μετα-ανάλυση. Για παράδειγμα, σε μια μετα-ανάλυση μελετών μικροσυστοιχιών, υπάρχουν περιπτώσεις στις οποίες χορηγούνται διαφορετικά φάρμακα, είτε καταγράφονται διαφορετικές δόσεις φαρμάκου. Σε αυτές τις περιπτώσεις η σύγκριση που ενδιαφέρει τον ερευνητή είναι η επίδραση (ποια φαρμακευτική αγωγή αποδίδει καλύτερα και ούτω καθεξής). Σε ένα άλλο παράδειγμα, μερικές (συναφείς) ασθένειες, διερευνώνται έναντι μιας κοινής υγιούς ομάδας, προκειμένου να βρεθούν διαφορικά εκφραζόμενα γονίδια κοινά σε όλες τις συνθήκες ή γονίδια που διαφέρουν μεταξύ των συνθηκών (π.χ. σκλήρυνση κατά πλάκας, ρευματοειδής αρθρίτιδα και σακχαρώδης διαβήτης τύπου I).

Τα προαναφερθέντα προβλήματα αντιμετωπίζονται στο πλαίσιο της πολυμεταβλητής μετα-ανάλυσης στην οποία γίνεται χρήση ενός κοινού μοντέλου για τις διαφορετικές ποσότητες ενδιαφέροντος (van Houwelingen, Arends et al. 2002, Jackson, Riley et al. 2011, Mavridis and Salanti 2012). Στην πολυμεταβλητή μετα-ανάλυση υπάρχουν επίσης τα μοντέλα σταθερών και τυχαίων επιδράσεων, τα οποία προσαρμόζονται στην πολυμεταβλητή φύση του προβλήματος.

1.6.4.1 Πλεονεκτήματα της πολυμεταβλητής μετα-ανάλυσης

Η πολυμεταβλητή μετα-ανάλυση επιτρέπει στον ερευνητή να διεξάγει συνολικούς ελέγχους (global tests) αποφεύγοντας πολλαπλές συγκρίσεις και, κατά συνέπεια, την συνεχή αύξηση του ποσοστού του σφάλματος τύπου I.

Επιπλέον, με τη χρήση της πολυμεταβλητής μετα-ανάλυσης μπορεί να υπολογιστεί ο πίνακας συνδιακύμανσης για όλες τις παραμέτρους ενδιαφέροντος. Αυτό μπορεί να είναι ενδιαφέρον από μόνο του, για παράδειγμα όταν πρόκειται να υπολογιστεί ο συσχετισμός μεταξύ ενός υποκατάστατου και ενός πραγματικού αποτελέσματος, αλλά μπορεί επίσης να είναι σημαντικό όταν πραγματοποιείται σύγκριση μεταξύ των δύο εκτιμητών και είναι εφικτό να δημιουργηθεί ένα διάστημα εμπιστοσύνης για τη διαφορά ή κάποια άλλη συνάρτησή τους (Altman and Bland 2003).

Τέλος, ένα μεγάλο πλεονέκτημα της πολυμεταβλητής μετα-ανάλυσης είναι ότι μπορεί να περιλαμβάνει, υπό την προϋπόθεση τυχαίας ελλείψεως (missing at random, MAR), μελέτες οι οποίες να αναφέρουν μόνο μία από τις παραμέτρους ενδιαφέροντος, με αποτέλεσμα τον δανεισμό ισχύος από εξωτερικές μελέτες (Higgins and Whitehead 1996).

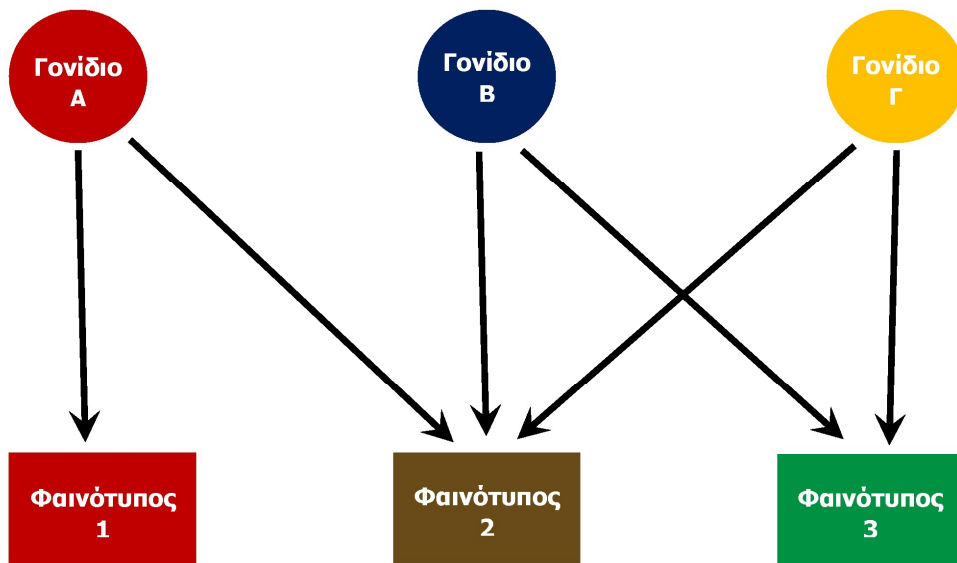
Σε κάθε περίπτωση, η εφαρμογή μιας πολυμεταβλητής μεθοδολογίας και η συγκέντρωση δεδομένων από πολλαπλές πηγές μπορεί να αποδειχθεί σημαντική για την αύξηση της ισχύος της ανάλυσης και στο να επιτρέπει, με μια ενοποιημένη ανάλυση, τον εντοπισμό γονιδίων που διαδραματίζουν κεντρικό ρόλο σε αυτές τις συνθήκες.

1.6.5 Πλειοτροπία

Η πλειοτροπία (προέρχεται από τις Ελληνικές λέξεις πλείον και τρόπος) είναι το γενετικό φαινόμενο στο οποίο ένα γονίδιο επηρεάζει, με την έκφρασή του, τον φαινότυπο πολλών άσχετων μεταξύ τους χαρακτήρων. Έτσι γονίδια τα οποία επηρεάζουν πολλαπλούς, μη σχετιζόμενους φαινομενικά, φαινότυπους θεωρούνται πλειοτροπικά γονίδια (Εικόνα 6) (Lobo 2008, Klug, Cummings et al. 2015). Έχουν βρεθεί αρκετά πλειοτροπικά γονίδια τόσο στον άνθρωπο όσο και σε άλλους οργανισμούς. Γνωστές ασθένειες του ανθρώπου οι οποίες οφείλονται σε πλειοτροπικά γονίδια είναι η Φαινυλκετονουρία (Paul 2000), το σύνδρομο Marfan και η παραλλάσσουσα πορφυρία ή διαλείπουσα πορφυρία (Klug, Cummings et al. 2015).

Εκτός από διάφορους φαινότυπους και ασθένειες, μέσω της πλειοτροπίας κάποιων γονιδίων μπορούμε να αποκτήσουμε πληροφορίες σχετικά με την εξέλιξη διαφορετικών γονιδίων

και γονιδιακών οικογενειών, καθώς κάποια γονίδια επιλέγονται μαζί, αποδίδοντάς σε αυτά νέες ιδιότητες πέραν αυτών που θεωρείτο ότι είχαν αρχικά (Hodgkin 1998). Πλειοτροπία θα μπορούσε να θεωρηθεί το φαινόμενο που παρατηρείται σε διαφορετικούς τύπους κυττάρων πολλές πρωτεΐνες να έχουν πολλαπλούς ρόλους. Γίνεται λοιπόν αντιληπτό ότι μια γενετική αλλαγή είτε στη λειτουργία είτε στην έκφραση κάποιου γονιδίου ενδέχεται να έχει διαφορετικές επιδράσεις σε πολλά είδη ιστών (Lobo 2008).



Εικόνα 6 Ένας απλός χάρτης γονότυπου-φαινοτύπου που δείχνει το φαινόμενο της πλειοτροπίας. Τα Γονίδια 1, 2, 3 είναι διαφορετικά γονίδια που συμβάλλουν στους φαινοτύπους 1, 2 και 3.

1.6.6 Μετα-ανάλυση σε δεδομένα μικροσυστοιχιών

Η μετα-ανάλυση έχει χρησιμοποιηθεί και για να συνοψίζονται αποτελέσματα πειραμάτων υψηλής απόδοσης όπως τα προερχόμενα από μελέτες συσχέτισης ολόκληρου του γονιδιώματος (Genome Wide Association Study, GWAS) (Zeggini and Ioannidis 2009) και από μελέτες γονιδιακής έκφρασης μικροσυστοιχιών ή αλληλούχισης νέας γενιάς (Next Generation Sequencing, NGS) (Hong and Breitling 2008). Στα πειράματα γονιδιακής έκφρασης, τα επίπεδα mRNA χιλιάδων γονιδίων μετριοούνται ταυτόχρονα από δείγματα ιστών. Ιδανικά, η εύρεση διαφορεικά εκφραζόμενων γονιδίων (ΔΕΓ), δηλαδή γονιδίων τα οποία υπερεκφράζονται ή υπόεκφράζονται σε διαφορετικές συνθήκες, είναι μια απλή μέθοδος που τελικά ανάγεται, όπως προαναφέρθηκε, σε ένα στατιστικό έλεγχο για την ισότητα των μέσων τιμών (π.χ. *t*-test).

Ωστόσο τα σύνολα δεδομένων (datasets) των μικροσυστοιχιών χαρακτηρίζονται από βασικά διακριτικά γνωρίσματα όπως μικρό αριθμό δειγμάτων, μεγάλο αριθμό τιμών και υπερβολική ποσότητα θορύβου· γι' αυτό έχουν προταθεί πολλές προηγμένες στατιστικές μέθοδοι, για τον αποτελεσματικό χειρισμό αυτών των ζητημάτων και για τον ακριβή υπολογισμό των p -values και των διαστημάτων εμπιστοσύνης.

Σχετικά με τη μετα-ανάλυση μελετών μικροσυστοιχιών και RNAseq υπάρχουν τρεις ευρείες κατηγορίες στατιστικών μεθόδων. Αυτές που χρησιμοποιούν το μέγεθος επίδρασης, τις p -values ή τους βαθμούς κατάταξης (rank product).

1.6.7 Μέθοδοι διόρθωσης πολλαπλών συγκρίσεων

Όπως έχει προαναφερθεί σε ένα πείραμα μικροσυστοιχιών υπολογίζεται ταυτόχρονα η έκφραση χιλιάδων γονιδίων υπό συγκεκριμένες συνθήκες. Από νωρίς στη βιβλιογραφία έγινε κατανοητό ότι απαιτείται μια στατιστική μεθοδολογία ώστε να εντοπιστούν τα γονίδια τα οποία εκφράζονται διαφορετικά και είναι στατιστικώς σημαντικά σε σύγκριση με το σύνολο των γονιδίων τα οποία εξετάζονται στο πείραμα της μικροσυστοιχίας. Οι περισσότερες μέθοδοι δίνουν για κάθε γονίδιο μια συγκεκριμένη p -value με βάση την οποία γίνεται κατανοητό το εάν υπάρχουν στατιστικώς σημαντικές διαφορές στην έκφρασή του.

Όταν αναλύονται τα αποτελέσματα μιας μετα-ανάλυσης από δεδομένα μικροσυστοιχιών, τα οποία είναι της τάξεως των 20.000 και περισσότερων γονιδίων γίνεται αντιληπτό ότι έστω και αν θέσουμε ένα αυστηρό επίπεδο σημαντικότητας της τάξης του 1% υπάρχει μεγάλη πιθανότητα να βρεθούν γονίδια στατιστικώς σημαντικά απλά και μόνο λόγω τύχης. Για τον λόγο αυτό χρειάζεται να γίνεται ένας έλεγχος στα διαφορετικά εκφραζόμενα γονίδια, ώστε να υπάρχει βεβαιότητα ότι τα στατιστικώς σημαντικά ευρήματα είναι όντως σημαντικά (διατηρείται δηλαδή το ονομαστικό επίπεδο σημαντικότητας) και ακόμη, ότι δεν απορρίφθηκαν γονίδια τα οποία δεν έπρεπε.

Οι μέθοδοι διόρθωσης πολλαπλών συγκρίσεων αναλαμβάνουν να διεξάγουν αυτούς τους ελέγχους. Οι μέθοδοι διόρθωσης πολλαπλών συγκρίσεων λαμβάνουν ως είσοδο μια λίστα με p -values, οι οποίες έχουν προκύψει από τους ελέγχους, και επιστρέφουν νέες διορθωμένες p -values για όλα τα γονίδια της λίστας εισόδου. Χωρίζονται σε δύο κύριες κατηγορίες, σε αυτές οι οποίες ελέγχουν το ποσοστό σφάλματος FWER (family-wise error rate) και σε αυτές οι οποίες ελέγχουν το ποσοστό σφάλματος FDR.

1.6.8 Μέθοδοι αναδειγματοληψίας (Resampling methods)

Τα πειράματα μικροσυστοιχιών, όπως έχει ήδη αναφερθεί, διαθέτουν συνήθως μικρό μέγεθος δείγματος και μη κανονική κατανομή στις τιμές έκφρασης των γονιδίων. Για το λόγο αυτό ο έλεγχος βασισμένος στο *t*-test ακολουθείται αρκετές φορές από μια μέθοδο αναδειγματοληψίας όπως το Bootstrap (Efron 1982, Efron and Tibshirani 1993) και τον έλεγχο μετάθεσης (permutation test) (Kontou, Pavlopoulou et al. 2018).

1.7 Προεκλαμψία

Η προεκλαμψία αποτελεί μια σοβαρή επιπλοκή της κύησης και μία από τις κυριότερες αιτίες νοσηρότητας και θνησιμότητας της μητέρας και του εμβρύου, καθώς επηρεάζει το 2-8% όλων των κυήσεων παγκοσμίως (Stegers, von Dadelszen et al. 2010). Οι υπερτασικές διαταραχές της κύησης (όπου περιλαμβάνεται και η προεκλαμψία) είναι οι συχνότερες αιτίες θανάτου κύησης (Henderson, Whitlock et al. 2014) και είχαν ως αποτέλεσμα 46.900 θανάτους το 2015 (Wang, Naghavi et al. 2016). Υπάρχουν δύο κύριοι υπότυποι της προεκλαμψίας, η ήπια και η σοβαρή ή βαρεία.

Τα κλινικά συμπτώματα της ήπιας προεκλαμψίας συνήθως είναι η παρουσία νεοεμφανιζόμενης υπέρτασης (διαστολική αρτηριακή πίεση ≥ 90 mmHg και συστολική αρτηριακή πίεση ≥ 140 mmHg), σε μετρήσεις που απέχουν τουλάχιστον 4 ως 6 ώρες η μια από την άλλη, σε έγκυο η οποία είχε, μέχρι πρότινος, φυσιολογική πίεση και η πρωτεϊνουρία (ποσότητα λευκώματος μεγαλύτερη από 0,3 g σε συλλογή 24ώρου) μετά την 20^η εβδομάδα της κύησης. Σε μερικές περιπτώσεις προεκλαμψία εμφανίζεται και χωρίς την παρουσία λευκωματουρίας. Στις περιπτώσεις αυτές η υπέρταση συνοδεύεται από επίμονη κεφαλαλγία, διαταραχές της όρασης, επιγαστραλγία ή άλγος του δεξιού υποχονδρίου με ναυτία και εμέτους, θρομβοπενία και παθολογικές τιμές των ηπατικών ενζύμων (Brown, Hague et al. 2000, Gifford, August et al. 2000, Sibai 2003, Longo and Harrison 2012).

Η σοβαρή προεκλαμψία, αντίθετα, χαρακτηρίζεται από ένα από τα ακόλουθα: συνεχής συστολική αρτηριακή πίεση ≥ 160 mmHg ή διαστολική αρτηριακή πίεση ≥ 110 mmHg, πρωτεϊνουρία νεφρωσικής έκτασης, αιφνίδια ολιγουρία, διαταραχές του κεντρικού νευρικού συστήματος (κεφαλαλγία, παροδική τύφλωση, διαταραχές όρασης), πνευμονικό οίδημα και γενικευμένο οίδημα (άνω άκρα και πρόσωπο), κυάνωση, πόνος επιγαστρικού ή δεξιού άνω

τεταρτημορίου, δυσλειτουργία του ήπατος, θρομβοπενία, εκλαμψία και περιορισμός εμβρυικής ανάπτυξης (Brown, Hague et al. 2000, Sibai 2003, Turner 2010).

Η προεκλαμψία μπορεί να ξεκινήσει από την ύπαρξη του πλακούντα στη μήτρα και τελειώνει με την αφαίρεση του πλακουντιακού ιστού κατά τον τοκετό (Louwen, Muschol-Steinmetz et al. 2012). Γενικά, η προεκλαμψία θεωρείται μια σύνθετη αλληλεπίδραση μεταξύ μίας σειράς μητρικών γονιδίων, περιβαλλοντικών παραγόντων (ρύπανση, παχυσαρκία) (Bodnar, Caton et al. 2007) και/ή απορυθμισμένης ανοσολογικής απόκρισης της μητέρας στα πατρικά συστατικά του γονότυπου του εμβρύου χωρίς σαφή αιτιολογία (Walker 2000).

Αρκετοί υποψήφιοι βιοδείκτες οι οποίοι έχουν συνδεθεί με την προεκλαμψία εμπλέκονται στην πλακουντιοποίηση, τη ρύθμιση της αρτηριακής πίεσης, τη φλεγμονή, τον αγγειακό σχηματισμό και τη λειτουργία των ενδοθηλιακών κυττάρων. Αυτοί οι βιοδείκτες περιλαμβάνουν τον EGF (παράγοντα επιδερμικής ανάπτυξης), PLGF, sFLT-1, HO-2, HIF1-a, Apo E, INHA, LEP, TLR4, προ-φλεγμονώδεις κυτοκίνες, απόπτωση, γονίδια επιβίωσης των κυττάρων και γονίδια διαφοροποίησης (Chappell and Morgan 2006), αλλά χωρίς απόλυτη προγνωστική αξία (Grill, Rusterholz et al. 2009, El-Sayed 2017).

1.8 Ιδιοπαθής Φλεγμονώδης Νόσος του Εντέρου

Η Ιδιοπαθής Φλεγμονώδης Νόσος του Εντέρου (ΙΦΝΕ) είναι μια χρόνια πολυπαραγοντική ανοσολογική διαταραχή η οποία χαρακτηρίζεται από υποτροπιάζουσα φλεγμονή του εντέρου. Η νόσος του Crohn και η ελκώδης κολίτιδα είναι οι δύο κύριες μορφές της ΙΦΝΕ (Podolsky 1991).

Κλινικά, η νόσος του Crohn και η ελκώδης κολίτιδα μοιράζονται παρόμοια συμπτώματα, αλλά διαφέρουν στη θέση και στην έκταση της φλεγμονής, καθώς επίσης στις επιπλοκές και την επίπτωση (Εικόνα 7, Πίνακας 2). Η ακριβής αιτιολογία της ΙΦΝΕ εξακολουθεί να είναι ασαφής, αλλά πιστεύεται ότι τόσο οι γενετικοί (de Lange, Moutsianas et al. 2017) όσο και περιβαλλοντικοί παράγοντες (Piovani, Danese et al. 2019) συμβάλλουν στην ανάπτυξή της.

1.8.1 Νόσος του Crohn

Η νόσος του Crohn είναι μια προοδευτική νόσος η οποία χαρακτηρίζεται από εστιακή διατοιχωματική φλεγμονή. Μπορεί να εμφανιστεί οπουδήποτε στον γαστρεντερικό σωλήνα, από το στόμα μέχρι τον πρωκτό (Εικόνα 7). Η φλεγμονή που εμφανίζεται κατά την νόσο του Crohn συνήθως είναι βαθιά και ασυνεχής, με ξεχωριστές περιοχές φυσιολογικού ιστού μεταξύ των

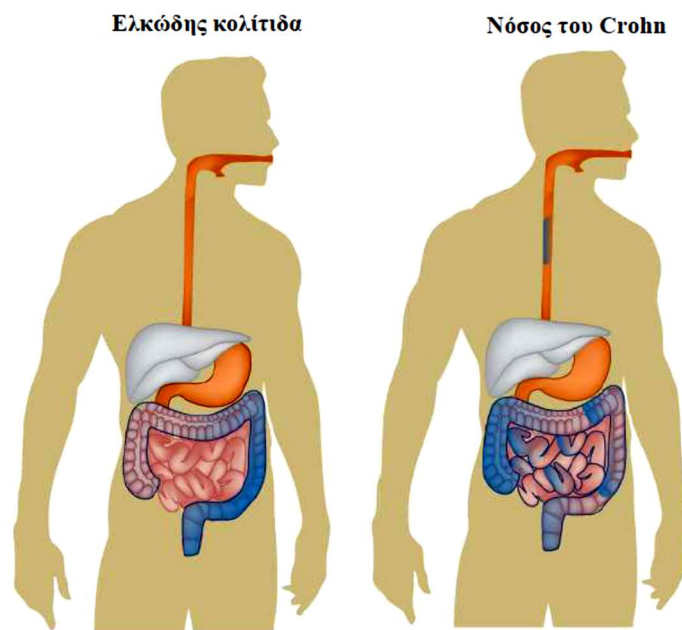
προσβεβλημένων περιοχών, την ονομαζόμενη «αποσπασματική βλάβη/κάκωση». Εξαιτίας της διατοιχωματικής φύσης της φλεγμονής της νόσου του Crohn, συχνά παρατηρείται η ανάπτυξη συριγγίου ή στενώσεων.

Τα κλινικά χαρακτηριστικά της νόσου συσχετίζονται με τη θέση, τη συμπεριφορά της και τη συνύπαρξη περιπρωκτικής ή εξεντερικής νόσου. Για παράδειγμα η οισοφαγική νόσος του Crohn παρουσιάζει οδυνοφαγία ή δυσφαγία, ή ενδεχομένως να μιμείται τη γαστροοισοφαγική παλινδρόμηση. Η γαστρική ή δωδεκαδακτυλική νόσος του Crohn ενδέχεται να εμφανίσει δυσπεψία, κοιλιακό άλγος και μεταγευματική πληρότητα ή μη θεραπευτικό έλκος (Ananthakrishnan, Xavier et al. 2017).

1.8.2 Ελκώδης Κολίτιδα

Βασικό χαρακτηριστικό της ελκώδους κολίτιδας είναι η επιφανειακή και διάχυτη φλεγμονή, η οποία συνήθως ξεκινά από το ορθό και εκτείνεται εγγύτατα, με συνεχή τμήματα του παχέος εντέρου (Εικόνα 7). Η φλεγμονή στην ελκώδη κολίτιδα είναι διάχυτη και επιφανειακή.

Τα κλινικά χαρακτηριστικά της νόσου εξαρτώνται γενικά από την έκταση της φλεγμονής (π.χ. ελκώδης ορθίτιδα, αριστερόπλευρη κολίτιδα, εκτεταμένη κολίτιδα κ.α.). Τα χαρακτηριστικότερα γνωρίσματα της ελκώδους κολίτιδας είναι η διάρροια και η αιμορραγία του ορθού. Η νόσος διαθέτει υποκατηγορίες και χαρακτηρίζεται ως ήπια, σοβαρή ή «κεραυνοβόλος» (fulminant) ελκώδης κολίτιδα με βάση τις κλινικές παραμέτρους της συχνότητας των κοπράνων, την παρουσία αίματος στα κόπρανα, τη θερμοκρασία του σώματος, τον παλμό και τις εργαστηριακές παραμέτρους, δηλαδή την ταχύτητα καθίζησης των ερυθροκυττάρων και τον αιματοκρίτη (Ananthakrishnan, Xavier et al. 2017).



Εικόνα 7 Περιοχές φλεγμονής της νόσου του Crohn και της ελκώδους κολίτιδας αντίστοιχα.
<https://www.augustaendoscopy.com/gac-condition/inflammatory-bowel-disease/>

Πίνακας 2 Ευρήματα που διαφοροποιούν τη νόσο του Crohn από την ελκώδη κολίτιδα

Χαρακτηριστικό	Νόσος του Crohn	Ελκώδης κολίτιδα
<u>Κλινικά χαρακτηριστικά</u>		
Γαστρεντερικά συμπτώματα	Διάρροια, κοιλιακό άλγος, επιβράδυνση ανάπτυξης, αναιμία	Αιμορραγία από τον ορθό, διάρροια, τεινεσμός
Εξεντερικές εκδηλώσεις	Σπάνια	Συχνά
Περιεδρική νόσος	Εμφανίζεται στο 1/5 των περιπτώσεων	Δεν εμφανίζεται
Στενώσεις	Ενδέχεται να εμφανιστούν	Δεν εμφανίζονται
Διατοιχωματικές προσβολές	Ενδέχεται να εμφανιστούν	Δεν εμφανίζονται
Περιοχή επιπλοκής	Από το στόμα έως και τον πρωκτό	Μόνο στο παχύ έντερο
<u>Ενδοσκοπικά ευρήματα</u>		
Κατανομή της νόσου	Τμηματική με ασυνεχείς περιοχές και επηρεάζοντας τον βλεννογόνο	Συνεχής συνήθως
Πρωκτική συμμετοχή	Συνήθως δεν εμφανίζεται	Σχεδόν συνολική και ομοιόμορφη συμμετοχή
Ειλεϊκή συμμετοχή	Συχνά. Εμφανίζονται άφθες, εξελκώσεις, στενώσεις	Σπάνια. Εμφανίζεται ειλεΐτιδα εκ παλινδρομήσεως (Backwash ileitis), σε ένα υποσύνολο των ασθενών
Τυπικά Χαρακτηριστικά	Ερπυστικά έλκη, βαθιά έλκη, πέτρες	Ερυθριματώδης βλεννογόνος με διάβρωση και ευθρυπτότητα, κοκκιωματώδης βλεννογόνος
<u>Ιστοπαθολογικά ευρήματα</u>		
Βάθος της φλεγμονής	Ενδέχεται να είναι διατοιχωματική	Περιορίζεται στον βλεννογόνο και τον υποβλεννογόνο
Κοκκιώματα	Εμφανίζονται στο 1/5 των ασθενών	Συνήθως δεν εμφανίζονται

(Ananthakrishnan, Xavier et al. 2017).

1.9 Σκοπός της Παρούσας Διατριβής

Η παρούσα διδακτορική διατριβή ασχολήθηκε με την εμπειριστατωμένη μελέτη και εφαρμογή των μεθόδων και στατιστικών τεχνικών της μετα-ανάλυσης, σε δεδομένα υψηλής απόδοσης, και πιο συγκεκριμένα σε δεδομένα γονιδιακής έκφρασης από πειράματα μικροσυστοιχιών.

Αρχικός στόχος της παρούσας διατριβής ήταν η διερεύνηση της συσχέτισης γονιδίων και βιοχημικών μονοπατιών με πολυπαραγοντικές ασθένειες. Γενικά θεωρείται ότι οι σύνθετες-πολυπαραγοντικές ασθένειες επηρεάζονται από πλήθος γονιδίων και περιβαλλοντικών παραγόντων οι οποίοι πιθανολογείται ότι έχουν προσθετική ή συνεργατική επίδραση στην εμφάνιση και εξέλιξή τους (Yagil and Yagil 2006). Η υπέρταση, η προεκλαμψία, η ελκώδης κολίτιδα, η νόσος του Crohn, η σκλήρυνση κατά πλάκας και άλλες είναι χαρακτηριστικά παραδείγματα σύνθετων ασθενειών με σημαντική επιβάρυνση στον παγκόσμιο πληθυσμό. Για πολλές από τις προαναφερθείσες ασθένειες υπάρχει μεγάλος αριθμός μεμονωμένων μελετών μικροσυστοιχιών οι οποίες όμως πολλές φορές αποκλίνουν στα αποτελέσματά τους και ακόμη μεγάλο μέρος των αποτελεσμάτων δεν μπορεί να αναπαραχθεί, είτε εν μέρει είτε πλήρως (Ioannidis, Allison et al. 2009). Αυτό οφείλεται κυρίως στην έλλειψη προτυποποίησης και στον ελλιπή σχολιασμό, είτε στις προδιαγραφές επεξεργασίας και ανάλυσης των δεδομένων. Έτσι δεν γίνεται εφικτό να διεξαχθεί ένα τελικό συμπέρασμα στο οποίο να σχετίζονται κάποια γονίδια είτε κάποια βιοχημικά μονοπάτια με την εμφάνιση ή την εξέλιξη της ασθένειας. Για τον λόγο αυτό χρησιμοποιείται συχνά η μεθοδολογία της μετα-ανάλυσης κατά την οποία συνδυάζονται τα αποτελέσματα πολλαπλών, ανεξάρτητων μελετών οι οποίες ερευνούν την ίδια ασθένεια, σε μελέτες ασθενών υγιών, με στόχο να διεξαχθεί ένα τελικό αποτέλεσμα. Στην παρούσα διατριβή πραγματοποιήθηκε διερεύνηση διαφόρων διορθώσεων και τροποποιήσεων για την εύρεση της βέλτιστης μεθοδολογίας μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών. Ακολουθώντας αυτό το μονοπάτι πραγματοποιήθηκαν διάφοροι έλεγχοι εκτελώντας μετα-αναλύσεις με χρήση διαφορετικών τρόπων βελτίωσης των αποτελεσμάτων (διόρθωση Hedges, μέθοδος αναδειγματοληψίας Bootstrap) σε μικρά αλλά και σε μεγάλα τυχαία σύνολα δεδομένων προερχόμενα από μικροσυστοιχίες.

Ακολουθώντας πραγματοποιήθηκε μετα-ανάλυση γονιδιακών δεδομένων από πειράματα μικροσυστοιχιών σχετικά με την ασθένεια της προεκλαμψίας, με σκοπό τη διερεύνηση και αποκρυπτογράφηση των γονιδίων και των βιοχημικών μονοπατιών, τα οποία παίζουν ρόλο στην εμφάνιση και εξέλιξη της νόσου.

Επόμενος στόχος της διατριβής ήταν η ανάπτυξη και η επέκταση των ήδη υπαρχουσών μεθόδων μετα-ανάλυσης γονιδιακών δεδομένων, ώστε να καταστεί δυνατή μια πολυμεταβλητή μετα-ανάλυση η οποία θα μπορεί να περιλαμβάνει δεδομένα πολλαπλών εκβάσεων ή πολλαπλών παραγόντων κινδύνου. Έτσι πραγματοποιήθηκε η δημιουργία ενός εύκολου και αξιόπιστου λογισμικού, με τη χρήση του στατιστικού πακέτου Stata, το οποίο επιτρέπει για πρώτη φορά στην βιβλιογραφία, τη διεξαγωγή πολυμεταβλητής μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών. Η υλοποίηση της μεθόδου είναι διαθέσιμη στο ευρύ κοινό στην ιστοσελίδα www.compgen.org/tools/multivariate-microarrays.

Στενά συνυφασμένος ήταν και ο τελικός στόχος της διατριβής σχετικά με τη διερεύνηση της συσχέτισης γονιδίων και βιοχημικών μονοπατιών μεταξύ πολυπαραγοντικών ασθενειών με παρόμοιες αιτιολογίες είτε παρόμοιες εκδηλώσεις. Για τον λόγο αυτό έγινε χρήση του προαναφερθέντος λογισμικού και πραγματοποιήθηκε πολυμεταβλητή μετα-ανάλυση γονιδιακών δεδομένων από πειράματα μικροσυστοιχιών τα οποία μελετούσαν άτομα τα οποία έπασχαν είτε από την νόσο του Crohn είτε από ελκώδη κολίτιδα και διερευνήθηκαν τα κοινά (και για τις δύο ασθένειες), αλλά και για κάθε ασθένεια ξεχωριστά, γονίδια και βιοχημικά μονοπάτια τα οποία βρέθηκαν στατιστικώς σημαντικά.

Κεφάλαιο 2^ο: Υλικά και Μέθοδοι

2.1 Ανάλυση δεδομένων μικροσυστοιχιών

2.1.1 Συλλογή Δεδομένων

Για τη συλλογή των κατάλληλων για την προκείμενη έρευνα δεδομένων πραγματοποιήθηκε αναζήτηση βιβλιογραφίας σε διεθνείς βάσεις δεδομένων.

Είναι αναγκαίο ωστόσο πριν από την διεξαγωγή της αναζήτησης και φυσικά της μετα-ανάλυσης να τεθούν τα κριτήρια επιλεξιμότητας των μελετών που θα συμπεριληφθούν στη μελέτη. Τα κριτήρια αυτά μπορεί να είναι βιολογικά, για παράδειγμα η ασθένεια, ο τύπος του ιστού που μελετάται, είτε τεχνικά, για παράδειγμα ο τύπος της μικροσυστοιχίας, η πλατφόρμα κ.α. Με βάση αυτά τα κριτήρια, φτιάχνονται οι όροι αναζήτησης και πραγματοποιείται η αναζήτηση κυρίως σε δημόσιες βάσεις δεδομένων μικροσυστοιχιών, αλλά και γενικά στη βιβλιογραφία ώστε να ανακτηθούν οι σχετικές μελέτες.

Συγκεκριμένα στην πρώτη μας μελέτη, για την ασθένεια της προεκλαμψίας διεξήγαμε μια εκτενή έρευνα για την εύρεση δεδομένων γονιδιακής έκφρασης σχετικά με την ασθένεια. Τα σύνολα δεδομένων ανακτήθηκαν από τη GEO χρησιμοποιώντας τους όρους αναζήτησης “preeclampsia” OR “pre-eclampsia” AND “Homo sapiens”. Επίσης διερευνήθηκε και η βιβλιογραφική βάση δεδομένων PubMed (Information 2007) για ερευνητικά άρθρα, χρησιμοποιώντας τους όρους αναζήτησης (“preeclampsia” OR “pre-eclampsia” OR “PE”) AND (“expressed genes” OR “microarray” OR “micro array” OR “micro-array”) χωρίς περιορισμούς ημερομηνίας δημοσίευσης (για τα σύνολα δεδομένων). Η τελική ηλεκτρονική έρευνα πραγματοποιήθηκε στις 15 Φεβρουαρίου 2017. Μελέτες που είτε περιελάμβαναν δεδομένα γονιδιακής έκφρασης σε άλλους ιστούς πλην του πλακούντα είτε ανέφεραν την επίδραση των φαρμάκων στην προεκλαμψία αποκλείστηκαν.

Στη δεύτερη μελέτη μας, εκτελέσαμε μια εκτενή αναζήτηση για την εύρεση δεδομένων γονιδιακής έκφρασης σχετικά με την ΙΦΝΕ. Ανακτήσαμε τα σύνολα δεδομένων από το δημόσιο αποθετήριο δεδομένων μικροσυστοιχιών GEO χρησιμοποιώντας τους όρους αναζήτησης (“Crohn's Disease” OR “CD” AND “Ulcerative colitis” OR “UC”) AND “Homo sapiens” χωρίς περιορισμούς ημερομηνίας δημοσίευσης (για τα σύνολα δεδομένων) και αντίστοιχα από την ArrayExpress. Η τελική ηλεκτρονική έρευνα πραγματοποιήθηκε στις 13 Μαΐου 2019. Μελέτες που είτε περιελάμβαναν δεδομένα γονιδιακής έκφρασης σε ιστούς διαφορετικούς του αίματος είτε δεν είχαν δεδομένα και για τις δύο ασθένειες είτε για υγιείς εξαιρέθηκαν από την ανάλυση μας.

Και για τις δύο μετα-αναλύσεις επιλέχθηκαν μελέτες ασθενών-μαρτύρων, αντίθετα όσες μελέτες ερευνούσαν διαφορετικό είδος (ζώα) και επεξεργαζόταν δείγματα είτε microRNA, είτε lncRNA, είτε μεθυλίωσης DNA, είτε RNAseq απορρίφθηκαν. Στις συστηματικές ανασκοπήσεις που διενεργήσαμε ακολουθήσαμε τις οδηγίες που υπάρχουν για τα προτιμώμενα στοιχεία αναφοράς για συστηματικές ανασκοπήσεις και μετα-αναλύσεις (Preferred Reporting Items for Systematic Reviews and Meta-Analyses, PRISMA) (Moher, Liberati et al. 2009) (βλέπε Εικόνες 20 και 26).

2.1.2 Ανάκτηση και καταγραφή δεδομένων γονιδιακής έκφρασης από τις μελέτες

Αρχικά για τις μελέτες που πληρούσαν τα κριτήρια επιλεξιμότητας πραγματοποιήθηκε καταγραφή, σε λογιστικά φύλλα, των βασικών τους πληροφοριών. Οι βασικές πληροφορίες κάθε μελέτης είναι ο μοναδικός κωδικός βάσης, το έτος, το σύνολο των υγιών ατόμων, το σύνολο των ασθενών ατόμων, ο ιστός, η πλατφόρμα που χρησιμοποιήθηκε στο πείραμα, ο αριθμός των ανιχνευτών και ο αριθμός των γονιδίων του τσιπ μικροσυστοιχίας.

Υπάρχουν διαφορετικοί τύποι δεδομένων που προσφέρουν τα πειράματα των μικροσυστοιχιών. Προσφέρεται η λίστα των γονιδίων τα οποία βρέθηκαν να εκφράζονται διαφορετικά, η δημοσιευμένη λίστα γονιδίων (published gene list, PGLs), οι μετρήσεις για κάθε ανιχνευτή σε κάθε υβριδισμό, τα μεμονωμένα δεδομένα επιπέδου ασθενούς (individual patient-level data, IPD), οι πίνακες που περιέχουν τις τιμές έκφρασης κάθε ανιχνευτή για κάθε δείγμα και οι πίνακες επιπέδων γονιδιακής έκφρασης (gene expression data matrix, GEDM). Ωστόσο οι PGLs και GEDM δεν μπορούν να χρησιμοποιηθούν εύκολα στις μετα-αναλύσεις καθώς κάθε μελέτη μπορεί να χρησιμοποιεί διαφορετικούς αλγόριθμους για την παραγωγή τους και ενδέχεται να υπάρχει πρόβλημα προκατάληψης (bias) (Ramasamy, Mondry et al. 2008).

2.1.3 Μεθοδολογία ανάλυσης δεδομένων μικροσυστοιχιών

Όπως αναφέρθηκε στην προηγούμενη παράγραφο τα αρχεία τύπου PGLs και GEDM δεν μπορούν να χρησιμοποιηθούν εύκολα στις μετα-αναλύσεις, και για αυτό το λόγο συστήνεται στη βιβλιογραφία η χρήση των μη επεξεργασμένων δεδομένων από τα σύνολα των δεδομένων (datasets), δηλαδή τα αρχεία εξόδου εξαγωγής σε επίπεδο χαρακτηριστικών (feature-level extraction output, FLEO), όπως είναι τα αρχεία CEL και GPR, και να μετατραπούν από τον ερευνητή σε GEDMs. Εάν στη μελέτη χρησιμοποιείται η ίδια πλατφόρμα τότε γίνεται χρήση του

ίδιου αλγορίθμου, ενώ εάν χρησιμοποιήθηκαν διαφορετικές πλατφόρμες, συνιστάται η χρήση συγκρίσιμων αλγορίθμων, ώστε να μπορούν να συνδυαστούν σωστά κατά τη μετα-ανάλυση τα δεδομένα και να ελαχιστοποιηθεί η προκατάληψη λόγω της διαφορετικής επεξεργασίας τους (Ramasamy, Mondry et al. 2008).

2.1.4 Ονοματολογία γονιδίων

Πριν τη χρήση των δεδομένων στη μετα-ανάλυση πρέπει να γίνει προσδιορισμός των γονιδίων μεταξύ των συνόλων δεδομένων. Τα σύνολα δεδομένων χρησιμοποιούν ως σύμβολα/αναγνωριστικά γονιδίων τα ονόματα από τη βάση δεδομένων είτε Entrez Gene είτε GenBank είτε RefSeq και αναφέρονται συλλογικά είτε ως GeneIDs είτε ως Gene Symbol. Δύο βασικά προβλήματα υπάρχουν κατά τον προσδιορισμό των γονιδίων.

Το πρώτο πρόβλημα έγκειται στη χρήση πολλών και διαφορετικών χρονολογικά μελετών κατά τη μετα-ανάλυση, με αποτέλεσμα τα ονόματα των γονιδίων τα οποία χρησιμοποιούσε κάθε πείραμα μικροσυστοιχίας, ενδέχεται (για κάποια γονίδια), να έχουν αλλάξει ή να μην ισχύουν και να έχουν καταργηθεί. Η λύση σε αυτό το πρόβλημα είναι ο ερευνητής να χρησιμοποιήσει το Entrez Gene ID κάθε ανιχνευτή και με τη βοήθεια της βάσης δεδομένων Gene της NCBI (<https://www.ncbi.nlm.nih.gov/gene>) καθώς και της HGNC (HUGO Gene Nomenclature Committee, <https://www.genenames.org/>) να αντιστοιχίσει το σωστό και ανανεωμένο GeneID (Εικόνα 8).

Στη συνέχεια το δεύτερο και πολύ βασικό πρόβλημα έγκειται είτε στο ότι πολλαπλοί ανιχνευτές μπορούν να υβριδοποιηθούν με το ίδιο ανανεωμένο GeneID (π.χ. επειδή τα τσιπ που χρησιμοποιούνται περιέχουν διπλούς ανιχνευτές) είτε ένας ανιχνευτής μπορεί να αντιστοιχεί σε περισσότερα του ενός GeneID (εάν η αλληλουχία του ανιχνευτή δεν είναι αρκετά συγκεκριμένη). Μία κατεύθυνση προς την επίλυση του προβλήματος «πολλά-προς-πολλά» (“many-to-many”) στην αντιστοίχιση των ανιχνευτών με τα γονίδια είναι να χρησιμοποιηθούν μόνο οι ανιχνευτές οι οποίοι συνδέονται με ένα μόνο γονίδιο κατά τη μετα-ανάλυση και να αποκλειστούν οι ανιχνευτές οι οποίοι συνδέονται με περισσότερα του ενός γονίδια. με τον τρόπο αυτό όμως υπάρχει φόβος να χαθεί σημαντικό τμήμα της πληροφορίας. Μια διαφορετική κατεύθυνση είναι η χρήση διαφόρων στατιστικών τεχνικών προκειμένου να μειωθεί η σχέση σε «ένα-προς-ένα» μεταξύ των ανιχνευτών και των μοναδικών γονιδίων για κάθε μελέτη. Είτε χρησιμοποιώντας τον μέσο όρο των τιμών έκφρασης των ανιχνευτών οι οποίοι αντιστοιχούν στο ίδιο γονίδιο, είτε τη μέγιστη τιμή έκφρασης

μεταξύ αυτών των ανιχνευτών, είτε μετα-αναλύοντας τις τιμές έκφρασης των ανιχνευτών που αντιστοιχούν στο ίδιο γονίδιο κ.α. (Warnat, Eils et al. 2005, Ramasamy, Mondry et al. 2008) (Εικόνα 9).

Εάν, ένας ανιχνευτής δεν έχει επαρκείς πληροφορίες για να υβριδοποιηθεί με κανένα GeneID τότε παραλείπεται από την περαιτέρω μελέτη.

ENTREZ_GENE_ID	Gene_Symbol	GSM76030	GSM76031	GSM76032	GSM76033	GSM76034	GSM76035	GSM76036	GSM76037
29974	A1CF	89.2408	72.55365	78.34144	78.21591	53.76302	52.44809	66.31123	68.48706
2	A2M	27.67372	5.720719	22.45506	26.12351	11.56055	14.66815	11.94915	3.387551
53947	A4GALT	34.72043	51.43071	36.70437	27.66871	24.31281	27.50267	21.62269	14.58852
51146	A4GNT	25.74486	30.88782	52.2123	16.49483	3.389503	11.14331	13.92067	38.91544
8086	AAAS	30.77449	12.661	61.32746	41.88271	44.19855	28.86434	29.16914	18.28437
65985	AACS	46.31213	65.77393	68.42493	56.12495	64.44863	77.92645	57.41499	72.69466
13	AADAC	2.201447	21.92679	2.139272	4.430514	5.326071	12.22125	4.153542	2.990068
79719	AAGAB	6.455825	6.343799	11.19733	6.358145	12.3579	21.14751	6.612822	10.61753
79719	AAGAB	51.178	68.12455	63.05055	63.52162	57.83913	69.15249	42.85292	29.96792
22848	AAK1	473.5273	469.1604	315.9598	355.0759	474.4536	525.6613	307.8606	574.4396
22848	AAK1	65.14898	9.515209	49.51562	57.78429	31.69544	43.42001	63.86556	47.08524
22848	AAK1	2.5463	4.5897	4.106212	2.539378	2.09778	2.355343	2.571502	4.315664
22848	AAK1	16.30959	46.90851	12.4402	35.81676	15.36039	7.744009	21.64518	37.19827
22848	AAK1	11.77509	36.73914	17.4938	9.415741	8.03341	3.45659	13.11809	5.832594
28971	AAMDC	2.045385	1.971407	1.194755	4.240792	1.302513	0.790407	8.911655	1.461081
28971	AAMDC	2.429358	18.91971	5.740527	2.656967	12.64153	19.72792	23.43904	5.880326
28971	AAMDC	36.45052	40.96142	48.70085	41.04791	32.37396	31.583	40.16014	17.3033

Εικόνα 8 Ένα τμήμα του πίνακα γονιδιακής έκφρασης, από μια μελέτη μικροσυστοιχίας. Εμφανίζονται πολλαπλοί ανιχνευτές του ίδιου γονιδίου με διαφορετικές τιμές έκφρασης.

gene_symbol	gsm76030	gsm76031	gsm76032	gsm76033	gsm76034	gsm76035	gsm76036	gsm76037
A1CF	89.2408	72.5536	78.3414	78.2159	53.763	52.4481	66.3112	68.4871
A2M	27.6737	5.72072	22.4551	26.1235	11.5605	14.6682	11.9491	3.38755
A4GALT	34.7204	51.4307	36.7044	27.6687	24.3128	27.5027	21.6227	14.5885
A4GNT	25.7449	30.8878	52.2123	16.4948	3.3895	11.1433	13.9207	38.9154
AAAS	30.7745	12.661	61.3275	41.8827	44.1986	28.8643	29.1691	18.2844
AACS	46.3121	65.7739	68.4249	56.125	64.4486	77.9264	57.415	72.6947
AADAC	2.20145	21.9268	2.13927	4.43051	5.32607	12.2212	4.15354	2.99007
AAGAB	28.8169	37.2342	37.1239	34.9399	35.0985	45.15	24.7329	20.2927
AAK1	113.861	113.383	79.9031	92.1264	106.328	116.527	81.8122	133.774
AAMDC	13.6418	20.6175	18.5454	15.9819	15.4393	17.3671	24.1703	8.2149
AAMP	83.6236	97.1336	139.591	128.112	94.3267	94.9446	118.662	79.3889
AANAT	3.85716	7.02272	6.11666	1.99253	3.79421	2.53652	3.58476	5.13116
AAR2	145.334	203.109	154.356	141.336	181.015	129.972	142.363	236.675
AARS	123.057	107.786	218.871	185.288	80.304	100.776	168.88	111.879
AASDHPPT	136.014	94.9354	75.8788	87.5413	101.888	95.7523	86.913	119.988
AASS	1.5277	1.43298	1.59631	0.603405	1.2238	0.854341	1.64734	1.2693

Εικόνα 9 Ένα τμήμα του πίνακα γονιδιακής έκφρασης, από μια μελέτη μικροσυστοιχίας. Η τιμή έκφρασης κάθε γονιδίου υπολογίστηκε από τον μέσο όρο των πολλαπλών ανιχνευτών.

Παρατηρώντας αυτά τα προβλήματα αναπτύξαμε ένα φιλικό προς τον χρήστη εργαλείο το οποίο να αντιστοιχίζει τους ανιχνευτές του πειράματος μικροσυστοιχίας στα ενημερωμένα σύμβολα γονιδίων και επιπλέον να προσφέρει στον χρήστη την επιλογή διαφόρων στατιστικών τεχνικών ώστε να υπάρχει η σχέση «ένα-προς-ένα» μεταξύ των ανιχνευτών και των μοναδικών γονιδίων για κάθε μελέτη. Το εργαλείο είναι διαθέσιμο στην ιστοσελίδα www.compgen.org/tools/GISU.

2.2 Μεθοδολογία μετα-ανάλυσης

2.2.1 Στατιστικό μοντέλο

Η επιλογή του κατάλληλου στατιστικού μοντέλου για μια μετα-ανάλυση εξαρτάται από τον τύπο των δεδομένων των μελετών που θα συμπεριληφθούν. Τα δεδομένα μπορεί να είναι είτε δυαδικά είτε συνεχή. Στα συνεχή δεδομένα, όπως είναι και η γονιδιακή έκφραση, μέγεθος επίδρασης θ [δηλαδή ο παράγοντας με τον οποίο θα πραγματοποιηθεί η μετα-ανάλυση και η διακύμανση (s_i^2) της εκάστοτε μελέτης], αποτελεί η διαφορά μέσων τιμών των δειγμάτων ελέγχου ως προς των δειγμάτων αναφοράς (Normand 1999). Η διαφορά τους μπορεί να είναι είτε τυποποιημένη (standardized mean difference) είτε μη τυποποιημένη (un standardized mean difference).

- Η τυποποιημένη διαφορά μέσων τιμών δίνεται από τον τύπο

$$d_i \equiv \frac{\bar{x}_{1i} - \bar{x}_{0i}}{sd_i} \quad (1)$$

και η θεωρούμενη κοινή τυπική απόκλιση (pooled standard deviation) δίνεται από τον τύπο

$$s_{pooled} = \sqrt{\frac{(n_{1i} - 1)s_{1i}^2 + (n_{0i} - 1)s_{0i}^2}{n_1 + n_0 - 2}} \quad (2)$$

- Η μη τυποποιημένη διαφορά μέσων τιμών δίνεται από τον τύπο

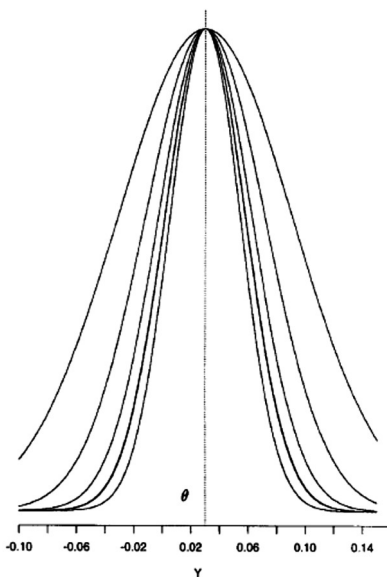
$$d_i = \bar{x}_{1i} - \bar{x}_{0i} \quad (3)$$

όπου \bar{x}_{1i} η μέση τιμή του δείγματος ελέγχου και \bar{x}_{0i} η μέση τιμή του δείγματος αναφοράς της, n_0 και n_1 το μέγεθος δείγματος κάθε ομάδας και S_{1i} και S_{0i} η τυπική απόκλιση της γονιδιακής έκφρασης κάθε ομάδας (Thakkinstian, McElduff et al. 2005).

2.2.1.1 Μοντέλο σταθερών επιδράσεων (*fixed effects model, FEM*)

Στο μοντέλο σταθερών επιδράσεων θεωρείται ότι όλες οι ανεξάρτητες μελέτες, Y_i , έχουν ένα μέσο κοινό μέγεθος επίδρασης, θ , το οποίο είναι ανεξάρτητο του αριθμού των μελετών και της διακύμανσής τους, s_i^2 . Ο τύπος που ακολουθεί αυτό το μοντέλο είναι (όπως φαίνεται και στην Εικόνα 10):

$$Y_i \sim N(\theta, s_i^2) \quad (4) \text{ για } i = 1, 2, 3, \dots, k.$$



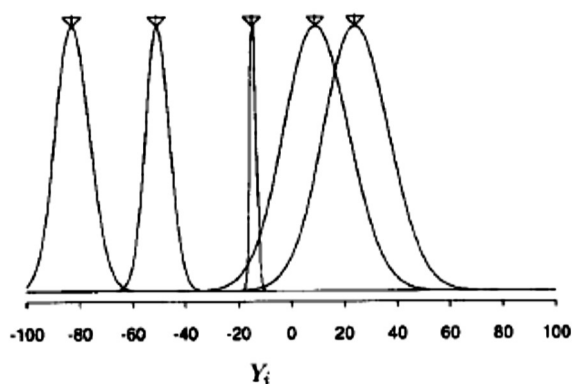
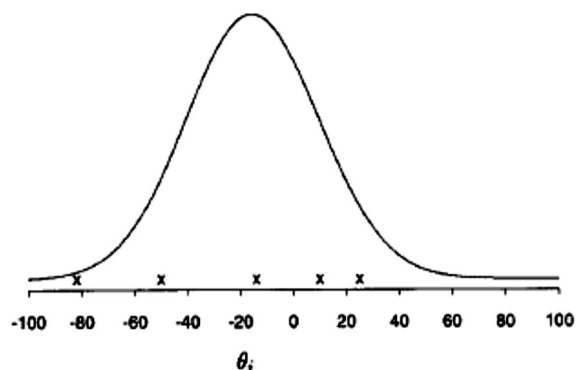
Εικόνα 10 Το μοντέλο σταθερών επιδράσεων.

Η κατανομή 5 υποθετικών επιδράσεων χρησιμοποιώντας το μοντέλο σταθερών επιδράσεων. Κάθε δείγμα Y_i έχει ένα κοινό μέγεθος επίδρασης θ . Η διαφορά ανάμεσα στις 5 μελέτες είναι η διαφορετική διακύμανση κάθε μελέτης s_i^2 , δηλαδή πόσο καλά υπολογίζει κάθε μελέτη το θ (Normand 1999).

2.2.1.2 Μοντέλο τυχαίων επιδράσεων (*random effects model, REM*)

Στο μοντέλο τυχαίων επιδράσεων θεωρείται ότι τα δείγματα της μελέτης μπορούν να προέλθουν από μια κατανομή πληθυσμών και με τον τρόπο αυτό επιτρέπεται η ύπαρξη ετερογένειας. Η καθεμία ανεξάρτητη μελέτη, Y_i , έχει ένα ξεχωριστό μέγεθος επίδρασης, θ_i , και μια τιμή διακύμανσης s_i^2 . Το κάθε θ_i θεωρείται ότι είναι δείγμα που προέρχεται από ένα σύνολο τιμών εκτιμητών με μέση τιμή (κοινό μέγεθος επίδρασης) θ και διακύμανση τ^2 . Γίνεται αντιληπτό ότι σε αυτό το μοντέλο εισάγεται μια νέα μεταβλητή τ^2 η οποία αντιπροσωπεύει την παρατηρούμενη ετερογένεια ανάμεσα στις ξεχωριστές μελέτες. Ο τύπος που ακολουθεί αυτό το μοντέλο είναι (όπως φαίνεται και στην Εικόνα 11):

$$Y_i \sim N(\theta_i, s_i^2) \quad (5) \text{ και } \theta_i \sim N(\theta, \tau^2) \quad (6) \text{ για } i = 1, 2, 3, \dots, k.$$



Εικόνα 11 Το μοντέλο τυχαίων επιδράσεων.

Κατανομή 5 υποθετικών δειγμάτων χρησιμοποιώντας το μοντέλο τυχαίων επιδράσεων. Κάθε μέγεθος επίδρασης θ_i προέρχεται από τον υπερπληθυσμό με μέγεθος επίδρασης θ και διακύμανση τ^2 . Στο παράδειγμα κάθε ένα από τα μεγέθη επίδρασης δημιούργησαν τα 5 αποτελέσματα των μοντέλων (Normand 1999).

2.1.2 Έλεγχος ετερογένειας

Όπως έχει ήδη αναφερθεί το μοντέλο των τυχαίων επιδράσεων λαμβάνει υπόψη την ετερογένεια μεταξύ των μελετών, αλλά αντίθετα αυτό δεν συμβαίνει στο μοντέλο των σταθερών επιδράσεων.

Ο έλεγχος για την παρουσία ετερογένειας μεταξύ των μελετών προσδιορίζεται από τον έλεγχο Q του Cochran. Το Q υπολογίζεται από τον τύπο:

$$Q = \sum_i^k W_i (d_i - D)^2 \sim \chi_{k-1}^2 \quad (7)$$

$$\text{με } D = \frac{\sum_{i=1}^k w_i d_i}{\sum_{i=1}^k w_i} \quad (8) \text{ και } w_i = \frac{1}{\text{var}(d_i)} \quad (9)$$

Το Q ακολουθεί κατανομή χ τετράγωνο (χ^2) με βαθμούς ελευθερίας $k-1$, όπου k είναι ο αριθμός των συνολικών μελετών και μπορεί να εφαρμοστεί για τον έλεγχο της ετερογένειας σε συνεχή και

διακριτά δεδομένα. Ο έλεγχος Q Cochran έχει καλύτερη απόδοση όταν στη μετα-ανάλυση συμπεριλαμβάνονται πολλές μελέτες (Villanueva and Zavarsek 2004).

Μια ακόμη στατιστική μέθοδος για τον έλεγχο της ετερογένειας είναι ο έλεγχος I^2 , ο οποίος κάνει χρήση του Q του έλεγχου του Cochran. Το I^2 υπολογίζεται από τον τύπο:

$$I^2 = \max\left(0, \frac{Q - (k-1)}{Q}\right) \times 100\% \quad (10)$$

Το I^2 παίρνει τιμές από 0% έως 100%, ανεξαρτήτως των τιμών που παίρνει η μετα-ανάλυση. Αρνητικές τιμές του I^2 θεωρούνται ίσες με το 0%. Τιμές μικρότερες του 25% φανερώνουν μικρή είτε αμελητέα ετερογένεια, τιμές μεταξύ 25% και 50% φανερώνουν μέτρια ετερογένεια, ενώ τιμές μεγαλύτερες του 50% υποδηλώνουν ότι η ετερογένεια είναι μεγάλη, με αποτέλεσμα η ετερογένεια να αποτελεί σοβαρό πρόβλημα κατά τη μετα-ανάλυση (Higgins, Thompson et al. 2003). Το I^2 δεν επηρεάζεται άμεσα από τον αριθμό μελετών της μετα-ανάλυσης.

Τέλος υπάρχει και ένας ακόμη εκτιμητής της ετερογένειας ο τ^2 , ο οποίος υπολογίζει τη μεταβλητότητα μεταξύ των δύο μελετών. Ο εκτιμητής τ^2 ορίστηκε αρχικά από τους Dersimonian και Laird με τον παρακάτω τύπο

$$\tau_{DL}^2 = \max\left\{0, \frac{Q_w - (k-1)}{\sum W_i - \frac{\sum w_i^2}{\sum w_i}}\right\} \quad (11)$$

Υψηλές τιμές των δεικτών ετερογένειας φανερώνουν την ύπαρξη ετερογένειας, ενώ όταν είναι ίσοι είτε πλησιάζουν το μηδέν φανερώνουν ότι δεν υπάρχει ετερογένεια μεταξύ των μελετών.

Κατά την διερεύνηση για την ύπαρξη ετερογένειας ορίζονται οι παρακάτω υποθέσεις:

- H_0 : Τα δείγματα είναι ομοιογενή
- H_1 : Τα δείγματα είναι ανομοιογενή

Όταν χρησιμοποιείται το μοντέλο τυχαίων επιδράσεων, ο υπολογισμός της ετερογένειας γίνεται από τους τύπους

$$\theta(\tau)_{MLE} = \frac{\sum_i W_i(\tau) Y_i}{\sum_k W_i(\tau)} \quad (12) \quad \text{με} \quad W_i(\tau) = \frac{1}{s_i^2 + \tau^2} \quad (13) \quad \text{και} \quad Y_i = \bar{x}_i - \bar{x}_0 \quad (14) \quad \text{ή οποιοδήποτε άλλο}$$

μέγεθος επίδρασης.

Αν η H_0 απορριφθεί, τότε θεωρούμε ότι υπάρχει ετερογένεια μεταξύ των μελετών (DerSimonian and Laird 1986).

2.3 Μεθοδολογία μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών

2.3.1 Μέθοδος βασισμένη στο t -test

Η στατιστική μέθοδος βασισμένη στο t -test, με χρήση σταθερών ή τυχαίων επιδράσεων, είναι ευρέως διαδεδομένη κατά τη μετα-ανάλυση δεδομένων μικροσυστοιχιών. Το μέγεθος επίδρασης, που χρησιμοποιείται σε αυτή τη μεθοδολογία είναι η τυποποιημένη μέση διαφορά των τιμών έκφρασης των δειγμάτων (ελέγχου και αναφοράς) (Choi, Yu et al. 2003). Δίνεται από τον τύπο

$$\hat{d}_i = \frac{\bar{X}_{1i} - \bar{X}_{0i}}{S_{pooled,i}} \quad (15)$$

όπου X_{1i} και X_{0i} είναι οι μέσες τιμές των δυο υπό μελέτη ομάδων (ελέγχου και αναφοράς) από τη σύγκριση της i μελέτης και $S_{pooled,i}$ είναι η τυπική απόκλιση, η οποία δίνεται από τον τύπο:

$$S_{pooled,i} = \sqrt{\frac{(n_{0i} - 1)S_{0i}^2 + (n_{1i} - 1)S_{1i}^2}{n_{0i} + n_{1i} - 2}} \quad (16)$$

με n_{1i} , και n_{0i} να είναι το μέγεθος δείγματος κάθε ομάδας και S_{1i} και S_{0i} να είναι η τυπική απόκλιση της γονιδιακής έκφρασης κάθε ομάδας.

Η διακύμανση του μεγέθους επίδρασης, το οποίο είναι d_{1i} , εκτιμάται με βάση την εξίσωση (Gleser and Olkin 2007):

$$\text{var}(\hat{d}_i) = s_{ii}^2 = \frac{1}{n_{1i}} + \frac{1}{n_{0i}} + \frac{\hat{d}_i^2}{2n_i} \quad (17)$$

Στη βιβλιογραφία, η εκτίμηση του δείγματος της τυποποιημένης διαφορά μέσω των τιμών αναφέρεται ως d του Cohen (Cohen 1988). Οι εκτιμήσεις του δείγματος της τυποποιημένης μέσης διαφοράς d έχουν την τάση να υπερεκτιμούν την απόλυτη τιμή σε μικρά δείγματα.

Λαμβάνοντας υπόψη ότι οι μελέτες μικροσυστοιχιών συνήθως διαθέτουν μικρό μέγεθος δείγματος σε κάθε ομάδα μελέτης και μη κανονική κατανομή των τιμών έκφρασης των γονιδίων αρκετές στατιστικές μέθοδοι έχουν προταθεί για την αποτελεσματική αντιμετώπιση αυτών των ζητημάτων. Αυτές οι μέθοδοι περιλαμβάνουν μεθόδους αναδειγματοληψίας όπως Bootstrap και permutation, Μπεϋζιανές μεθόδους, κανονικοποιημένες ή διορθωμένες τιμές t , τη διόρθωση του Hedges κ.α. Τα Bootstrap και permutation εφαρμόζονται εύκολα, αλλά απαιτούν τεράστιο

υπολογιστικό χρόνο. Οι κανονικοποιημένοι έλεγχοι t από την άλλη απαιτούν πρόσθετες τροποποιήσεις για την πολυμεταβλητή ανάλυση. Ωστόσο η παλιά αλλά αξιόπιστη εναλλακτική λύση, της διόρθωσης του Hedges, έχει αποδειχθεί ότι έχει καλές επιδόσεις, τουλάχιστον για μέτρια μεγέθη δείγματος (δηλαδή ~ 10).

Η χρήση του λεγόμενου g του Hedges καταφέρνει να διορθώσει τη μεροληψία που παρουσιάζεται στο d , δημιουργώντας μια αμερόληπτη εκτίμηση. Ένας συντελεστής διόρθωσης, που ονομάζεται J , χρησιμοποιείται για τη μετατροπή από d σε g .

$$\hat{g}_i = J(\nu_i) \hat{d}_i \quad (18)$$

Αν και υπάρχουν κατά προσέγγιση τύποι για τον J , με ποιο γνωστό τον $J(\nu_i) \approx \frac{3}{4\nu_i - 1}$ (19),

ωστόσο ο ακριβής τύπος για τον J , δίνεται από τον τύπο:

$$J(\nu_i) = \frac{\Gamma\left(\frac{\nu_i}{2}\right)}{\sqrt{\frac{\nu_i}{2}} \Gamma\left(\frac{\nu_i - 1}{2}\right)} \quad (20)$$

Όπου $\nu_i = n_{0i} + n_{1i} - 2$. Όταν χρησιμοποιείται το g , η διακύμανση θα αλλάξει επίσης σε:

$$\text{var}(\hat{g}_i) = s_i^2 = (J(\nu_i))^2 \text{var}(\hat{d}_i) \quad (21)$$

(Hedges 1981).

Υπάρχουν και άλλες προσεγγίσεις για τη βελτίωση των αποτελεσμάτων. Όπως οι μέθοδοι αναδειγματοληψίας, στις οποίες θα αναφερθούμε εκτενέστερα σε επόμενη ενότητα και άλλες, όπως μια προσέγγιση βασισμένη στη μέθοδο *limma*, η οποία χρησιμοποιεί γραμμικά μοντέλα για να αναλύσει πειράματα μικροσυστοιχιών. Είναι μια μέθοδος δύο βημάτων. Το αρχικό βήμα αποτελεί μια παραλλαγή του ελέγχου t , όπου το μέγεθος επίδρασης είναι $d_{Limma} = t_{Limma} / \sqrt{n}$ (Marot, Foulley et al. 2009) και το επόμενο βήμα είναι η πραγματοποίηση μετα-ανάλυσης με τη χρήση του μοντέλου τυχαίων επιδράσεων και είναι διαθέσιμη στο πακέτο Bioconductor στην R (<http://bioinf.wehi.edu.au/limma>).

2.3.2 Μέθοδος συνδυασμού των p -values

Η μέθοδος συνδυασμού των p -values είναι μια εύχρηστη μέθοδος κατά την οποία μπορούν να συνδυαστούν οι p -values από ανεξάρτητες μεταβλητές, αρκεί να δίνονται οι p -values όλων των

γονιδίων που μελετώνται ανεξαρτήτως μεθόδου ανάλυσης. Η πιο συχνά χρησιμοποιούμενη μέθοδος είναι η μέθοδος αθροίσματος των λογαρίθμων των p -values. Σε αυτή τη μέθοδο ως μέγεθος επίδρασης χρησιμοποιείται το S_i , το οποίο δίνεται από τον τύπο

$$S_i = -2 \sum_{k=1}^k \log(p_{ik}) \quad (22)$$

όπου S_i , είναι το άθροισμα των λογαρίθμων των p -values για κάθε γονίδιο i μεταξύ των μελετών k (Fisher 1934, Campain and Yang 2010).

Η μέθοδος συνδυασμού των p -values παρουσιάζει ορισμένα προβλήματα κατά τον συνδυασμό των μεγεθών επίδρασης. Ένα πρόβλημα της μεθόδου είναι ότι η κατεύθυνση της σύνδεσης δεν λαμβάνεται υπόψη, με αποτέλεσμα όλες οι p -values να πρέπει να είναι μονόπλευρες, αλλιώς τα υπερ- και υποεκφραζόμενα γονίδια πρέπει να συνδυαστούν ξεχωριστά. Επιπλέον, με αυτή τη μέθοδο δεν μπορεί να ποσοτικοποιηθεί το εύρος της συσχέτισης (δηλαδή το μέγεθος της επίδρασης) και να υπολογιστεί η ετερογένεια μεταξύ των μελετών. Λόγω αυτών των προβλημάτων χρησιμοποιούνται συνήθως διαφορετικές μέθοδοι στις οποίες, για παράδειγμα, συνδυάζονται ισοδύναμα z -scores αντί των p -values (Kontou, Pavlopoulou et al. 2018).

2.3.3 Μέθοδος υπολογισμού του γινομένου των βαθμών κατάταξης (*rank product*)

Η μέθοδος υπολογισμού του γινομένου των βαθμών κατάταξης χρησιμοποιείται κατά κόρον κατά τη μετα-ανάλυση δεδομένων γονιδιακής έκφρασης. Η μέθοδος αυτή υπολογίζει την αλλαγή πτύχωσης (fold change) κάθε γονιδίου, το οποίο ισοδυναμεί με τον λόγο έκφρασης των γονιδίων του δείγματος ελέγχου ως προς των γονιδίων του δείγματος αναφοράς, και στη συνέχεια αναλόγως της τιμής του λόγου των εκφράσεων, κατατάσσονται κατά φθίνουσα σειρά, δίνοντας έτσι σε κάθε γονίδιο έναν αριθμό (βαθμό κατάταξης) ο οποίος αντιπροσωπεύει την κατάταξή του στην συνολική λίστα. Η διαδικασία αυτή γίνεται σε όλα τα γονίδια, όλων των ατόμων και όλων των μελετών που χρησιμοποιούνται στη μετα-ανάλυση. Σε αυτή τη μέθοδο μέγεθος επίδρασης αποτελεί το γινόμενο των βαθμών κατάταξης, το οποίο προκύπτει από τον πολλαπλασιασμό των βαθμών κατάταξης του κάθε γονιδίου για όλες τις μελέτες που χρησιμοποιήθηκαν στη μετα-ανάλυση, και δίνεται από τον τύπο

$$RP_g = \left(\prod_{i=1}^k r_{gik} \right)^{\frac{1}{k}} \quad (23)$$

όπου r_{gik} είναι η i -οστή θέση, δηλαδή ο αριθμός κατάταξης, του γονιδίου g στη λίστα των συνολικών γονιδίων και k ο αριθμός των επαναλήψεων του γονιδίου στις μελέτες (Breitling, Armengaud et al. 2004).

Υπάρχει και μια άλλη παραλλαγή της μεθόδου, στην οποία χρησιμοποιείται ο μέσος όρος των βαθμών κατάταξης των γονιδίων μεταξύ των μελετών, ως μέγεθος επίδρασης, και δίνεται από τον τύπο (Zintzaras and Ioannidis 2008)

$$R^* = \frac{\sum_i^k R_i}{k} \quad (24).$$

2.3.4 Στατιστικά πακέτα μετα-ανάλυσης

2.3.4.1 R

Η R είναι μια γλώσσα προγραμματισμού ανοικτού κώδικα και περιβάλλον που προσφέρει τη δυνατότητα στον χρήστη να πραγματοποιήσει υπολογιστική στατιστική και διάφορα γραφήματα (Team 2015).

Για την εκτέλεση μιας μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών υπάρχουν διάφορα πακέτα, κυρίως στην R, για σχεδόν όλες τις προαναφερθείσες μεθοδολογίες όπως για παράδειγμα το πακέτο *metaMA* το οποίο προτείνει μεθόδους υπολογισμού των μεγεθών επίδρασης ή p -values από σταθερούς και ρυθμιζόμενους (moderated) ελέγχους t , χρησιμοποιώντας τις μεθόδους *limma* *SMVar* ή ελέγχους t σε διαφορετικά σύνολα δεδομένων τα οποία υπάρχουν σε διάφορες βάσεις δεδομένων μικροσυστοιχιών. Προσφέρει στον ερευνητή τη δυνατότητα χρήσης μεθόδων διόρθωσης πολλαπλών συγκρίσεων και δημιουργίας διαγραμμάτων Venn (Marot and Bruyere 2015). Το πακέτο *GeneMeta* εστιάζει στον συνδυασμό συνόλων δεδομένων σε συγκρίσεις δύο δειγμάτων, κυρίως με έμφαση στον έλεγχο t (Gentleman, Ruschhaupt et al. 2008). Το πακέτο *MAMA* το οποίο υλοποιεί εννέα διαφορετικές μεθόδους μετα-ανάλυσης μικροσυστοιχιών για την ταυτοποίηση διαφορετικά εκφραζόμενων γονιδίων όπως μεγεθών επίδρασης (FEM, REM), συνδυασμού p -values (Fisher), μέγιστη p -value, ελάχιστη p -value, z -scores, rank product κ.α. Προσφέρει ακόμη τη δημιουργία διαγραμμάτων Venn και την περαιτέρω ανάλυσή τους (Meta-Analysis Pattern Matches) (Ihnatova 2013). Το πακέτο *MetaDE* το οποίο εφαρμόζει δώδεκα κύριες μεθόδους μετα-ανάλυσης για την ανάλυση της διαφορικής έκφρασης όπως μεγεθών επίδρασης (FEM, REM), συνδυασμού p -values (Fisher), rank product, z -scores κ.α. Παρέχει στον ερευνητή τις επιλογές επεξεργασίας των δεδομένων πριν τη μετα-ανάλυση καθώς και γραφικές απεικόνισες των αποτελεσμάτων (Wang, Kang et al. 2012).

2.3.4.2 Stata

Μετα-ανάλυση σε δεδομένα μικροσυστοιχιών μπορεί να εκτελεσθεί και με τη χρήση του στατιστικού πακέτου Stata (StataCorp 2013).

Το Stata είναι ένα στατιστικό πακέτο λογισμικού γενικής χρήσης που δημιουργήθηκε το 1985 από την StataCorp. Χρησιμοποιείται σε διάφορους τομείς έρευνας όπως της οικονομίας, της επιδημιολογίας, της βιοϊατρικής, της βιοστατιστικής κ.α. Το Stata παρέχει στον χρήστη διάφορες δυνατότητες όπως διαχείριση δεδομένων, στατιστική ανάλυση, δημιουργία γραφικών παραστάσεων και προσαρμοσμένο προγραμματισμό. Το Stata επιτρέπει εντολές που έχουν γραφτεί από τον χρήστη και υπάρχει μια ενεργή κοινότητα επιστημόνων που υποστηρίζει την ανάπτυξη νέων μεθόδων.

Αν και το Stata δεν προσφέρει μια έτοιμη εντολή για μετα-ανάλυση σε δεδομένα μικροσυστοιχιών, ωστόσο είναι εφικτό να δημιουργηθεί ένας κώδικας για την εκτέλεσή της. Στην παρούσα διατριβή τροποποιήσαμε μια ήδη υπάρχουσα (από την ομάδα του εργαστηρίου μας, EMYBG) προσέγγιση για μετα-ανάλυση δεδομένων μικροσυστοιχιών με χρήση του μοντέλου των τυχαίων επιδράσεων. Η προσέγγιση μας θα επιτρέψει την πραγματοποίηση μετα-ανάλυσης (με τη χρήση της εντολής *metan*) με μέγεθος επίδρασης το g του Hedges, σε ένα μόνο βήμα, με ταυτόχρονη εκτέλεση της προσομοίωσης της μεθόδου Bootstrap. Η μέθοδος Bootstrap εκτελείται με ένα στρωματοποιημένο τρόπο στον οποίο οι μελέτες αντιμετωπίζονται ως στρώματα (ο κώδικας υπάρχει στο Παράρτημα).

2.4 Μεθοδολογία πολυμεταβλητής μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών

Όπως αναφέραμε στη θεωρία υπάρχουν μελέτες στις οποίες έχουμε πολλαπλές καταστάσεις (π.χ. πολλαπλά αποτελέσματα ή πολλαπλές θεραπείες) και που επίσης μας ενδιαφέρουν οι ίδιες συγκρίσεις όπως και στη μονομεταβλητή μετα-ανάλυση. Για παράδειγμα, σε μια μετα-ανάλυση μελετών μικροσυστοιχιών, υπάρχουν περιπτώσεις στις οποίες χορηγούνται διαφορετικά φάρμακα είτε καταγράφονται διαφορετικές δόσεις φαρμάκου. Σε αυτές τις περιπτώσεις η σύγκριση που ενδιαφέρει τον ερευνητή είναι η επίδραση (ποια φαρμακευτική αγωγή αποδίδει καλύτερα και ούτω καθεξής). Σε ένα άλλο παράδειγμα, μερικές (συναφείς) ασθένειες, διερευνώνται εναντίον μιας κοινής υγιούς ομάδας, προκειμένου να βρεθούν διαφορικά εκφραζόμενα γονίδια κοινά σε όλες τις συνθήκες ή γονίδια που διαφέρουν μεταξύ των συνθηκών

(π.χ. σκλήρυνση κατά πλάκας, ρευματοειδής αρθρίτιδα και σακχαρώδης διαβήτης τύπου I). Στην πολυμεταβλητή μετα-ανάλυση υπάρχουν επίσης τα μοντέλα σταθερών και τυχαίων επιδράσεων, τα οποία προσαρμόζονται στην πολυμεταβλητή φύση του προβλήματος.

Στο βασικό μοντέλο της πολυμεταβλητής μετα-ανάλυσης τυχαίων επιδράσεων, θεωρείται ότι το $g_i=(g_{1i}, g_{2i})$ κατανέμεται ακολουθώντας μια πολυμεταβλητή κανονική κατανομή γύρω από τις πραγματικές μέσες τιμές, σύμφωνα με το περιθώριο μοντέλο:

$$g_i \sim MVN(g, C + \Sigma_i) \quad (25)$$

Με το Σ_i δηλώνουμε τον πίνακα εντός των μελετών συνδιακύμανσης:

$$\Sigma_i = \begin{bmatrix} s_{1i}^2 & \rho_w s_{1i} s_{2i} \\ \rho_w s_{1i} s_{2i} & s_{2i}^2 \end{bmatrix} \quad (26)$$

Τα διαγώνια στοιχεία του Σ_i είναι οι ειδικές για τη μελέτη εκτιμήσεις της διακύμανσης οι οποίες θεωρούνται γνωστές, ενώ τα εκτός διαγωνίου στοιχεία αντιστοιχούν, στις ζευγαρωτές εντός των μελετών συνδιακυμάνσεις, για παράδειγμα, $\text{cov}(g_{1i}, g_{2i}) = \rho_w s_{1i} s_{2i}$. Με το C υποδηλώνουμε τον πίνακα συνδιακύμανσης μεταξύ των μελετών, ο οποίος πρέπει να εκτιμηθεί κατά τη διάρκεια της διαδικασίας προσαρμογής του μοντέλου:

$$C = \begin{bmatrix} \tau_1^2 & \rho_B \tau_1 \tau_2 \\ \rho_B \tau_1 \tau_2 & \tau_2^2 \end{bmatrix} \quad (27)$$

Μια χρήσιμη, μερικές φορές, απλοποίηση μπορεί να γίνει στην εξίσωση 27 εάν υποθέσουμε ότι οι διακυμάνσεις μεταξύ των μελετών είναι ίσες (Higgins and Whitehead 1996). Σε αυτή την περίπτωση αφήνοντας $\tau = \tau_1 = \tau_2$, ο C γίνεται:

$$C = \begin{bmatrix} \tau^2 & \rho \tau^2 \\ \rho \tau^2 & \tau^2 \end{bmatrix} \quad (28)$$

Αυτό είναι ιδιαίτερα χρήσιμο σε σύγκριση πολλών θεραπειών και σε περιπτώσεις όπου οι άμεσες συγκρίσεις είναι σπάνιες, καθώς με αυτήν την παραμετροποίηση ο συνολικός αριθμός των εκτιμώμενων παραμέτρων μειώνεται. Σε κάθε περίπτωση, το τελικό περιθώριο μοντέλο στο οποίο βασίζονται τα συμπεράσματα είναι:

$$\begin{bmatrix} g_{1i} \\ g_{2i} \end{bmatrix} \sim MVN \left\{ \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}, \begin{bmatrix} s_{1i}^2 + \tau_1^2 & \rho_w s_{1i} s_{2i} + \rho_B \tau_1 \tau_2 \\ \rho_w s_{1i} s_{2i} + \rho_B \tau_1 \tau_2 & s_{2i}^2 + \tau_2^2 \end{bmatrix} \right\} \quad (29)$$

2.4.1 Συνολικοί έλεγχοι (global tests)

Ένα σημαντικό πλεονέκτημα αυτού του μοντέλου πολυμεταβλητής μετα-ανάλυσης είναι ότι έχοντας υπολογιστεί ο πίνακας διακύμανσης-συνδιακύμανσης, μπορεί να χρησιμοποιηθεί ώστε να εκτελεσθούν συνολικοί έλεγχοι για τα μεγέθη επίδρασης.

Για παράδειγμα, μπορούμε να ελέγξουμε τη μηδενική υπόθεση ότι και οι δύο εκτιμητές των g_1 και g_2 είναι ίσοι με το μηδέν. Το στατιστικό του ελέγχου ακολουθεί κατανομή χ^2 με 2 βαθμούς ελευθερίας και μπορεί να ληφθεί χρησιμοποιώντας τη μέθοδο δέλτα:

$$W = \begin{bmatrix} \hat{g}_1 \\ \hat{g}_2 \end{bmatrix}^T \begin{bmatrix} \text{vâr}(\hat{g}_1) & \text{côv}(\hat{g}_1, \hat{g}_2) \\ \text{côv}(\hat{g}_1, \hat{g}_2) & \text{vâr}(\hat{g}_2) \end{bmatrix}^{-1} \begin{bmatrix} \hat{g}_1 \\ \hat{g}_2 \end{bmatrix} \sim \chi_2^2 \quad (30)$$

Με αυτή τη στατιστική μέθοδο μπορεί να γίνει έλεγχος πλειοτροπίας, δηλαδή έλεγχος για γονίδια τα οποία επηρεάζουν και τις δύο εκβάσεις. Πιο σημαντικά, ο συνολικός έλεγχος μπορεί να ανακαλύψει μια μεταβολή η οποία να μην είναι ανιχνεύσιμη κατά τη μονομεταβλητή ανάλυση. Ένας παρόμοιος έλεγχος, με τον παραπάνω έλεγχο, έχει προταθεί για την από κοινού εκτίμηση των κύριων γονιδιακών επιδράσεων και των αλληλεπιδράσεων γονιδίου-περιβάλλοντος σε μελέτες συσχέτισης ολόκληρου του γονιδιώματος (Genome Wide Association Study, GWAS) και για τον εντοπισμό πλειοτροπικών επιδράσεων στη μετα-ανάλυση των GWAS (Manning, LaValley et al. 2011) και από μέλη του εργαστηρίου μας (EMYBF) για τον εντοπισμό πλειοτροπικών επιδράσεων στην διμεταβλητή μετα-ανάλυση των GWAS (Medina-Gomez, Kemp et al. 2017).

Ένας άλλος χρήσιμος έλεγχος είναι ο έλεγχος της ισότητας των εκτιμητών. Σε αυτή την περίπτωση, η μηδενική υπόθεση θα είναι: $D=g_1-g_2=0$ και μπορεί να ελεγχθεί χρησιμοποιώντας τη δοκιμή Wald (ή ισοδύναμη δοκιμή χ^2):

$$\frac{\hat{D}}{\sqrt{\text{vâr}(\hat{D})}} = \frac{\hat{g}_1 - \hat{g}_2}{\sqrt{\text{vâr}(\hat{g}_1 - \hat{g}_2)}} \sim N(0,1) \quad (31)$$

Η διακύμανση του D θα μπορούσε να βρεθεί εύκολα με τη μέθοδο δέλτα:

$$\text{vâr}(\hat{g}_1 - \hat{g}_2) = \text{vâr}(\hat{g}_1) + \text{vâr}(\hat{g}_2) - 2\text{côv}(\hat{g}_1, \hat{g}_2) \quad (32)$$

Ο έλεγχος αυτός είναι παρόμοιος με τον έλεγχο για το γενετικό μοντέλο κληρονομικότητας σε μελέτες γενετικής συσχέτισης (Bagos 2008, Bagos 2013) και μπορεί επίσης, να χρησιμοποιηθεί για να γίνει σύγκριση των δύο εκτιμητών και να κατασκευαστεί ένα διάστημα εμπιστοσύνης είτε για τη διαφορά είτε για κάποια άλλη συνάρτηση των εκτιμητών.

Στην παρούσα διατριβή δημιουργήσαμε μια προσέγγιση για τη μέθοδο της πολυμεταβλητής μετα-ανάλυσης σε δεδομένα γονιδιακής έκφρασης, για χρήση στο στατιστικό πακέτο Stata, καθώς απ' όσο γνωρίζουμε είναι η πρώτη προσπάθεια προς αυτή την κατεύθυνση στη βιβλιογραφία. Όλα τα στοιχεία τα οποία είναι απαραίτητα για την εκτέλεση των εντολών αυτών είναι διαθέσιμα στο ευρύ κοινό μέσω της ιστοσελίδας www.compgen.org/tools/multivariate-microarrays και το θεωρητικό υπόβαθρο καθώς και η μεθοδολογία που χρησιμοποιήσαμε υπάρχουν στο άρθρο Vennou et al. 2020 (Vennou, Piovani et al. 2020).

2.5 Μέθοδοι διόρθωσης πολλαπλών συγκρίσεων

Οι μέθοδοι διόρθωσης πολλαπλών συγκρίσεων ελέγχουν εάν τα στατιστικώς σημαντικά διαφορικά εκφραζόμενα γονίδια είναι όντως σημαντικά και ότι δεν απορρίφθηκαν γονίδια τα οποία ήταν στατιστικώς σημαντικά.

Οι μέθοδοι διόρθωσης πολλαπλών συγκρίσεων λαμβάνουν ως είσοδο μια λίστα με p -values, οι οποίες έχουν προκύψει από τους ελέγχους, και επιστρέφουν νέες διορθωμένες p -values για όλα τα γονίδια της λίστας εισόδου. Οι δύο κατηγορίες είναι αυτές οι οποίες ελέγχουν το ποσοστό σφάλματος FWER και αυτές οι οποίες ελέγχουν το ποσοστό σφάλματος FDR.

Ο Bonferroni πρότεινε, αντί να ελέγχεται η πιθανότητα σφάλματος Τύπου I για κάθε μεμονωμένη δοκιμή, να ελέγχεται το συνολικό ποσοστό σφάλματος FWER που ορίζεται ως η πιθανότητα να εμφανιστεί τουλάχιστον ένα σφάλμα τύπου I. Υπάρχουν αρκετές μέθοδοι για τον έλεγχο FWER οι οποίες σαν αποτέλεσμα παρέχουν μια διορθωμένη p -value για ένα σύνολο στο οποίο ισχύει η μηδενική υπόθεση. Η πιο γνωστή μέθοδος είναι η διόρθωση Bonferroni στην οποία το επίπεδο σημαντικότητας α για κάθε έλεγχο υπολογίζεται από τη διαίρεση του FWER με τον αριθμό των ελέγχων (στην περίπτωση των μικροσυστοιχιών το πλήθος των γονιδίων). Έτσι η διορθωμένη p -value δίνεται από τον τύπο:

$$p_{cor(i)} = \frac{P_{(i)}}{n} \quad (33)$$

Εάν η διορθωμένη p -value ενός γονιδίου βρίσκεται κάτω από το όριο σφάλματος τότε το γονίδιο αυτό είναι στατιστικώς σημαντικό (Dudoit, Yang et al. 2000).

Υπάρχουν κι άλλες μέθοδοι διόρθωσης FWER, όπως η παρόμοια με τη μέθοδο Bonferroni αλλά πιο αυστηρή, ως προς τη νέα p -value, διόρθωση Bonferroni Step-Down ή Holm (Holm

1979), η διόρθωση Holland η οποία είναι μια παραλλαγή της διόρθωσης Holm (Holland and Copenhagen 1987) και η συχνά χρησιμοποιούμενη διόρθωση Sidak (Šidák 1967).

Η μέθοδος ελέγχου του FDR μας προσφέρει πιο κατανοητή και διαισθητική έννοια για τις πολλαπλές δοκιμές, αποδίδοντας την πιθανότητα να απορριφθεί εσφαλμένα μια δοκιμή (ψευδώς θετικό), καθώς το FDR ορίζεται ως το αναμενόμενο ποσοστό θετικών ελέγχων οι οποίοι απορρίπτονται λανθασμένα. Στις μελέτες ανίχνευσης γονιδίων, το FDR είναι η αναμενόμενη αναλογία των γονιδίων τα οποία ανιχνεύονται διαφορετικά εκφραζόμενα αλλά στην πραγματικότητα δεν είναι. Οι διαδικασίες ελέγχου του FWER παρέχουν πιο αυστηρό έλεγχο των σφαλμάτων τύπου I σε σχέση με τις διαδικασίες ελέγχου του FDR. Αυτό συνεπάγεται ότι οι διαδικασίες ελέγχου του FDR έχουν μεγαλύτερη στατιστική ισχύ (μικρότερο σφάλμα τύπου II), αλλά με κόστος την αύξηση του ποσοστού των σφαλμάτων τύπου I. Στην αρχή της μεθόδου ελέγχου του FDR οι p -values των γονιδίων ταξινομούνται κατά αύξουσα σειρά και στη συνέχεια η μεγαλύτερη p -value παραμένει ίδια και οι επόμενες p -values διορθώνονται με τον ακόλουθο τύπο (Benjamini and Hochberg 1995)

$$p_{cor(i)} = \frac{i}{n} p_{(i)} \quad (34)$$

2.6 Μέθοδοι αναδειγματοληψίας

Έχει ήδη αναφερθεί, ότι τα πειράματα μικροσυστοιχιών διαθέτουν συνήθως μικρό μέγεθος δείγματος και μη κανονική κατανομή στις τιμές έκφρασης των γονιδίων. Για τον λόγο αυτό ο βασισμένος στο t -test έλεγχος ακολουθείται αρκετές φορές από μια μέθοδο αναδειγματοληψίας.

Η μέθοδος Bootstrap βασίζεται στη δημιουργία νέων δειγμάτων ίδιου πλήθους με το αρχικό δείγμα. Τα νέα δείγματα δημιουργούνται με δειγματοληψία με επανάθεση από το αρχικό δείγμα. Για καθένα από τα νέα δείγματα Bootstrap υπολογίζεται η εκτιμήτρια της παραμέτρου που μας ενδιαφέρει και παράγεται η κατανομή Bootstrap της εκτιμήτριας. Βασική προϋπόθεση είναι ότι οι παρατηρήσεις του αρχικού δείγματος απεικονίζουν όλον τον πληθυσμό ώστε τελικά να παραχθούν, εκ νέου, πιο αξιόπιστα το τυπικό σφάλμα και το διαστήματα εμπιστοσύνης. Μια μικρή δειγματοληψία είναι επαρκής για να υπολογίσει το τυπικό σφάλμα αλλά χρειάζεται μεγαλύτερο δείγμα για την κατασκευή ενός 95% διαστήματος εμπιστοσύνης (Efron 1982, Efron and Tibshirani 1993). Επαναλήψεις της τάξης των 1.000 παράγουν πολύ ακριβείς εκτιμήσεις, ωστόσο συνήθως 50 ως 200 επαναλήψεις απαιτούνται για την εκτίμηση των τυπικών σφαλμάτων, κατά τις μεθόδους

μετα-ανάλυσης. Η μέθοδος Bootstrap έχει εφαρμοστεί σε διάφορων ειδών πειράματα, μελέτες αποδεικνύουν ότι παράγει ακριβείς εκτιμήσεις και σε πειράματα μικροσυστοιχιών, τουλάχιστον για μέτρια μεγέθη δειγμάτων (Meuwissen and Goddard 2004). Για πολύ μικρά μεγέθη δείγματος (κάτω των 10 ατόμων), διάφορες τροποποιήσεις στην τυποποιημένη μέθοδο Bootstrap έχουν προταθεί (Neuhauser and Jockel 2006, Jiang and Simon 2007) για καλύτερα αποτελέσματα.

Μια διαφορετική μέθοδος αναδειγματοληψίας είναι ο έλεγχος μετάθεσης (permutation test). Είναι ένας έλεγχος της στατιστικής σημαντικότητας όπου η κατανομή της στατιστικής δοκιμής υπό τη μηδενική υπόθεση λαμβάνεται υπολογίζοντας όλες τις πιθανές τιμές της στατιστικής δοκιμής μετά από ανακατατάξεις των ετικετών των παρατηρήσεων. Αν οι ετικέτες είναι ανταλλάξιμες, κάτω από τη μηδενική υπόθεση, τότε παράγονται ακριβή επίπεδα σημαντικότητας από τους προκύπτοντες ελέγχους (Kontou, Pavlopoulou et al. 2018). Γενικά ο έλεγχος μετάθεσης δεν είναι κατάλληλος για πολύ μικρά μεγέθη δείγματος (π.χ. τρία ανά ομάδα) (Neuhauser and Jockel 2006).

Η μέθοδος Bootstrap καθώς επίσης και ο έλεγχος μετάθεσης μπορούν να χρησιμοποιηθούν κατά τη μετα-ανάλυση με δύο διαφορετικούς τρόπους. Είτε να εκτελεστεί αρχικά για κάθε μελέτη η μέθοδος της αναδειγματοληψίας ξεχωριστά, να ληφθούν οι διορθωμένες εκτιμήσεις της διακύμανσης και κατόπιν αυτές να χρησιμοποιηθούν για τον υπολογισμό των βαρών της μετα-ανάλυσης, σε μια διαδικασία δύο βημάτων (two step), είτε να εκτελεστεί η ανάλυση σε ένα μόνο βήμα (one step) χρησιμοποιώντας τη μέθοδο αναδειγματοληψίας με ένα στρωματοποιημένο τρόπο, όπου οι μελέτες αντιμετωπίζονται ως στρώματα κατά τη διαδικασία της δειγματοληψίας.

2.7 IDR και IRR

Δύο σημαντικές αναλογίες που χρησιμοποιούνται για την επικύρωση της μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών είναι τα ποσοστά Integration-driven discovery, IDR (Ανακάλυψη με γνώμονα την ενσωμάτωση) και Integration-Driven Revision, IRR (Επανεξέταση με γνώμονα την ενσωμάτωση).

Οι Choi και συνεργάτες (Choi, Yu et al. 2003) το 2003 έθεσαν τον όρο της αναλογίας Integration-driven discovery, όταν γονίδια στατιστικώς σημαντικά διαφορεικά εκφραζόμενα κατά τη μετα-ανάλυση, μπορεί να μην θεωρούνται στατιστικώς σημαντικά σε καμία από τις μεμονωμένες μελέτες. Ο τύπος του ποσοστού IDR είναι:

$$IDR(x) = \frac{\# \text{γονιδίων}[(D_i \geq x) \text{ στην } _ \text{μετα} - \text{ανάλυση}] + [(D_i < x) \text{ σε } _ \text{τουλάχιστον } _ 1 _ \text{μελέτη}]}{\# \text{γονιδίων}[(D_i \geq x) \text{ στην } _ \text{μετα} - \text{ανάλυση}]} \quad (35)$$

Ο λόγος IDR μας δείχνει ότι συλλέγοντας ξεχωριστά σύνολα δεδομένων επιτυγχάνουμε το αποτέλεσμα της αύξησης του μεγέθους του δείγματος. Ακόμη αποτελεί μια διαβεβαίωση ότι τα ξεχωριστά αποτελέσματα έχουν επικυρωθεί μέσω μιας συγκριτικής ανάλυσης. Δηλαδή ότι συνολικά, η ενσωμάτωση των συνόλων δεδομένων μικροσυστοιχιών συνέβαλε στην επίτευξη πιο αξιόπιστων αποτελεσμάτων με αυξημένη ευαισθησία ξεπερνώντας τα τεχνητά αποτελέσματα των μεμονωμένων αναλύσεων.

Οι Stevens και Doerge (Stevens and Doerge 2005) χρησιμοποίησαν για πρώτη φορά τον όρο Integration-Driven Revision, για να περιγράψουν γονίδια τα οποία θεωρήθηκαν στατιστικώς σημαντικά διαφορεικά εκφραζόμενα από πολλαπλές μελέτες, αλλά στα αποτελέσματα της μετα-ανάλυσης δεν εμφανίζονται στατιστικώς σημαντικά. Ενώ πολλές μελέτες θα προωθούσαν τέτοια γονίδια για περαιτέρω μελέτη λόγω του μεγάλου και σημαντικού μεγέθους του αποτελέσματος, η μετα-ανάλυση θα κατέληγε στο συμπέρασμα ότι, λόγω των ασυνεπειών στα μεγέθη των αποτελεσμάτων στο σύνολο των μελετών, τα γονίδια εμφανίζονται στατιστικώς σημαντικά. Ο τύπος που χρησιμοποιείται για την εύρεση του ποσοστού IRR είναι:

$$IRR(x) = \frac{\# \text{γονιδίων}[(D_i \geq x) \text{ σε } _ \text{τουλάχιστον } _ 1 _ \text{μελέτη}] + [(D_i < x) \text{ στην } _ \text{μετα} - \text{ανάλυση}]}{\# \text{γονιδίων}[(D_i \geq x) \text{ σε } _ \text{τουλάχιστον } _ 1 _ \text{μελέτη}]} \quad (36)$$

Μεγάλες τιμές IRR τείνουν να εμφανίζονται όταν συνδυάζονται μεγάλες αλλά αντιφατικές εκτιμήσεις μεγέθους αποτελέσματος, το οποίο συμβαίνει συχνά στη μετα-ανάλυση μικροσυστοιχιών. Αντίθετα από το IDR, το οποίο καταδεικνύει την υπεροχή μιας μετα-ανάλυσης, το IRR μπορεί να καταδείξει την ανεπάρκεια μιας μετα-ανάλυσης, που ωστόσο συμβαίνει λόγω της μεγάλης αυστηρότητας της ως προς την ανάδειξη στατιστικώς σημαντικών αποτελεσμάτων.

2.8 Εργαλεία εμπλουτισμού

Μετά το πέρας της μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών και όταν είναι πλέον γνωστά τα στατιστικώς σημαντικά διαφορεικά εκφραζόμενα γονίδια χρειάζεται να βρεθεί η βιολογική σημασία αυτού του αποτελέσματος. Για τον λόγο αυτό χρησιμοποιούνται διάφορα εργαλεία βιοπληροφορικής ανάλυσης τα οποία πολλές φορές χρησιμοποιούν μεθόδους εμπλουτισμού. Υπάρχουν αρκετά εργαλεία είτε διαδικτυακά είτε τοπικά στον υπολογιστή του χρήστη.

2.8.1 PANTHER

Η PANTHER (Protein ANalysis THrough Evolutionary Relationships, πρωτεϊνική ανάλυση μέσω εξελικτικών σχέσεων) (<http://www.pantherdb.org/>) (Thomas, Kejariwal et al. 2003) είναι μια βιολογική βάση δεδομένων η οποία περιέχει πληροφορίες για τις οικογένειες και τις λειτουργικά συγγενείς υποοικογένειες γονιδίων/πρωτεϊνών, και προσφέρει στον χρήστη την εύρεση της ταξινόμησης και της λειτουργίας γονιδιακών προϊόντων. Η PANTHER αποτελεί κομμάτι του έργου Gene Ontology Reference Genome (Οντολογίας γονιδίων αναφοράς γονιδιώματος) το οποίο σχεδιάστηκε για να ταξινομεί τις πρωτεΐνες και τα γονίδιά τους με αναλύσεις υψηλής απόδοσης. Συνδυάζοντας τις γονιδιακές λειτουργίες, την οντολογία, τα μονοπάτια και τα εργαλεία στατιστικής ανάλυσης, η PANTHER δίνει τη δυνατότητα ανάλυσης σε γενετικά δεδομένα μεγάλης κλίμακας τα οποία λαμβάνονται από πειράματα υψηλής απόδοσης, όπως τα πειράματα γονιδιακής έκφρασης.

2.8.2 STRING

Η STRING (Search Tool for the Retrieval of INteracting Genes/proteins) (<https://string-db.org/>) (Szklarczyk, Gable et al. 2019) είναι μια βιολογική βάση δεδομένων και διαδικτυακή πηγή γνωστών και προβλεπόμενων αλληλεπιδράσεων πρωτεϊνών-πρωτεϊνών. Η STRING περιέχει πληροφορίες από πολλές πηγές, συμπεριλαμβανομένων πειραματικών δεδομένων, μεθόδων υπολογιστικής πρόβλεψης και δημόσιων βιβλιοθηκών. Χρησιμοποιεί ένα σύνολο λειτουργικών συστημάτων ταξινόμησης όπως είναι γονιδιακές οντολογίες (GO), βάσεις πρωτεϊνών (Pfam) και βιοχημικά μονοπάτια (KEGG) ώστε να εμπλουτίσει τις διάφορες λειτουργίες στις λίστες των πρωτεϊνών που δέχεται ως είσοδο από τον χρήστη. Περιέχει πληροφορίες για περίπου 9,6 εκατομμύρια πρωτεΐνες για περισσότερους από 2.000 οργανισμούς. Η STRING αναπτύχθηκε από μια κοινοπραξία διάφορων ακαδημαϊκών ιδρυμάτων (CPR, EMBL, KU, SIB, TUD και UZH). Η STRING χρησιμοποιώντας τις λίστες των πρωτεϊνών που εισάγει ο χρήστης επιστρέφει δίκτυα αλληλεπίδρασης πρωτεϊνών-πρωτεϊνών, τα οποία είναι ένα σημαντικό βήμα για την κατανόηση των κυτταρικών διεργασιών σε επίπεδο συστήματος. Η διερεύνηση των προβλεπόμενων δικτύων αλληλεπίδρασης μπορεί να χρησιμοποιηθεί για την αξιολόγηση των δομικών, λειτουργικών και εξελικτικών ιδιοτήτων των πρωτεϊνών και να προτείνει νέες κατευθύνσεις για μελλοντική πειραματική έρευνα (Schwartz, Yu et al. 2009).

2.8.3 MGI

Όπως έχει αναφερθεί κατά τη μετα-ανάλυση γονιδιακών δεδομένων από μικροσυστοιχίες ανακαλύπτονται στατιστικώς σημαντικά διαφορικά εκφραζόμενα γονίδια για μια συγκεκριμένη κατάσταση (π.χ. ασθένεια) σε σύγκριση με μια άλλη κατάσταση αναφοράς (π.χ. υγεία). Αναφέραμε στις δύο προηγούμενες παραγράφους εργαλεία, τα οποία βοηθούν στην καλύτερη κατανόηση των βιολογικών και κυτταρικών διεργασιών, των βιοχημικών μονοπατιών και των πρωτεϊνικών αλληλεπιδράσεων των ΔΕΓ. Ακόμη ένα πολύ σημαντικό βήμα στην καλύτερη κατανόηση μιας κατάστασης και του πώς επηρεάζουν τον φαινότυπο αυτά τα ΔΕΓ θα ήταν η στοχευμένη πειραματική μελέτη στα συγκεκριμένα γονίδια. Το ιδανικό σενάριο βεβαίως είναι η εκ νέου διερεύνηση αυτών των ΔΕΓ, ωστόσο κάτι τέτοιο δεν είναι πάντα εφικτό για διάφορους λόγους (κυρίως οικονομικούς, ελλιπών εγκαταστάσεων, τεχνογνωσίας, χρονικούς περιορισμούς). Υπάρχει εντούτοις η δυνατότητα να βρεθούν χρήσιμες πληροφορίες για αυτά τα γονίδια και ειδικά για τον τρόπο επιρροής τους στον φαινότυπο με τη βοήθεια της βάσης δεδομένων του γονιδιώματος του ποντικού MGI (Mouse Genome Informatics) (<http://www.informatics.jax.org/>) και πιο συγκεκριμένα ενός τμήματός της, της βάσης δεδομένων της έκφρασης γονιδίων του ποντικού (Gene Expression Database, GXD) (Smith, Hayamizu et al. 2019). Η GXD περιέχει δεδομένα από *in situ* υβριδοποιήσεις RNA, ανοσοϊστοχημεία, RT-PCR, πειράματα στυπώματος Northern και Western σε ποντίκια άγριου τύπου και μεταλλαγμένα (knock out, knock in κ.α.). Συνδυάζοντας αυτές τις πληροφορίες με δεδομένα της MGI και άλλων βάσεων δεδομένων παρέχει στον ερευνητή πληροφορίες σχετικά με γονίδια τα οποία έχει αποδειχθεί ότι εκφράζονται αποκλειστικά σε μια συγκεκριμένη δομή και/ή αναπτυξιακό στάδιο, με δεδομένα φαινοτύπων σε συγκεκριμένες ανατομικές δομές, με δεδομένα γονιδιακής έκφρασης άγριου τύπου για τους ιστούς οι οποίοι επηρεάζονται από μια συγκεκριμένη μετάλλαξη, και πίνακες σύγκρισης ιστών ανάμεσα σε ένα γονίδιο το οποίο εκφράζεται κανονικά και τις μεταλλάξεις σε αυτό καθώς και τι είδους ανωμαλίες αυτές προκαλούν.

Κεφάλαιο 3^ο:

Αποτελέσματα

3.1 Διερεύνηση της μεθοδολογίας της μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών

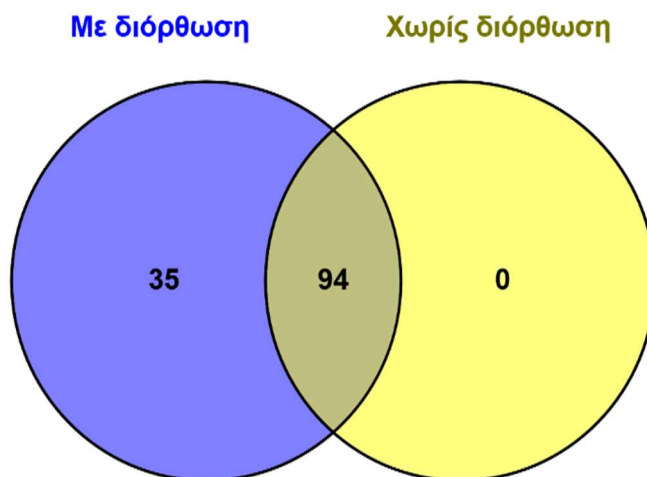
Πραγματοποιήσαμε διάφορους ελέγχους για την εύρεση βέλτιστου τρόπου μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών σε ένα δοκιμαστικό και τυχαίο υποσύνολο των 99 ατόμων (49 ασθενείς 50 υγιείς) και 10 μελετών το οποίο αποτελούνταν από 27.867 μοναδικά γονίδια/γονιδιακούς τόπους. Κάναμε σύγκριση των αποτελεσμάτων της μετα-ανάλυσης χωρίς και με τη διόρθωση του Hedges, ακόμη χωρίς και με τη μέθοδο αναδειγματοληψίας Bootstrap και επιπλέον μεταξύ διαφορετικών αριθμών επαναλήψεων της αναδειγματοληψίας. Συγκεκριμένα η μέθοδος Bootstrap εκτελέστηκε σε ένα βήμα με ένα στρωματοποιημένο τρόπο, όπου οι μελέτες χρησιμοποιούνται ως στρώματα. Τα ΔΕΓ από όλες τις διαφορετικές μεθοδολογίες είχαν τιμή FDR 1%,

3.1.1 Σύγκριση εφαρμογής και μη εφαρμογής διόρθωσης Hedges

Αρχικά θέλαμε να ελέγξουμε τη σημαντικότητα της εφαρμογής της διόρθωσης του Hedges στο μέγεθος επίδρασης. Για τον λόγο αυτό χρησιμοποιήσαμε το ίδιο σύνολο δεδομένων και εκτελέσαμε τη μετα-ανάλυση με τη χρήση του στατιστικού πακέτου Stata (ο κώδικας υπάρχει στο Παράρτημα). Χρονικά δεν παρατηρήσαμε κάποια διαφορά καθώς σχεδόν ο ίδιο χρόνος χρειαζόταν για την εκτέλεση της μετα-ανάλυσης. Στα αποτελέσματα ωστόσο παρατηρήσαμε αρκετή διαφορά, καθώς όπως φαίνεται και στην Εικόνα 12 όλα τα ΔΕΓ που βρέθηκαν στατιστικώς σημαντικά με τη μετα-ανάλυση χωρίς την διόρθωση Hedges βρέθηκαν και όταν έγινε χρήση της διόρθωσης, συν 35 ακόμη γονίδια τα οποία δεν είχαν θεωρηθεί στατιστικώς σημαντικά. Ακόμη η τιμή του συντελεστή συσχέτισης μεταξύ τους είναι 0.9965. Είναι λογικό να είναι τόσο μεγάλη καθώς δεν υπάρχουν μεγάλες αποκλίσεις στις δύο μεθόδους, αλλά παρατηρούμε ότι με τη διόρθωση κερδίζουμε πληροφορία.

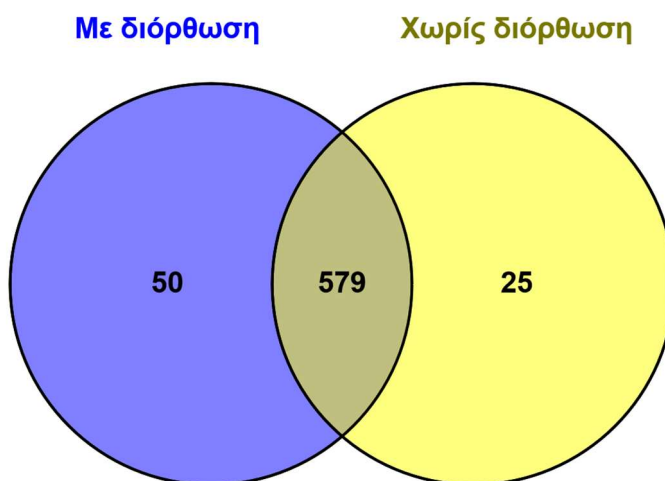
Τον ίδιο έλεγχο πραγματοποιήσαμε και σε ένα μεγαλύτερο σύνολο δεδομένων 426 ατόμων (195 ασθενείς και 231 υγιείς) και 10 μελετών το οποίο αποτελούνταν από 27.867 μοναδικά γονίδια/γονιδιακούς τόπους. Παρατηρούμε στην Εικόνα 13 ότι ένα μεγάλο ποσοστό των στατιστικώς σημαντικών ΔΕΓ είναι κοινά. Με τη διόρθωση Hedges θεωρήθηκαν στατιστικώς σημαντικά 50 διαφορετικά ΔΕΓ σε σχέση με τη μετα-ανάλυση χωρίς διόρθωση, η οποία βρήκε 25 στατιστικώς σημαντικά ΔΕΓ που δεν βρέθηκαν κατά τη μετα-ανάλυση με τη διόρθωση. Καταλαβαίνουμε ότι υπάρχουν κάποιες αποκλίσεις μεταξύ των δύο τρόπων και ότι ενδεχομένως

να χάνεται ένα μικρό ποσοστό πληροφορίας της τάξης του 4 %, αλλά με τη διόρθωση Hedges βρίσκεται εξίσου το 96% των ΔΕΓ της απλής μετα-ανάλυση συν 8% επιπλέον ΔΕΓ τα οποία διαφορετικά δεν θα μπορούσαν να βρεθούν. Παρατηρούμε ότι όσο μεγαλύτερο είναι το δείγμα, η επίδραση της διόρθωσης γίνεται μικρότερη, κάτι που αναμενόταν.



Εικόνα 12 Διάγραμμα Venn των στατιστικώς σημαντικών γονιδίων τα οποία προέκυψαν από τη μετα-ανάλυση με και χωρίς διόρθωση Hedges σε ένα μικρό σύνολο δεδομένων (99 άτομα).

Παρατηρούμε λοιπόν και στις δύο περιπτώσεις (μεγάλα και μικρά σύνολα δεδομένων) ότι η χρήση της διόρθωσης Hedges, η οποία δημιουργεί αμερόληπτες εκτιμήσεις, είναι μια αναγκαία διαδικασία διότι έχει καλύτερα αποτελέσματα καθώς δεν χάνεται στατιστικώς σημαντική πληροφορία (ειδικά σε μικρά σύνολα δεδομένων), ενώ αντίθετα δίνει επιπλέον πληροφορία η οποία θα χανόταν κατά την απλή μετα-ανάλυση. Γίνεται έτσι κατανοητό ότι σε μικρά σύνολα δεδομένων είναι αναγκαία η χρήση της διόρθωσης Hedges.

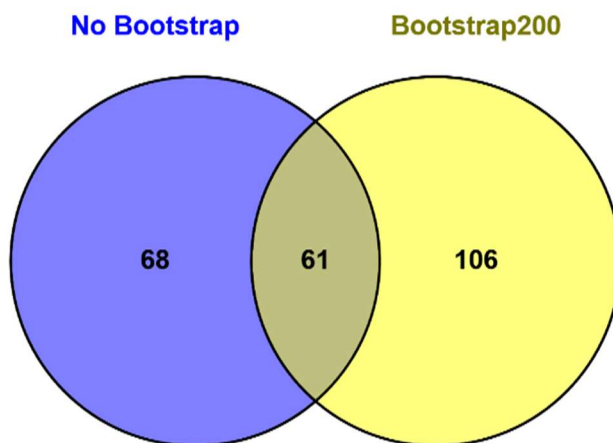


Εικόνα 13 Διάγραμμα Venn των στατιστικώς σημαντικών γονιδίων τα οποία προέκυψαν από τη μετα-ανάλυση με και χωρίς διόρθωση Hedges σε ένα μεγάλο σετ δεδομένων (426 άτομα).

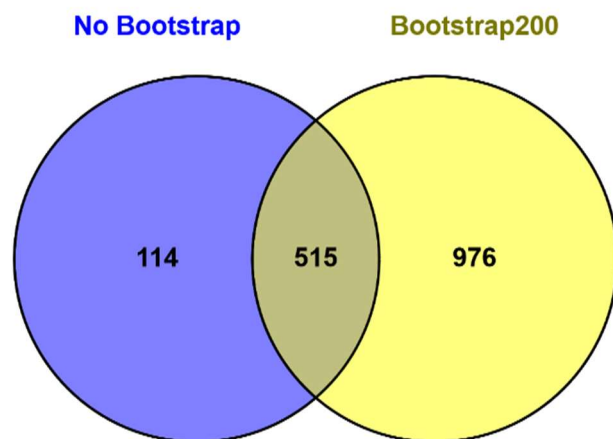
3.1.2 Σύγκριση εφαρμογής διόρθωσης Hedges χωρίς και με 200 επαναλήψεις Bootstrap

Πραγματοποιήσαμε στη συνέχεια συγκρίσεις μεταξύ της διόρθωσης Hedges με και χωρίς τη μεθοδολογία του Bootstrap, σε μικρό και σε μεγαλύτερο σύνολο δεδομένων.

Σε μικρό σύνολο δεδομένων (99 ατόμων, περίπου 10 ατόμων ανά μελέτη) η απλή μετα-ανάλυση με μόνο τη διόρθωση του Hedges χρειάζεται 1,5 ώρα ενώ με τη χρήση της αναδειγματοληψίας Bootstrap των 200 επαναλήψεων περίπου 9 ημέρες. Αντίστοιχα σε μεγάλο σύνολο δεδομένων (426 ατόμων) η απλή μετα-ανάλυση με μόνο τη διόρθωση του Hedges χρειάζεται 7,5 ώρες ενώ με τη χρήση της αναδειγματοληψίας Bootstrap των 200 επαναλήψεων περίπου 17 ημέρες. Όλα αυτά, με χρήση συμβατικού προσωπικού ηλεκτρονικού υπολογιστή (Intel Pentium Processor G3240, 3.10 GHz, 4GB RAM, Σκληρός δίσκος 1TB, Windows 7 professional). Οι συντελεστές συσχέτισης ήταν 0.8586 και 0.8331 για το κάθε σύνολο αντίστοιχα και οι αρνητικοί λογάριθμοι των p -values των γονιδίων κατανέμονται κανονικά (Εικόνα 16 και 17).

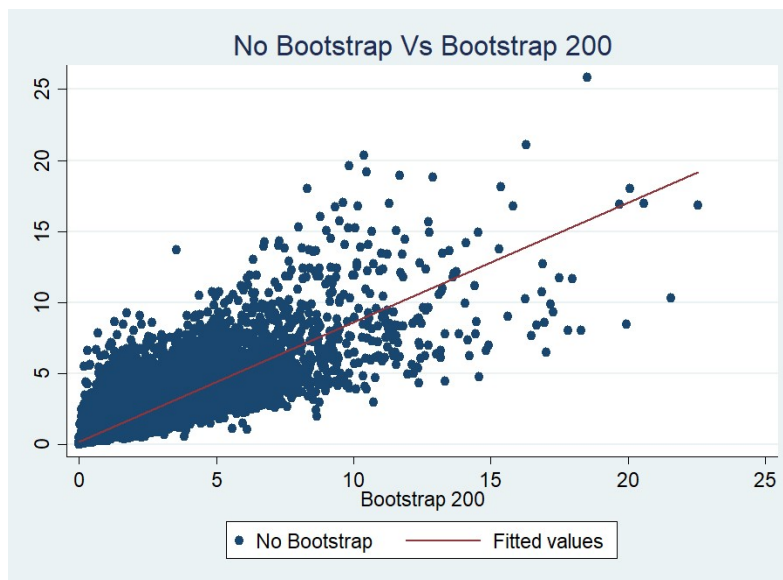


Εικόνα 14 Διάγραμμα Venn των στατιστικώς σημαντικών γονιδίων τα οποία προέκυψαν από τη μετα-ανάλυση με και χωρίς bootstrap 200 επαναλήψεων σε ένα μικρό σύνολο δεδομένων (99 άτομα).

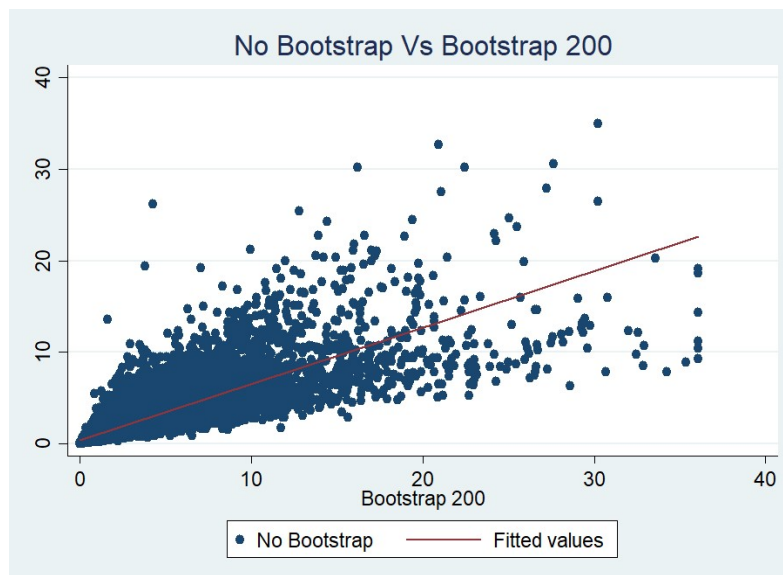


Εικόνα 15 Διάγραμμα Venn των στατιστικώς σημαντικών γονιδίων τα οποία προέκυψαν από τη μετα-ανάλυση με και χωρίς bootstrap 200 επαναλήψεων σε ένα μεγάλο σύνολο δεδομένων (426 άτομα).

Θεωρητικά γνωρίζουμε ότι η διόρθωση με τη χρήση της μεθόδου Bootstrap παράγει πάντα καλύτερα αποτελέσματα, ειδικά σε περιπτώσεις μικρού δείγματος ή μη κανονικής κατανομής των τιμών (της έκφρασης). Τα αποτελέσματά μας δείχνουν ότι η μέθοδος Bootstrap θα δώσει πάντα περισσότερα γονίδια ως στατιστικώς σημαντικά (βλέπε Εικόνα 14 και 15), πιθανώς σε περιπτώσεις που η δειγματική κατανομή δεν ήταν η θεωρητικά αναμενόμενη. Από την άλλη, ο σημαντικός υπολογιστικός χρόνος τον οποίο απαιτεί μια τέτοια ανάλυση αποτελεί ένα σημαντικό πρόβλημα. Φαίνεται λοιπόν, ότι η διόρθωση του Hedges αποτελεί έναν αναγκαίο συμβιβασμό.



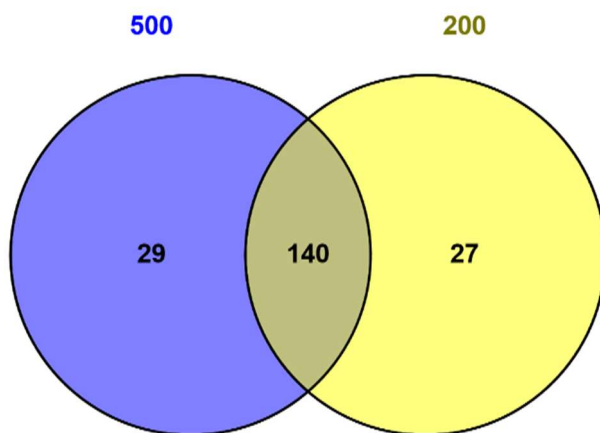
Εικόνα 16 Διάγραμμα του αρνητικού λογάριθμου των p -values όλων των γονιδίων σε ένα μικρό σύνολο δεδομένων (99 άτομα) σε σχέση με την χρήση ή όχι της μεθόδου Bootstrap με 200 επαναλήψεις.



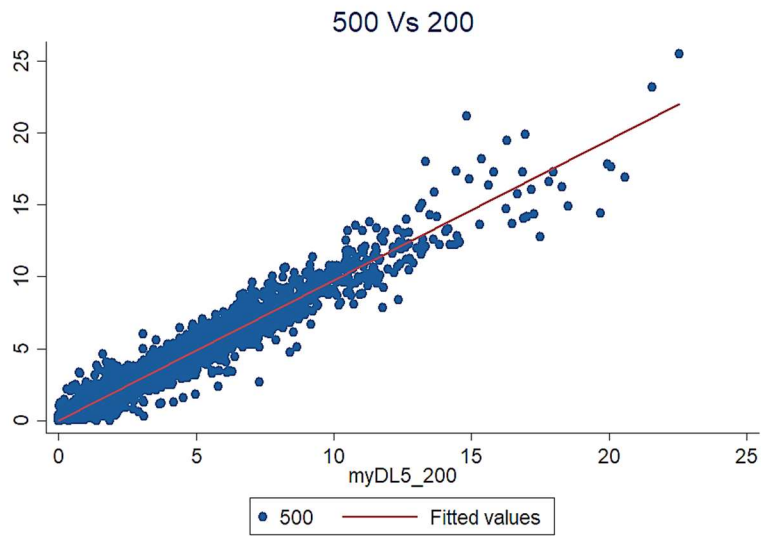
Εικόνα 17 Διάγραμμα του αρνητικού λογάριθμου των p -values όλων των γονιδίων σε ένα μεγάλο σύνολο δεδομένων (426 άτομα) σε σχέση με την χρήση ή όχι της μεθόδου Bootstrap με 200 επαναλήψεις.

3.1.3 Έλεγχος αριθμών επαναλήψεων Bootstrap με εφαρμογή διόρθωσης Hedges

Πραγματοποιήσαμε συγκρίσεις για διάφορους αριθμούς επαναλήψεων της αναδειγματοληψίας Bootstrap μετά την εφαρμογή της διόρθωσης Hedges. Όπως φαίνεται και στις Εικόνες 18 και 19 δεν υπάρχουν μεγάλες διαφορές μεταξύ των στατιστικώς σημαντικών ΔΕΓ (σε FDR 1%) συγκριτικά με τις 200 και 500 επαναλήψεις. Ωστόσο υπάρχει αρκετή διαφορά στον χρόνο εκτέλεσης. Σε ένα μικρό σύνολο δεδομένων <100 ατόμων (περίπου 10 ατόμων ανά μελέτη) που χρησιμοποιήσαμε για τον έλεγχο είδαμε ότι Bootstrap 200 επαναλήψεων χρειάζονταν χρονικό διάστημα περίπου 9 ημερών ενώ αντίθετα οι 500 επαναλήψεις χρειάζονταν περίπου 24 ημέρες, με χρήση συμβατικού προσωπικού ηλεκτρονικού υπολογιστή, χωρίς να προσφέρουν μεγάλες αποκλίσεις ούτε στις *p-values* των γονιδίων, ούτε και στο σύνολο των ΔΕΓ (στα γονίδια και στον αριθμό τους), όπως φαίνεται και στην Εικόνα 18 και στην Εικόνα 19 και από τον συντελεστή συσχέτισης μεταξύ τους ο οποίος είναι 0.9860. Παρατηρούμε από την Εικόνα 19 ότι οι αρνητικοί λογάριθμοι των *p-values* των γονιδίων ακολουθούν γραμμική σχέση σχεδόν τέλεια, με λίγες παρεκκλίσεις, οι οποίες προκαλούν μείωση του συντελεστή συσχέτισης από το 1 στο 0.9860. Θεωρητικά, λοιπόν, θα μπορούσαμε να πούμε ότι η χρήση των 200 επαναλήψεων για τη μεθοδολογία Bootstrap είναι ένας ικανοποιητικός αριθμός τόσο σε χρόνο όσο σε σημαντικότητα και ως προς την ορθότητα του αποτελέσματος.



Εικόνα 18 Διάγραμμα Venn των στατιστικώς σημαντικών γονιδίων τα οποία προέκυψαν από τη μετα-ανάλυση με διαφορετικό αριθμό επαναλήψεων. Συγκεκριμένα 200 εναντίον 500 επαναλήψεων.

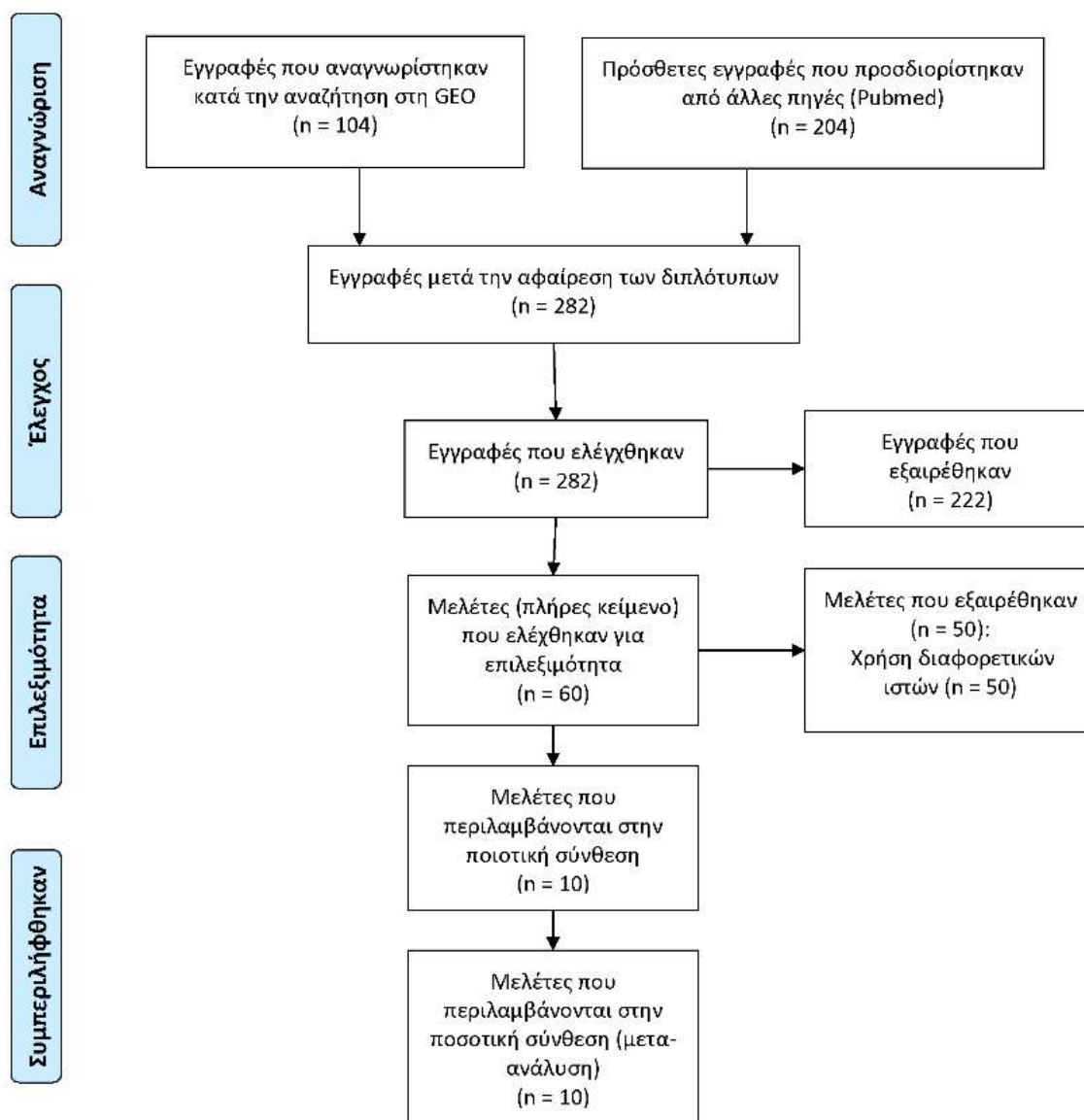


Εικόνα 19 Διάγραμμα του αρνητικού λογάριθμου των p -values όλων των γονιδίων σε σχέση με τον αριθμό των επαναλήψεων της μεθόδου Bootstrap. Συγκεκριμένα 200 εναντίον 500 επαναλήψεων.

3.2 Μετα-ανάλυση δεδομένων γονιδιακής έκφρασης στην προεκλαμψία

3.2.1 Αποτελέσματα

Από τις βάσεις δεδομένων PubMed και GEO ανακτήθηκαν και διερευνήθηκαν 60 σύνολα δεδομένων και άρθρα. Μόνο δέκα πληρούσαν τα κριτήρια επιλεξιμότητας και στη συνέχεια συμπεριλήφθηκαν στη μετα-ανάλυση (Εικόνα 20). Τα σύνολα δεδομένων παρείχαν δεδομένα για 27.867 μοναδικά γονίδια/γονιδιακούς τόπους από 195 εγκύους με προεκλαμψία και 231 υγιείς εγκύους (Πίνακας 3). Ο αριθμός των γονιδίων που εμφανίστηκαν στις μελέτες εξετάστηκε και παρουσιάζεται επίσης στον Πίνακα Παραρτήματος 1.



Εικόνα 20 Διάγραμμα ροής της μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών για την ασθένεια της προεκλαμψίας

Πίνακας 3 Χαρακτηριστικά των μελετών που συμπεριλήφθηκαν στη μετα-ανάλυση για την ασθένεια της προεκλαμψίας

Αναφορά	Κωδικός GEO	Πλατφόρμα	Ασθενείς	Μ.Ο. ηλικίας ασθενών	Εβδομάδα κύησης ασθενών	Θηλυκοί απόγονοι ασθενών	Υγιείς	Μ.Ο. ηλικίας υγιών	Εβδομάδα κύησης υγιών	Θηλυκοί απόγονοι υγιών	Κριτήρια	Συνολικός αριθμός ανιχνευτών	Συνολικός αριθμός γονιδίων
(Nishizawa, Pryor-Koishi et al. 2007)	GSE4707 ^a	Agilent-012391 Whole Human Genome Oligo Microarray G4112A (Feature Number version)	10 (σοβαρή PE ^β) (5 πρόωρη-, 5 καθυστερημένα -έναρξη)	30,3	32,9 ± 4,0	ΔΥΠ ^γ	4	29,2	32,9 ± 5,6	ΔΥΠ	Σοβαρή PE θεωρούταν όταν ΣΑΠ ^δ >160 mmHg/6h και ΔΑΠ ^ε >110 mmHg/6h and Πρωτεϊνουρία >2 g/24h. Η ομάδα αναφοράς ήταν γυναίκες με κανονική πίεση και φυσιολογική κύηση	43.931	18.153
(Sitras, Paulssen et al. 2009)	GSE10588	ABI Human Genome Survey Microarray Version 2	17 (σοβαρή PE)	30,5	34 ± 3,6	ΔΥΠ	26	30,2	39,6 ± 1,3	ΔΥΠ	Σοβαρή PE θεωρούταν όταν ΣΑΠ ≥160 mmHg/6h και/ή ΔΑΠ ≥110 mmHg/6h and Πρωτεϊνουρία ≥2+ g/24h σε στικ ούρων σε κύηση μετά την 20 ^η εβδομάδα. Η ομάδα αναφοράς ήταν γυναίκες χωρίς PE και φυσιολογική κύηση	32.879	16.360

(Winn, Gormley et al. 2009)	GSE14722	[HG-U133A] Affymetrix Human Genome U133A Array	12 (σοβαρή PE)	30,7	32,1 ± 3,3	ΔΥΠ	11	30,2	31,0 ± 4,6	ΔΥΠ	Δεν δίνονταν οι πληροφορίες για τα κριτήρια κατάταξης της PE. Η ομάδα αναφοράς ήταν γυναίκες με κανονική πίεση και φυσιολογική κύηση με πρόωρο τοκετό χωρίς καμία ένδειξη λοίμωξης. Σοβαρή PE θεωρούταν όταν ΣΑΠ >160 mmHg και ΔΑΠ >110 mmHg και/ή Πρωτεϊνουρία ≥2 g/24h. Η ομάδα αναφοράς ήταν γυναίκες με κανονική πίεση και φυσιολογική κύηση	22.115	12.347
(Nishizawa, Ota et al. 2011)	GSE24129	[HuGene-1_0-st] Affymetrix Human Gene 1.0 ST Array [transcript (gene) version]	8 (σοβαρή PE)	31,0	34,4 ± 1,8	ΔΥΠ	8	31,5	38,1 ± 0,8	ΔΥΠ	PE οριζόταν όταν ΣΑΠ ≥140 mmHg και ΔΑΠ ≥90 mmHg και Πρωτεϊνουρία ≥0.3 g/24h. Η ομάδα αναφοράς ήταν γυναίκες με κανονική	24.805	20.761
(Tsai, Hardison et al. 2011)	GSE25906	Illumina human-6 v2.0 expression beadchip	23	≥18	33,6 ± 3,7	56,5%	37	≥18	37,6 ± 1,93	43,2%		48.701	20.007

(Meng, Chen et al. 2012)	GSE30186	Illumina HumanHT-12 V4.0 expression beadchip	6	26,0	36,4 ± 0,9	ΔΥΠ	6	28,5	39 ± 0,7	ΔΥΠ	πίεση και φυσιολογική κύηση PE οριζόταν όταν ΣΑΠ >140 mmHg/6h και ΔΑΠ >90 mmHg/6h και Πρωτεϊνουρία ≥0.3 g/24h ή >2+ σε στικ ούρων σε κύηση μετά την 20 ^η εβδομάδα. Η ομάδα αναφοράς ήταν γυναίκες με κανονική πίεση και φυσιολογική κύηση	47.323	23.474
(Jebbink, Boot et al. 2015)	GSE54618	Illumina HumanHT-12 V4.0 expression beadchip	5	30	34,61	51,7%	12	30	35,5	46,9%	PE οριζόταν όταν ΣΑΠ ≥140 mmHg/6h ή ΔΑΠ ≥90 mmHg/6h και Πρωτεϊνουρία ≥0.3 g/24h or >1+ σε στικ ούρων σε κύηση μετά την 20 ^η εβδομάδα. Η ομάδα αναφοράς ήταν γυναίκες με κανονική πίεση και φυσιολογική κύηση	47.231	23.470

(Liang, Niu et al. 2016)	Agilent-039494 SurePrint G3 Human GE v2 8x60K Microarray 039381 (Feature Number version)	15 (7 πρόωρη- 8 καθυστερημένη -έναρξη)	17.1 (πρόωρη- έναρξη), 28.2 (καθυστερ ημένη- έναρξη)	31,7 (πρόωρη- έναρξη), 38,4 (καθυστερημέν η-έναρξη)	ΔΥΠ	10	29,8 (πρόωρη ομάδα αναφοράς) , 28,6 (καθυστερ ημένη ομάδα αναφοράς)	30,2 (πρόωρη ομάδα αναφοράς), 40,6 (καθυστερη μένη ομάδα αναφοράς)	ΔΥΠ	PE οριζόταν όταν ΣΑΠ=140 mmHg/6h και/ή ΔΑΠ=90 mmHg/6h και Πρωτεϊνουρία ≥0.3 g/24h. Πρόωρη έναρξη της PE <34 εβδομάδες κύησης, Καθυστερημένη- έναρξη της PE >34 εβδομάδες κύησης. Η ομάδα αναφοράς ήταν γυναίκες με κανονική πίεση και πρόωρο ή φυσιολογικό τοκετό	62.976	21.459
(Guo, Tsai et al. 2013)	Illumina human-6 v2.0 expression beadchip	19	ΔΥΠ	33,6 ± 3,7	57,9%	40	ΔΥΠ	37.7 ± 1.9	42,9%	PE οριζόταν όταν ΣΑΠ >140 mmHg/6h και ΔΑΠ >90 mmHg/6h και Πρωτεϊνουρία ≥0.03 g/4h ή >1+ σε στικ ούρων σε κύηση μετά την 20 ^η εβδομάδα. Η ομάδα αναφοράς ήταν γυναίκες με μη προεκλαμπτική κύηση	48.700	20.007

(Leavey, Benton et al. 2016)	GSE75010	[HuGene-1_0-st] Affymetrix Human Gene 1.0 ST Array [transcript (gene) version]	80	32,9	31,9 ± 3,5	ΔΥΠ	77	32,8	34,9 ± 4,9	ΔΥΠ	PE οριζόταν όταν ΣΑΠ ≥140 mmHg/6h και/ή ΔΑΠ ≥90 mmHg/6h και Πρωτεϊνουρία ≥0.3 g/24h ή ≥2+ σε στικ ούρων σε κύηση μετά την 20 ^η εβδομάδα. Η ομάδα αναφοράς ήταν γυναίκες με μη προεκλαμπτική κύηση	32.321	23.016
------------------------------	----------	--	----	------	------------	-----	----	------	------------	-----	--	--------	--------

^γΣΑΠ =

Συστολική αρτηριακή πίεση

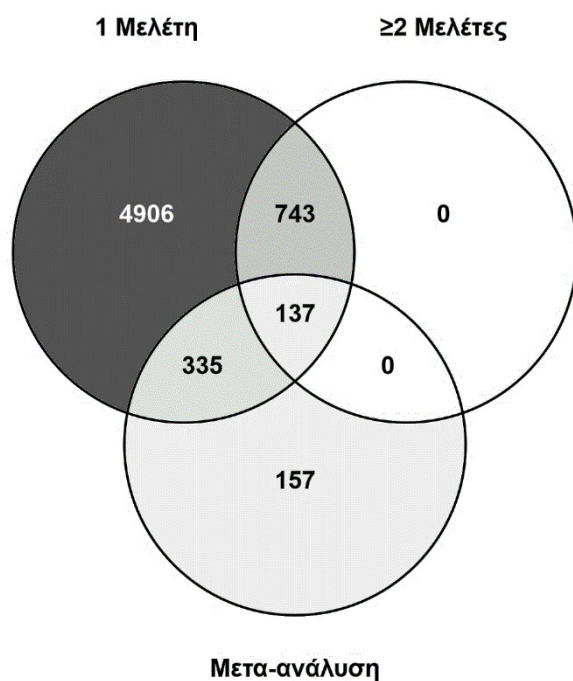
^δΔΑΠ = Διαστολική αρτηριακή πίεση

^αGSE = GEO αριθμός πρόσβασης

^βPE = Προεκλαμψία

ΔΥΠ^γ Δεν υπήρχαν οι πληροφορίες

Διεξήχθη μετα-ανάλυση τυχαίων επιδράσεων ακολουθούμενη από πολλαπλούς ελέγχους μεθόδων διόρθωσης (Bonferoni, Sidak, Holm και Holland) προκειμένου να περιοριστεί ο αριθμός των ψευδώς θετικών αποτελεσμάτων (Πίνακας Παραρτήματος 2). Βρέθηκαν 629 στατιστικώς σημαντικά ΔΕΓ σε ασθενείς με προεκλαμψία σε σύγκριση με υγιή άτομα, σε FDR 1%. Το IDR βρέθηκε ίσο με 0,234 κάτι που υποδεικνύει ότι το 23,4% των ΔΕΓ προσδιορίστηκε μόνο μέσω της μετα-ανάλυσης και όχι από τις πρωτογενείς μελέτες (δηλαδή 147 γονίδια από τα συνολικά 629), απεικονίζοντας την ευρωστία και την αποτελεσματικότητα της μετα-ανάλυσης. Η τιμή του IRR ήταν 0,92, υποδεικνύοντας ότι τα αναγνωρισμένα ΔΕΓ στις μεμονωμένες μελέτες αλλά όχι στη μετα-ανάλυση ήταν 92,1% περισσότερα (5.640 από 6.122 γονίδια σε όλες τις μεμονωμένες μελέτες) (Εικόνα 21). Αυξημένη τιμή IRR είναι κάτι που αναμένεται σε τέτοιες αναλύσεις, καθώς η μεταβλητότητα και η μη αναπαραγωγιμότητα είναι αυξημένες σε δεδομένα μικροσυστοιχιών.



Εικόνα 21 Διάγραμμα Venn σύγκρισης των ΔΕΓ που προέκυψαν από τις μελέτες μεμονωμένα και από τη μετα-ανάλυση. Τα αποτελέσματα που προέκυψαν από τη μετα-ανάλυση (629 ΔΕΓ) συγκρίνονται με τα ΔΕΓ τα οποία προσδιορίστηκαν σε τουλάχιστον μία μελέτη (6.121 ΔΕΓ) και σε τουλάχιστον δύο μελέτες (880 ΔΕΓ).

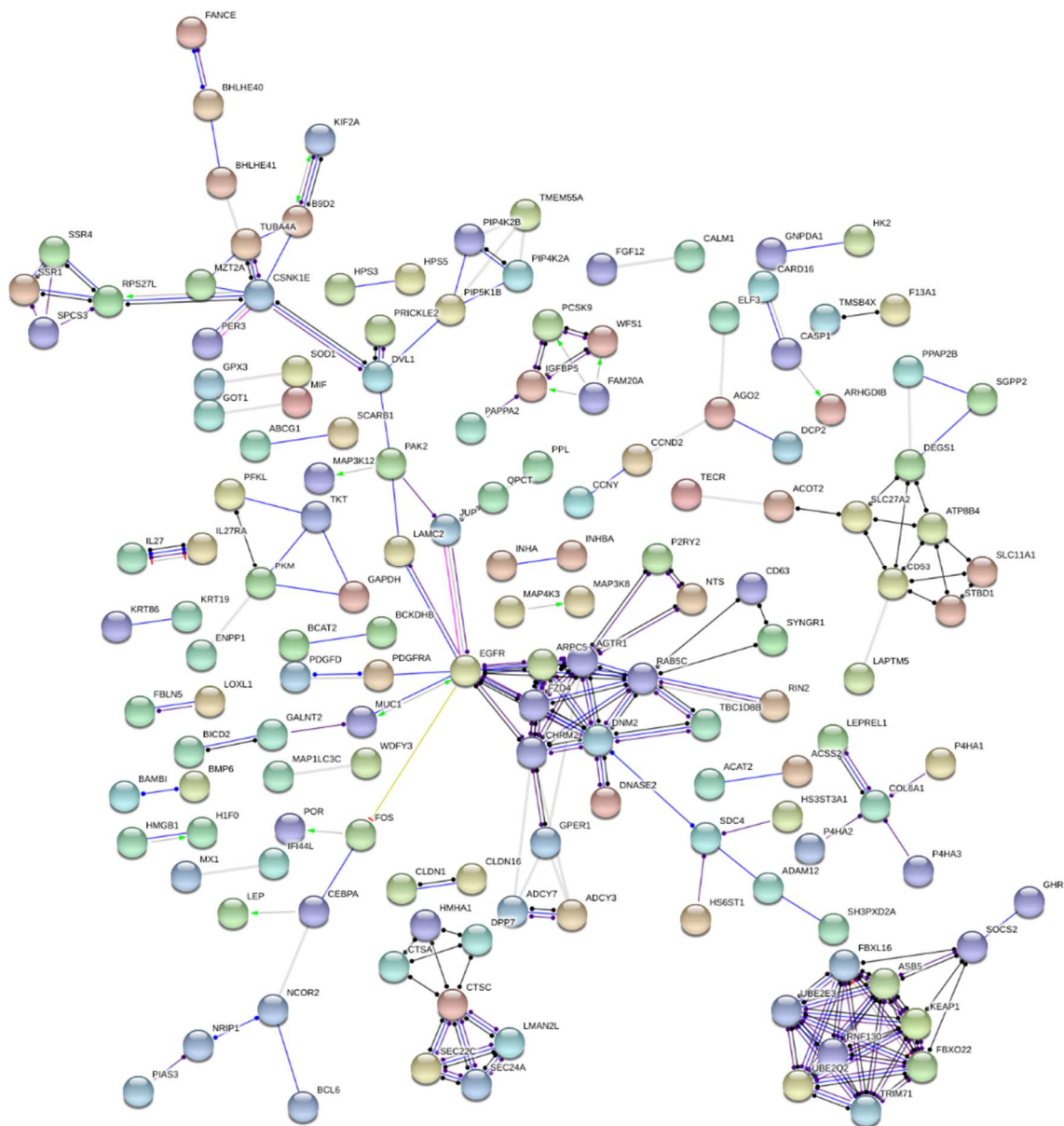
Τα 629 ΔΕΓ που βρέθηκαν στατιστικώς σημαντικά κατά τη μετα-ανάλυση μαζί με τις *p*-values τους παρουσιάζονται στον Πίνακα Παραρτήματος 3. Συγκεκριμένα, από τα 629 ΔΕΓ 336 υπερεκφράστηκαν και 293 υποεκφράστηκαν (Πίνακας Παραρτήματος 3). Αν και τα

υπερεκφραζόμενα γονίδια είναι ελαφρώς πιο πολλά, δεν μπορούμε να εξαγάγουμε συμπεράσματα για γενικευμένη καταστολή των γονιδίων.

Οι πιθανές αλληλεπιδράσεις μεταξύ των 629 ΔΕΓ διερευνήθηκαν και απεικονίστηκαν με τη χρήση της διαδικτυακής βάσης δεδομένων STRING. Η STRING αναγνώρισε 525 γονίδια από τα συνολικά 629 ΔΕΓ. Από αυτά μόνο 149 ήταν πολύ συνδεδεμένα (με τιμή ορίου αλληλεπίδρασης 0,9) (Εικόνα 22). Το γεγονός ότι υπάρχουν ξέχωροι κόμβοι γονιδιακών προϊόντων στο δίκτυο υποδηλώνει ότι διαφορετικά γονιδιακά μονοπάτια εμπλέκονται στην παθογένεση της προεκλαμψίας. Η εμφάνιση του EGFR και σε μικρότερη έκταση του CTSC ως κόμβοι υποδεικνύει ότι τα μονοπάτια μεταγωγής σήματος βρίσκονται να ρυθμίζουν τη μεσολαβούμενη από κυστίδια μεταφορά. Τα RPS27L, CSNK1E, DVL1 και IGFBP5 συμμετέχουν στην απόκριση της εξωκυτταρικής σηματοδότησης μέσω μονοπατιών μεταγωγής σήματος ή ελέγχου της μετάφρασης. Σημαντικό εύρημα είναι ότι οι κόμβοι ATP8B4 και RNF130 που υποδεικνύουν ότι τα μονοπάτια της φυσικής ανοσίας τα οποία σχετίζονται είτε με τη μεταφορά ιόντων είτε με την παρουσίαση αντιγόνου σχετίζονται με την ανάπτυξη προεκλαμψίας.

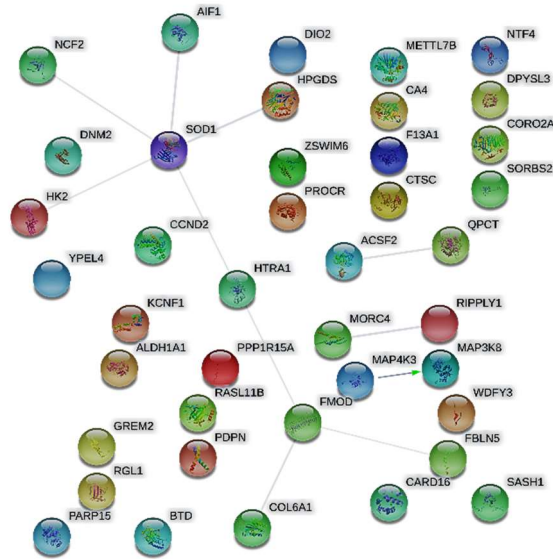
Προσπαθήσαμε να δημιουργήσουμε κι ένα δίκτυο αλληλεπίδρασης για τα κορυφαία 20 γονίδια (από άποψη *p*-values) αλλά δεν παρουσιάστηκε καμία σύνδεση, ενώ για τα κορυφαία 50 γονίδια σχηματίστηκε μόνο ένα μικρό δίκτυο εννέα γονιδίων (Εικόνα 23). Το κύριο δίκτυο φάνηκε να αποτελείται από γονίδια τα οποία συμμετέχουν κυρίως στην ανοσοαπόκριση.

Λόγω της μεγάλης ανομοιογένειας στην αιτιολογία της προεκλαμψίας και των μη συνδεδεμένων γονιδιακών υποδικτύων, διερευνήσαμε εάν αυτή η ετερογένεια οφειλόταν στα ξεχωριστά αίτια μεταξύ σοβαρής και ήπιας προεκλαμψίας (τα ΔΕΓ για κάθε υπότυπο υπάρχουν αναλυτικά στους Πίνακες Παραρτήματος 4 και 5). Η στρωματοποιημένη μετα-ανάλυση σε σχέση με τους υποτύπους οδήγησε σε δεκαοκτώ γονίδια κοινά μεταξύ ήπιας και σοβαρής προεκλαμψίας (Εικόνα 24). Σημαντικότερα, 216 από τα 365 ΔΕΓ στην ήπια προεκλαμψία βρέθηκαν να είναι στατιστικώς σημαντικά τόσο στην ήπια όσο και στην συνολική προεκλαμψία. Από τα 287 ΔΕΓ στην σοβαρή προεκλαμψία μόνο 89 ήταν κοινά με τα ΔΕΓ που ανακαλύφθηκαν στην συνολική προεκλαμψία. Αξιοσημείωτο ήταν ότι η συνολική μετα-ανάλυση αποκάλυψε 306 επιπρόσθετα στατιστικώς σημαντικά ΔΕΓ που δεν εντοπίστηκαν στις δύο υποομάδες.

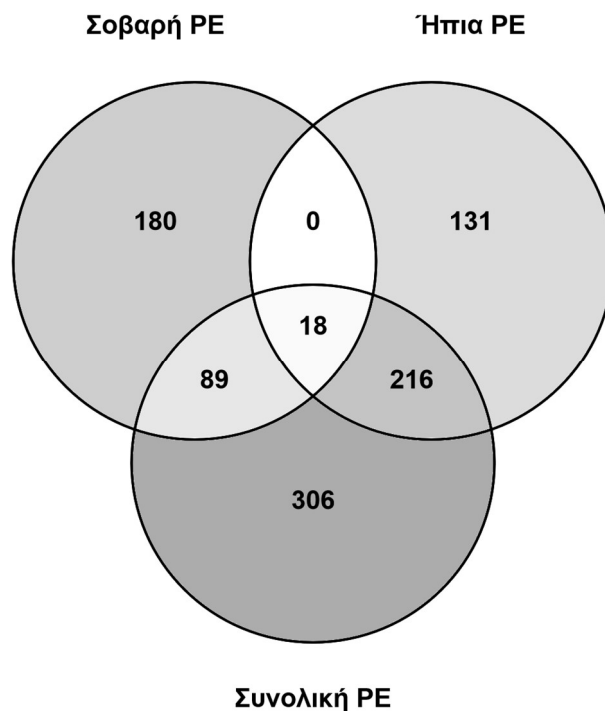


Εικόνα 22 Δίκτυο συσχέτισης γονιδίων/πρωτεϊνών των 525 ΔΕΓ (από τα 629) για την προεκλαμμία, σε FDR 1%, με τιμή ορίου αλληλεπίδρασης 0,9.

Η τιμή δεν υποδεικνύει τη δύναμη ή την ιδιαιτερότητα της αλληλεπίδρασης, αλλά μάλλον αποτελεί δείκτη εμπιστοσύνης. Το όριο του 0,9 αντιπροσωπεύει πολύ υψηλή εμπιστοσύνη. Γραμμές διαφορετικών χρωμάτων υποδεικνύουν τους προβλεπόμενους τρόπους δράσης. Οι πράσινες γραμμές αντιπροσωπεύουν την ενεργοποίηση, οι μπλε τη δέσμευση, οι γαλάζιες τον φαινόμενο, οι μαύρες αντίδραση, οι κόκκινες αναστολή, οι μωβ κατάλυση, οι ροζ τροποποιήσεις και οι κίτρινες μεταγραφικές ρυθμίσεις. Οι γραμμές που καταλήγουν σε βέλος υποδεικνύουν θετική αλληλεπίδραση, σε ορθογώνιο υποδεικνύουν αρνητική αλληλεπίδραση και σε κύκλο υποδεικνύουν μη καθορισμένη αλληλεπίδραση. Το δίκτυο κατασκευάστηκε χρησιμοποιώντας την βάση STRING. Οι ασύνδετοι κόμβοι του δικτύου δεν εμφανίζονται στο σχήμα.



Εικόνα 23 Δίκτυο συσχέτισης γονιδίων/πρωτεϊνών των τοπ 50 (41 αναγνωρίστηκαν) ΔΕΓ τα οποία αναγνωρίστηκαν για την συνολική προεκλαμψία, με τιμή ορίου αλληλεπίδρασης 0,4. Η τιμή δεν υποδεικνύει τη δύναμη ή την ιδιαιτερότητα της αλληλεπίδρασης, αλλά μάλλον αποτελεί δείκτη εμπιστοσύνης. Το όριο του 0,4 αντιπροσωπεύει μέτρια εμπιστοσύνη. Γραμμές διαφορετικών χρωμάτων υποδεικνύουν τους προβλεπόμενους τρόπους δράσης. Οι πράσινες γραμμές αντιπροσωπεύουν την ενεργοποίηση, οι μπλε τη δέσμευση, οι γαλάζιες τον φαινότυπο, οι μαύρες αντίδραση, οι κόκκινες αναστολή, οι μωβ κατάλυση, οι ροζ τροποποιήσεις και οι κίτρινες μεταγραφικές ρυθμίσεις. Οι γραμμές που καταλήγουν σε βέλος υποδεικνύουν θετική αλληλεπίδραση, σε ορθογώνιο υποδεικνύουν αρνητική αλληλεπίδραση και σε κύκλο υποδεικνύουν μη καθορισμένη αλληλεπίδραση. Το δίκτυο κατασκευάστηκε χρησιμοποιώντας την βάση STRING. Οι ασύνδετοι κόμβοι του δικτύου δεν εμφανίζονται στο σχήμα.



Εικόνα 24 Διάγραμμα Venn της σύγκρισης των ΔΕΓ που προέκυψαν από τη μετα-ανάλυση της συνολικής προεκλαμψίας και των υποτύπων της.

Τα δεκαοκτώ κοινά ΔΕΓ στην ήπια και σοβαρή προεκλαμψία συμμετέχουν: α) στην κατανάλωση ενέργειας (ACSF2, APMAP, BTM, HK2 και NR2F1), β) σε αναπτυξιακές διεργασίες (DNM2, GREM2, CA4, PPP1R15A και HTRA1), γ) στη ρύθμιση της ανοσολογικής απόκρισης (HTRA1, CTSC, PROCR, SASH1 και RASL11B), δ) σε διαδικασίες διατήρησης της εγκυμοσύνης (CGB3 και MORC4), ε) στη ρύθμιση της απόπτωσης (SOD1) είτε ζ) είναι μεταγραφικοί παράγοντες (MORC4 και ZNF555).

Τα 629 ΔΕΓ εξετάστηκαν περαιτέρω για τις βιολογικές διεργασίες και τα βιοχημικά μονοπάτια στα οποία εμπλέκονται με τη χρήση της διαδικτυακής βάσης δεδομένων PANTHER. Η PANTHER αναγνώρισε 545 από τα συνολικά 629 ΔΕΓ. Στη PANTHER χρησιμοποιήθηκε το εργαλείο το οποίο πραγματοποιεί ORA (αναλύσεις υπερ-εκπροσώπησης, Over-representation analysis) (Πίνακας 4). Το μονοπάτι μεταγωγής σήματος εμφανίστηκε σημαντικό (εννέα γονίδια) για τα 629 ΔΕΓ. Ο μεταβολισμός των υδατανθράκων και το μονοπάτι κατανάλωσης της ενέργειας εμφανίστηκαν επίσης σημαντικά με επτά γονίδια (GNPDA1, NPL, HK2, TKT, PFKL, GAPDH και PKM)· αυτά τα μονοπάτια δεν έχουν διερευνηθεί αρκετά στη βιβλιογραφία. Τρία γονίδια (NT5E, DPYSL3 και DPYD) που εμπλέκονται στο μεταβολισμό της πυριμιδίνης και δύο γονίδια (BCAT2 και ILVBL) τα οποία εμπλέκονται στον μεταβολισμό των αμινοξέων, βρέθηκαν να είναι ΔΕΓ στην προεκλαμψία. Ενδιαφέρον είναι και το ότι βρήκαμε δύο γονίδια τα οποία εμπλέκονται στο σύστημα του κίρκαδικού ρολογιού (PER3 και CSNK1E) να είναι ΔΕΓ στην προεκλαμψία. Στα βιοχημικά μονοπάτια της PANTHER για τη συνολική προεκλαμψία, 15 ΔΕΓ υπερεκφράζονται και 9 υποεκφράζονται (Πίνακας 4).

Δεδομένου ότι η σοβαρή και η ήπια προεκλαμψία αποτελούν διακριτές οντότητες της ασθένειας, ερευνήσαμε επίσης τα εμπλουτισμένα βιοχημικά μονοπάτια σύμφωνα με την PANTHER για καθέναν από τους δύο υποτύπους της προεκλαμψίας, για τη διερεύνηση διακριτών μηχανισμών οι οποίοι να υποκρύπτουν τους δύο υποτύπους. Όπως φαίνεται στον Πίνακα 6 και 7, τα μονοπάτια του κίρκαδικού ρολογιού και του κυτταρικού κύκλου είναι σημαντικά στην ήπια αλλά όχι στην σοβαρή προεκλαμψία. Αντιστρόφως, υπάρχουν βιοχημικά μονοπάτια τα οποία βρέθηκαν σημαντικά στην σοβαρή αλλά όχι στην ήπια προεκλαμψία, όπως τα μονοπάτια του ορμονικού υποδοχέα απελευθέρωσης γοναδοτροπίνης (Gonadotropin-releasing hormone receptor, GnRHR) (με οκτώ γονίδια), της βιοσύνθεσης αμινοξέων και της αποικοδόμησης του αμινοβουτυρικού. Κοινά μονοπάτια μεταξύ ήπιας και σοβαρής προεκλαμψίας είναι αυτά τα οποία περιλαμβάνουν μονοπάτια αξιοποίησης ενέργειας και μεταγωγής σήματος. Όλα τα εμπλουτισμένα

βιοχημικά μονοπάτια που βρέθηκαν στους ξεχωριστούς υποτύπους της προεκλαμψίας βρέθηκαν και στη συνολική προεκλαμψία, εκτός από αυτά που σχετίζονται με το μονοπάτι της GnRHR και την αποικοδόμηση του αμινοβουτυρικού.

Πίνακας 4 Τα βιοχημικά μονοπάτια που συμμετέχουν τα στατιστικώς σημαντικά ΔΕΓ όπως προέκυψαν από την PANTHER, για την συνολική προεκλαμψία

PANTHER βιοχημικό μονοπάτι	Γονίδιο	p-value
Συνολική προεκλαμψία		
Γλυκόλυσης	HK2 (↑), PFKL (↑), GAPDH (↑), PKM (↑)	0.00288
Μεταβολισμού της Πυριμιδίνης	DPYSL3 (↓), NT5E (↑), DPYD (↓)	0.00381
Βιοσύνθεσης της Βαλίνης	BCAT2 (↑), ILVBL (↑)	0.00607
Βιοσύνθεσης της Ισολευκίνης	BCAT2 (↑), ILVBL (↑)	0.00896
Μεταβολισμού της N-ακετυλογλυκοζαμίνης	GNPDA1 (↓), NPL (↓)	0.0204
Μονοπάτι της φωσφορικής πεντόζης	HK2 (↑), TKT (↓)	0.0251
Συστήματος Κιρκαδικού ρολογιού	PER3 (↓), CSNK1E (↑)	0.0302
Ετεροτριμερές G-πρωτεΐνης μονοπάτι μεταγωγής σήματος Gi alpha και Gs alpha	CALM1 (↑), SSR4 (↑), ADCY7 (↓), GRK3 (↑), ADCY3 (↓), SSR1 (↑), RGS10 (↓), KCNJ6 (↑), CHRM2 (↑)	0.0471
Μεταβολισμού της φρουκτόζης γαλακτόζης	HK2 (↑), GALK1 (↑)	0.0476

↓ υποεκφράζεται, ↑ υπερεκφράζεται

Πίνακας 5 Τα βιοχημικά μονοπάτια που συμμετέχουν τα στατιστικώς σημαντικά ΔΕΓ όπως προέκυψαν από την PANTHER, για την ήπια προεκλαμψία

PANTHER βιοχημικό μονοπάτι	Γονίδιο	p-value
Ήπια Προεκλαμψία		
Γλυκόλυσης	HK2 (↑), PFKL (↑), PKM (↑)	0.00291
Συστήματος Κιρκαδικού ρολογιού	PER3 (↓), CSNK1E (↑)	0.00807
Ετεροτριμερές G-πρωτεΐνης μονοπάτι μεταγωγής σήματος Gq alpha και Go alpha	SSR4 (↑), PRKCB (↓), CACNA1A (↑), RGS10 (↓), KCNJ6 (↑)	0.026
Κυτταρικού Κύκλου	CCND2 (↓), APC2 (↑)	0.0336
Σηματοδοτικό μονοπάτι Toll υποδοχέα	TLR8 (↓), TLR6 (↓), MAP3K8 (↑)	0.0384
Χρήσης του οξικού	ACSS2 (↓)	0.0495

↓ υποεκφράζεται, ↑ υπερεκφράζεται

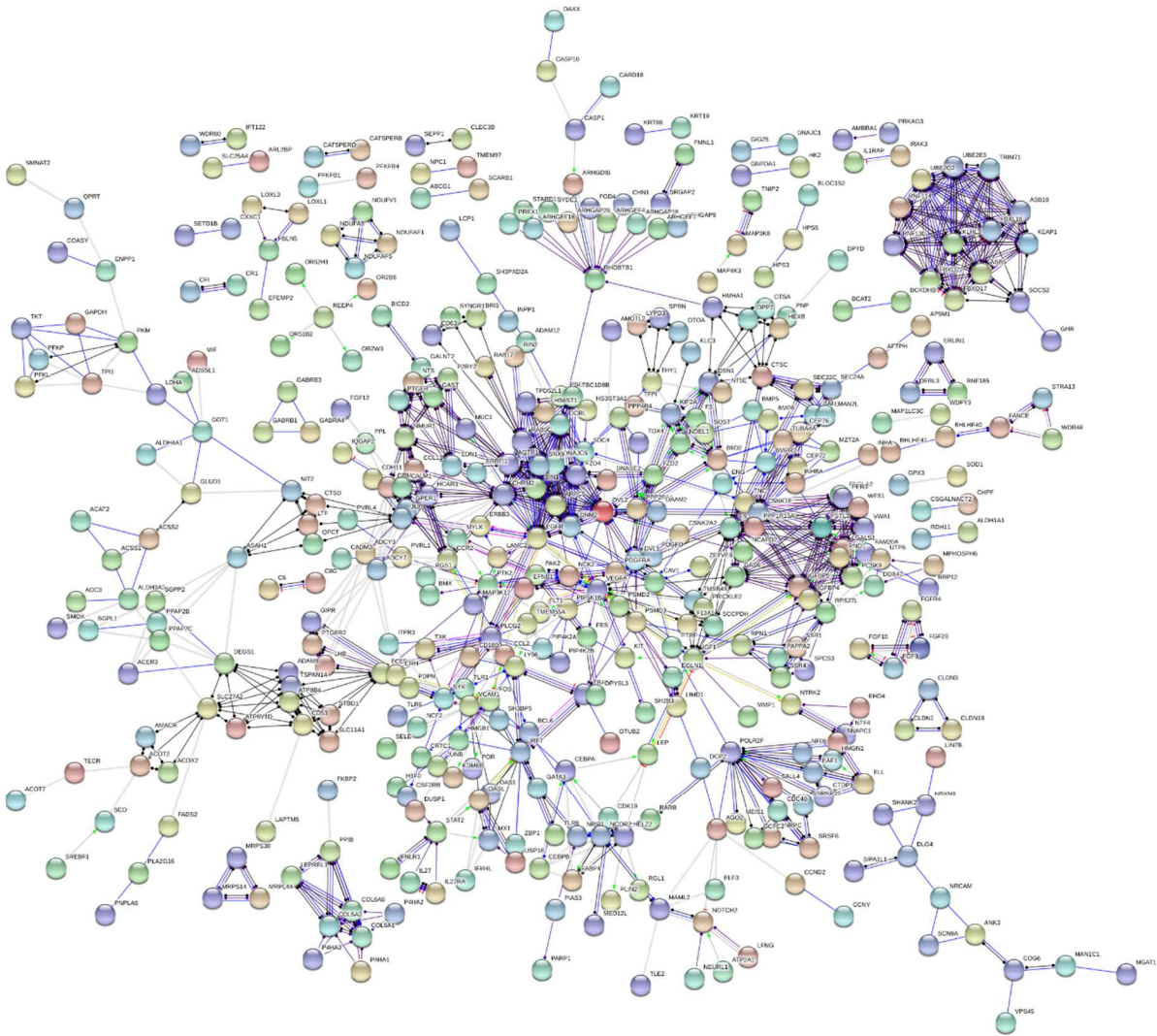
Πίνακας 6 Τα βιοχημικά μονοπάτια που συμμετέχουν τα στατιστικώς σημαντικά ΔΕΓ όπως προέκυψαν από την PANTHER, για τη σοβαρή προεκλαμψία

PANTHER βιοχημικό μονοπάτι	Γονίδιο	p-value
Σοβαρή Προεκλαμψία		
TGF-β σηματοδοτικό μονοπάτι	SMAD6 (↑), FKBP1C (↓), FKBP2 (↑), INHBA (↑), TGFB1 (↑), BMP5 (↓), JUNB (↑)	0.000629
Γλυκόλυσης	HK2 (↑), GAPDH (↑), TPI1 (↑)	0.00383
Μονοπάτι της φωσφορικής πεντόζης	HK2 (↑), TKT (↓)	0.00805
Μεταβολισμού του Πυροσταφυλικού	PC (↓), PDHA1 (↓)	0.0136
Μεταβολισμού της φρουκτόζης γαλακτόζης	GALE (↑), HK2 (↑)	0.0157
Ορμονικού υποδοχέα απελευθέρωσης γοναδοτροπίνης	INHA (↑), SLC2A1 (↑), DUSP1 (↑), INHBA (↑), LHB (↑), TGFB1 (↑), JUNB (↑), EGFR (↑)	0.0186
Μονοπάτι νόσου Alzheimer - πρεσενιλίνης	MME (↑), JUP (↑), FZD10 (↑), MMP1 (↓), CDH1 (↓)	0.0363
Αποικοδόμησης Αμινοβουτυρικού	ALDH5A1 (↓)	0.0411
Βιοσύνθεσης Αλανίνης	BCAT2 (↑)	0.0411
Βιοσύνθεσης Λευκίνης	BCAT2 (↑)	0.0411

↓ υποεκφράζεται, ↑ υπερεκφράζεται

Για να αποσαφηνιστεί εάν το αυστηρό όριο του FDR 1% θα μπορούσε πιθανώς να χάσει κάποια αληθώς θετικά ΔΕΓ, πραγματοποιήσαμε μια νέα μετα-ανάλυση με λιγότερο αυστηρό FDR του 5% (Πίνακας Παραρτήματος 6) και προχωρήσαμε επίσης σε ανάλυση εμπλουτισμού. Βρήκαμε 1.342 ΔΕΓ τα οποία σχημάτιζαν ένα περισσότερο διασυνδεδεμένο δίκτυο σε σύγκριση με αυτό που σχηματίστηκε σε FDR 1% (Εικόνα 24). Τέσσερα πρόσθετα βιοχημικά μονοπάτια βρέθηκαν, η βιοσύνθεση των ορμονών του φύλου, η αγγειογένεση και τα μονοπάτια

σηματοδότησης Notch, ενώ το μονοπάτι της GnRH βρέθηκε σε αυτή την περίπτωση και στην ολική προεκλαμψία και όχι μόνο στην σοβαρή (Εικόνα 24 και Πίνακας 7).



Εικόνα 25 Δίκτυο συσχέτισης γονιδίων/πρωτεϊνών των 1.071 ΔΕΓ (από τα 1.342) για την προεκλαμψία, σε FDR 5%, με τιμή ορίου αλληλεπίδρασης 0,9.

Η τιμή δεν υποδεικνύει τη δύναμη ή την ιδιαιτερότητα της αλληλεπίδρασης, αλλά μάλλον αποτελεί δείκτη εμπιστοσύνης. Το όριο του 0,9 αντιπροσωπεύει πολύ υψηλή εμπιστοσύνη. Γραμμές διαφορετικών χρωμάτων υποδεικνύουν τους προβλεπόμενους τρόπους δράσης. Οι πράσινες γραμμές αντιπροσωπεύουν την ενεργοποίηση, οι μπλε τη δέσμευση, οι γαλάζιες τον φαινότυπο, οι μαύρες αντίδραση, οι κόκκινες αναστολή, οι μωβ κατάλυση, οι ροζ τροποποιήσεις και οι κίτρινες μεταγραφικές ρυθμίσεις. Οι γραμμές που καταλήγουν σε βέλος υποδεικνύουν θετική αλληλεπίδραση, σε ορθογώνιο υποδεικνύουν αρνητική αλληλεπίδραση και σε κύκλο υποδεικνύουν μη καθορισμένη αλληλεπίδραση. Το δίκτυο κατασκευάστηκε χρησιμοποιώντας την βάση STRING. Οι ασύνδετοι κόμβοι του δικτύου δεν εμφανίζονται στο σχήμα.

Πίνακας 7 Τα βιοχημικά μονοπάτια που συμμετέχουν τα στατιστικώς σημαντικά ΔΕΓ όπως προέκυψαν από την PANTHER, για την προεκλαμψία σε FDR 5%

PANTHER βιοχημικό μονοπάτι	Γονίδιο	p-value
Ορμονικού υποδοχέα απελευθέρωσης γοναδοτροπίνης	PTGFR (↓), PTK2 (↓), INHA (↑), PTGER2 (↓), SLC2A1 (↑), MAP3K12 (↓), DUSP1 (↑), BMPR2 (↑), INHBA (↑), ITPR3 (↑), CRTC1 (↑), IGF1 (↓), LHB (↑), FOS (↑), BMP6 (↓), MAP3K8 (↑), MAP4K3 (↓), MIF (↑), SDF4 (↑), JUNB (↑), PTGER2 (↓), EGFR (↑), CAV1 (↓)	0.0055
Γλυκόλυσης	HK2 (↑), PFKL (↑), GAPDH (↑), TPI1 (↑), PKM (↑)	0.00723
Αγγειογένεσης	PRKD3 (↑), PTK2 (↓), DVL1 (↑), PLCG2 (↓), DLK1 (↓), EFNB1 (↑), FZD2 (↓), FOS (↑), DVL1 (↑), PDGFD (↓), PAK2 (↓), DVL2 (↑), F3 (↓), FRZB (↓), PDGFRA (↓), VEGFA (↑), NOTCH2 (↓), NCK2 (↓)	0.011
Σηματοδοτικό μονοπάτι Toll υποδοχέα	LY96 (↓), TLR6 (↓), TLR8 (↓), IRF7 (↑), MAP3K8 (↑), TLR1 (↓), TLR9 (↑), IRAK3 (↓)	0.0143
Βιοσύνθεσης της Βαλίνης	BCAT2 (↑), ILVBL (↑)	0.0227
Χρήσης του οξικού	ACSS1 (↓), ACSS2 (↓)	0.0227
Μεταβολισμού της Πυριμιδίνης	DPYSL3 (↓), NT5E (↑), DPYD (↓)	0.0247
Βιοσύνθεσης της Ισολευκίνης	BCAT2 (↑), ILVBL (↑)	0.033
Μονοπάτι σηματοδότησης Notch	LFNG (↑), NEURL1 (↑), DLK1 (↓), LNX2 (↓), NCOR2 (↑), NOTCH2 (↓)	0.035
Βιοσύνθεσης	ACAT2 (↓), SOAT1 (↓), HSD17B1 (↑)	0.0365
Ανδογόνων/Οιστρογόνων/Προγεστερόνης		
Μεταβολισμού της Πουρίνης	NT5E (↑), PNP (↓)	0.0446

↓ υποεκφράζεται, ↑ υπερεκφράζεται

Τα 24 γονίδια (Πίνακας 4) τα οποία συμμετέχουν στα εμπλουτισμένα μονοπάτια της PANTHER στην ολική προεκλαμψία διερευνήθηκαν για τη βιολογική τους λειτουργία στη βάση δεδομένων MGI σε διαγονιδιακά knockout μοντέλα ποντικού. Διαπιστώσαμε ότι η διαγραφή

ορισμένων γονιδίων οδήγησε στη θνησιμότητα σε διαφορετικά στάδια ζωής των αναπτυσσόμενων ποντικών. Τα GNPD (Smith, Hayamizu et al. 2019), HK2 (Heikkinen, Pietila et al. 1999) και TKT (Xu, Wawrousek et al. 2002) οδήγησαν σε εμβρυϊκή θνησιμότητα. Τα ADCY3, ADCY7 και RGS10 οδήγησαν σε νεογνική θνησιμότητα (Smith, Hayamizu et al. 2019). Τέλος, το PFKL και το PKM οδήγησαν σε πρόωρη θνησιμότητα (Smith, Hayamizu et al. 2019). Η διαγραφή των CHR2 (Smith, Hayamizu et al. 2019) και NT5E (Koszalka, Ozuyaman et al. 2004, Zukowska, Kutryb-Zajac et al. 2017) προκάλεσε προβλήματα στο καρδιαγγειακό σύστημα, μια κατάσταση η οποία σχετίζεται με την υπέρταση. Τα knockout ποντίκια για τα γονίδια ADCY7 (Duan, Davis et al. 2010) και RGS10 (Yang and Li 2007) είχαν προβλήματα στο ανοσοποιητικό τους σύστημα, ένα εύρημα που επιβεβαιώνει την άποψη που υποστηρίζεται από άλλους (Chappell and Morgan 2006, Grill, Rusterholz et al. 2009, El-Sayed 2017) και από τα αποτελέσματά μας ότι η μη ρυθμισμένη ανοσοαπόκριση εμπλέκεται στην ανάπτυξη της προεκλαμψίας. Η διαγραφή του GNPD1 οδήγησε σε μη φυσιολογική μορφολογία της μήτρας (Smith, Hayamizu et al. 2019) ένα αποτέλεσμα το οποίο είναι σύμφωνο με τα συμπεράσματά μας, ότι δηλαδή η μη φυσιολογική έκφραση των γονιδίων που εμπλέκονται στις αναπτυξιακές διεργασίες συνδέεται με την εμφάνιση προεκλαμψίας.

Η συγκεκριμένη μελέτη δεν είναι η πρώτη προσπάθεια να παρουσιαστεί μια τέτοια ανάλυση, καθώς αρκετές μελέτες έχουν δοκιμάσει στο παρελθόν να συνδυάσουν, με μετα-ανάλυση, δεδομένα μικροσυστοιχιών για την προεκλαμψία. Η παρούσα μετα-ανάλυση πλεονεκτεί σε αρκετές πτυχές σε σχέση με τις προηγούμενες μελέτες (Πίνακας 8). Πρώτον, οι παλαιότερες μελέτες (Moslehi, Mills et al. 2013, Kawasaki, Kondoh et al. 2015, Sitras, Fenton et al. 2015, van Uitert, Moerland et al. 2015, Brew, Sullivan et al. 2016, Vaiman and Miralles 2016) περιέχουν λιγότερα σύνολα δεδομένων με λιγότερα άτομα ελέγχου συγκριτικά με την παρούσα εργασία (3-7 συμπεριλαμβανόμενες μελέτες έναντι 10 σε αυτή την εργασία), επομένως η τρέχουσα μελέτη έχει μεγαλύτερη στατιστική ισχύ. Δεύτερον, εφαρμόσαμε αυστηρά κριτήρια συμπερίληψης χρησιμοποιώντας μόνο δείγματα πλακούντα. Τρίτον, χρησιμοποιήσαμε μια αυστηρή μεθοδολογία μετα-ανάλυσης εφαρμόζοντας τη διόρθωση του Hedges και χρησιμοποιήσαμε μια μετα-ανάλυση τυχαίων επιδράσεων, με διάφορες μεθόδους διόρθωσης και χαμηλή τιμή FDR (1%) προκειμένου να προσδιοριστούν οι πιο αξιόπιστες μετρήσεις. Τέταρτον, χρησιμοποιήσαμε σύγχρονα εργαλεία βιοπληροφορικής (STRING και PANTHER) για να αποκαλύψουμε κυρίως τα ισχυρά γονιδιακά μονοπάτια. Οι προηγουμένως δημοσιευμένες μετα-αναλύσεις, χρησιμοποιώντας μια ποικιλία

μεθόδων, κατέδειξαν αρκετά γενικά βιοχημικά μονοπάτια να συνδέονται με την προεκλαμψία, συμπεριλαμβανομένων αυτά της φλεγμονής, της αγγειογένεσης και του μεταβολισμού των αμινοξέων (van Uiter, Moerland et al. 2015, Brew, Sullivan et al. 2016, Vaiman and Miralles 2016). Αν και η προεκλαμψία δεν μπορεί να αποδοθεί σε ένα μόνο βιοχημικό μονοπάτι, η πρόταση από τις προηγούμενες μετα-αναλύσεις πως ένα πλήθος ΔΕΓ και μονοπατιών συνδέονται με τους προεκλαμπτικούς πλακούντες είναι φτωχής διαγνωστικής αξίας καθώς το σύνολο όλων αυτών των μονοπατιών δεν μπορεί να είναι ούτε συγκεκριμένο ούτε χαρακτηριστικό της προεκλαμψίας. Αντίθετα εμείς, εντοπίσαμε λίγα μονοπάτια και υπογραμμίσαμε ότι το μονοπάτι αξιοποίησης της ενέργειας χρειάζεται περαιτέρω έρευνα.

Εν κατακλείδι τα ευρήματά μας, συνολικά, με τη χρήση βιοπληροφορικών μεθοδολογιών και εργαλείων, προτείνουν σθεναρά τον μεταβολισμό των υδατανθράκων και την αξιοποίηση της ενέργειας ως νέα μονοπάτια τα οποία μπορεί να απορυθμίζονται κατά την εμφάνιση της προεκλαμψίας. Αν και τα παραπάνω μονοπάτια πιθανώς να φαίνονται μάλλον μη σχετικά με τους έως τώρα γνωστούς βιολογικούς δείκτες [EGF, PLGF, sFLT-1, προ-φλεγμονώδεις κυτοκίνες, απόπτωση, γονίδια επιβίωσης και γονίδια διαφοροποίησης (Grill, Rusterholz et al. 2009, El-Sayed 2017)], χρειάζεται να λάβουμε υπόψη ότι οι γνωστοί βιοδείκτες ελέγχονται σε δείγματα αίματος της μητέρας. Αντιθέτως τα αποτελέσματά της έρευνάς μας αφορούν πλακουντιακό ιστό και για τον λόγο αυτό θεωρούμε ότι χρειάζεται περαιτέρω πειραματική έρευνα για τη μοριακή λειτουργία των γονιδίων που σχετίζονται με τη χρήση ενέργειας στον πλακούντα, επειδή θα μπορούσε άμεσα να καταδείξει εάν υπάρχει απορυθμισμένη έκφραση γονιδίων προτού αυτά ασκήσουν τις συστημικές τους δράσεις. Ωστόσο, μελλοντική ανάλυση των λειτουργιών για τα υπόλοιπα γονίδια/μονοπάτια, τα οποία βρέθηκαν με τις μεθόδους εμπλουτισμού, θα ήταν απαραίτητη για την επαλήθευση των αποτελεσμάτων μας και για την ανακάλυψη περισσότερων μηχανισμών στους οποίους υπόκειται η εμφάνιση της προεκλαμψίας.

Πίνακας 8 Σύγκριση μεταξύ προηγούμενων μετα-αναλύσεων και της συγκεκριμένης μελέτης

Μελέτη	Έτος	Συμπεριλαμβανόμενες Μελέτες	Πλήθος ασθενών ατόμων	Πλήθος υγιών ατόμων	ΔΕΓ	Πλήθος Βιοχημικών Μονοπατιών
Παρούσα μετα-ανάλυση	2018	GSE4707, GSE10588, GSE14722, GSE24129, GSE25906, GSE30186, GSE54618, GSE74341, GSE35574, GSE75010	195	231	629	9 PANTHER μονοπάτια (με χρήση PANTHER ORA)
(Vaiman and Miralles 2016)	2016	GSE10588, GSE14722, GSE24129, GSE25906, GSE54618, GSE44711 (χοριακές λάχνες), GSE43942	85	109	369	28 (αναφερόταν στο άρθρο) KEGG& Wiki μονοπάτια (με χρήση WebGestat)
(Brew, Sullivan et al. 2016)	2016	GSE47187, GSE43942, GSE30186, GSE25906, GSE4707, GSE35574	68	99	9540	207 KEGG μονοπάτια (με χρήση WebGestat) δεν αναφερόταν όλα στο άρθρο

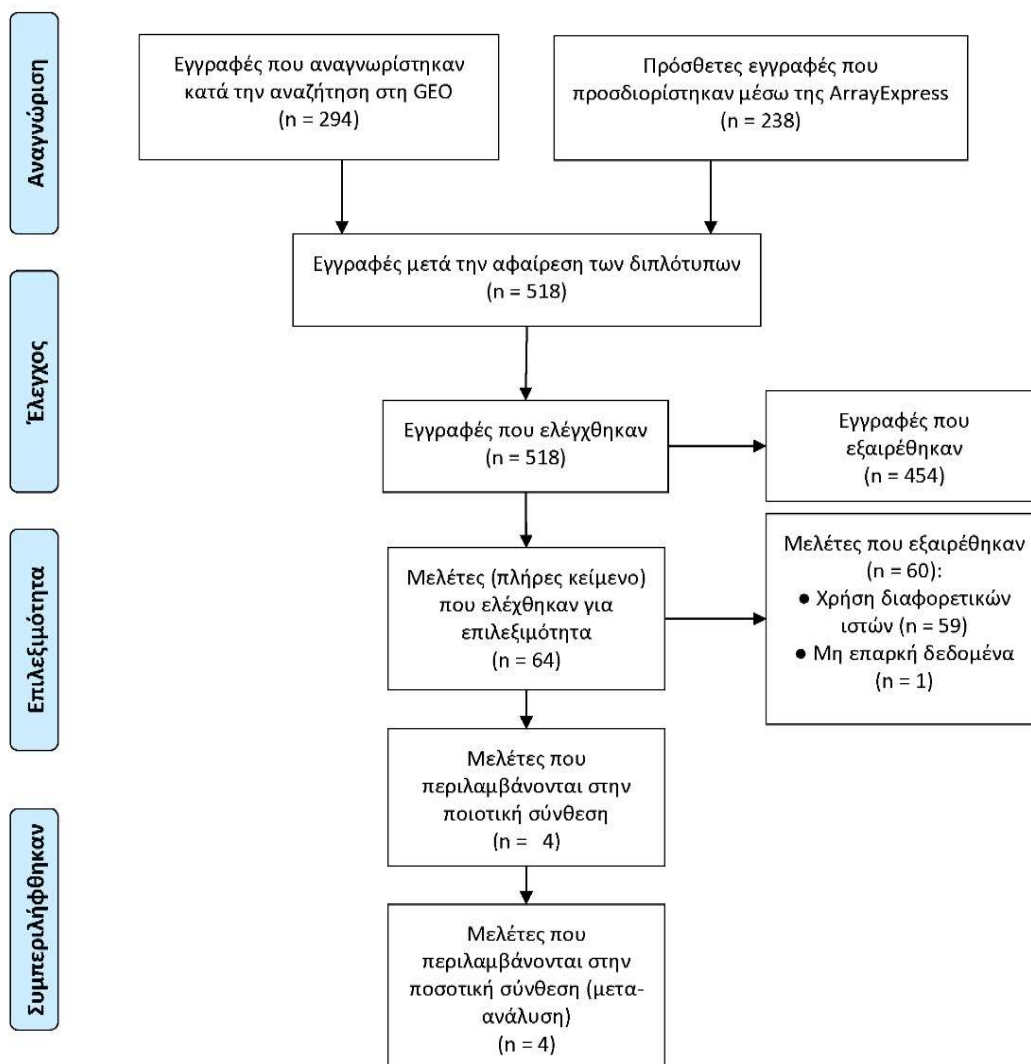
(van Uitert, Moerland et al. 2015)	2015	GSE5461, GSE30186, GSE4707, GSE24129, GSE10588, GSE25906, GSE14722, Author ¹ , Author ² , Author ³ , Author ⁴	116	139	388	71 KEGG μονοπάτια (αναφερόταν στο άρθρο, αλλά δεν υπήρχαν πληροφορίες, με χρήση Graphite)
(Sitras, Fenton et al. 2015)	2014	GSE4707, GSE10588, GSE12767, (χοριακές λάχνες), GSE25906, GSE30186	63	85	947	22 PANTHER μονοπάτια (αναφερόταν στο άρθρο)
(Kawasaki, Kondoh et al. 2015)	2014	GSE10588, GSE14722, GSE25906, GSE30186, GSE25906, GSE30186, GSE4711 (χοριακές λάχνες)	66	88	Δεν αναφερόταν στο άρθρο	21 αναφερόταν στο άρθρο, αλλά υπήρχαν πληροφορίες για 10
(Moslehi, Mills et al. 2013)	2013	GSE10588, GSE14722, GSE4707, GSE24129	47	49	419	8 (δεν αναφερόταν όλα στο άρθρο, με χρήση Biobase)

^{1, 2, 3, 4}: δεδομένα λήφθηκαν από το αρχείο των συγγραφέων μετά από επικοινωνία μαζί τους

3.3 Πολυμεταβλητή μετα-ανάλυση δεδομένων γονιδιακής έκφρασης της ιδιοπαθούς φλεγμονώδους νόσου του εντέρου

3.3.1 Αποτελέσματα

Ανακτήσαμε και εξετάσαμε, 518 σύνολα δεδομένων αθροιστικά από τις βάσεις δεδομένων GEO και ArrayExpress. Από αυτά, μόνο τέσσερα πληρούσαν τα κριτήρια επιλεξιμότητας και συμπεριλήφθηκαν στη μετα-ανάλυση που πραγματοποιήθηκε στη συνέχεια (Εικόνα 26). Τα σύνολα δεδομένων περιείχαν δεδομένα για 25.409 μοναδικά γονίδια/γονιδιακούς τόπους από 251 ασθενείς με νόσο του Crohn, 175 ασθενείς με ελκώδη κολίτιδα και 159 υγιή άτομα (Πίνακας 9).



Εικόνα 26 Διάγραμμα ροής της πολυμεταβλητής μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών για τις ασθένειες της ΙΦΝΕ.

Πίνακας 9 Χαρακτηριστικά των μελετών που χρησιμοποιήθηκαν στην πολυμεταβλητή μετα-ανάλυση

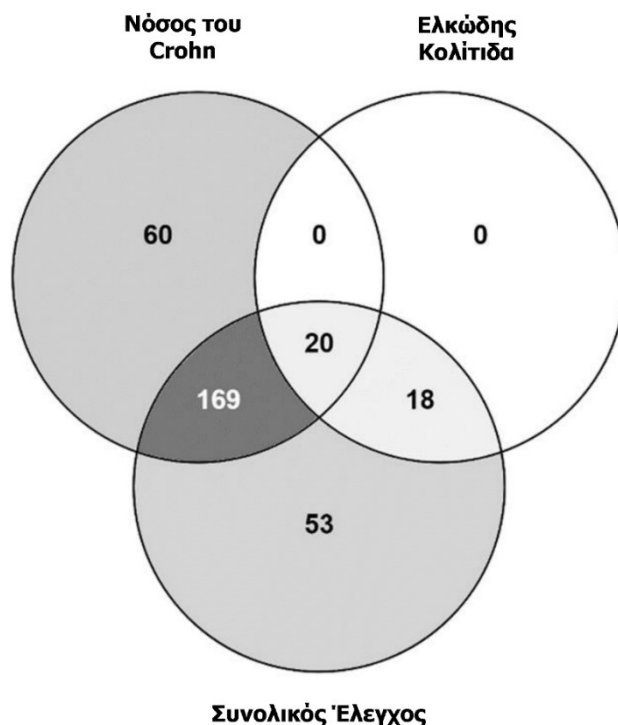
Αναφορά	Κωδικός GEO	Ασθενείς με νόσο του Crohn	Ασθενείς με ελκώδη κολίτιδα	Υγιή άτομα	Συνολικός αριθμός ανιχνευτών	Συνολικός αριθμός μοναδικών γονιδίων
(Buczynski, Peterson et al. 2006)	GSE3365	59	26	42	22283	12393
(Ventham, Kennedy et al. 2016)	GSE86434	94	79	48	46835	23619
(Planel, Masamunt et al. 2017)	GSE94648	50	25	22	19674	19358
(Ostrowski, Goryca et al. 2019)	GSE119600	48	45	47	47230	23774

Η μετα-ανάλυση εντόπισε 249 στατιστικώς σημαντικά ΔΕΓ σε ασθενείς με νόσο του Crohn και 38 σε ασθενείς με ελκώδη κολίτιδα, σε σύγκριση με υγιή άτομα, σε FDR 1%. 20 από τα ΔΕΓ ήταν κοινά και στις δύο ασθένειες.

Με τη χρήση του συνολικού ελέγχου (W), ο οποίος περιγράφηκε στη μεθοδολογία, προσδιορίστηκαν 260 στατιστικώς σημαντικά ΔΕΓ που απέρριπταν τη μηδενική υπόθεση ότι τα g_1 και g_2 είναι και τα δύο ίσα με το μηδέν. 53 από τα 260 στατιστικώς σημαντικά ΔΕΓ, δεν βρέθηκαν στατιστικώς σημαντικά σε καμία από τις δύο ασθένειες (Εικόνα 27) (τα ΔΕΓ υπάρχουν αναλυτικά στον Πίνακα Παραρτήματος 7).

Η τιμή του IDR για τη νόσο του Crohn ήταν 0,341, για την ελκώδη κολίτιδα ήταν 0,711 και για τον συνολικό έλεγχο ήταν 0,377. Αυτές οι τιμές υποδηλώνουν ότι το 34,1%, 71,1% και 37,7%, αντίστοιχα, των ΔΕΓ ταυτοποιήθηκαν μέσω της μετα-ανάλυσης αλλά όχι από τις πρωτογενείς μελέτες, επισημαίνοντας έτσι την ισχύ και την αποτελεσματικότητα της μετα-ανάλυσης. Οι τιμές του IRR όπως αναμενόταν ήταν υψηλότερες, 0,973 για τη νόσο του Crohn, 0,996 για την ελκώδη κολίτιδα και 0,974 για τον συνολικό έλεγχο, δείχνοντας ότι τα ΔΕΓ που βρέθηκαν στις μεμονωμένες μελέτες αλλά όχι στη μετα-ανάλυση ήταν 97,3%, 99,6% και 97,4% περισσότερα αντίστοιχα, υπογραμμίζοντας την αυστηρότητα της μετα-ανάλυσης στα αποτελέσματα και το χαμηλό ψευδώς θετικό ποσοστό. Ωστόσο, αυτό είναι κάτι αναμενόμενο σε

τέτοιες αναλύσεις, λαμβάνοντας υπόψη τη μεταβλητότητα και τη μη αναπαραγωγιμότητα των δεδομένων των μικροσυστοιχιών.



Εικόνα 27 Διάγραμμα Venn ταξινόμησης των ΔΕΓ που προέκυψαν από τη μετα-ανάλυση. Τα αποτελέσματα που προέκυψαν από τον συνολικό έλεγχο (260 γονίδια) συγκρίνονται με τα ΔΕΓ που εντοπίστηκαν σε κάθε μια νόσο (249 γονίδια στη νόσο του Crohn και 38 γονίδια στην ελκώδη κολίτιδα).

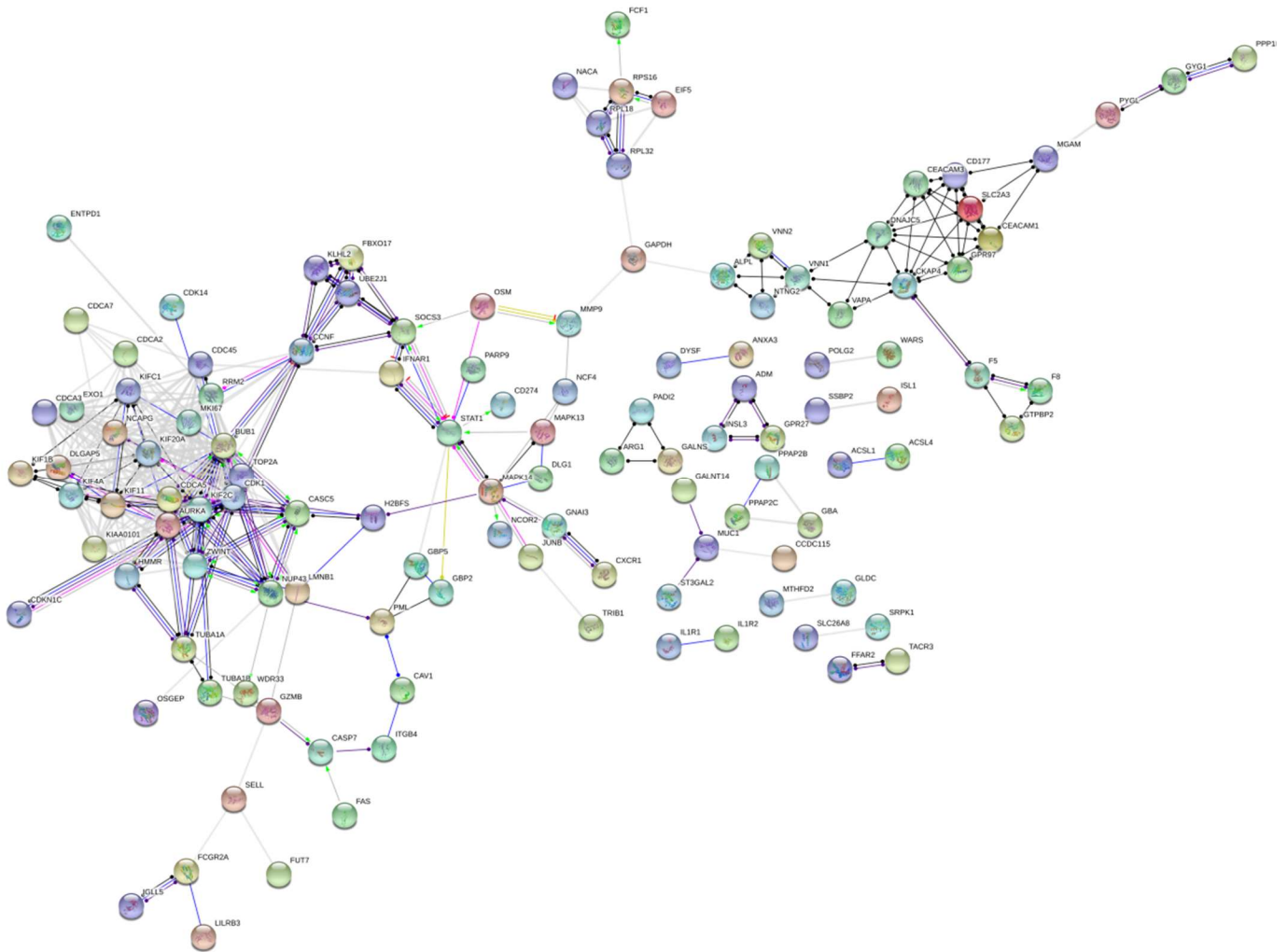
Όλα τα ΔΕΓ τα οποία εντοπίστηκαν στη μετα-ανάλυση μαζί με το μέγεθος επίδρασης τους και τις *p*-values τους παρουσιάζονται στον Πίνακα Παραρτήματος 7. Ερευνήσαμε και απεικόνισαμε τις αλληλεπιδράσεις μεταξύ των ΔΕΓ, χρησιμοποιώντας τη βάση δεδομένων STRING, όπως φαίνεται και στις Εικόνες 28-30. Στην Εικόνα 28 τα ΔΕΓ τα οποία εντοπίστηκαν από τον συνολικό έλεγχο φαίνεται να σχηματίζουν τέσσερις κύριες συστάδες (clusters). Η μεγαλύτερη και ιδιαίτερα συνδεδεμένη συστάδα αποτελείται κυρίως από γονίδια τα οποία εμπλέκονται στην κυτταρική διαίρεση και στην καρκινογένεση όπως CDKs, κυκλίνες, κινεσίνες, κοντεσίνες, τουμπουλίνες, λαμίνες, νουκλεοπορίνες, ρυθμιστές της μίτωσης κ.ο.κ. (BUB1, CDK1, CDKN1C, CDK14, CASC5, AURKA, NUP43, KIF2C, CDCA5, ZWINT, CCNF, TOP2A κ.ο.κ.). Η πρώτη συστάδα συνδέεται μέσω του CCNF με μια πιο χαλαρά συνδεδεμένη συστάδα η οποία αποτελείται κυρίως από πρωτεΐνες σχετιζόμενες με την ουμπικιτίνη (UBE2J1, KLHL2, FBXO17 και TRIB1) και πρωτεΐνες οι οποίες εμπλέκονται σε εξωκυτταρική σηματοδότηση όπως ιντερφερόνες, ενεργοποιούμενες από μιτογόνα κινάσες, μεταλλοπρωτεϊνάσες της εξωκυττάριας

ουσίας (matrix metalloproteinases), υποδοχείς χημειοκίνης, ογκοστατίνη και άλλες (STAT1, IFNAR1, SOCS3, OSM, MAPK13, DLG1, MAPK14, NCOR2, GNAI3 και CXCR1). Αυτή η δεύτερη συστάδα είναι συνδεδεμένη μέσω του GAPDH με δύο επιπλέον απομονωμένες συστάδες. Η πρώτη περιλαμβάνει ριβοσωμικές πρωτεΐνες και πρωτεΐνες που εμπλέκονται στη μετάφραση και επεξεργασία του rRNA (RPL32, RPL18, RPS16, EIF5, NACA και FCF1), ενώ η άλλη αποτελείται από πρωτεΐνες οι οποίες εμπλέκονται στη μεταφορά και το μεταβολισμό της γλυκόζης (PP1R3B, GYG1, PYGL, MGAM, CD177 και SLC2A3) και από πρωτεΐνες οι οποίες εμπλέκονται στην κυτταρική προσκόλληση (cell-adhesion) και στην εξωκύττωση (CEACAM1, CEACAM3, DNAJC5, GPR97, CKAP4, VAPA, VNN1 και VNN2). Από αυτά τα ΔΕΓ, με τη χρήση του Cytoscape (Shannon, Markiel et al. 2003) μιας πλατφόρμας λογισμικού ανοιχτού κώδικα για την απεικόνιση δικτύων μοριακής αλληλεπίδρασης, βρήκαμε τα 10 κορυφαία γονίδια με τις περισσότερες αλληλεπιδράσεις. Αυτά τα γονίδια ήταν τα CDK1, KIF2C, KIF11, BUB1, KIF20A, RRM2, TOP2A, AURKA, KIF4A και NCAPG.

Η συνδυασμένη λίστα των ΔΕΓ που βρέθηκαν (συμπεριλαμβανομένων των ΔΕΓ που εντοπίστηκαν στη νόσο του Crohn, στην ελκώδη κολίτιδα και στον συνολικό έλεγχο) διερευνήθηκαν επίσης για τις βιολογικές διεργασίες και τα βιοχημικά μονοπάτια στα οποία εμπλέκονται με τη χρήση της βάσης δεδομένων PANTHER. Αξιολογήσαμε το σύνολο των γονιδίων με την PANTHER χρησιμοποιώντας το εργαλείο εμπλουτισμού ORA (Πίνακας 10, Εικόνα 32). Τα περισσότερα από τα μονοπάτια περιλαμβάνουν γονίδια που προσδιορίζονται από όλους τους ελέγχους (συνολικός, νόσος του Crohn κι ελκώδης κολίτιδα), εκτός από τα μονοπάτια βιοσύνθεσης Λευκίνης και Αλανίνης τα οποία περιέχουν ένα γονίδιο (BCAT1) που ήταν στατιστικώς σημαντικό μόνο στον συνολικό έλεγχο και το μονοπάτι των JAK/STAT, το οποίο εντοπίστηκε στον συνολικό έλεγχο και στη νόσο του Crohn.

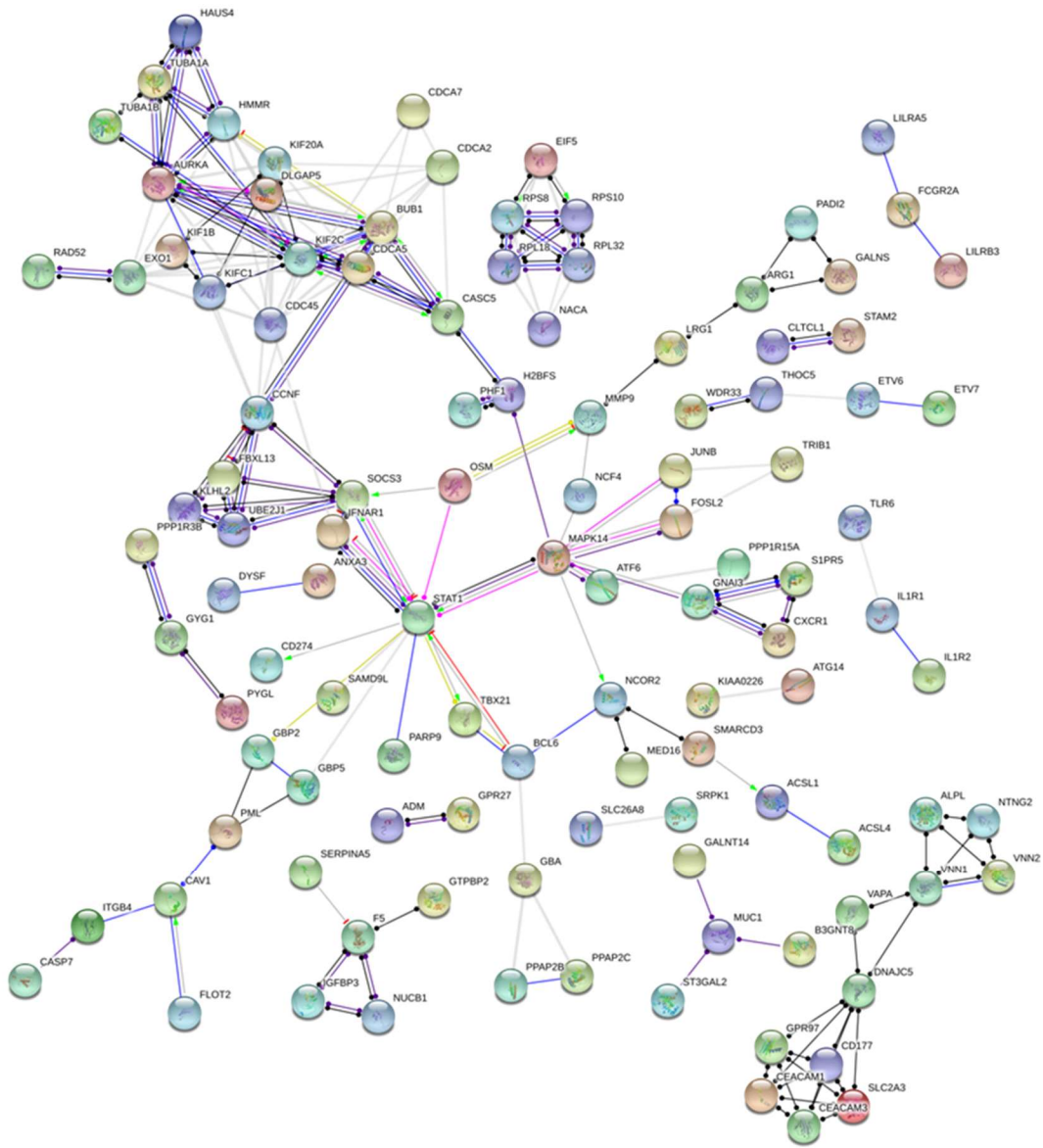
Συμπερασματικά, πραγματοποιήσαμε μια ενδιαφέρουσα εφαρμογή της πολυμεταβλητής μετα-ανάλυσης δεδομένων γονιδιακής έκφρασης της Ιδιοπαθούς Φλεγμονώδους Νόσου του Εντέρου. Η ανάλυση των μονοπατιών επιβεβαίωσε σημαντικά μοριακά μονοπάτια τα οποία εμπλέκονται στην παθογένεση της νόσου, όπως τα μονοπάτια σηματοδότησης των JAK/STAT και της IFN- γ , γονίδια τα οποία εμπλέκονται στην κυτταρική προσκόλληση, την απόπτωση και την

καρκινογένεση. Μεταξύ άλλων, τα γονίδια BCAT1 και GZMB είναι ενδιαφέροντα νέα ΔΕΓ και αξίζουν περισσότερη έρευνα σε πειραματικά μοντέλα κολίτιδας.



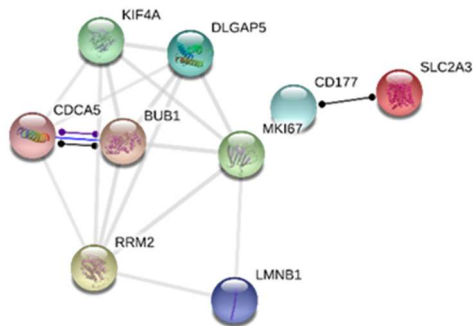
Εικόνα 28 Δίκτυο συσχέτισης γονιδίων/πρωτεϊνών των 232 (από τα 260) ΔΕΓ τα οποία αναγνωρίστηκαν από τον συνολικό έλεγχο, με τιμή ορίου αλληλεπίδρασης 0,7.

Η τιμή δεν υποδεικνύει τη δύναμη ή την ιδιαιτερότητα της αλληλεπίδρασης, αλλά μάλλον αποτελεί δείκτη εμπιστοσύνης. Το όριο του 0,7 αντιπροσωπεύει υψηλή εμπιστοσύνη. Γραμμές διαφορετικών χρωμάτων υποδεικνύουν τους προβλεπόμενους τρόπους δράσης. Οι πράσινες γραμμές αντιπροσωπεύουν την ενεργοποίηση, οι μπλε τη δέσμευση, οι γαλάζιες τον φαινότυπο, οι μαύρες αντίδραση, οι κόκκινες αναστολή, οι μωβ κατάλυση, οι ροζ τροποποιήσεις και οι κίτρινες μεταγραφικές ρυθμίσεις. Οι γραμμές που καταλήγουν σε βέλος υποδεικνύουν θετική αλληλεπίδραση, σε ορθογώνιο υποδεικνύουν αρνητική αλληλεπίδραση και σε κύκλο υποδεικνύουν μη καθορισμένη αλληλεπίδραση. Το δίκτυο κατασκευάστηκε χρησιμοποιώντας την βάση STRING. Οι ασύνδετοι κόμβοι του δικτύου δεν εμφανίζονται στο σχήμα.



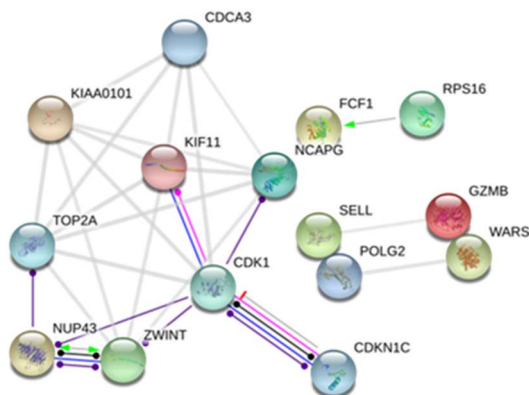
Εικόνα 29 Δίκτυο συσχέτισης γονιδίων/πρωτεϊνών των 224 (από τα 249) ΔΕΓ για την νόσο του Crohn, με τιμή ορίου αλληλεπίδρασης 0,7.

Η τιμή δεν υποδεικνύει τη δύναμη ή την ιδιαιτερότητα της αλληλεπίδρασης, αλλά μάλλον αποτελεί δείκτη εμπιστοσύνης. Το όριο του 0,7 αντιπροσωπεύει υψηλή εμπιστοσύνη. Γραμμές διαφορετικών χρωμάτων υποδεικνύουν τους προβλεπόμενους τρόπους δράσης. Οι πράσινες γραμμές αντιπροσωπεύουν την ενεργοποίηση, οι μπλε τη δέσμευση, οι γαλάζιες τον φαινότυπο, οι μαύρες αντίδραση, οι κόκκινες αναστολή, οι μωβ κατάλυση, οι ροζ τροποποιήσεις και οι κίτρινες μεταγραφικές ρυθμίσεις. Οι γραμμές που καταλήγουν σε βέλος υποδεικνύουν θετική αλληλεπίδραση, σε ορθογώνιο υποδεικνύουν αρνητική αλληλεπίδραση και σε κύκλο υποδεικνύουν μη καθορισμένη αλληλεπίδραση. Το δίκτυο κατασκευάστηκε χρησιμοποιώντας την βάση STRING. Οι ασύνδετοι κόμβοι του δικτύου δεν εμφανίζονται στο σχήμα.



Εικόνα 30 Δίκτυο συσχέτισης γονιδίων/πρωτεϊνών των 29 (από τα 38) ΔΕΓ για την ελκώδη κολίτιδα, με τιμή ορίου αλληλεπίδρασης 0,7.

Η τιμή δεν υποδεικνύει τη δύναμη ή την ιδιαιτερότητα της αλληλεπίδρασης, αλλά μάλλον αποτελεί δείκτη εμπιστοσύνης. Το όριο του 0,7 αντιπροσωπεύει υψηλή εμπιστοσύνη. Γραμμές διαφορετικών χρωμάτων υποδεικνύουν τους προβλεπόμενους τρόπους δράσης. Οι πράσινες γραμμές αντιπροσωπεύουν την ενεργοποίηση, οι μπλε τη δέσμευση, οι γαλάζιες τον φαινότυπο, οι μαύρες αντίδραση, οι κόκκινες αναστολή, οι μωβ κατάλυση, οι ροζ τροποποιήσεις και οι κίτρινες μεταγραφικές ρυθμίσεις. Οι γραμμές που καταλήγουν σε βέλος υποδεικνύουν θετική αλληλεπίδραση, σε ορθογώνιο υποδεικνύουν αρνητική αλληλεπίδραση και σε κύκλο υποδεικνύουν μη καθορισμένη αλληλεπίδραση. Το δίκτυο κατασκευάστηκε χρησιμοποιώντας την βάση STRING. Οι ασύνδετοι κόμβοι του δικτύου δεν εμφανίζονται στο σχήμα.



Εικόνα 31 Δίκτυο συσχέτισης γονιδίων/πρωτεϊνών των 49 (από τα 53) ΔΕΓ τα οποία εντοπίστηκαν μόνο από τον συνολικό έλεγχο, με τιμή ορίου αλληλεπίδρασης 0,7.

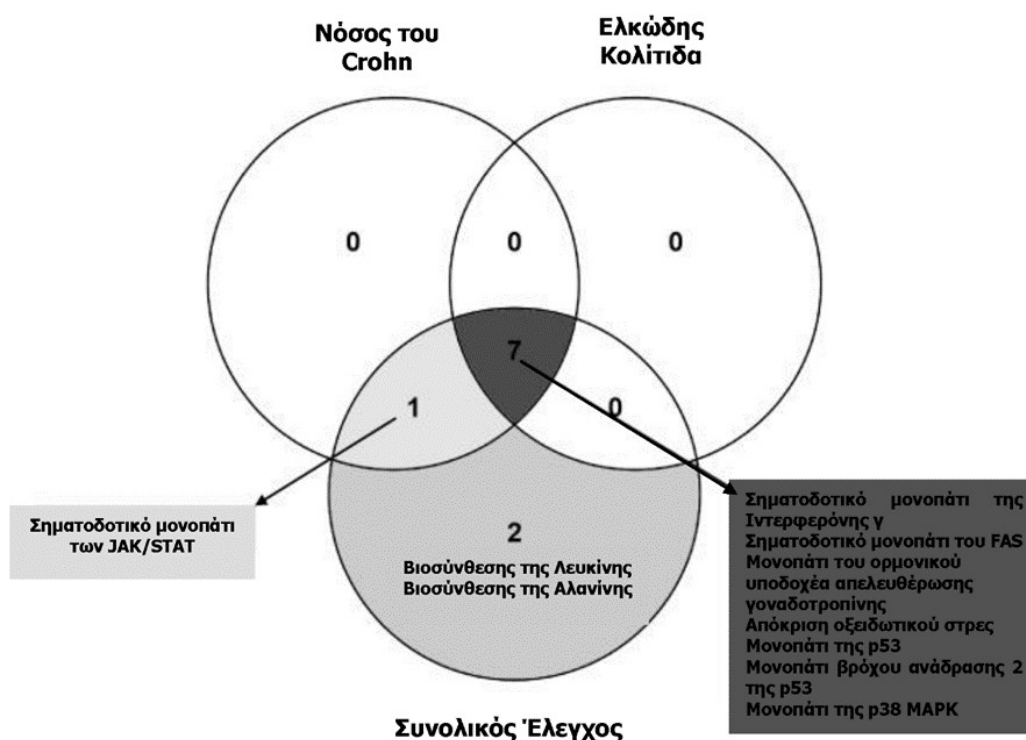
Η τιμή δεν υποδεικνύει τη δύναμη ή την ιδιαιτερότητα της αλληλεπίδρασης, αλλά μάλλον αποτελεί δείκτη εμπιστοσύνης. Το όριο του 0,7 αντιπροσωπεύει υψηλή εμπιστοσύνη. Γραμμές διαφορετικών χρωμάτων υποδεικνύουν τους προβλεπόμενους τρόπους δράσης. Οι πράσινες γραμμές αντιπροσωπεύουν την ενεργοποίηση, οι μπλε τη δέσμευση, οι γαλάζιες τον φαινότυπο, οι μαύρες αντίδραση, οι κόκκινες αναστολή, οι μωβ κατάλυση, οι ροζ τροποποιήσεις και οι κίτρινες μεταγραφικές ρυθμίσεις. Οι γραμμές που καταλήγουν σε βέλος υποδεικνύουν θετική αλληλεπίδραση, σε ορθογώνιο υποδεικνύουν αρνητική αλληλεπίδραση και σε κύκλο υποδεικνύουν μη καθορισμένη αλληλεπίδραση. Το δίκτυο κατασκευάστηκε χρησιμοποιώντας την βάση STRING. Οι ασύνδετοι κόμβοι του δικτύου δεν εμφανίζονται στο σχήμα.

Πίνακας 10 Στατιστικώς σημαντικά PANTHER βιοχημικά μονοπάτια στα οποία συμμετέχουν τα ΔΕΓ της μετα-ανάλυσης

PANTHER βιοχημικό μονοπάτι	Γονίδιο	p-value	
Σηματοδοτικό μονοπάτι της Ιντερφερόνης-γ	STAT1 (↑ CD & UC), SOCS3 (↑ CD & UC), MAPK13 (↑ CD & UC), MAPK14 (↑ CD & UC)	0.00114	Συνολικός έλεγχος, νόσος του Crohn & ελκώδης κολίτιδα
Μονοπάτι του ορμονικού υποδοχέα απελευθέρωσης γοναδοτροπίνης	GNAI3 (↑ CD & UC), TUBA1A (↑ CD & UC), TUBA1B (↑ CD & UC), MAPK13 (↑ CD & UC), MAPK14 (↑ CD & UC), NFATC4 (↑ CD & UC), JUNB (↑ CD & UC), CAV1 (↑ CD & UC), ID3 (↓ CD & UC)	0.00701	Συνολικός έλεγχος, νόσος του Crohn & ελκώδης κολίτιδα
Μονοπάτι της p53	FAS (↑ CD & UC), RRM2 (↑ CD & UC), CDK1 (↑ CD & UC), PML (↑ CD & UC), IGFBP3 (↓ CD & UC)	0.00974	Συνολικός έλεγχος, νόσος του Crohn & ελκώδης κολίτιδα
Σηματοδοτικό μονοπάτι του FAS	FAS (↑ CD & UC), CASP7 (↑ CD & UC), LMNB1 (↑ CD & UC)	0.014	Συνολικός έλεγχος, νόσος του Crohn & ελκώδης κολίτιδα
Μονοπάτι της p38 MAPK	MAPK13 (↑ CD & UC), MAPK14 (↑ CD & UC), IL1R1 (↑ CD & UC)	0.0254	Συνολικός έλεγχος, νόσος του Crohn & ελκώδης κολίτιδα
Σηματοδοτικό μονοπάτι των JAK/STAT	STAT1 (↑ CD & UC), MAPK14 (↑ CD & UC)	0.0289	Συνολικός έλεγχος & νόσος του Crohn
Μονοπάτι βρόχου ανάδρασης 2 της p53	STAT1 (↑ CD & UC), MAPK13 (↑ CD & UC), MAPK14 (↑ CD & UC)	0.0405	Συνολικός έλεγχος, νόσος του Crohn & ελκώδης κολίτιδα

Βιοσύνθεσης της Αλανίνης	BCAT1 (↑ CD & UC)	0.0417	Συνολικός έλεγχος
Βιοσύνθεσης της Λευκίνης	BCAT1 (↑ CD & UC)	0.0417	Συνολικός έλεγχος
Απόκρισης οξειδωτικού στρες	STAT1 (↑ CD & UC), MAPK13 (↑ CD & UC), MAPK14 (↑ CD & UC)	0.0484	Συνολικός έλεγχος, νόσος του Crohn & ελκώδης κολίτιδα

↑ υπερεκφράζονται; ↓ υπόεκφράζονται



Εικόνα 32 Διάγραμμα Venn σύγκρισης των στατιστικώς σημαντικών μονοπατιών της PANTHER μεταξύ των ασθενειών και του συνολικού ελέγχου.

3.4 Δημοσιεύσεις σε επιστημονικά περιοδικά

Σε αυτό το κεφάλαιο παρουσιάστηκαν τα αποτελέσματα από τη μετα-ανάλυση δεδομένων μικροσυστοιχιών, με σκοπό την εύρεση διαφορικά εκφρασμένων γονιδίων, από μελέτες οι οποίες μελετούσαν την ασθένεια της προεκλαμψίας σε πλακούντες. Τα αποτελέσματα της μετα-ανάλυσης και η υλοποίησή της, που παρουσιάστηκαν στο παρόν καθώς και στο προηγούμενο κεφάλαιο έχουν δημοσιευθεί στο διεθνές επιστημονικό περιοδικό *Pregnancy Hypertension*

(Vennou, Kontou et al. 2020). Επιπλέον στο κεφάλαιο αυτό παρουσιάστηκαν τα αποτελέσματα από την πολυμεταβλητή μετα-ανάλυση δεδομένων μικροσυστοιχιών, με σκοπό την εύρεση διαφορεικά εκφρασμένων γονιδίων, από μελέτες οι οποίες μελετούσαν την νόσο του Crohn και την ελκώδη κολίτιδα σε δείγματα αίματος. Τα αποτελέσματα της μετα-ανάλυσης, τα οποία παρουσιάστηκαν στο παρόν κεφάλαιο έχουν δημοσιευθεί στο διεθνές επιστημονικό περιοδικό *Genomics* (Vennou, Piovani et al. 2020). Η υλοποίηση της μεθοδολογίας της πολυμεταβλητής μετα-ανάλυσης, η οποία παρουσιάστηκε στο προηγούμενο κεφάλαιο έχει δημοσιευτεί στο διεθνές επιστημονικό περιοδικό *MethodsX* (Vennou, Piovani et al. 2020).

Κεφάλαιο 4^ο:
Συζήτηση - Συμπεράσματα

Η παρούσα διδακτορική διατριβή ασχολήθηκε με ερωτήματα και θέματα της Βιολογίας και της Βιοπληροφορικής, τα οποία συνέκλιναν στη μεθοδολογία της υπολογιστικής ανάλυσης βιολογικών δεδομένων και στην προσπάθεια εντοπισμού γονιδίων και των βιοχημικών μονοπατιών τα οποία εμπλέκονται στην αιτιολογία και εκδήλωση πολυπαραγοντικών ασθενειών. Κυρίως χρησιμοποιήθηκαν γονιδιακά δεδομένα πειραμάτων μικροσυστοιχιών τα οποία διερευνήθηκαν με τη μεθοδολογία της μετα-ανάλυσης για την εύρεση στατιστικώς σημαντικών διαφορικά εκφραζόμενων γονιδίων και στη συνέχεια τα βιοχημικά μονοπάτια, στα οποία συμμετέχουν. Ακόμη έγινε μια προσπάθεια διερεύνησης των βέλτιστων πρακτικών στη μεθοδολογία μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών.

4.1 Συμπεράσματα της διερεύνησης της μεθοδολογίας της μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών

Όπως αναφέρεται και στο κεφάλαιο 3.1 πραγματοποιήθηκε μια προσπάθεια διερεύνησης των βέλτιστων πρακτικών στη μεθοδολογία της μετα-ανάλυσης σε δεδομένα μικροσυστοιχιών. Ερευνήσαμε διάφορες πρακτικές που αναφέρονται στη βιβλιογραφία σχετικά με τη μετα-ανάλυση. Αρχικά πραγματοποιήθηκε ένας έλεγχος σε μικρό δοκιμαστικό και τυχαίο υποσύνολο δεδομένων των 99 ατόμων και στη συνέχεια σε ένα πολύ μεγαλύτερο των 426 ατόμων. Και τα δύο προέρχονταν από 10 μελέτες και αποτελούνταν από 27.867 μοναδικά γονίδια. Έπειτα έγινε διερεύνηση της παλιάς και κλασσικής μεθόδου διόρθωσης Hedges η οποία δίνει αμερόληπτες εκτιμήσεις (Hedges and Olkin 1985, Ramasamy, Mondry et al. 2008). Παρατηρήσαμε ότι σε μικρά σύνολα δεδομένων η διόρθωση βελτιώνει τα αποτελέσματα καθώς δεν χανόταν καθόλου πληροφορία, τουναντίον υπήρχε κέρδος, που υποδεικνύει την επιτακτική ανάγκη της εφαρμογής της διόρθωσης Hedges σε μικρά σύνολα. Στο μεγαλύτερο σύνολο υπήρχε μια διαφοροποίηση καθώς υπήρχε και ένα μικρό ποσοστό γονιδίων της τάξης του 4% από το σύνολο των ΔΕΓ το οποίο δεν είχε βρεθεί με τη χρήση της διόρθωσης. Παρατηρείται γενικά ότι όσο μεγαλύτερο είναι το δείγμα, η επίδραση της διόρθωσης γίνεται μικρότερη, που ήταν αναμενόμενο.

Στη συνέχεια εφαρμόσαμε τη μεθοδολογία του Bootstrap με χρήση 200 επαναλήψεων, που θεωρείται γενικά ένας αξιόλογος αριθμός επαναλήψεων στην βιβλιογραφία ώστε να παράγει ακριβείς μετρήσεις σε μέτριου μεγέθους σύνολα δεδομένων (Meuwissen and Goddard 2004) ύστερα από εφαρμογή της διόρθωσης του Hedges και συγκρίναμε τα αποτελέσματα αυτά με τα

αποτελέσματα μετα-ανάλυσης με χρήση μόνο της διόρθωσης του Hedges. Τα αποτελέσματά μας δείχνουν ότι η μέθοδος Bootstrap επιστρέφει περισσότερα γονίδια ως στατιστικώς σημαντικά, πιθανώς σε περιπτώσεις που η δειγματική κατανομή δεν ήταν η θεωρητικά αναμενόμενη. Ωστόσο απαιτείται σημαντικός υπολογιστικός χρόνος και σε μια τέτοια ανάλυση αποτελεί ένα αποτρεπτικό παράγοντα και για αυτό παρατηρείται ότι η μέθοδος αυτή δεν χρησιμοποιείται στην πράξη στις περισσότερες μετα-αναλύσεις. Φαίνεται λοιπόν, ότι η διόρθωση του Hedges αποτελεί έναν αναγκαίο συμβιβασμό. Πιθανές εναλλακτικές, θα μπορούσαν να αποτελέσουν οι «μικτές στρατηγικές», στις οποίες θα ελέγχονται με Bootstrap μόνο τα γονίδια που εμφανίζουν p -value κάτω από ένα δεδομένο σημείο, π.χ. p -value < 0.05 . Με αυτόν τον τρόπο, στη δική μας περίπτωση θα πραγματοποιούσαμε έλεγχο μόνο σε 3.168 (για το μικρό σετ δεδομένων)/4.582 (για το μεγάλο σετ δεδομένων) γονίδια ο οποίος θα απαιτούσε περίπου 19 και 42 ώρες αντίστοιχα.

Διερευνήσαμε επίσης τον βέλτιστο αριθμό επαναλήψεων του Bootstrap. Πραγματοποιήσαμε σύγκριση των 200 επαναλήψεων που προτείνεται στην βιβλιογραφία, όπως προαναφέραμε για μετρίου μεγέθους δείγματα και των 500 επαναλήψεων που θεωρείται ότι παράγουν ακριβείς εκτιμήσεις, αν και φυσικά ενδέχεται να χρειαστούν περισσότερες για την ακριβή εκτίμηση των p -values. Ωστόσο κάτι τέτοιο θα ήταν υπερβολικά χρονοβόρο και θα μπορούσε να καθυστερήσει αρκετά μια μελέτη. Στα αποτελέσματα που πήραμε από τη σύγκριση παρατηρήσαμε ότι και με τις 200 επαναλήψεις προκύπτουν αρκετά ακριβείς εκτιμήσεις πολύ κοντά στις εκτιμήσεις των 500 επαναλήψεων (τιμή συσχέτισης 0.9860) και με αρκετά μεγάλο ποσοστό κοινών στατιστικώς σημαντικών ΔΕΓ και σε πολύ λιγότερο χρόνο. Καταλήγοντας έτσι στο συμπέρασμα ότι μετα-ανάλυση με διόρθωση Hedges και ακολουθούμενη από Bootstrap με χρήση 200 επαναλήψεων θα επιστέφει πολύ καλά αποτελέσματα.

4.2 Συμπεράσματα της μετα-ανάλυσης δεδομένων γονιδιακής έκφρασης στην προεκλαμψία

Στο επόμενο κομμάτι της διατριβής, όπως αναφέρεται και στο κεφάλαιο 3.2 πραγματοποιήθηκε εφαρμογή των μεθόδων σε μια μετα-ανάλυση γονιδιακών προφίλ έκφρασης για την ασθένεια της προεκλαμψίας. Η παρούσα μετα-ανάλυση διεξήχθη σε όλα τα διαθέσιμα δεδομένα μικροσυστοιχίας σε πλακούντες γυναικών με προεκλαμψία και ταυτοποίησε 629 στατιστικώς σημαντικά ΔΕΓ σημαντικά συνδεδεμένα με την προεκλαμψία. Αυτά τα ΔΕΓ μπορούν δυνητικά να χρησιμοποιηθούν στην αποκρυπτογράφηση της αιτιολογίας της νόσου, ακόμη και ως δείκτες

πρόβλεψης του κινδύνου εμφάνισής της. Το δίκτυο αλληλεπίδρασης για τα 629 ΔΕΓ προσδιορίζει λίγα διασυνδεδεμένα κατά πολύ μεταξύ τους γονιδιακά προϊόντα, υποδηλώνοντας ότι πολλαπλά μεταβολικά μονοπάτια εμπλέκονται στον σχηματισμό του προεκλαμπτικού πλακούντα, γεγονός το οποίο είναι ενδεικτικό της ετερογένειας της παθολογίας της ασθένειας. Η ετερογένεια στις αιτίες της νόσου υποστηρίζεται περαιτέρω από το γεγονός ότι τα κορυφαία 20 ΔΕΓ (σύμφωνα με την *p*-value) δεν αλληλεπιδρούν και από τα κορυφαία 50 ΔΕΓ μόνο λίγα αλληλεπιδρούν. Η άποψη πως η ασθένεια διέπεται από ετερογένεια έχει ήδη προταθεί και από άλλους ερευνητές [βλέπε (Williams and Broughton Pipkin 2011, Louwen, Muschol-Steinmetz et al. 2012, El-Sayed 2017)] δίνοντας έτσι την εγκυρότητα της ανάλυσής μας.

Τα άκρως διασυνδεδεμένα γονίδια της προεκλαμψίας (Εικόνα 22) συμμετέχουν σε διάφορα μονοπάτια μεταγωγής σήματος και στην έμφυτη ανοσοαπόκριση. Η έρευνα για τα 18 κοινά ΔΕΓ στην ήπια και την σοβαρή προεκλαμψία αποκάλυψε γονίδια τα οποία συμμετέχουν στις διαδικασίες ανάπτυξης/εγκυμοσύνης, στην ανοσοαπόκριση και στην αξιοποίηση των υδατανθράκων και της ενέργειας. Με τη χρήση βιοπληροφορικών αναλύσεων εμπλουτισμού (ORA-PANTHER) τα μονοπάτια μεταβολισμού υδατανθράκων/ενεργειακής αξιοποίησης εμφανίστηκαν ως τα πιο ισχυρά και κοινά στους δύο υποτύπους της προεκλαμψίας. Το γεγονός ότι οκτώ γονίδια αυτού του μονοπατιού υπερεκπροσωπούνται στα μονοπάτια που συμμετέχουν τα ΔΕΓ της μετα-ανάλυσης, υποδηλώνει έντονα την πραγματική συμμετοχή των γονιδίων της ενεργειακής αξιοποίησης στην ανάπτυξη της προεκλαμψίας. Αυτό το εύρημα έχει μεγάλη σημασία καθώς η απορυθμισμένη λειτουργία αυτών των γονιδίων μπορεί να οδηγήσει σε απορυθμισμένη αξιοποίηση της ενέργειας και ίσως στην παχυσαρκία. Είναι γνωστό ότι ο αυξημένος δείκτης μάζας σώματος (ΔΜΣ) και το μη φυσιολογικό λιπιδιακό προφίλ είναι χαρακτηριστικά της προεκλαμψίας (Lopez-Jaramillo, Barajas et al. 2018). Παρόλο που πολλές μελέτες οι οποίες διερευνούν υποψήφιους βιοδείκτες δεν έχουν αναφερθεί ποτέ σε γονίδια ενεργειακής αξιοποίησης, έχει ωστόσο διερευνηθεί η παχυσαρκία ως παράγοντας κινδύνου και έχει βρεθεί ότι σχετίζεται με την προεκλαμψία [αναφέρεται στα (Chappell and Morgan 2006, Bodnar, Catov et al. 2007, Grill, Rusterholz et al. 2009, El-Sayed 2017)]. Αναφέρεται επίσης ότι η δυσλιπιδαιμία συμβάλλει στη δυσλειτουργία των ενδοθηλιακών κυττάρων (Belo, Gaffney et al. 2004) και σχετίζεται με αυξημένη φλεγμονή, αγγειογένεση και αλλοιώσεις των κυκλοφορούντων ρυθμιστών ανάπτυξης (Fontana, Eagon et al. 2007). Μεταλλάξεις στα γονίδια LPL και ApoE έχουν προταθεί ως παράγοντες κινδύνου (Williams and Broughton Pipkin 2011). Έτσι, δεν μπορεί

να αποκλειστεί η πιθανότητα η απορυθμισμένη έκφραση των γονιδίων που εμπλέκονται στην κατανάλωση ενέργειας να οδηγήσει σε παχυσαρκία και εν συνεχεία σε προεκλαμψία.

Η ανάλυση εμπλουτισμού των μονοπατιών οδήγησε επίσης σε μερικά ενδιαφέροντα ευρήματα. Πρώτον, το μονοπάτι βιοσύνθεσης αμινοξέων βρέθηκε στατιστικώς σημαντικό στην προεκλαμψία, μια πρόσφατη μεταβολομική ανάλυση έδειξε ότι η συγκέντρωση ορισμένων αμινοξέων στο αίμα μεταβάλλεται σε ασθενείς με προεκλαμψία (Nobakht 2018). Επιπροσθέτως, έχει προταθεί ότι η λήψη συμπληρωμάτων L-αργινίνης από ασθενείς με προεκλαμψία μπορεί να μειώσει τα επίπεδα της αρτηριακής πίεσης (Vadillo-Ortega, Perichart-Perera et al. 2011). Δεύτερον, το μονοπάτι του κίρκαδικού ρολογιού βρέθηκε επίσης στατιστικώς σημαντικό στην ολική και στην ήπια προεκλαμψία, το οποίο αντιπροσωπεύεται από δύο γονίδια. Αυτό το εύρημα είναι αναμενόμενο καθώς η αρτηριακή πίεση στη φυσιολογική εγκυμοσύνη ακολουθεί έναν κίρκαδικό ρυθμό με χαμηλότερες τιμές κατά τη διάρκεια της νύχτας, ενώ σε πολλές ασθενείς με προεκλαμψία δεν παρατηρείται τέτοια πτώση (Pears, Makris et al. 2018). Τρίτον, βρήκαμε τρία ΔΕΓ τα οποία συμμετέχουν στα μονοπάτια μεταβολισμού πυριμιδίνης. Έχει αναφερθεί ότι τα επίπεδα κυκλοφορίας του ελεύθερου εμβρυϊκού DNA (cell free fetal DNA, cff DNA) είναι υψηλότερα στην προεκλαμψία και το cff DNA έχει διερευνηθεί ως πιθανός προγνωστικός βιοδείκτης για αυτήν (Konecna, Vlkova et al. 2015, Sifakis, Koukou et al. 2015, Vlkova, Turna et al. 2015, Contro, Bernabini et al. 2017). Συνεπώς, η εμπλοκή γονιδίων τα οποία συμμετέχουν στο μονοπάτι μεταβολισμού πυριμιδίνης φαίνεται να έχει μια λογική συσχέτιση. Τέλος, το μονοπάτι του GnRHR βρέθηκε στατιστικώς σημαντικό μόνο στην σοβαρή προεκλαμψία. Είναι γνωστό ότι ο GnRH εκφράζεται στον πλακούντα και δρα μέσω του συγκεκριμένου GnRHR τοπικά για να ασκήσει αυτοκρινείς λειτουργίες (Votolini and Petraglia 2014) είναι απολύτως εύλογο λοιπόν να βρεθεί στατιστικώς σημαντικό. Ενδιαφέρον είναι ακόμη το ότι όταν πραγματοποιήθηκε μετα-ανάλυση με λιγότερο αυστηρό FDR (FDR <0.05), το μονοπάτι του GnRHR εμφανίστηκε στατιστικώς σημαντικό και στη συνολική προεκλαμψία.

Αξιοσημείωτο είναι ότι στο αρχικό δίκτυο που δημιουργήθηκε με τα 629 ΔΕΓ (Εικόνα 22) εμφανίστηκαν δύο κύριοι κόμβοι με γονίδια που εμπλέκονται σε μονοπάτια μεταγωγής σήματος τα οποία επίσης βρέθηκαν σημαντικά στην συνολική καθώς και στους δύο υποτύπους της προεκλαμψίας. Η αφθονία των γονιδίων μεταγωγής σήματος σε όλα τα αποτελέσματα είναι αναμενόμενη και απαραίτητη καθώς σε όλα σχεδόν τα ειδικά μονοπάτια τα οποία ανταποκρίνονται σε εξωτερικά σήματα, εμπλέκονται τέτοια γονίδια και επομένως αναμένεται η

αφθονία τους. Από την άλλη πλευρά, η αφθονία των γονιδίων τα οποία εμπλέκονται στη μεταγωγή σήματος χρησιμεύουν ως επικύρωση της μεθόδου μας, δείχνοντας ότι η ανάλυση πραγματοποιήθηκε σωστά.

Ωστόσο, μελλοντική ανάλυση των λειτουργιών για τα υπόλοιπα γονίδια/μονοπάτια, τα οποία βρέθηκαν με τις μεθόδους εμπλουτισμού, θα ήταν απαραίτητη για την επαλήθευση των αποτελεσμάτων μας και για την ανακάλυψη περισσότερων μηχανισμών στους οποίους υπόκειται η εμφάνιση της προεκλαμψίας.

4.3 Συμπεράσματα της μεθοδολογίας της πολυμεταβλητής μετα-ανάλυσης δεδομένων γονιδιακής έκφρασης

Ένα σημαντικό κομμάτι αυτής της διατριβής ήταν η ανάπτυξη μεθοδολογίας και η δημιουργία λογισμικού, το οποίο να επιτρέπει την πολυμεταβλητή μετα-ανάλυση σε δεδομένα γονιδιακής έκφρασης και πιο συγκεκριμένα σε δεδομένα προερχόμενα από μικροσυστοιχίες, όπως αναφέρεται και στο κεφάλαιο 2.4. Οι μέθοδοι πολυμεταβλητής μετα-ανάλυσης είναι ευρέως διαδεδομένες σε άλλους τομείς, όπως για παράδειγμα σε μελέτες γενετικής συσχέτισης και σε συγκρίσεις διαφορετικών θεραπειών, αλλά απ' ό,τι γνωρίζουμε, αυτή είναι η πρώτη προσπάθεια για την πραγματοποίηση πολυμεταβλητής μετα-ανάλυσης σε δεδομένα γονιδιακής έκφρασης. Όπως αναφέρθηκε ήδη, η μεθοδολογία της πολυμεταβλητής μετα-ανάλυσης παρέχει ένα ευρύ φάσμα πλεονεκτημάτων στον αναλυτή. Στην συγκεκριμένη περίπτωση, χρησιμοποιήσαμε το μοντέλο πολλαπλών μεταβλητών και την εκτίμηση του πίνακα συνδιακύμανσης για την πραγματοποίηση συνολικών ελέγχων αποφεύγοντας πολλαπλές συγκρίσεις και κατά συνέπεια αποφεύγοντας την συνεχή αύξηση του ρυθμού του σφάλματος τύπου I. Ένα βασικό πλεονέκτημα είναι ότι επιτρέπει διάφορους συνολικούς ελέγχους. Χαρακτηριστικό παράδειγμα είναι ο W ο οποίος ελέγχει τη μηδενική υπόθεση, ότι και οι δύο εκτιμητές των g_1 και g_2 είναι ίσοι με το μηδέν. Με αυτή τη στατιστική μέθοδο μπορεί να γίνει έλεγχος πλειοτροπίας, δηλαδή έλεγχος για γονίδια τα οποία επηρεάζουν και τις δύο εκβάσεις τα οποία δεν θα μπορούσαν να έχουν βρεθεί σε μια μονομεταβλητή ανάλυση. Ένας παρόμοιος έλεγχος έχει προταθεί για την από κοινού εκτίμηση των κύριων επιδράσεων των γονιδίων και των αλληλεπιδράσεων γονιδίου-περιβάλλοντος σε μελέτες GWAS (Manning, LaValley et al. 2011) και για τον εντοπισμό πλειοτροπικών επιδράσεων στη διμεταβλητή μετα-ανάλυση των GWAS (Medina-Gomez, Kemp et al. 2017). Άλλο ένα παράδειγμα συνολικού ελέγχου είναι ο έλεγχος για την ισοδυναμία των μεγεθών επίδρασης D . Ο

οποίος είναι παρόμοιος με τον έλεγχο για το γενετικό μοντέλο κληρονομικότητας σε μελέτες γενετικής συσχέτισης (Bagos 2008, Bagos 2013) και μπορεί να χρησιμοποιηθεί για τη σύγκριση δύο εκτιμητών και για την κατασκευή ενός διαστήματος εμπιστοσύνης για τη διαφορά ή κάποια άλλη συνάρτησή τους.

Ακόμη ένα μεγάλο πλεονέκτημα της πολυμεταβλητής μετα-ανάλυσης είναι ότι μπορεί να περιλαμβάνει, υπό την προϋπόθεση τυχαίας ελλείψεως, μελέτες οι οποίες αναφέρουν μόνο μια από τις παραμέτρους ενδιαφέροντος, καταλήγοντας σε εξωτερικό δανεισμό ισχύος από τις άλλες μελέτες (Higgins and Whitehead 1996). Αυτή η ιδιότητα θα φανεί αρκετά χρήσιμη σε περιπτώσεις όπου κάποιες μελέτες αναφέρουν μόνο μία από τις παραμέτρους ενδιαφέροντος, και θεωρείται κάτι σημαντικό και για αυτό είναι καίριας σημασίας η σωστή εξειδίκευση του πολυμεταβλητού μοντέλου ώστε να πραγματοποιείται σωστή εκτίμηση της εντός των μελετών συνδιακύμανσης (Riley, Abrams et al. 2007, Riley, Abrams et al. 2007, Riley 2009). Η υλοποίηση της μεθόδου είναι διαθέσιμη στην ιστοσελίδα: www.compgen.org/tools/multivariate-microarrays.

Ενδιαφέρον θα αποτελούσε η εφαρμογή της πολυμεταβλητής μεθοδολογίας σε άλλες περιπτώσεις δεδομένων γονιδιακής έκφρασης καθώς και η επέκταση της μεθόδου ώστε να επιτρέπει πιο δύσκολες διαδικασίες, όπως τη μετα-ανάλυση δικτύων για συγκρίσεις μεικτής θεραπείας (Lu and Ades 2004), όπου θα χρειαστούν πρόσθετες τροποποιήσεις, ειδικά όταν πρόκειται για προδιαγραφές μοντέλων.

4.4 Συμπεράσματα της πολυμεταβλητής μετα-ανάλυσης σε δεδομένα γονιδιακής έκφρασης για την Ιδιοπαθή Φλεγμονώδη Νόσο του Εντέρου

Τέλος σε αυτή τη διατριβή, όπως αναφέρεται και στο κεφάλαιο 3.3, πραγματοποιήθηκε πολυμεταβλητή μετα-ανάλυση δεδομένων γονιδιακής έκφρασης της Ιδιοπαθούς Φλεγμονώδους Νόσου του Εντέρου. Οι γενετικές μελέτες έχουν εντοπίσει πάνω από 200 παραλλαγές οι οποίες σχετίζονται με την ΙΦΝΕ, διασαφηνίζοντας έτσι βασικές απόψεις σχετικά με τη βιολογία της ασθένειας (Parkes, Barrett et al. 2007, Anderson, Boucher et al. 2011, Kenny, Pe'er et al. 2012, Yamazaki, Umeno et al. 2013, Julia, Domenech et al. 2014, Yang, Hong et al. 2014, Liu, van Sommeren et al. 2015, Ellinghaus, Jostins et al. 2016, de Lange, Moutsianas et al. 2017). Στην συγκεκριμένη πολυμεταβλητή μετα-ανάλυση σε δεδομένα γονιδιακής έκφρασης προσδιορίστηκαν 260 ΔΕΓ σε ασθενείς με ΙΦΝΕ (συνολικός έλεγχος, W) σε σύγκριση με υγιή

άτομα επιβεβαιώνοντας διάφορα μονοπάτια τα οποία ενδέχεται να εμπλέκονται στη νόσο. Ο συνολικός αυτός έλεγχος επιπλέον αναγνώρισε 53 γονίδια τα οποία δεν θα μπορούσαν να έχουν βρεθεί σε μια μονομεταβλητή ανάλυση. Ακόμη ο έλεγχος για την ισοδυναμία των μεγεθών επίδρασης (*D*) ανίχνευσε μόνο ένα μικρό αριθμό γονιδίων (2) με ίδια επίπεδα έκφρασης (τα TRIP6 και IGHG1). Επιπροσθέτως επισημάνθηκαν κάποια γονίδια τα οποία δεν έχουν ποτέ συσχετιστεί με την ΙΦΝΕ.

Η ανάλυσή που πραγματοποιήθηκε επιβεβαίωσε τον σημαντικό ρόλο των μονοπατιών των JAK/STAT, της ιντερφερόνης- γ (IFN- γ) και των μονοπατιών απόκρισης σε οξειδωτικό στρες (STAT1, MAPKs) στην παθογένεια της ΙΦΝΕ (Monteleone, Biancone et al. 1997, Ma 2001, Waetzig, Seegert et al. 2002, Trinchieri, Pflanz et al. 2003, Hunter 2005, Scaldaferrri, Sans et al. 2009, De Vries, Wildenberg et al. 2017). Η σηματοδότηση των JAK/STAT εμπλέκεται στον έλεγχο πολλών προ-φλεγμονωδών κυτοκινών [ιντερφερόνες και ιντερλευκίνες (De Vries, Wildenberg et al. 2017)]. Η παθογόνος περίσσεια αυτών των κυτοκινών έχει αποδοθεί σε ενεργοποιημένα αντογίνοπαρουσιαστικά κύτταρα και σε Τ-βοηθητικά λεμφοκύτταρα μέσω της αύξησης της ρύθμισης των ιντερλευκινών (interleukin, IL) 12 (IL-12) και 27 (IL-27) (Monteleone, Biancone et al. 1997, Trinchieri, Pflanz et al. 2003, Hunter 2005). Συγκεκριμένα, η IL-12 συμμετέχει στη διαφοροποίηση των παρθένων CD4⁺ Τ-λεμφοκυττάρων σε Τ-βοηθητικά λεμφοκύτταρα τα οποία παράγουν IFN- γ και επηρεάζουν το επίπεδο του παράγοντα νέκρωσης όγκων α (TNF- α), με αποτέλεσμα να επηρεάζουν τις λειτουργίες των μακροφάγων στην επίκτητη και στη φυσική ανοσία (Ma 2001). Οι ενεργοποιούμενες από μιτογόνα πρωτεϊνικές κινάσες (Mitogen-activated protein kinases, MAPKs) συνδέονται με τη σηματοδότηση του TNF- α και έχει αποδειχθεί ότι ελέγχουν την προσκόλληση των λεμφοκυττάρων και τη μετανάστευση κατά μήκος των μικροαγγείων (microvasculature) (Waetzig, Seegert et al. 2002, Scaldaferrri, Sans et al. 2009). Επιβεβαιώθηκε επίσης ο ρόλος ορισμένων πρωτεϊνών οι οποίες εμπλέκονται στην κυτταρική προσκόλληση και την εξωκύττωση, ιδιαίτερα της οικογένειας του καρκινοεμβρυικού αντιγόνου (Carcinoembryonic antigen-related cell adhesion molecule, CEACAM). Αυτές οι πρωτεΐνες έχουν συσχετιστεί με την ΙΦΝΕ και είναι γνωστό ότι επηρεάζουν ιδιαίτερος την ανοσία του βλεννογόνου και ρυθμίζουν τη λειτουργία των Τ- και Β-λεμφοκυττάρων (Gray-Owen and Blumberg 2006, Roda, Dahan et al. 2009). Με τη χρήση της μετα-ανάλυσης βρέθηκαν διάφορα ΔΕΓ και μονοπάτια τα οποία έχουν βρεθεί να εμπλέκονται στην καρκινογένεση, όπως είναι η απόπτωση που προκαλείται από την επίδραση της p53 [FAS (Bennett, Macdonald et al. 1998),

PML (Guo, Salomoni et al. 2000), CDK1 (Innocente, Abrahamson et al. 1999)], ο/η σχηματισμός/σταθεροποίηση των μικροσωληνίσκων [AURKA (Goos, Coupe et al. 2013), KIF2C (Gnjatic, Cao et al. 2010)], όπως επίσης το πρωτοογκογονίδιο BCL3 (Brocke-Heidrich, Ge et al. 2006) και η τοποϊσομεράση TOP2A (Coss, Tosetto et al. 2009). Η ΙΦΝΕ σχετίζεται με υψηλότερο κίνδυνο ανάπτυξης νεοπλασμάτων της γαστρεντερικής οδού (Danese and Mantovani 2010, Olen, Askling et al. 2017) και ειδικότερα του καρκίνου του παχέος εντέρου (Rubin, Shaker et al. 2012). Αυτό έχει συνδεθεί με τη ζωτική συνεισφορά της δυσπλασίας και της χρόνιας φλεγμονής στην καρκινογένεση (Mantovani, Allavena et al. 2008, Danese and Mantovani 2010, Rubin, Shaker et al. 2012). Ωστόσο, η εμφάνιση παιδιατρικής ΙΦΝΕ έχει συνδεθεί με διπλάσιο κίνδυνο ανάπτυξης κακοήθειας (Olen, Askling et al. 2017), υποδεικνύοντας ότι ορισμένα γενετικά χαρακτηριστικά της ΙΦΝΕ μπορεί να είναι κοινά ή να προδιαθέτουν καρκίνο. Το κατά πόσο αυτές οι δυσλειτουργίες μπορεί να είναι γενετικά χαρακτηριστικά της ΙΦΝΕ ή να είναι συνέπεια του φαινότυπου της εξακολουθεί να αποτελεί θέμα συζήτησης.

Το γονίδιο MMP9 βρέθηκε να εκφράζεται διαφορετικά στη νόσο του Crohn, στην ελκώδη κολίτιδα και στην ΙΦΝΕ (συνολικό έλεγχο) και έδειξε αρκετές λειτουργικές συνδέσεις με άλλα γονίδια. Οι μεταλλοπρωτεϊνάσες της εξωκυττάριας ουσίας και συγκεκριμένα η MMP9 είναι βασικά ένζυμα τα οποία εμπλέκονται στην αναδόμηση του ιστού και συμβάλλουν στην καταστροφή του εντερικού φράγματος (Shimshoni, Yablecovitch et al. 2015). Ωστόσο, δεν είναι σαφές εάν ο ρόλος του MMP9 είναι παθογόνος ή είναι απλώς το αποτέλεσμα ενεργοποίησης ουδετερόφιλων κατά τη διάρκεια της πορείας της εκτεταμένης φλεγμονής (extensive inflammation process) (de Bruyn and Ferrante 2018). Διαπιστώσαμε ότι ορισμένες πρωτεΐνες οι οποίες εμπλέκονται στη μεταφορά και τον μεταβολισμό της γλυκόζης εκφράστηκαν διαφορετικά σε ασθενείς με ΙΦΝΕ. Συγκεκριμένα, τα δεδομένα μας επιβεβαιώνουν τις πρόσφατα ανακαλυφθείσες σχέσεις των SLC2A3 και GYG1 με την ΙΦΝΕ (Uniken Venema, Voskuil et al. 2019). Αυτές οι πρωτεΐνες εμπλέκονται σε μεταβολικές διεργασίες οι οποίες συνδέονται με την έμφυτη ανοσολογική απόκριση (Smith, Dickinson et al. 2014).

Ένα σχετικό κλινικό ζήτημα της ΙΦΝΕ είναι η ίνωση και οι σχετιζόμενες με αυτή εντερικές στενώσεις και εμπόδια τα οποία συχνά απαιτούν χειρουργική διόρθωση (D'Haens, Rieder et al. 2019). Μια πολύ πρόσφατη μελέτη έχει βρει ότι η υπερέκφραση του BUB1 μπορεί να προκαλέσει ίνωση του εντέρου στη νόσο του Crohn (Garlatti, Ungaro et al. 2019). Η ανάλυσή μας έδειξε δυσλειτουργία αυτού του γονιδίου σε ασθενείς με νόσο του Crohn και ελκώδη κολίτιδα. Το BUB1

είναι ένας ρυθμιστής της απράκτου και η γενετική του μετάλλαξη έχει αναγνωριστεί ως παράγοντας κινδύνου για τον καρκίνο του παχέος εντέρου, υποδηλώνοντας την αλληλοεπικάλυψη των γενετικών χαρακτηριστικών μεταξύ της ΙΦΝΕ και του καρκίνου (de Voer, Geurts van Kessel et al. 2013).

Βρήκαμε αρκετά διαφορετικά εκφραζόμενα γονίδια είτε μόνο στη νόσο του Crohn (STAT1, MAPK14, SOCS3, TLR6, IFNAR1, GNAI3, CXCR1, BCL3, JUNB και άλλα) είτε μόνο στην ελκώδη κολίτιδα (RRM2, MAPK13, CAMKK1 και άλλα). Αυτή η ειδικότητα μπορεί να αντικατοπτρίζει τη διαφορετική παθογένεια και τα διαφορετικά χαρακτηριστικά μεταξύ των δύο ασθενειών (Khor, Gardet et al. 2011). Ωστόσο, όλα αυτά τα γονίδια εκτός από το TLR6, βρέθηκαν στατιστικώς σημαντικά και με την ΙΦΝΕ από τον συνολικό έλεγχο. Η υπερέκφραση του 6^{ου} Toll-like υποδοχέα έχει ήδη βρεθεί ότι εμπλέκεται στην προαγωγή κολίτιδας σε ποντίκια μέσω αυξημένων αποκρίσεων Τύπου 1 (Th1) και Τύπου 17 (Th17) βοηθητικών Τ-λεμφοκυττάρων (Type x T helper cells, Thx) σε σχετιζόμενο με το έντερο λεμφικό ιστό (Morgan, Koelink et al. 2014). Τα αποτελέσματά μας επιβεβαιώνουν ότι αυτό το γονίδιο θα μπορούσε να αποτελέσει έναν πιθανό στόχο για τη νόσο του Crohn.

Επίσης βρέθηκε στατιστικώς σημαντική διαφορική έκφραση κάποιων γονιδίων μέσω του συνολικού ελέγχου, η οποία δεν ήταν εμφανής όταν γινόταν ο έλεγχος για τη νόσο του Crohn και την ελκώδη κολίτιδα ξεχωριστά. Αυτό είναι ένα σημαντικό ατού της πολυμεταβλητής μετα-ανάλυσης, όπως έχει ήδη αναφερθεί. Το BCAT1 ήταν ένα από τα πιο σημαντικά γονίδια που προσδιορίστηκαν με αυτόν τον έλεγχο. Το BCAT1 ελέγχει τον μεταβολικό επαναπρογραμματισμό των ενεργοποιημένων (ανθρώπινων) μακροφάγων, με σημαντικές επιπτώσεις στις φλεγμονώδεις καταστάσεις (Parathanassiu, Ko et al. 2017). Άλλα γονίδια τα οποία εντοπίστηκαν από τον συνολικό έλεγχο βρέθηκαν να σχετίζονται με το σηματοδοτικό μονοπάτι της απόπτωσης μέσω της p53 (FAS (Bennett, Macdonald et al. 1998) και CDK1 (Innocente, Abrahamson et al. 1999) και άλλων μηχανισμών κυτταρικού θανάτου που διαμεσολαβούνται από το GZMB (Lord, Rajotte et al. 2003) και το GAPDH [αυτοφαγία (Colell, Ricci et al. 2007)]]. Επιπροσθέτως, το GZMB εκφράζεται σε ρυθμιστικά Τ-λεμφοκύτταρα και διαδραματίζει κρίσιμο ρόλο στον έλεγχο της φλεγμονής των πνευμόνων (Loebbermann, Thornton et al. 2012). Όλα αυτά τα γονίδια αξίζουν περαιτέρω διερεύνηση σε πειραματικά μοντέλα της ΙΦΝΕ καθώς θα μπορούσαν να είναι υποψήφια για την ανεύρεση φαρμάκων και για διάφορους ελέγχους για πιθανές θεραπείες. Στο γονίδιο F8 (παράγοντας θρομβώσεως VIII) βρέθηκε επίσης στατιστικώς σημαντική η διαφορική

έκφρασή του μόνο στον συνολικό έλεγχο. Είναι γνωστή η αλληλεπίδραση μεταξύ της ενεργοποίησης της πήξης του αίματος και της φλεγμονής (Danese, Para et al. 2007, Novacek, Weltermann et al. 2010). Ασθενείς με ΙΦΝΕ εμφανίζουν συχνά προθρομβωτική πάθηση και για τον λόγο αυτό διατρέχουν αυξημένο κίνδυνο επαναλαμβανόμενου φλεβικού θρομβοεμβολισμού (Novacek, Weltermann et al. 2010, Singh, Kullo et al. 2015) και ενδέχεται να παρουσιάσουν αυξημένο κίνδυνο συνοδών καρδιαγγειακών νοσημάτων (Singh, Kullo et al. 2015).

Τέλος, σε μακροσκοπικό επίπεδο, υπάρχει ένα επιπλέον εύρημα το οποίο αξίζει να αναφερθεί. Η πλειοψηφία των ΔΕΓ είναι ρυθμισμένη προς τα επάνω, δηλαδή η έκφρασή τους αυξάνεται στην ΙΦΝΕ (203/249 στη νόσο του Crohn, 33/38 στην ελκώδη κολίτιδα, 216/260 στον συνολικό έλεγχο). Η κατεύθυνση της αλλαγής είναι επίσης ταυτόσημη, τόσο στη νόσο του Crohn όσο και στην ελκώδη κολίτιδα, σε όλα τα γονίδια εκτός από τρία (TACR3, MATN4 και IGHG1) στα οποία το μεγάλο μέγεθος επίδρασης της νόσου του Crohn επηρέασε τον συνολικό έλεγχο.

Βιβλιογραφία

- Alberts, B., D. Bray, K. Hopkin, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter (2004). Essential cell biology, Second edition. New York, Garland Science Pub.
- Alberts, B., J. H. Wilson, A. Johnson, T. Hunt, J. Lewis, K. Roberts, M. Raff and P. Walter (2008). Molecular Biology of the Cell, Garland Science.
- Altman, D. G. and J. M. Bland (2003). "Interaction revisited: the difference between two estimates." BMJ **326**(7382): 219.
- Ananthakrishnan, A. N., R. J. Xavier and D. K. Podolsky (2017). Clinical Features and Diagnosis of Crohn's Disease. Inflammatory Bowel Diseases: A Clinician's Guide. J. W. S. Ltd.: 17-31.
- Ananthakrishnan, A. N., R. J. Xavier and D. K. Podolsky (2017). Clinical Features and Diagnosis of Ulcerative Colitis. Inflammatory Bowel Diseases: A Clinician's Guide. J. W. S. Ltd: 33-47.
- Anderson, C. A., G. Boucher, C. W. Lees, A. Franke, M. D'Amato, K. D. Taylor, J. C. Lee, P. Goyette, M. Imielinski, A. Latiano, C. Lagace, R. Scott, L. Amininejad, S. Bumpstead, L. Baidoo, R. N. Baldassano, M. Barclay, T. M. Bayless, S. Brand, C. Buning, J. F. Colombel, L. A. Denson, M. De Vos, M. Dubinsky, C. Edwards, D. Ellinghaus, R. S. Fehrmann, J. A. Floyd, T. Florin, D. Franchimont, L. Franke, M. Georges, J. Glas, N. L. Glazer, S. L. Guthery, T. Haritunians, N. K. Hayward, J. P. Hugot, G. Jobin, D. Laukens, I. Lawrance, M. Lemann, A. Levine, C. Libioulle, E. Louis, D. P. McGovern, M. Milla, G. W. Montgomery, K. I. Morley, C. Mowat, A. Ng, W. Newman, R. A. Ophoff, L. Papi, O. Palmieri, L. Peyrin-Biroulet, J. Panes, A. Phillips, N. J. Prescott, D. D. Proctor, R. Roberts, R. Russell, P. Rutgeerts, J. Sanderson, M. Sans, P. Schumm, F. Seibold, Y. Sharma, L. A. Simms, M. Seielstad, A. H. Steinhart, S. R. Targan, L. H. van den Berg, M. Vatn, H. Verspaget, T. Walters, C. Wijmenga, D. C. Wilson, H. J. Westra, R. J. Xavier, Z. Z. Zhao, C. Y. Ponsioen, V. Andersen, L. Torkvist, M. Gazouli, N. P. Anagnou, T. H. Karlsen, L. Kupcinkas, J. Sventoraityte, J. C. Mansfield, S. Kugathasan, M. S. Silverberg, J. Halfvarson, J. I. Rotter, C. G. Mathew, A. M. Griffiths, R. Gearry, T. Ahmad, S. R. Brant, M. Chamaillard, J. Satsangi, J. H. Cho, S. Schreiber, M. J. Daly, J. C. Barrett, M. Parkes, V. Annese, H. Hakonarson, G. Radford-Smith, R. H. Duerr, S. Vermeire, R. K. Weersma and J. D. Rioux (2011). "Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47." Nat Genet **43**(3): 246-252.
- Athar, A., A. Fullgrabe, N. George, H. Iqbal, L. Huerta, A. Ali, C. Snow, N. A. Fonseca, R. Petryszak, I. Papatheodorou, U. Sarkans and A. Brazma (2019). "ArrayExpress update - from bulk to single-cell expression data." Nucleic Acids Res **47**(D1): D711-D715.
- Avery, O. T., C. M. Macleod and M. McCarty (1944). "STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES : INDUCTION OF TRANSFORMATION BY A DESOXYRIBONUCLEIC ACID FRACTION ISOLATED FROM PNEUMOCOCCUS TYPE III." The Journal of experimental medicine **79**(2): 137-158.
- Babu, M. M. (2004). An introduction to microarray data analysis Computational Genomics: Theory and Application R. P. Grant. Cambridge, UK, Horizon Bioscience 225-249.

- Bagos, P. G. (2008). "A unification of multivariate methods for meta-analysis of genetic association studies." Stat Appl Genet Mol Biol **7**(1): Article31.
- Bagos, P. G. (2013). "Genetic model selection in genome-wide association studies: robust methods and the use of meta-analysis." Stat Appl Genet Mol Biol **12**(3): 285-308.
- Baird, D. (2012). Analysis of Microarray Data. GenStat for Windows. GenSta, VSN International, 5 The Waterhouse, Waterhouse Street, Hemel Hempstead, Hertfordshire HP1 1ES, UK.
- Baldi, P. and A. D. Long (2001). "A Bayesian framework for the analysis of microarray expression data: regularized t -test and statistical inferences of gene changes." Bioinformatics **17**(6): 509-519.
- Baltimore, D. (1970). "RNA-dependent DNA polymerase in virions of RNA tumour viruses." Nature **226**(5252): 1209-1211.
- Barrett, T. and R. Edgar (2006). "Mining Microarray Data at NCBI's Gene Expression Omnibus (GEO)." Methods in molecular biology (Clifton, N.J.) **338**: 175-190.
- Belo, L., D. Gaffney, M. Caslake, A. Santos-Silva, L. Pereira-Leite, A. Quintanilha and I. Rebelo (2004). "Apolipoprotein E and cholesteryl ester transfer protein polymorphisms in normal and preeclamptic pregnancies." Eur J Obstet Gynecol Reprod Biol **112**(1): 9-15.
- Benjamini, Y. and Y. Hochberg (1995). "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." Journal of the Royal Statistical Society. Series B (Methodological) **57**(1): 289-300.
- Bennett, M., K. Macdonald, S. W. Chan, J. P. Luzio, R. Simari and P. Weissberg (1998). "Cell surface trafficking of Fas: a rapid mechanism of p53-mediated apoptosis." Science **282**(5387): 290-293.
- Bodnar, L. M., J. M. Catov and J. M. Roberts (2007). "Racial/ethnic differences in the monthly variation of preeclampsia incidence." Am J Obstet Gynecol **196**(4): 324 e321-325.
- Brazma, A. and J. Vilo (2000). "Gene expression data analysis." FEBS Lett **480**(1): 17-24.
- Breitling, R., P. Armengaud, A. Amtmann and P. Herzyk (2004). "Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments." FEBS Lett **573**(1-3): 83-92.
- Brew, O., M. H. Sullivan and A. Woodman (2016). "Comparison of Normal and Pre-Eclamptic Placental Gene Expression: A Systematic Review with Meta-Analysis." PLoS One **11**(8): e0161504.
- Brocke-Heidrich, K., B. Ge, H. Cvijic, G. Pfeifer, D. Loffler, C. Henze, T. W. McKeithan and F. Horn (2006). "BCL3 is induced by IL-6 via Stat3 binding to intronic enhancer HS4 and represses its own transcription." Oncogene **25**(55): 7297-7304.
- Brown, M. A., W. M. Hague, J. Higgins, S. Lowe, L. McCowan, J. Oats, M. J. Peek, J. A. Rowan, B. N. Walters and P. Australasian Society of the Study of Hypertension in (2000). "The

- detection, investigation and management of hypertension in pregnancy: full consensus statement." Aust N Z J Obstet Gynaecol **40**(2): 139-155.
- Burczynski, M. E., R. L. Peterson, N. C. Twine, K. A. Zuberek, B. J. Brodeur, L. Casciotti, V. Maganti, P. S. Reddy, A. Strahs, F. Immermann, W. Spinelli, U. Schwertschlag, A. M. Slager, M. M. Cotreau and A. J. Dorner (2006). "Molecular classification of Crohn's disease and ulcerative colitis patients using transcriptional profiles in peripheral blood mononuclear cells." J Mol Diagn **8**(1): 51-61.
- Campain, A. and Y. H. Yang (2010). "Comparison study of microarray meta-analysis methods." BMC Bioinformatics **11**(1): 408.
- Chappell, S. and L. Morgan (2006). "Searching for genetic clues to the causes of pre-eclampsia." Clin Sci (Lond) **110**(4): 443-458.
- Choi, J. K., U. Yu, S. Kim and O. J. Yoo (2003). "Combining multiple microarray studies and modeling interstudy variation." Bioinformatics **19 Suppl 1**: i84-90.
- Choudhuri, S. (2004). "Microarrays in biology and medicine." J Biochem Mol Toxicol **18**(4): 171-179.
- Cobb, M. (2017). "60 years ago, Francis Crick changed the logic of biology." PLoS biology **15**(9): e2003243-e2003243.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. 2nd edn. Hillsdale, New Jersey: L, Erlbaum.
- Colell, A., J. E. Ricci, S. Tait, S. Milasta, U. Maurer, L. Bouchier-Hayes, P. Fitzgerald, A. Guio-Carrion, N. J. Waterhouse, C. W. Li, B. Mari, P. Barbry, D. D. Newmeyer, H. M. Beere and D. R. Green (2007). "GAPDH and autophagy preserve survival after apoptotic cytochrome c release in the absence of caspase activation." Cell **129**(5): 983-997.
- Contro, E., D. Bernabini and A. Farina (2017). "Cell-Free Fetal DNA for the Prediction of Pre-Eclampsia at the First and Second Trimesters: A Systematic Review and Meta-Analysis." Mol Diagn Ther **21**(2): 125-135.
- Coss, A., M. Tosetto, E. J. Fox, B. Sapetto-Rebow, S. Gorman, B. N. Kennedy, A. T. Lloyd, J. M. Hyland, D. P. O'Donoghue, K. Sheahan, D. T. Leahy, H. E. Mulcahy and J. N. O'Sullivan (2009). "Increased topoisomerase IIalpha expression in colorectal cancer is associated with advanced disease and chemotherapeutic resistance via inhibition of apoptosis." Cancer Lett **276**(2): 228-238.
- Crick, F. (1970). "Central Dogma of Molecular Biology." Nature **227**(5258): 561-563.
- D'Haens, G., F. Rieder, B. G. Feagan, P. D. R. Higgins, J. Panes, C. Maaser, G. Rogler, M. Lowenberg, R. van der Voort, M. Pinzani, L. Peyrin-Biroulet and S. Danese (2019). "Challenges in the Pathophysiology, Diagnosis and Management of Intestinal Fibrosis in Inflammatory Bowel Disease." Gastroenterology.
- Danese, S. and A. Mantovani (2010). "Inflammatory bowel disease and intestinal cancer: a paradigm of the Yin-Yang interplay between inflammation and cancer." Oncogene **29**(23): 3313-3323.

- Danese, S., A. Papa, S. Saibeni, A. Repici, A. Malesci and M. Vecchi (2007). "Inflammation and coagulation in inflammatory bowel disease: The clot thickens." Am J Gastroenterol **102**(1): 174-186.
- de Bruyn, M. and M. Ferrante (2018). "Failure of MMP-9 antagonists in IBD: demonstrating the importance of molecular biology and well-controlled preclinical studies." J Crohns Colitis.
- de Lange, K. M., L. Moutsianas, J. C. Lee, C. A. Lamb, Y. Luo, N. A. Kennedy, L. Jostins, D. L. Rice, J. Gutierrez-Achury, S. G. Ji, G. Heap, E. R. Nimmo, C. Edwards, P. Henderson, C. Mowat, J. Sanderson, J. Satsangi, A. Simmons, D. C. Wilson, M. Tremelling, A. Hart, C. G. Mathew, W. G. Newman, M. Parkes, C. W. Lees, H. Uhlig, C. Hawkey, N. J. Prescott, T. Ahmad, J. C. Mansfield, C. A. Anderson and J. C. Barrett (2017). "Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease." Nat Genet **49**(2): 256-261.
- de Voer, R. M., A. Geurts van Kessel, R. D. Weren, M. J. Ligtenberg, D. Smeets, L. Fu, L. Vreede, E. J. Kamping, E. T. Verwiel, M. M. Hahn, M. Ariaans, L. Spruijt, T. van Essen, G. Houge, H. K. Schackert, J. Q. Sheng, H. Venselaar, C. M. van Ravenswaaij-Arts, J. H. van Krieken, N. Hoogerbrugge and R. P. Kuiper (2013). "Germline mutations in the spindle assembly checkpoint genes BUB1 and BUB3 are risk factors for colorectal cancer." Gastroenterology **145**(3): 544-547.
- De Vries, L. C. S., M. E. Wildenberg, W. J. De Jonge and G. R. D'Haens (2017). "The Future of Janus Kinase Inhibitors in Inflammatory Bowel Disease." J Crohns Colitis **11**(7): 885-893.
- DerSimonian, R. and N. Laird (1986). "Meta-analysis in clinical trials." Control Clin Trials **7**(3): 177-188.
- Dombroski, B. A., Q. Feng, S. L. Mathias, D. M. Sassaman, A. F. Scott, H. H. Kazazian, Jr. and J. D. Boeke (1994). "An in vivo assay for the reverse transcriptase of human retrotransposon L1 in *Saccharomyces cerevisiae*." Mol Cell Biol **14**(7): 4485-4492.
- Duan, B., R. Davis, E. L. Sadat, J. Collins, P. C. Sternweis, D. Yuan and L. I. Jiang (2010). "Distinct roles of adenylyl cyclase VII in regulating the immune responses in mice." J Immunol **185**(1): 335-344.
- Dudoit, S., Y. H. Yang, M. J. Callow and T. P. Speed (2000). "Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments." Statistica Sinica **12**(1): 111-139.
- Efron, B. (1982). The jackknife, the bootstrap and other resampling plans, SIAM.
- Efron, B. and R. Tibshirani (1993). An Introduction to the Bootstrap. Boca Raton, FL, Chapman & Hall/CRC.
- El-Sayed, A. A. F. (2017). "Preeclampsia: A review of the pathogenesis and possible management strategies based on its pathophysiological derangements." Taiwan J Obstet Gynecol **56**(5): 593-598.
- Ellinghaus, D., L. Jostins, S. L. Spain, A. Cortes, J. Bethune, B. Han, Y. R. Park, S. Raychaudhuri, J. G. Pouget, M. Hubenthal, T. Folseraas, Y. Wang, T. Esko, A. Metspalu, H. J. Westra, L. Franke, T. H. Pers, R. K. Weersma, V. Collij, M. D'Amato, J. Halfvarson, A. B. Jensen,

- W. Lieb, F. Degenhardt, A. J. Forstner, A. Hofmann, S. Schreiber, U. Mrowietz, B. D. Juran, K. N. Lazaridis, S. Brunak, A. M. Dale, R. C. Trembath, S. Weidinger, M. Weichenthal, E. Ellinghaus, J. T. Elder, J. N. Barker, O. A. Andreassen, D. P. McGovern, T. H. Karlsen, J. C. Barrett, M. Parkes, M. A. Brown and A. Franke (2016). "Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci." Nat Genet **48**(5): 510-518.
- Fisher, R. A. (1934). "Statistical methods for research workers." Statistical methods for research workers.(5th Ed).
- Fontana, L., J. C. Eagon, M. E. Trujillo, P. E. Scherer and S. Klein (2007). "Visceral fat adipokine secretion is associated with systemic inflammation in obese humans." Diabetes **56**(4): 1010-1013.
- Fox, R. J. and M. W. Dimmic (2006). "A two-sample Bayesian t-test for microarray data." BMC Bioinformatics **7**: 126.
- Freeman, W. M., D. J. Robertson and K. E. Vrana (2000). "Fundamentals of DNA hybridization arrays for gene expression analysis." Biotechniques **29**(5): 1042-1046, 1048-1055.
- Garlatti, V., F. Ungaro, S. Spanò, S. D'Alessio and S. Danese (2019). "OP33 BUB1: a new player in the development of Crohn's disease (CD)-associated fibrosis." Journal of Crohn's and Colitis **13**(Supplement_1): S612-S612.
- Gentleman, R., M. Ruschhaupt, W. Huber and L. Lusa (2008). Meta-analysis for microarray experiments, Bioconductor.
- Gifford, R. W., P. A. August and G. Cunningham (2000). "Report of the National High Blood Pressure Education Program Working Group on High Blood Pressure in Pregnancy." Am J Obstet Gynecol **183**(1): S1-S22.
- Gleser, L. J. and I. Olkin. (2007). "Stochastically Dependent Effect Sizes.", from <https://statistics.stanford.edu/research/stochastically-dependent-effect-sizes>.
- Gnjatic, S., Y. Cao, U. Reichelt, E. F. Yekebas, C. Nolker, A. H. Marx, A. Erbersdobler, H. Nishikawa, Y. Hildebrandt, K. Bartels, C. Horn, T. Stahl, I. Gout, V. Filonenko, K. L. Ling, V. Cerundolo, T. Luetkens, G. Ritter, K. Friedrichs, R. Leuwer, S. Hegewisch-Becker, J. R. Izbicki, C. Bokemeyer, L. J. Old and D. Atanackovic (2010). "NY-CO-58/KIF2C is overexpressed in a variety of solid tumors and induces frequent T cell responses in patients with colorectal cancer." Int J Cancer **127**(2): 381-393.
- Goos, J. A., V. M. Coupe, B. Diosdado, P. M. Delis-Van Diemen, C. Karga, J. A. Belien, B. Carvalho, M. P. van den Tol, H. M. Verheul, A. A. Geldof, G. A. Meijer, O. S. Hoekstra and R. J. Fijneman (2013). "Aurora kinase A (AURKA) expression in colorectal cancer liver metastasis is associated with poor prognosis." Br J Cancer **109**(9): 2445-2452.
- Gray-Owen, S. D. and R. S. Blumberg (2006). "CEACAM1: contact-dependent control of immunity." Nat Rev Immunol **6**(6): 433-446.
- Greenland, S. (1998). Meta-analysis. Modern Epidemiology. K. J. Rothman and S. Greenland, Lippincott Williams & Wilkins: 643-673.

- Grill, S., C. Rusterholz, R. Zanetti-Dallenbach, S. Tercanli, W. Holzgreve, S. Hahn and O. Lapaire (2009). "Potential markers of preeclampsia--a review." Reprod Biol Endocrinol **7**: 70.
- Guo, A., P. Salomoni, J. Luo, A. Shih, S. Zhong, W. Gu and P. P. Pandolfi (2000). "The function of PML in p53-dependent apoptosis." Nat Cell Biol **2**(10): 730-736.
- Guo, L., S. Q. Tsai, N. E. Hardison, A. H. James, A. A. Motsinger-Reif, B. Thames, E. A. Stone, C. Deng and J. A. Piedrahita (2013). "Differentially expressed microRNAs and affected biological pathways revealed by modulated modularity clustering (MMC) analysis of human preeclamptic and IUGR placentas." Placenta **34**(7): 599-605.
- Hayes, D. N. and M. Meyerson (2005). Microarray Approaches to Gene Expression Analysis. Molecular Diagnostics: For the Clinical Laboratorian. W. B. Coleman and G. J. Tsongalis. Totowa, NJ, Humana Press: 121-148.
- Hedges, L. V. (1981). "Distribution Theory for Glass's Estimator of Effect Size and Related Estimators." Journal of Educational Statistics **6**(2): 107-128.
- Hedges, L. V. and I. Olkin (1985). Statistical methods for meta-analysis, Orlando: Academic Press.
- Heikkinen, S., M. Pietila, M. Halmekyto, S. Suppola, E. Pirinen, S. S. Deeb, J. Janne and M. Laakso (1999). "Hexokinase II-deficient mice. Prenatal death of homozygotes without disturbances in glucose tolerance in heterozygotes." J Biol Chem **274**(32): 22517-22523.
- Henderson, J. T., E. P. Whitlock, E. O'Connor, C. A. Senger, J. H. Thompson and M. G. Rowland (2014). "Low-dose aspirin for prevention of morbidity and mortality from preeclampsia: a systematic evidence review for the U.S. Preventive Services Task Force." Ann Intern Med **160**(10): 695-703.
- Hershey, A. D. and M. Chase (1952). "Independent functions of viral protein and nucleic acid in growth of bacteriophage." J Gen Physiol **36**(1): 39-56.
- Higgins, J. P., S. G. Thompson, J. J. Deeks and D. G. Altman (2003). "Measuring inconsistency in meta-analyses." Bmj **327**(7414): 557-560.
- Higgins, J. P. and A. Whitehead (1996). "Borrowing strength from external trials in a meta-analysis." Stat Med **15**(24): 2733-2749.
- Hodgkin, J. (1998). "Seven types of pleiotropy." Int J Dev Biol **42**(3): 501-505.
- Hoheisel, J. D. (2006). "Microarray technology: beyond transcript profiling and genotype analysis." Nat Rev Genet **7**(3): 200-210.
- Holden, M. J. and L. Wang (2008). Quantitative Real-Time PCR: Fluorescent Probe Options and Issues. Standardization and Quality Assurance in Fluorescence Measurements II: Bioanalytical and Biomedical Applications. U. Resch-Genger. Berlin, Heidelberg, Springer Berlin Heidelberg: 489-508.
- Holland, B. and M. Copenhaver (1987). "An Improved Sequentially Rejective Bonferroni Test Procedure." Biometrics **43**(2): 417.
- Holm, S. (1979). "A simple sequentially rejective multiple test procedure." Scandinavian Journal of Statistics **6**: 65-70.

- Hong, F. and R. Breitling (2008). "A comparison of meta-analysis methods for detecting differentially expressed genes in microarray experiments." *Bioinformatics* **24**(3): 374-382.
- Hunter, C. A. (2005). "New IL-12-family members: IL-23 and IL-27, cytokines with divergent functions." *Nat Rev Immunol* **5**(7): 521-531.
- Ihnatova, I. (2013). "MAMA: an R package for Meta-Analysis of MicroArray." *R package version* **2**(1).
- Ikeo, K., J. Ishi-i, T. Tamura, T. Gojobori and Y. Tateno (2003). "CIBEX: center for information biology gene expression database." *C R Biol* **326**(10-11): 1079-1082.
- Information, N. C. f. B. (2007). "PubMed help."
- Innocente, S. A., J. L. Abrahamson, J. P. Cogswell and J. M. Lee (1999). "p53 regulates a G2 checkpoint through cyclin B1." *Proc Natl Acad Sci U S A* **96**(5): 2147-2152.
- Inouye, S., P. J. Herzer and M. Inouye (1990). "Two independent retrons with highly diverse reverse transcriptases in *Myxococcus xanthus*." *Proc Natl Acad Sci U S A* **87**(3): 942-945.
- Ioannidis, J. P., D. B. Allison, C. A. Ball, I. Coulibaly, X. Cui, A. C. Culhane, M. Falchi, C. Furlanello, L. Game and G. Jurman (2009). "Repeatability of published microarray gene expression analyses." *Nature genetics* **41**(2): 149-155.
- Jackson, D., R. Riley and I. R. White (2011). "Multivariate meta-analysis: Potential and promise." *Stat Med*.
- Jebbink, J. M., R. G. Boot, R. Keijser, P. D. Moerland, J. Aten, G. J. Veenboer, M. van Wely, M. Buimer, E. Ver Loren van Themaat, J. M. Aerts, J. A. van der Post, G. B. Afink and C. Ris-Stalpers (2015). "Increased glucocerebrosidase expression and activity in preeclamptic placenta." *Placenta* **36**(2): 160-169.
- Jiang, W. and R. Simon (2007). "A comparison of bootstrap methods and an adjusted bootstrap approach for estimating the prediction error in microarray classification." *Stat Med* **26**(29): 5320-5334.
- Julia, A., E. Domenech, M. Chaparro, V. Garcia-Sanchez, F. Gomollon, J. Panes, M. Manosa, M. Barreiro-De Acosta, A. Gutierrez, E. Garcia-Planella, M. Aguas, F. Munoz, M. Esteve, J. L. Mendoza, M. Vera, L. Marquez, R. Tortosa, M. Lopez-Lasanta, A. Alonso, J. L. Gelpi, A. C. Garcia-Montero, J. Bertranpetit, D. Absher, R. M. Myers, J. P. Gisbert and S. Marsal (2014). "A genome-wide association study identifies a novel locus at 6q22.1 associated with ulcerative colitis." *Hum Mol Genet* **23**(25): 6927-6934.
- Kawasaki, K., E. Kondoh, Y. Chigusa, M. Ujita, R. Murakami, H. Mogami, J. B. Brown, Y. Okuno and I. Konishi (2015). "Reliable pre-eclampsia pathways based on multiple independent microarray data sets." *Mol Hum Reprod* **21**(2): 217-224.
- Kenny, E. E., I. Pe'er, A. Karban, L. Ozelius, A. A. Mitchell, S. M. Ng, M. Erazo, H. Ostrer, C. Abraham, M. T. Abreu, G. Atzmon, N. Barzilai, S. R. Brant, S. Bressman, E. R. Burns, Y. Chowers, L. N. Clark, A. Darvasi, D. Doheny, R. H. Duerr, R. Eliakim, N. Giladi, P. K. Gregersen, H. Hakonarson, M. R. Jones, K. Marder, D. P. McGovern, J. Mulle, A. Orr-Urtreger, D. D. Proctor, A. Pulver, J. I. Rotter, M. S. Silverberg, T. Ullman, S. T. Warren,

- M. Waterman, W. Zhang, A. Bergman, L. Mayer, S. Katz, R. J. Desnick, J. H. Cho and I. Peter (2012). "A genome-wide scan of Ashkenazi Jewish Crohn's disease suggests novel susceptibility loci." PLoS Genet **8**(3): e1002559.
- Khor, B., A. Gardet and R. J. Xavier (2011). "Genetics and pathogenesis of inflammatory bowel disease." Nature **474**(7351): 307-317.
- Klug, W. S., M. R. Cummings, C. A. Spencer and M. A. Palladino (2015). Concepts of Genetics, 11th Edition.
- Konecna, B., B. Vlkova and P. Celec (2015). "Role of fetal DNA in preeclampsia (review)." Int J Mol Med **35**(2): 299-304.
- Kontou, P. I., A. Pavlopoulou and P. G. Bagos (2018). "Methods of Analysis and Meta-Analysis for Identifying Differentially Expressed Genes." Methods Mol Biol **1793**: 183-210.
- Koszalka, P., B. Ozuyaman, Y. Huo, A. Zerneck, U. Flogel, N. Braun, A. Buchheiser, U. K. Decking, M. L. Smith, J. Sevigny, A. Gear, A. A. Weber, A. Molojavyi, Z. Ding, C. Weber, K. Ley, H. Zimmermann, A. Godecke and J. Schrader (2004). "Targeted disruption of cd73/ecto-5'-nucleotidase alters thromboregulation and augments vascular inflammatory response." Circ Res **95**(8): 814-821.
- Leavey, K., S. J. Benton, D. Grynspan, J. C. Kingdom, S. A. Bainbridge and B. J. Cox (2016). "Unsupervised Placental Gene Expression Profiling Identifies Clinically Relevant Subclasses of Human Preeclampsia." Hypertension **68**(1): 137-147.
- Lee, H. K., A. K. Hsu, J. Sajdak, J. Qin and P. Pavlidis (2004). "Coexpression analysis of human genes across many microarray data sets." Genome Res **14**(6): 1085-1094.
- Liang, M., J. Niu, L. Zhang, H. Deng, J. Ma, W. Zhou, D. Duan, Y. Zhou, H. Xu and L. Chen (2016). "Gene expression profiling reveals different molecular patterns in G-protein coupled receptor signaling pathways between early- and late-onset preeclampsia." Placenta **40**: 52-59.
- Liang, P. and A. B. Pardee (1992). "Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction." Science **257**(5072): 967-971.
- Liu, J. Z., S. van Sommeren, H. Huang, S. C. Ng, R. Alberts, A. Takahashi, S. Ripke, J. C. Lee, L. Jostins, T. Shah, S. Abedian, J. H. Cheon, J. Cho, N. E. Dayani, L. Franke, Y. Fuyuno, A. Hart, R. C. Juyal, G. Juyal, W. H. Kim, A. P. Morris, H. Poustchi, W. G. Newman, V. Midha, T. R. Orchard, H. Vahedi, A. Sood, J. Y. Sung, R. Malekzadeh, H. J. Westra, K. Yamazaki, S. K. Yang, J. C. Barrett, B. Z. Alizadeh, M. Parkes, T. Bk, M. J. Daly, M. Kubo, C. A. Anderson and R. K. Weersma (2015). "Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations." Nat Genet **47**(9): 979-986.
- Lobo, I. (2008). "Pleiotropy: One Gene Can Affect Multiple Traits." Nature Education **1**(1): 10.
- Loebbermann, J., H. Thornton, L. Durant, T. Sparwasser, K. E. Webster, J. Sprent, F. J. Culley, C. Johansson and P. J. Openshaw (2012). "Regulatory T cells expressing granzyme B play a critical role in controlling lung inflammation during acute viral infection." Mucosal Immunol **5**(2): 161-172.

- Longo, D. L. and T. R. Harrison (2012). Harrison's principles of internal medicine. New York, N.Y., McGraw-Hill Medical.
- Lopez-Jaramillo, P., J. Barajas, S. M. Rueda-Quijano, C. Lopez-Lopez and C. Felix (2018). "Obesity and Preeclampsia: Common Pathophysiological Mechanisms." Front Physiol **9**: 1838.
- Lord, S. J., R. V. Rajotte, G. S. Korbitt and R. C. Bleackley (2003). "Granzyme B: a natural born killer." Immunol Rev **193**: 31-38.
- Louwen, F., C. Muschol-Steinmetz, J. Reinhard, A. Reitter and J. Yuan (2012). "A lesson for cancer research: placental microarray gene analysis in preeclampsia." Oncotarget **3**(8): 759-773.
- Lu, G. and A. Ades (2004). "Combination of direct and indirect evidence in mixed treatment comparisons." Statistics in medicine **23**(20): 3105-3124.
- Lukk, M., M. Kapushesky, J. Nikkilä, H. Parkinson, A. Goncalves, W. Huber, E. Ukkonen and A. Brazma (2010). "A global map of human gene expression." Nature biotechnology **28**(4): 322-324.
- Ma, X. (2001). "TNF-alpha and IL-12: a balancing act in macrophage functioning." Microbes Infect **3**(2): 121-129.
- Manning, A. K., M. LaValley, C. T. Liu, K. Rice, P. An, Y. Liu, I. Miljkovic, L. Rasmussen-Torvik, T. B. Harris, M. A. Province, I. B. Borecki, J. C. Florez, J. B. Meigs, L. A. Cupples and J. Dupuis (2011). "Meta-analysis of gene-environment interaction: joint estimation of SNP and SNP x environment regression coefficients." Genet Epidemiol **35**(1): 11-18.
- Mantovani, A., P. Allavena, A. Sica and F. Balkwill (2008). "Cancer-related inflammation." Nature **454**(7203): 436-444.
- Marot, G. and R. Bruyere (2015). Using metaMA for differential gene expression analysis from multiple studies.
- Marot, G., J.-L. Foulley, C.-D. Mayer and F. Jaffrézic (2009). "Moderated effect size and P-value combinations for microarray meta-analyses." Bioinformatics **25**(20): 2692-2699.
- Mavridis, D. and G. Salanti (2012). "A practical introduction to multivariate meta-analysis." Stat Methods Med Res.
- McCarthy, B. J. and J. J. Holland (1965). "Denatured DNA as a direct template for in vitro protein synthesis." Proc Natl Acad Sci U S A **54**(3): 880-886.
- Medina-Gomez, C., J. P. Kemp, N. L. Dimou, E. Kreiner, A. Chesi, B. S. Zemel, K. Bonnelykke, C. G. Boer, T. S. Ahluwalia, H. Bisgaard, E. Evangelou, D. H. M. Heppel, L. F. Bonewald, J. P. Gorski, M. Ghanbari, S. Demissie, G. Duque, M. T. Maurano, D. P. Kiel, Y. H. Hsu, C. J. v. d. E. B, C. Ackert-Bicknell, S. Reppe, K. M. Gautvik, T. Raastad, D. Karasik, J. van de Peppel, V. W. V. Jaddoe, A. G. Uitterlinden, J. H. Tobias, S. F. A. Grant, P. G. Bagos, D. M. Evans and F. Rivadeneira (2017). "Bivariate genome-wide association meta-analysis of pediatric musculoskeletal traits reveals pleiotropic effects at the SREBF1/TOM1L2 locus." Nat Commun **8**(1): 121.

- Meng, T., H. Chen, M. Sun, H. Wang, G. Zhao and X. Wang (2012). "Identification of differential gene expression profiles in placentas from preeclamptic pregnancies versus normal pregnancies by DNA microarrays." OMICS **16**(6): 301-311.
- Meuwissen, T. H. and M. E. Goddard (2004). "Bootstrapping of gene-expression data improves and controls the false discovery rate of differentially expressed genes." Genet Sel Evol **36**(2): 191-205.
- Meuwissen, T. H. E. and M. E. Goddard (2004). "Bootstrapping of gene-expression data improves and controls the false discovery rate of differentially expressed genes." Genetics Selection Evolution **36**(2): 191.
- Moher, D., A. Liberati, J. Tetzlaff and D. G. Altman (2009). "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement." PLoS Med **6**(7): e1000097.
- Monteleone, G., L. Biancone, R. Marasco, G. Morrone, O. Marasco, F. Luzzza and F. Pallone (1997). "Interleukin 12 is expressed and actively released by Crohn's disease intestinal lamina propria mononuclear cells." Gastroenterology **112**(4): 1169-1178.
- Morgan, M. E., P. J. Koelink, B. Zheng, M. H. den Brok, H. J. van de Kant, H. W. Verspaget, G. Folkerts, G. J. Adema and A. D. Kraneveld (2014). "Toll-like receptor 6 stimulation promotes T-helper 1 and 17 responses in gastrointestinal-associated lymphoid tissue and modulates murine experimental colitis." Mucosal Immunol **7**(5): 1266-1277.
- Moslehi, R., J. L. Mills, C. Signore, A. Kumar, X. Ambroggio and A. Dzutsev (2013). "Integrative transcriptome analysis reveals dysregulation of canonical cancer molecular pathways in placenta leading to preeclampsia." Sci Rep **3**: 2407.
- Neuhauser, M. and K. H. Jockel (2006). "A bootstrap test for the analysis of microarray experiments with a very small number of replications." Appl Bioinformatics **5**(3): 173-179.
- Nishizawa, H., S. Ota, M. Suzuki, T. Kato, T. Sekiya, H. Kurahashi and Y. Udagawa (2011). "Comparative gene expression profiling of placentas from patients with severe pre-eclampsia and unexplained fetal growth restriction." Reprod Biol Endocrinol **9**: 107.
- Nishizawa, H., K. Pryor-Koishi, T. Kato, H. Kowa, H. Kurahashi and Y. Udagawa (2007). "Microarray analysis of differentially expressed fetal genes in placental tissue derived from early and late onset severe pre-eclampsia." Placenta **28**(5-6): 487-497.
- Nobakht, M. G. B. F. (2018). "Application of metabolomics to preeclampsia diagnosis." Syst Biol Reprod Med **64**(5): 324-339.
- Normand, S. L. (1999). "Meta-analysis: formulating, evaluating, combining, and reporting." Stat Med **18**(3): 321-359.
- Novacek, G., A. Weltermann, A. Sobala, H. Tilg, W. Petritsch, W. Reinisch, A. Mayer, T. Haas, A. Kaser, T. Feichtenschlager, H. Fuchssteiner, P. Knoflach, H. Vogelsang, W. Miehsler, R. Platzer, W. Tillinger, B. Jaritz, A. Schmid, B. Blaha, C. Dejaco and S. Eichinger (2010). "Inflammatory bowel disease is a risk factor for recurrent venous thromboembolism." Gastroenterology **139**(3): 779-787, 787.e771.

- Olen, O., J. Askling, M. C. Sachs, P. Frumento, M. Neovius, K. E. Smedby, A. Ekbom, P. Malmberg and J. F. Ludvigsson (2017). "Childhood onset inflammatory bowel disease and risk of cancer: a Swedish nationwide cohort study 1964-2014." *Bmj* **358**: j3951.
- Ostrowski, J., K. Goryca, I. Lazowska, A. Rogowska, A. Paziewska, M. Dabrowska, F. Ambrozkiwicz, J. Karczmariski, A. Balabas, A. Kluska, M. Piatkowska, N. Zeber-Lubecka, M. Kulecka, A. Habor, M. Mikula, B. Walewska-Zielecka, M. Krawczyk, H. Cichoz-Lach, P. Milkiewicz, A. Kowalik, K. Mucha, J. Raczynska, J. Musialik, G. Boryczka, M. Wasilewicz, I. Cieccko-Michalska, M. Ferenc, M. Janiak, A. Kanikowska, R. Stankiewicz, M. Hartleb, T. Mach, M. Grzymislawski, J. Raszeja-Wyszomirska, E. Wunsch, T. Bobinski, J. Kierkus, P. Socha, M. Lodyga, M. Klopocka, B. Iwanczak, K. Bak-Drabik, J. Walkowiak, P. Radwan, U. Grzybowska-Chlebowczyk, B. Korczowski, T. Starzynska, P. B. C. s. G. The Polish and I. B. D. s. G. The Polish (2019). "Common functional alterations identified in blood transcriptome of autoimmune cholestatic liver and inflammatory bowel diseases." *Scientific Reports* **9**(1): 7190.
- Papathanassiou, A. E., J. H. Ko, M. Imprialou, M. Bagnati, P. K. Srivastava, H. A. Vu, D. Cucchi, S. P. McAdoo, E. A. Ananieva, C. Mauro and J. Behmoaras (2017). "BCAT1 controls metabolic reprogramming in activated human macrophages and is associated with inflammatory diseases." *Nat Commun* **8**: 16040.
- Parkes, M., J. C. Barrett, N. J. Prescott, M. Tremelling, C. A. Anderson, S. A. Fisher, R. G. Roberts, E. R. Nimmo, F. R. Cummings, D. Soars, H. Drummond, C. W. Lees, S. A. Khawaja, R. Bagnall, D. A. Burke, C. E. Todhunter, T. Ahmad, C. M. Onnie, W. McArdle, D. Strachan, G. Bethel, C. Bryan, C. M. Lewis, P. Deloukas, A. Forbes, J. Sanderson, D. P. Jewell, J. Satsangi, J. C. Mansfield, L. Cardon and C. G. Mathew (2007). "Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility." *Nat Genet* **39**(7): 830-832.
- Paul, D. (2000). "A double-edged sword." *Nature* **405**(6786): 515-515.
- Pears, S., A. Makris and A. Hennessy (2018). "The chronobiology of blood pressure in pregnancy." *Pregnancy Hypertens* **12**: 104-109.
- Piovani, D., S. Danese, L. Peyrin-Biroulet, G. K. Nikolopoulos, T. Lytras and S. Bonovas (2019). "Environmental Risk Factors for Inflammatory Bowel Diseases: an Umbrella Review of Meta-analyses." *Gastroenterology*.
- Planell, N., M. C. Masamunt, R. F. Leal, L. Rodriguez, M. Esteller, J. J. Lozano, A. Ramirez, M. L. S. Ayrizono, C. S. R. Coy, I. Alfaro, I. Ordas, S. Visvanathan, E. Ricart, J. Guardiola, J. Panes and A. Salas (2017). "Usefulness of Transcriptional Blood Biomarkers as a Non-invasive Surrogate Marker of Mucosal Healing and Endoscopic Response in Ulcerative Colitis." *J Crohns Colitis* **11**(11): 1335-1346.
- Podolsky, D. K. (1991). "Inflammatory bowel disease (1)." *N Engl J Med* **325**(13): 928-937.
- Poole, J. C., L. G. Andrews and T. O. Tollefsbol (2001). "Activity, function, and gene regulation of the catalytic subunit of telomerase (hTERT)." *Gene* **269**(1-2): 1-12.
- Quackenbush, J. (2001). "Computational analysis of microarray data." *Nat Rev Genet* **2**(6): 418-427.

- Ramasamy, A., A. Mondry, C. C. Holmes and D. G. Altman (2008). "Key issues in conducting a meta-analysis of gene expression microarray datasets." PLoS Med **5**(9): e184.
- Rhodes, D. R., J. Yu, K. Shanker, N. Deshpande, R. Varambally, D. Ghosh, T. Barrette, A. Pandey and A. M. Chinnaiyan (2004). "ONCOMINE: a cancer microarray database and integrated data-mining platform." Neoplasia **6**(1): 1-6.
- Riley, R. D. (2009). "Multivariate meta-analysis: the effect of ignoring within-study correlation." Journal of the Royal Statistical Society: Series A (Statistics in Society) **172**(4): 789-811.
- Riley, R. D., K. R. Abrams, P. C. Lambert, A. J. Sutton and J. R. Thompson (2007). "An evaluation of bivariate random-effects meta-analysis for the joint synthesis of two correlated outcomes." Stat Med **26**(1): 78-97.
- Riley, R. D., K. R. Abrams, A. J. Sutton, P. C. Lambert and J. R. Thompson (2007). "Bivariate random-effects meta-analysis and the estimation of between-study correlation." BMC Med Res Methodol **7**: 3.
- Roda, G., S. Dahan, L. Mezzanotte, A. Caponi, F. Roth-Walter, D. Pinn and L. Mayer (2009). "Defect in CEACAM family member expression in Crohn's disease IECs is regulated by the transcription factor SOX9." Inflamm Bowel Dis **15**(12): 1775-1783.
- Rubin, D. C., A. Shaker and M. S. Levin (2012). "Chronic intestinal inflammation: inflammatory bowel disease and colitis-associated colon cancer." Front Immunol **3**: 107.
- Sacks, H. S., J. Berrier, D. Reitman, V. A. Ancona-Berk and T. C. Chalmers (1987). "Meta-analyses of randomized controlled trials." N Engl J Med **316**(8): 450-455.
- Scaldaferri, F., M. Sans, S. Vetrano, C. Correale, V. Arena, N. Pagano, G. Rando, F. Romeo, A. E. Potenza, A. Repici, A. Malesci and S. Danese (2009). "The role of MAPK in governing lymphocyte adhesion to and migration across the microvasculature in inflammatory bowel disease." Eur J Immunol **39**(1): 290-300.
- Schwartz, A. S., J. Yu, K. R. Gardenour, R. L. Finley, Jr. and T. Ideker (2009). "Cost-effective strategies for completing the interactome." Nature methods **6**(1): 55-61.
- Seidel, C. (2008). Introduction to DNA Microarrays, in Analysis of Microarray Data: A Network-Based Approach (eds F. Emmert-Streib and M. Dehmer).
- Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker (2003). "Cytoscape: a software environment for integrated models of biomolecular interaction networks." Genome research **13**(11): 2498-2504.
- Shimshoni, E., D. Yablecovitch, L. Baram, I. Dotan and I. Sagi (2015). "ECM remodelling in IBD: innocent bystander or partner in crime? The emerging role of extracellular molecular events in sustaining intestinal inflammation." Gut **64**(3): 367-372.
- Sibai, B. M. (2003). "Diagnosis and management of gestational hypertension and preeclampsia." Obstet Gynecol **102**(1): 181-192.
- Šidák, Z. (1967). "Rectangular Confidence Regions for the Means of Multivariate Normal Distributions." Journal of the American Statistical Association **62**(318): 626-633.

- Sifakis, S., Z. Koukou and D. A. Spandidos (2015). "Cell-free fetal DNA and pregnancy-related complications (review)." Mol Med Rep **11**(4): 2367-2372.
- Singh, S., I. J. Kullo, D. S. Pardi and E. V. Loftus, Jr. (2015). "Epidemiology, risk factors and management of cardiovascular diseases in IBD." Nat Rev Gastroenterol Hepatol **12**(1): 26-35.
- Sitras, V., C. Fenton and G. Acharya (2015). "Gene expression profile in cardiovascular disease and preeclampsia: a meta-analysis of the transcriptome based on raw data from human studies deposited in Gene Expression Omnibus." Placenta **36**(2): 170-178.
- Sitras, V., R. H. Paulssen, H. Gronaas, J. Leirvik, T. A. Hanssen, A. Vartun and G. Acharya (2009). "Differential placental gene expression in severe preeclampsia." Placenta **30**(5): 424-433.
- Smith, C. L., P. Dickinson, T. Forster, M. Craigon, A. Ross, M. R. Khondoker, R. France, A. Ivens, D. J. Lynn, J. Orme, A. Jackson, P. Lacaze, K. L. Flanagan, B. J. Stenson and P. Ghazal (2014). "Identification of a human neonatal immune-metabolic network associated with bacterial infection." Nat Commun **5**: 4649.
- Smith, C. M., T. F. Hayamizu, J. H. Finger, S. M. Bello, I. J. McCright, J. Xu, R. M. Baldarelli, J. S. Beal, J. Campbell, L. E. Corbani, P. J. Frost, J. R. Lewis, S. C. Giannatto, D. Miers, D. R. Shaw, J. A. Kadin, J. E. Richardson, C. L. Smith and M. Ringwald (2019). "The mouse Gene Expression Database (GXD): 2019 update." Nucleic acids research **47**(D1): D774-D779.
- StataCorp (2013). *Stata 13 Base Reference Manual*, College Station, TX: Stata Press. .
- Stegers, E. A., P. von Dadelszen, J. J. Duvekot and R. Pijnenborg (2010). "Pre-eclampsia." Lancet **376**(9741): 631-644.
- Stevens, J. R. and R. W. Doerge (2005). "Combining Affymetrix microarray results." BMC Bioinformatics **6**: 57.
- Szklarczyk, D., A. L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N. T. Doncheva, J. H. Morris, P. Bork, L. J. Jensen and C. v. Mering (2019). "STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets." Nucleic acids research **47**(D1): D607-D613.
- Team, R. (2015). "R Core Team. R: a language and environment for statistical computing."
- Temin, H. M. and S. Mizutani (1970). "RNA-dependent DNA polymerase in virions of Rous sarcoma virus." Nature **226**(5252): 1211-1213.
- Thakkinstian, A., P. McElduff, C. D'Este, D. Duffy and J. Attia (2005). "A method for meta-analysis of molecular association studies." Stat Med **24**(9): 1291-1306.
- Thomas, P. D., A. Kejariwal, M. J. Campbell, H. Mi, K. Diemer, N. Guo, I. Ladunga, B. Ulitsky-Lazareva, A. Muruganujan, S. Rabkin, J. A. Vandergriff and O. Doremieux (2003). "PANTHER: a browsable database of gene products organized by biological function, using curated protein family and subfamily classification." Nucleic acids research **31**(1): 334-341.

- Trikalinos, T. A., G. Salanti, E. Zintzaras and J. P. Ioannidis (2008). "Meta-analysis methods." *Adv Genet* **60**: 311-334.
- Trinchieri, G., S. Pflanz and R. A. Kastelein (2003). "The IL-12 family of heterodimeric cytokines: new players in the regulation of T cell responses." *Immunity* **19**(5): 641-644.
- Tropp, B. E. (2012). *Principles of Molecular Biology*, Jones & Bartlett Learning.
- Tsai, C. A., Y. J. Chen and J. J. Chen (2003). "Testing for differentially expressed genes with microarray data." *Nucleic Acids Res* **31**(9): e52.
- Tsai, S., N. E. Hardison, A. H. James, A. A. Motsinger-Reif, S. R. Bischoff, B. H. Thames and J. A. Piedrahita (2011). "Transcriptional profiling of human placentas from pregnancies complicated by preeclampsia reveals dysregulation of sialic acid acetyltransferase and immune signalling pathways." *Placenta* **32**(2): 175-182.
- Turner, J. A. (2010). "Diagnosis and management of pre-eclampsia: an update." *Int J Womens Health* **2**: 327-337.
- Uniken Venema, W. T., M. D. Voskuil, A. V. Vila, G. van der Vries, B. H. Jansen, B. Jabri, K. N. Faber, G. Dijkstra, R. J. Xavier, C. Wijmenga, D. B. Graham, R. K. Weersma and E. A. Festen (2019). "Single-Cell RNA Sequencing of Blood and Ileal T Cells From Patients With Crohn's Disease Reveals Tissue-Specific Characteristics and Drug Targets." *Gastroenterology* **156**(3): 812-815.e822.
- Uzawa, T., A. Yamagishi and T. Oshima (2002). "Polypeptide synthesis directed by DNA as a messenger in cell-free polypeptide synthesis by extreme thermophiles, *Thermus thermophilus* HB27 and *Sulfolobus tokodaii* strain 7." *J Biochem* **131**(6): 849-853.
- Vadillo-Ortega, F., O. Perichart-Perera, S. Espino, M. A. Avila-Vergara, I. Ibarra, R. Ahued, M. Godines, S. Parry, G. Macones and J. F. Strauss (2011). "Effect of supplementation during pregnancy with L-arginine and antioxidant vitamins in medical food on pre-eclampsia in high risk population: randomised controlled trial." *BMJ* **342**: d2901.
- Vaiman, D. and F. Miralles (2016). "An Integrative Analysis of Preeclampsia Based on the Construction of an Extended Composite Network Featuring Protein-Protein Physical Interactions and Transcriptional Relationships." *PLoS One* **11**(11): e0165849.
- van Houwelingen, H. C., L. R. Arends and T. Stijnen (2002). "Advanced methods in meta-analysis: multivariate approach and meta-regression." *Stat Med* **21**(4): 589-624.
- van Uiter, M., P. D. Moerland, D. A. Enquobahrie, H. Laivuori, J. A. van der Post, C. Ris-Stalpers and G. B. Afink (2015). "Meta-Analysis of Placental Transcriptome Data Identifies a Novel Molecular Pathway Related to Preeclampsia." *PLoS One* **10**(7): e0132468.
- Velculescu, V. E., L. Zhang, B. Vogelstein and K. W. Kinzler (1995). "Serial analysis of gene expression." *Science* **270**(5235): 484-487.
- Vennou, K. E., P. I. Kontou, G. G. Braliou and P. G. Bagos (2020). "Meta-analysis of gene expression profiles in preeclampsia." *Pregnancy Hypertens* **19**: 52-60.

- Vennou, K. E., D. Piovani, P. I. Kontou, S. Bonovas and P. G. Bagos (2020). "Methods for multiple outcome meta-analysis of gene-expression data." MethodsX **7**: 100834.
- Vennou, K. E., D. Piovani, P. I. Kontou, S. Bonovas and P. G. Bagos (2020). "Multiple outcome meta-analysis of gene-expression data in inflammatory bowel disease." Genomics **112**(2): 1761-1767.
- Ventham, N. T., N. A. Kennedy, A. T. Adams, R. Kalla, S. Heath, K. R. O'Leary, H. Drummond, D. C. Wilson, I. G. Gut, E. R. Nimmo and J. Satsangi (2016). "Integrative epigenome-wide analysis demonstrates that DNA methylation may mediate genetic risk in inflammatory bowel disease." Nat Commun **7**: 13507.
- Villanueva, E. V. and S. Zavarsek (2004). "Evaluating heterogeneity in cumulative meta-analyses." BMC Medical Research Methodology **4**(1): 18.
- Vlkova, B., J. Turna and P. Celec (2015). "Fetal DNA in maternal plasma in preeclamptic pregnancies." Hypertens Pregnancy **34**(1): 36-49.
- Voltolini, C. and F. Petraglia (2014). "Neuroendocrinology of pregnancy and parturition." Handb Clin Neurol **124**: 17-36.
- Waetzig, G. H., D. Seegert, P. Rosenstiel, S. Nikolaus and S. Schreiber (2002). "p38 mitogen-activated protein kinase is activated and linked to TNF-alpha signaling in inflammatory bowel disease." J Immunol **168**(10): 5342-5351.
- Walker, J. J. (2000). "Pre-eclampsia." Lancet **356**(9237): 1260-1265.
- Wang, H., M. Naghavi, C. Allen, R. M. Barber, Z. A. Bhutta, A. Carter, D. C. Casey, F. J. Charlson, A. Z. Chen, M. M. Coates, M. Coggeshall, L. Dandona, D. J. Dicker, H. E. Erskine, A. J. Ferrari, C. Fitzmaurice, K. Foreman, M. H. Forouzanfar, M. S. Fraser, N. Fullman, P. W. Gething, E. M. Goldberg, N. Graetz, J. A. Haagsma, S. I. Hay, C. Huynh, C. O. Johnson, N. J. Kassebaum, Y. Kinfu, X. R. Kulikoff, M. Kutz, H. H. Kyu, H. J. Larson, J. Leung, X. Liang, S. S. Lim, M. Lind, R. Lozano, N. Marquez, G. A. Mensah, J. Mikesell, A. H. Mokdad, M. D. Mooney, G. Nguyen, E. Nsoesie, D. M. Pigott, C. Pinho, G. A. Roth, J. A. Salomon, L. Sandar, N. Silpakit, A. Sligar, R. J. D. Sorensen, J. Stanaway, C. Steiner, S. Teeple, B. A. Thomas, C. Troeger, A. VanderZanden, S. E. Vollset, V. Wang, H. A. Whiteford, T. Wolock, L. Zoeckler, K. H. Abate, C. Abbafati, K. M. Abbas, F. Abd-Allah, S. F. Abera, D. M. X. Abreu, L. J. Abu-Raddad, G. Y. Abyu, T. Achoki, A. L. Adelekan, Z. Ademi, A. K. Adou, J. C. Adsuar, K. A. Afanvi, A. Afshin, E. E. Agardh, A. Agarwal, A. Agrawal, A. A. Kiadaliri, O. N. Ajala, A. S. Akanda, R. O. Akinyemi, T. F. Akinyemiju, N. Akseer, F. H. A. Lami, S. Alabed, Z. Al-Aly, K. Alam, N. K. M. Alam, D. Alasfoor, S. F. Aldhahri, R. W. Aldridge, M. A. Alegretti, A. V. Aleman, Z. A. Alemu, L. T. Alexander, S. Alhabib, R. Ali, A. a. Alkerwi, F. Alla, P. Allebeck, R. Al-Raddadi, U. Alsharif, K. A. Altirkawi, E. A. Martin, N. Alvis-Guzman, A. T. Amare, A. K. Amegah, E. A. Ameh, H. Amini, W. Ammar, S. M. Amrock, H. H. Andersen, B. O. Anderson, G. M. Anderson, C. A. T. Antonio, A. F. Aregay, J. Ärnlöv, V. S. A. Arsenijevic, A. Artaman, H. Asayesh, R. J. Asghar, S. Atique, E. F. G. A. Avokpaho, A. Awasthi, P. Azzopardi, U. Bacha, A. Badawi, M. C. Bahit, K. Balakrishnan, A. Banerjee, A. Barac, S. L. Barker-Collo, T. Bärnighausen, L. Barregard, L. H. Barrero, A. Basu, S. Basu, Y. T. Bayou, S. Bazargan-Hejazi, J. Beardsley, N. Bedi, E. Beghi, H. A. Belay, B. Bell, M. L. Bell, A. K. Bello, D.

A. Bennett, I. M. Bensenor, A. Berhane, E. Bernabé, B. D. Betsu, A. S. Beyene, N. Bhala, A. Bhalla, S. Biadgilign, B. Bikbov, A. A. B. Abdulhak, B. J. Biroscak, S. Biryukov, E. Bjertness, J. D. Blore, C. D. Blosser, M. A. Bohensky, R. Borschmann, D. Bose, R. R. A. Bourne, M. Brainin, C. E. G. Brayne, A. Brazinova, N. J. K. Breitborde, H. Brenner, J. D. Brewer, A. Brown, J. Brown, T. S. Brugha, G. C. Buckle, Z. A. Butt, B. Calabria, I. R. Campos-Nonato, J. C. Campuzano, J. R. Carapetis, R. Cárdenas, D. O. Carpenter, J. J. Carrero, C. A. Castañeda-Orjuela, J. C. Rivas, F. Catalá-López, F. Cavalleri, K. Cercy, J. Cerda, W. Chen, A. Chew, P. P.-C. Chiang, M. Chibalabala, C. E. Chibueze, O. Chimed-Ochir, V. H. Chisumpa, J.-Y. J. Choi, R. Chowdhury, H. Christensen, D. J. Christopher, L. G. Ciobanu, M. Cirillo, A. J. Cohen, V. Colistro, M. Colomar, S. M. Colquhoun, C. Cooper, L. T. Cooper, M. Cortinovis, B. C. Cowie, J. A. Crump, J. Damsere-Derry, H. Danawi, R. Dandona, F. Daoud, S. C. Darby, P. I. Dargan, J. das Neves, G. Davey, A. C. Davis, D. V. Davitoiu, E. F. de Castro, P. de Jager, D. D. Leo, L. Degenhardt, R. P. Dellavalle, K. Deribe, A. Deribew, S. D. Dharmaratne, P. K. Dhillon, C. Diaz-Torné, E. L. Ding, K. P. B. dos Santos, E. Dossou, T. R. Driscoll, L. Duan, M. Dubey, B. B. Duncan, R. G. Ellenbogen, C. L. Ellingsen, I. Elyazar, A. Y. Endries, S. P. Ermakov, B. Eshrati, A. Esteghamati, K. Estep, I. D. A. Faghmous, S. Fahimi, E. J. A. Faraon, T. A. Farid, C. S. e. S. Farinha, A. Faro, M. S. Farvid, F. Farzadfar, V. L. Feigin, S.-M. Fereshtehnejad, J. G. Fernandes, J. C. Fernandes, F. Fischer, J. R. A. Fitchett, A. Flaxman, N. Foigt, F. G. R. Fowkes, E. B. Franca, R. C. Franklin, J. Friedman, J. Frostad, T. Fürst, N. D. Futran, S. L. Gall, K. Gambashidze, A. Gamkrelidze, P. Ganguly, F. G. Gankpé, T. Gebre, T. T. Gebrehiwot, A. T. Gebremedhin, A. A. Gebru, J. M. Geleijnse, B. D. Gessner, A. G. Ghoshal, K. B. Gibney, R. F. Gillum, S. Gilmour, A. Z. Giref, M. Giroud, M. D. Gishu, G. Giussani, E. Glaser, W. W. Godwin, H. Gomez-Dantes, P. Gona, A. Goodridge, S. V. Gopalani, R. A. Gosselin, C. C. Gotay, A. Goto, H. N. Gouda, F. Greaves, H. C. Gugrani, R. Gupta, R. Gupta, V. Gupta, R. A. Gutiérrez, N. Hafezi-Nejad, D. Haile, A. D. Hailu, G. B. Hailu, Y. A. Halasa, R. R. Hamadeh, S. Hamidi, J. Hancock, A. J. Handal, G. J. Hankey, Y. Hao, H. L. Harb, S. Harikrishnan, J. M. Haro, R. Havmoeller, S. R. Heckbert, I. B. Heredia-Pi, P. Heydarpour, H. B. M. Hilderink, H. W. Hoek, R. S. Hogg, M. Horino, N. Horita, H. D. Hosgood, P. J. Hotez, D. G. Hoy, M. Hsairi, A. S. Htet, M. M. T. Htike, G. Hu, C. Huang, H. Huang, L. Huiart, A. Husseini, I. Huybrechts, G. Huynh, K. M. Iburg, K. Innos, M. Inoue, V. J. Iyer, T. A. Jacobs, K. H. Jacobsen, N. Jahanmehr, M. B. Jakovljevic, P. James, M. Javanbakht, S. P. Jayaraman, A. U. Jayatilleke, P. Jeemon, P. N. Jensen, V. Jha, G. Jiang, Y. Jiang, T. Jibat, A. Jimenez-Corona, J. B. Jonas, T. K. Joshi, Z. Kabir, R. Kamal, H. Kan, S. Kant, A. Karch, C. K. Karema, C. Karimkhani, D. Karletsos, G. Karthikeyan, A. Kasaeian, M. Katibeh, A. Kaul, N. Kawakami, J. F. Kayibanda, P. N. Keiyoro, L. Kemmer, A. H. Kemp, A. P. Kengne, A. Keren, M. Kereselidze, C. N. Kesavachandran, Y. S. Khader, I. A. Khalil, A. R. Khan, E. A. Khan, Y.-H. Khang, S. Khera, T. A. M. Khoja, C. Kieling, D. Kim, Y. J. Kim, B. M. Kissela, N. Kissoon, L. D. Knibbs, A. K. Knudsen, Y. Kokubo, D. Kolte, J. A. Kopec, S. Kosen, P. A. Koul, A. Koyanagi, N. H. Krog, B. K. Defo, B. K. Bicer, A. A. Kudom, E. J. Kuipers, V. S. Kulkarni, G. A. Kumar, G. F. Kwan, A. Lal, D. K. Lal, R. Laloo, T. Lallukka, H. Lam, J. O. Lam, S. M. Langan, V. C. Lansingh, A. Larsson, D. O. Laryea, A. A. Latif, A. E. B. Lawrynnowicz, J. Leigh, M. Levi, Y. Li, M. P. Lindsay, S. E. Lipshultz, P. Y. Liu, S. Liu, Y. Liu, L.-T. Lo, G. Logroscino, P. A. Lotufo, R. M. Lucas, R. Lunevicius, R. A. Lyons, S. Ma, V. M. P. Machado, M. T. Mackay, J. H. MacLachlan, H. M. A. E. Razek, M. Magdy,

A. E. Razek, M. Majdan, A. Majeed, R. Malekzadeh, W. A. A. Manamo, J. Mandisarisa, S. Mangalam, C. C. Mapoma, W. Marcenes, D. J. Margolis, G. R. Martin, J. Martinez-Raga, M. B. Marzan, F. Masiye, A. J. Mason-Jones, J. Massano, R. Matzopoulos, B. M. Mayosi, S. T. McGarvey, J. J. McGrath, M. McKee, B. J. McMahan, P. A. Meaney, A. Mehari, M. M. Mehndiratta, F. Mejia-Rodriguez, A. B. Mekonnen, Y. A. Melaku, P. Memiah, Z. A. Memish, W. Mendoza, A. Meretoja, T. J. Meretoja, F. A. Mhimbira, R. Micha, A. Milllear, T. R. Miller, M. Mirarefin, A. Misganaw, C. N. Mock, K. A. Mohammad, A. Mohammadi, S. Mohammed, V. Mohan, G. L. D. Mola, L. Monasta, J. C. M. Hernandez, P. Montero, M. Montico, T. J. Montine, M. Moradi-Lakeh, L. Morawska, K. Morgan, R. Mori, D. Mozaffarian, U. O. Mueller, G. V. S. Murthy, S. Murthy, K. I. Musa, J. B. Nachege, G. Nagel, K. S. Naidoo, N. Naik, L. Naldi, V. Nangia, D. Nash, C. Nejjari, S. Neupane, C. R. Newton, J. N. Newton, M. Ng, F. N. Ngalesoni, J. de Dieu Ngirabega, Q. L. Nguyen, M. I. Nisar, P. M. N. Pete, M. Nomura, O. F. Norheim, P. E. Norman, B. Norrving, L. Nyakarahuka, F. A. Ogbo, T. Ohkubo, F. A. Ojelabi, P. R. Olivares, B. O. Olusanya, J. O. Olusanya, J. N. Opio, E. Oren, A. Ortiz, M. Osman, E. Ota, R. Ozdemir, M. Pa, A. Pain, J. D. Pandian, P. R. Pant, C. Papachristou, E.-K. Park, J.-H. Park, C. D. Parry, M. Parsaeian, A. J. P. Caicedo, S. B. Patten, G. C. Patton, V. K. Paul, N. Pearce, J. M. Pedro, L. P. Stokic, D. M. Pereira, N. Perico, K. Pesudovs, M. Petzold, M. R. Phillips, F. B. Piel, J. D. Pillay, D. Plass, J. A. Platts-Mills, S. Polinder, C. A. Pope, S. Popova, R. G. Poulton, F. Pourmalek, D. Prabhakaran, M. Qorbani, J. Quame-Amaglo, D. A. Quistberg, A. Rafay, K. Rahimi, V. Rahimi-Movaghar, M. Rahman, M. H. U. Rahman, S. U. Rahman, R. K. Rai, Z. Rajavi, S. Rajsic, M. Raju, I. Rakovac, S. M. Rana, C. L. Ranabhat, T. Rangaswamy, P. Rao, S. R. Rao, A. H. Refaat, J. Rehm, M. B. Reitsma, G. Remuzzi, S. Resnikoff, A. L. Ribeiro, S. Ricci, M. J. R. Blancas, B. Roberts, A. Roca, D. Rojas-Rueda, L. Ronfani, G. Roshandel, D. Rothenbacher, A. Roy, N. K. Roy, G. M. Ruhago, R. Sagar, S. Saha, R. Sahathevan, M. M. Saleh, J. R. Sanabria, M. D. Sanchez-Niño, L. Sanchez-Riera, I. S. Santos, R. Sarmiento-Suarez, B. Sartorius, M. Satpathy, M. Savic, M. Sawhney, M. P. Schaub, M. I. Schmidt, I. J. C. Schneider, B. Schöttker, A. E. Schutte, D. C. Schwebel, S. Seedat, S. G. Sepanlou, E. E. Servan-Mori, K. A. Shackelford, G. Shaddick, A. Shaheen, S. Shahraz, M. A. Shaikh, M. Shakh-Nazarova, R. Sharma, J. She, S. Sheikhabaei, J. Shen, Z. Shen, D. S. Shepard, K. N. Sheth, B. P. Shetty, P. Shi, K. Shibuya, M.-J. Shin, R. Shiri, I. Shiue, M. G. Shrimel, I. D. Sigfusdottir, D. H. Silberberg, D. A. S. Silva, D. G. A. Silveira, J. I. Silverberg, E. P. Simard, A. Singh, G. M. Singh, J. A. Singh, O. P. Singh, P. K. Singh, V. Singh, S. Soneji, K. Søreide, J. B. Soriano, L. A. Sposato, C. T. Sreeramareddy, V. Stathopoulou, D. J. Stein, M. B. Stein, S. Stranges, K. Stroupoulis, B. F. Sunguya, P. Sur, S. Swaminathan, B. L. Sykes, C. E. I. Szoeki, R. Tabarés-Seisdedos, K. M. Tabb, K. Takahashi, J. S. Takala, R. T. Talongwa, N. Tandon, M. Tavakkoli, B. Taye, H. R. Taylor, B. J. T. Ao, B. A. Tedla, W. M. Tefera, M. T. Have, A. S. Terkawi, F. H. Tesfay, G. A. Tessema, A. J. Thomson, A. L. Thorne-Lyman, A. G. Thrift, G. D. Thurston, T. Tillmann, D. L. Tirschwell, M. Tonelli, R. Topor-Madry, F. Topouzis, J. A. Towbin, J. Traebert, B. X. Tran, T. Truelsen, U. Trujillo, A. K. Tura, E. M. Tuzcu, U. S. Uchendu, K. N. Ukwaja, E. A. Undurraga, O. A. Uthman, R. V. Dingenen, A. van Donkelaar, T. Vasankari, A. M. N. Vasconcelos, N. Venketasubramanian, R. Vidavalur, L. Vijayakumar, S. Villalpando, F. S. Violante, V. V. Vlassov, J. A. Wagner, G. R. Wagner, M. T. Wallin, L. Wang, D. A. Watkins, S. Weichenthal, E. Weiderpass, R. G. Weintraub, A. Werdecker, R. Westerman, R. A. White, T. Wijeratne, J. D. Wilkinson,

- H. C. Williams, C. S. Wiysonge, S. M. Woldeyohannes, C. D. A. Wolfe, S. Won, J. Q. Wong, A. D. Woolf, D. Xavier, Q. Xiao, G. Xu, B. Yakob, A. Z. Yalew, L. L. Yan, Y. Yano, M. Yaseri, P. Ye, H. G. Yebyo, P. Yip, B. D. Yirsaw, N. Yonemoto, G. Yonga, M. Z. Younis, S. Yu, Z. Zaidi, M. E. S. Zaki, F. Zannad, D. E. Zavala, H. Zeeb, B. M. Zeleke, H. Zhang, S. Zodpey, D. Zonies, L. J. Zuhlke, T. Vos, A. D. Lopez and C. J. L. Murray (2016). "Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980–2013;2015: a systematic analysis for the Global Burden of Disease Study 2015." *The Lancet* **388**(10053): 1459-1544.
- Wang, X., D. D. Kang, K. Shen, C. Song, S. Lu, L. C. Chang, S. G. Liao, Z. Huo, S. Tang, Y. Ding, N. Kaminski, E. Sibille, Y. Lin, J. Li and G. C. Tseng (2012). "An R package suite for microarray meta-analysis in quality control, differentially expressed gene analysis and pathway enrichment detection." *Bioinformatics* **28**(19): 2534-2536.
- Wang, Z., M. Gerstein and M. Snyder (2009). "RNA-Seq: a revolutionary tool for transcriptomics." *Nature reviews. Genetics* **10**(1): 57-63.
- Warnat, P., R. Eils and B. Brors (2005). "Cross-platform analysis of cancer microarray data improves gene expression based classification of phenotypes." *BMC bioinformatics* **6**(1): 265.
- Watson, J. D. and F. H. Crick (1953). "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid." *Nature* **171**(4356): 737-738.
- Weinrich, S. L., R. Pruzan, L. Ma, M. Ouellette, V. M. Tesmer, S. E. Holt, A. G. Bodnar, S. Lichtsteiner, N. W. Kim, J. B. Trager, R. D. Taylor, R. Carlos, W. H. Andrews, W. E. Wright, J. W. Shay, C. B. Harley and G. B. Morin (1997). "Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT." *Nat Genet* **17**(4): 498-502.
- Williams, P. J. and F. Broughton Pipkin (2011). "The genetics of pre-eclampsia and other hypertensive disorders of pregnancy." *Best Pract Res Clin Obstet Gynaecol* **25**(4): 405-417.
- Winn, V. D., M. Gormley, A. C. Paquet, K. Kjaer-Sorensen, A. Kramer, K. K. Rumer, R. Haimov-Kochman, R. F. Yeh, M. T. Overgaard, A. Varki, C. Oxvig and S. J. Fisher (2009). "Severe preeclampsia-related changes in gene expression at the maternal-fetal interface include sialic acid-binding immunoglobulin-like lectin-6 and pappalysin-2." *Endocrinology* **150**(1): 452-462.
- Xu, Z. P., E. F. Wawrousek and J. Piatigorsky (2002). "Transketolase haploinsufficiency reduces adipose tissue and female fertility in mice." *Mol Cell Biol* **22**(17): 6142-6147.
- Yagil, Y. and C. Yagil (2006). "Integration--a key to success in the genetic dissection of complex diseases?" *Trends Cardiovasc Med* **16**(2): 35-38.
- Yamazaki, K., J. Umeno, A. Takahashi, A. Hirano, T. A. Johnson, N. Kumasaka, T. Morizono, N. Hosono, T. Kawaguchi, M. Takazoe, T. Yamada, Y. Suzuki, H. Tanaka, S. Motoya, M. Hosokawa, Y. Arimura, Y. Shinomura, T. Matsui, T. Matsumoto, M. Iida, T. Tsunoda, Y. Nakamura, N. Kamatani and M. Kubo (2013). "A Genome-Wide Association Study

- Identifies 2 Susceptibility Loci for Crohn's Disease in a Japanese Population." Gastroenterology **144**(4): 781-788.
- Yang, S. and Y. P. Li (2007). "RGS10-null mutation impairs osteoclast differentiation resulting from the loss of [Ca²⁺]_i oscillation regulation." Genes Dev **21**(14): 1803-1816.
- Yang, S. K., M. Hong, W. Zhao, Y. Jung, J. Baek, N. Tayebi, K. M. Kim, B. D. Ye, K. J. Kim, S. H. Park, I. Lee, E. J. Lee, W. H. Kim, J. H. Cheon, Y. H. Kim, B. I. Jang, H. S. Kim, J. H. Choi, J. S. Koo, J. H. Lee, S. A. Jung, Y. J. Lee, J. Y. Jang, H. D. Shin, D. Kang, H. S. Youn, J. Liu and K. Song (2014). "Genome-wide association study of Crohn's disease in Koreans revealed three new susceptibility loci and common attributes of genetic susceptibility across ethnic populations." Gut **63**(1): 80-87.
- Zeggini, E. and J. P. Ioannidis (2009). "Meta-analysis in genome-wide association studies." Pharmacogenomics **10**(2): 191-201.
- Zintzaras, E. and J. P. Ioannidis (2008). "Meta-analysis for ranked discovery datasets: theoretical framework and empirical demonstration for microarrays." Comput Biol Chem **32**(1): 38-46.
- Zukowska, P., B. Kutryb-Zajac, A. Jaształ, M. Toczek, M. Zabielska, T. Borkowski, Z. Khalpey, R. T. Smolenski and E. M. Slominska (2017). "Deletion of CD73 in mice leads to aortic valve dysfunction." Biochim Biophys Acta Mol Basis Dis **1863**(6): 1464-1472.
- Πανάρετος, Ι. and Ε. Ξεκαλάκη (2000). Εισαγωγή στη στατιστική σκέψη. Αθήνα.

Παράρτημα

Πίνακας 1 Γονίδια που υπάρχουν σε μία, δύο, τρεις και ούτω καθεξής μελέτες (που συμπεριλήφθηκαν στη μετα-ανάλυση για την ασθένεια της προεκλαμψίας)

Πλήθος Μελετών	Πλήθος κοινών γονιδίων
1 μελέτη	1.966 γονίδια
2 μελέτες	3.707 γονίδια
3 μελέτες	866 γονίδια
4 μελέτες	1.144 γονίδια
5 μελέτες	949 γονίδια
6 μελέτες	706 γονίδια
7 μελέτες	1.151 γονίδια
8 μελέτες	1.645 γονίδια
9 μελέτες	5.028 γονίδια
10 μελέτες (όλες)	10.705 γονίδια
Συνολικός αριθμός γονιδίων	27.867

Πίνακας 2 Το πλήθος των στατιστικώς σημαντικών ΔΕΓ για κάθε μέθοδο διόρθωσης (στη μετα-ανάλυση για την ασθένεια της προεκλαμψίας)

	p-value<0.01	FDR<0.01	Bonferroni	Sidak	Holm	Holland
GSE4707	1.152	1	2	2	2	2
GSE74341	4.376	1.872	72	76	74	76
GSE54618	512	0	1	1	1	1
GSE30186	877	0	0	0	0	0
GSE14722	245	1	3	3	3	3
GSE10588	2.834	1,372	282	283	283	287
GSE24129	980	0	1	1	1	1
GSE25906	1.490	267	68	68	68	68
GSE35574	2.202	697	131	132	131	133
GSE75010	5.060	2.891	866	868	870	872
Μετα-ανάλυση	2.256	629	167	169	168	169

Πίνακας 3 Τα ΔΕΓ που προέκυψαν από τη μετα-ανάλυση για τη συνολική προεκλαμψία σε FDR 1%

Γονίδιο	Μέγεθος Επίδρασης	p-value	Γονίδιο	Μέγεθος Επίδρασης	p-value	Γονίδιο	Μέγεθος Επίδρασης	p-value
Υποεκφραζόμενα								
TUG8	-1.98133	0.0000231	TMEM176A	-0.622181	0.0001149	FADS3	-0.478946	0.0000204
TUG3	-1.93585	0.0000308	MTFMT	-0.621754	0.0001843	AFF3	-0.477314	0.0000117
TUG6	-1.8361	0.0000585	MIR222	-0.621633	9.72E-06	LANCL1	-0.477205	0.0001848
UBE2Q2L	-1.80618	0.0000711	HIGD1B	-0.618748	3.12E-06	USP27X	-0.477015	4.59E-06
LOC200772	-1.77806	0.0000854	MIR199A2	-0.616842	0.0000111	CD180	-0.476163	9.14E-07
LOC100132977	-1.50769	0.0001128	FBLN5	-0.613367	3.39E-10	DPYD	-0.475624	3.86E-06
CRSP8P	-1.45113	0.0001809	IPO13	-0.613066	1.69E-08	MMP1	-0.474952	8.85E-06
HSD17B1P1	-1.11226	4.19E-12	RSPO2	-0.609081	7.01E-08	FERMT1	-0.470954	2.14E-04
RN7SKP258	-1.08675	1.43E-10	NR2F1	-0.605999	1.39E-08	AGO2	-0.4708	8.45E-06
SOD1	-1.06177	1.26E-10	COL6A1	-0.602973	6.94E-10	LIPG	-0.469584	1.20E-06
CALHM4	-0.938845	4.51E-06	RN7SKP56	-0.601659	7.96E-05	SYTL2	-0.466943	9.36E-05
MORC4	-0.928422	7.66E-13	LOC642366	-0.599849	1.05E-04	MUM1	-0.466713	1.39E-06
RNU6-1320P	-0.92602	3.91E-09	LACC1	-0.598805	4.50E-06	CCDC149	-0.466065	6.32E-05
RNU6-703P	-0.899494	5.98E-08	SPX	-0.597637	7.26E-05	ERV3-1	-0.464334	4.96E-06
F13A1	-0.898481	1.11E-10	FGF12	-0.594438	1.48E-08	PDGFRA	-0.4639	1.73E-06
APMAP	-0.877938	4.40E-07	RNU6-842P	-0.594362	9.75E-05	LINC00597	-0.463536	1.73E-06
CHURC1-FNTB	-0.872453	2.29E-08	MAP4K3	-0.594242	1.20E-09	PRICKLE2	-0.462999	3.52E-06
ACOXL	-0.872424	3.42E-06	ALDH1A1	-0.591606	1.40E-09	SPATA7	-0.460208	2.92E-06
SLAMF1	-0.868695	0.0000877	FMOD	-0.590378	1.42E-09	DLK1	-0.459388	0.0002208
RIPPLY1	-0.868602	1.31E-10	RGL1	-0.59031	1.51E-09	DBNDD2	-0.457212	9.32E-05

FBN2	-0.867898	9.01E-06	MIR214	-0.588373	2.57E-05	AOC3	-0.457096	2.35E-06
RNU6-732P	-0.866327	0.0000213	MRPS31P5	-0.587717	0.0001148	MAP1LC3C	-0.457032	0.0000738
P3H2	-0.865597	0.0000875	SORBS2	-0.583106	2.11E-09	BCKDHB	-0.455933	0.0000974
RNU6-1229P	-0.861726	1.87E-07	LARP1	-0.580238	1.28E-05	SOCS2	-0.455402	2.51E-06
CLDN1	-0.846953	1.11E-08	NCF2	-0.577983	3.17E-09	TMEM37	-0.455055	1.65E-05
RNU2-47P	-0.823676	5.73E-07	ZBTB20	-0.576815	6.39E-06	PAMR1	-0.454996	2.51E-06
CARD16	-0.823393	2.76E-11	OR51B2	-0.576111	1.05E-04	ZDHHC2	-0.452721	5.78E-06
BHLHE41	-0.823021	0.000018	ASPH	-0.576097	4.58E-06	THAP12	-0.447858	0.0000723
SRGAP2	-0.822746	0.0001248	PAK2	-0.574952	2.04E-08	PCDH18	-0.447431	0.0002208
RASL11B	-0.819225	7.99E-14	ISM1	-0.572429	9.90E-06	PIP4K2B	-0.443465	4.68E-06
SLC23A2	-0.813707	6.60E-06	SLC27A2	-0.572121	1.58E-05	RASSF4	-0.443033	4.52E-06
OLFML3	-0.809505	0.0000926	RNU6-984P	-0.570733	0.0001754	GJC1	-0.441533	0.0000052
GSTA3	-0.792495	1.99E-08	EID3	-0.568122	1.63E-04	BAMBI	-0.439314	5.49E-06
RNU4-15P	-0.789702	1.51E-06	PIP5K1B	-0.562422	2.01E-04	MEIS1	-0.439103	5.83E-06
ZNF555	-0.786665	1.11E-07	XPO4	-0.558649	1.57E-05	TACC2	-0.438358	1.86E-04
GOT1	-0.786162	0.0000385	YTHDC2	-0.555597	0.000037	ACOT2	-0.437752	0.0000848
BMP6	-0.78094	1.36E-07	TBC1D14	-0.553007	1.04E-05	TMEM74B	-0.437733	9.38E-05
HMGNI	-0.780088	0.00001	ZNF10	-0.552893	0.0001117	TSHZ3	-0.437347	0.0000117
RN7SKP283	-0.772073	5.99E-07	BMP5	-0.552268	1.61E-04	VPS39	-0.434813	6.51E-06
CTSC	-0.771709	6.66E-15	FRZB	-0.551714	2.78E-05	ZNF215	-0.433389	3.65E-05
HS3ST3A1	-0.768802	0.0002229	ARHGAP18	-0.549896	0.0000258	FOXF1	-0.431746	8.03E-06
TKT	-0.766917	0.00005	GRAMD1A	-0.549393	1.6E-07	NOX5	-0.430119	8.18E-06
CATSPERB	-0.750329	0.0000346	PDGFD	-0.549192	0.0000348	GIMAP7	-0.429713	0.0000162
RNU6-466P	-0.747633	1.30E-06	MS4A4A	-0.548903	3.91E-05	CDKAL1	-0.427613	6.47E-05

CCND2	-0.743952	5.13E-14	CXCL14	-0.546196	2.10E-06	ZNF625	-0.426452	1.02E-04
HMGB1	-0.743207	7.17E-06	DYRK4	-0.54577	4.95E-08	SYNPO2	-0.425258	1.96E-05
RCAN1	-0.741787	0.0000571	MSRB3	-0.545731	2.79E-07	MS4A6A	-0.423654	0.0000117
RNU6-1303P	-0.741206	5.80E-06	RBMS3	-0.54303	3.95E-05	CLMN	-0.423453	1.17E-05
ARPC5	-0.740788	6.38E-06	ANTXR2	-0.539875	1.30E-04	ACSS2	-0.422836	2.22E-05
ENPP1	-0.73339	1.01E-06	KEAP1	-0.538076	1.42E-05	KCTD17	-0.422356	1.26E-05
RNU6-1328P	-0.733194	7.20E-06	MCCC1	-0.535433	9.25E-06	GNPDA1	-0.420989	1.32E-05
CADM3	-0.728007	0.0000581	MCTP2	-0.535232	0.0000237	ABCA6	-0.420254	0.0000147
RAP1GAP2	-0.72473	2.29E-06	CD53	-0.534695	1.33E-05	UBE2E3	-0.419196	1.46E-05
MPEG1	-0.723699	4.37E-06	SEC22C	-0.533109	1.16E-05	SATB1	-0.418456	1.47E-05
TRAK2	-0.71959	1.60E-06	CYTL1	-0.530127	4.97E-08	ZNF214	-0.417968	2.40E-05
RNA5SP385	-0.718357	0.0000764	ZSWIM5	-0.529456	6.19E-07	SPATA20	-0.417684	0.0001065
METTL7B	-0.710229	2.29E-09	ARID5B	-0.52938	4.07E-06	MAP3K12	-0.415175	1.73E-05
SLC45A4	-0.709582	0.0000135	TLR6	-0.529258	5.32E-08	PRRX1	-0.414927	0.000018
DPYSL3	-0.703506	5.36E-11	CYP4F35P	-0.528484	1.85E-04	TBC1D19	-0.413475	1.89E-05
AIF1	-0.700113	1.10E-12	CFI	-0.527372	6.75E-06	ZNF738	-0.412973	9.10E-05
RNF138P1	-0.698555	0.0001503	RIN2	-0.52524	6.57E-08	ASCC1	-0.41273	0.0000201
TCF21	-0.696225	0.0000223	CMKLR1	-0.524899	0.0001625	FKBP7	-0.409783	0.0000385
CALHM5	-0.686522	0.000103	VPS45	-0.52475	0.0001267	CTDSP2	-0.408739	0.0000232
SYNPO2L	-0.682006	4.60E-06	HAPLN1	-0.522719	2.00E-06	FZD4	-0.40785	2.38E-05
ZNF185	-0.681512	0.0001128	PIP4P2	-0.521525	0.0001053	GCFC2	-0.407686	0.0000245
RMC1	-0.680204	0.0000355	NCMAP	-0.521177	0.0000404	CYB5B	-0.406288	0.0000263
LOC102725104	-0.677667	9.92E-06	ARL14EP	-0.519574	2.22E-07	APPBP2	-0.405716	1.81E-04
RGS10	-0.675043	6.43E-08	DHX57	-0.519431	1.50E-05	NMNAT2	-0.402162	4.91E-05

RPS15P9	-0.673746	0.000068	KANK2	-0.517908	0.0001299	TM4SF18	-0.39906	0.0000597
PLPP3	-0.671342	0.0000888	MID1IP1	-0.517024	1.34E-06	MMD	-0.395774	0.0000418
PER3	-0.669829	4.31E-07	SYT17	-0.516979	2.92E-05	PIAS3	-0.394639	1.43E-04
RN7SKP174	-0.669517	0.0000379	ADCY3	-0.516625	9.71E-08	TBC1D8B	-0.394162	0.0000442
CLDN16	-0.669176	5.97E-07	ADCK1	-0.512512	3.17E-07	GSTZ1	-0.393063	4.69E-05
GHR	-0.668687	9.44E-06	RNF130	-0.512019	1.48E-04	HPS5	-0.389351	9.03E-05
ATXN7L3B	-0.665767	5.60E-07	HNMT	-0.509495	1.68E-05	SMIM11A	-0.386933	1.01E-04
ATP8B4	-0.664724	0.0001642	ZMYND11	-0.508455	9.27E-06	CD209	-0.385477	0.0000627
ZSWIM6	-0.664706	1.20E-09	MTCH2	-0.508027	2.25E-04	NTS	-0.385208	8.12E-05
PARP6	-0.664482	0.0000205	DLG4	-0.50429	0.0001348	SLIT3	-0.384985	0.000066
ADCY7	-0.662373	0.0000724	RFX5	-0.502695	2.34E-07	IGF2BP1	-0.382174	0.0001229
PDZRN3	-0.659427	0.000132	TMSB4X	-0.502591	0.0000738	CEP19	-0.381985	0.0001235
RPL7AP28	-0.656951	0.0000181	ADGRD1	-0.501731	8.5E-07	VIM	-0.381826	0.0000766
CABLES1	-0.656947	0.0001313	TMEM131	-0.500704	0.0000546	OR2B6	-0.38136	0.0001772
PDPN	-0.654089	2.74E-09	NUDT9P1	-0.498731	8.69E-05	ANKRD36B	-0.380518	7.86E-05
MYO1B	-0.653043	0.0002223	ZFH3	-0.498214	2.89E-07	LPIN1	-0.379845	0.0001056
HPGDS	-0.649324	5.20E-09	FAM20A	-0.495176	7.58E-07	KCTD1	-0.378595	1.39E-04
C10ORF25	-0.648595	1.40E-07	PARM1	-0.494393	6.13E-07	ARHGDI1	-0.376184	9.49E-05
RAD54L2	-0.642809	0.0000471	PLAC1	-0.49184	0.0001782	CCDC68	-0.373117	0.0001091
PALLD	-0.63764	0.0001131	NDN	-0.491201	0.0000385	C9ORF72	-0.370721	0.0001894
RNU6-1321P	-0.637474	0.0000841	RNASE1	-0.489863	1.53E-06	MPHOSPH6	-0.367318	0.0001399
NPL	-0.635723	0.0000236	FAM172A	-0.489538	0.0001076	MFS12	-0.367196	0.0001402
RNU4-53P	-0.635033	0.0000892	CASP1	-0.488698	0.0000443	LOXL1	-0.365852	0.0001448
ADIRF	-0.634317	7.43E-06	WFDC1	-0.488678	1.69E-04	CUBN	-0.36554	1.51E-04

PRDM6	-0.634231	6.20E-09	ZNF780A	-0.488147	1.80E-04	ACAT2	-0.361928	1.77E-04
TMPRSS11F	-0.633674	0.0000842	SENP8	-0.487404	1.12E-06	SNRPC	-0.360098	0.0001887
LOC644584	-0.633449	0.0001684	AGTR1	-0.486378	5.4E-07	URGCP	-0.35991	0.0001901
CHN1	-0.631238	1.30E-08	RAB5C	-0.484687	1.09E-04	LAPTM5	-0.357885	2.04E-04
WDFY3	-0.628632	1.35E-10	ZNF610	-0.483814	9.21E-05	PCDH17	-0.357656	2.06E-04
PARP15	-0.624671	5.06E-09	SNX24	-0.482684	6.18E-06			
Υπερεκφραζόμενα								
ABCG1	0.35664	0.0002146	LRRC32	0.55259	4.98E-07	CNPY2	0.769836	4.48E-06
ZNF710	0.3571	0.0002082	POR	0.560437	0.0000242	TECR	0.770765	0.0001638
FOXJ1	0.358154	0.000201	SEMA3G	0.564207	1.05E-06	ILVBL	0.774879	0.0000294
FGF20	0.359557	0.0001926	RIMBP3	0.564225	0.0001381	GLOD5	0.778398	5.92E-09
PIGZ	0.360512	0.0001845	VSTM2B	0.564551	0.0000182	BAD	0.778673	0.000196
IGFBP5	0.362608	0.0001665	CHPF	0.566734	0.000015	ARRDC3	0.782862	0.0000566
SLC39A4	0.362657	0.0001686	ZNF525	0.568273	9.08E-08	GPR32P1	0.783211	4.29E-07
CDKL2	0.365196	0.000208	RNU1-85P	0.568522	0.0001858	UBE2Q2	0.784336	1.52E-07
GRK3	0.367906	0.000137	RN7SL241P	0.569524	0.0001834	PPL	0.784752	0.0001
PAK6	0.368712	0.0001289	SLC6A6	0.571715	0.0000462	EAF1	0.784776	0.0000843
FAM3B	0.368934	0.000206	PREX1	0.571816	0.0000107	RNU4-26P	0.784977	1.73E-06
BASP1-AS1	0.371276	0.0001846	MUC1	0.574069	7.64E-07	PLIN2	0.791447	0.0001363
IL17C	0.372041	0.0002192	RARRES1	0.574188	0.0001873	RN7SKP296	0.792505	3.13E-07
SLC41A2	0.375995	0.0001515	SNHG12	0.5751	6.88E-06	GAPDH	0.794812	0.0000194
ZBED6CL	0.378825	0.0001386	RN7SL381P	0.577166	0.0001505	GPX3	0.79647	2.21E-06
C17ORF77	0.379333	0.0001855	RNA5SP484	0.579303	0.0001431	CEBPA	0.797692	0.0001511
ANO1	0.380329	0.0000815	GJD3	0.57932	0.0000111	HMGB3P1	0.802927	0.0001308

NIM1K	0.380512	0.0001313	ELF3	0.579535	0.0000251	JUP	0.805305	0.0000512
HYAL2	0.380975	0.0000799	FBXL16	0.583233	5.16E-06	ARID3A	0.807129	0.0000305
DPP7	0.380983	0.0002249	PPP1R15A	0.583988	2.17E-09	RNU6-185P	0.808614	8.84E-07
FOS	0.380999	0.0000794	LSR	0.584086	0.0000147	ERGIC1	0.813718	0.0000185
KANSL1L	0.381788	0.0000769	DVL1	0.587307	0.0000373	DLX5	0.818182	0.0000168
BFAR	0.383524	0.0000706	PP7080	0.589523	0.0001368	CCSAP	0.819301	0.0001175
ZNF614	0.386623	0.0000608	RN7SL797P	0.590172	0.0001079	KRT17P2	0.822043	1.23E-07
SIRT6	0.387212	0.0000606	SSR4	0.590387	5.94E-08	PHYHIP	0.824171	4.56E-07
ELMOD1	0.389754	0.0000904	LIMS4	0.592557	0.0000298	CA4	0.83306	0
FGD4	0.39203	0.0000829	PRSS40B	0.594931	0.0001678	PRKD3	0.833505	0.0001081
PPP4R4	0.398675	0.0000366	LINC00189	0.595312	0.0002185	CCDC183	0.833706	1.58E-06
NUDT1	0.405064	0.0000268	IGHV3-35	0.596557	0.0000906	NEURL1	0.8362	4.81E-06
NUCB2	0.405807	0.0001438	HPS3	0.596612	0.0000874	P4HA1	0.836603	0.0000895
CDC42EP1	0.406313	0.0000562	SEC24A	0.599586	0.0000334	KDM3A	0.847723	1.15E-07
SALL4	0.408115	0.0000416	AKAP12	0.599974	0.0000304	CT47A10	0.84913	5.04E-08
MOB2	0.408571	0.0000523	RN7SL180P	0.601506	0.0000797	IGSF8	0.865063	0.000167
C1QTNF12	0.408748	0.0001052	OR52K3P	0.601731	0.0000401	GPER1	0.870534	1.22E-07
TMEM38B	0.410477	0.0000216	RN7SKP63	0.602319	0.000078	TMEM91	0.871012	1.80E-06
IGSF9B	0.414529	0.0000629	PFKL	0.604889	3.80E-08	KIF2A	0.871121	0.0000216
CHRM2	0.41576	0.0001435	RNU6-504P	0.604939	0.0001834	SYDE1	0.874717	0.0000572
APOBEC2	0.420949	0.000082	PRRG1	0.605065	0.0001875	BTD	0.881932	6.66E-16
C9ORF47	0.426162	0.0000533	WDR86	0.605704	1.43E-06	SLC2A1	0.882441	1.81E-06
GCHFR	0.428986	9.25E-06	CLN3	0.607566	0.0001167	GREM2	0.882543	3.09E-12
TOR1B	0.430111	0.0000284	GLRX	0.607589	8.72E-09	RNU6-996P	0.887723	8.56E-08

PHLDB3	0.431928	0.00005	ZNF175	0.609888	3.15E-06	ACSF2	0.888762	1.65E-09
ASB5	0.433824	0.0000245	MLPH	0.6108	1.42E-06	EPS8L1	0.891898	7.54E-06
NCAPD2	0.434537	0.000207	RNA5SP292	0.612677	0.0001527	FAM120AOS	0.896029	5.85E-07
BRI3BP	0.435413	0.0000128	SSR1	0.614645	2.20E-06	DIO2	0.901835	2.33E-10
SYNGR1	0.437966	5.88E-06	PLBD1	0.617132	0.0000604	MME	0.905447	0.0001369
CTU1	0.438482	0.0000108	CRIP1P4	0.617279	0.0000526	RNA5SP227	0.912723	3.98E-08
RPS27L	0.439734	5.28E-06	LOC399886	0.617938	0.000064	BHLHE40	0.915197	5.53E-06
ONECUT1	0.440231	5.44E-06	TMEM234	0.618539	4.74E-08	RN7SL854P	0.920112	4.60E-09
FGD6	0.441269	0.0000285	HIST2H2AA3	0.619928	0.0000138	RNU6-1151P	0.920714	3.11E-08
CIART	0.442883	0.000039	GALK1	0.621054	0.0000773	NPNT	0.923281	0.0001395
MZT2A	0.444772	0.000037	B9D2	0.623695	0.0000412	DNM2	0.9262	0
H1F0	0.445445	0.0000532	CALM1	0.626369	0.0000766	ASTE1	0.932076	0.0002047
C8G	0.445475	4.17E-06	DCP2	0.627289	0.0001476	MIR320A	0.938815	0.0000579
RECQL4	0.445748	4.14E-06	CCNY	0.627827	0.0002153	TREM1	0.941062	1.15E-06
RNF151	0.446598	0.0000235	CSF2RB	0.628613	2.24E-07	KRT19	0.942306	0.0000383
ADGRG6	0.446877	3.82E-06	MIR375	0.629078	7.55E-06	HILPDA	0.944923	3.82E-06
CPTP	0.447416	0.0002164	FTH1P4	0.630633	0.0000366	LIMCH1	0.946326	2.46E-06
KRT86	0.448674	7.14E-06	NTF4	0.634296	2.21E-09	SLCO2A1	0.953066	0.0000713
PLPPR4	0.45237	2.91E-06	GLRXP3	0.635007	6.09E-06	PPP1R12C	0.961512	0.0000152
TLE2	0.457136	0.0001664	ORMDL2	0.635489	0.0001688	CORO2A	0.961659	3.74E-11
KCNJ6	0.457341	2.26E-06	SMARCA1	0.635991	0.0001889	MIF	0.966552	6.01E-07
P4HA3	0.45805	0.0000145	LINGO3	0.637139	1.52E-06	RN7SL265P	0.971506	6.30E-09
BICD2	0.463556	1.71E-06	ACBD4	0.639035	0.0000374	FTH1P18	0.975719	6.34E-10
MX1	0.463859	0.0001147	RN7SL364P	0.642038	0.0000269	SLC6A10P	0.976456	2.10E-07

DNASE2	0.464849	0.0000145	BNIP3L	0.643525	0.0002165	RNU6-694P	0.980143	4.78E-09
OLAH	0.465879	1.48E-06	WFS1	0.645453	0.0000277	SCARB1	0.98188	0.0000692
PTPRF	0.474342	0.0001318	SLC11A1	0.646314	9.33E-06	SFXN3	1.00123	0.0000308
CARD10	0.475554	0.0001379	P2RY2	0.650092	1.52E-08	NEK11	1.007	0.0002226
TMEM184B	0.477775	2.98E-06	SLC11A2	0.651698	0.0000242	TBC1D22A	1.00755	2.27E-08
LRP2BP	0.478376	8.19E-07	MAP3K8	0.655373	2.46E-11	CSNK1E	1.01025	3.42E-08
SDC4	0.478957	7.92E-07	TXNP4	0.656033	0.0000532	RDH13	1.01499	0.0000126
OBSL1	0.481917	0.0000706	KCNF1	0.657517	1.89E-11	TPBG	1.02291	0.0000576
HS6ST1	0.483378	0.0000248	DNAJC1	0.661683	4.46E-08	SH3PXD2A	1.0271	5.45E-06
TVP23A	0.486435	0.0002074	GADD45G	0.668721	0.000046	HK2	1.04648	0
LAMC2	0.488317	0.0000184	OCEL1	0.669677	2.07E-07	ERO1A	1.04907	0.0000167
PIP4K2A	0.489779	0.0000244	EGFR	0.675995	7.74E-06	ARHGAP45	1.04946	0.0002079
NT5E	0.491285	4.15E-07	SDF2L1	0.677915	4.24E-07	CGB3	1.05149	0.0000335
SNAPC1	0.493052	1.06E-06	TENM4	0.67821	3.49E-08	HTRA1	1.05296	4.61E-10
P4HA2	0.495621	7.41E-07	COMTD1	0.679511	0.0000285	INHA	1.05391	0.0000296
SH2D3A	0.495798	0.0000525	ARNT2	0.682055	6.34E-07	LOC729870	1.08425	0.0000799
FBXO22	0.496587	0.0000271	HERPUD1	0.685157	0.0000588	RNU6-1150P	1.09086	9.27E-12
IFI44L	0.500178	2.64E-07	TUBA4A	0.685475	1.55E-06	SNORA18	1.09691	0.0000692
PCSK9	0.504483	0.0000309	RNU6-104P	0.686144	0.0000248	PLA2G16	1.12703	0.0000137
GNPTG	0.506396	4.34E-07	PAM	0.69086	0.0001867	BCL6	1.1438	0.000016
DEGS1	0.509491	0.0000304	TPI1P1	0.701412	9.09E-07	KMT2E-AS1	1.14927	0.0001642
RABAC1	0.510139	8.48E-06	RN7SKP130	0.704568	4.57E-06	NECTIN4	1.15054	0.0000227
TRIM71	0.510146	0.0001118	STBD1	0.709227	0.0000225	SIGLEC6	1.15527	2.66E-06
KRTCAP3	0.513477	2.87E-07	KIT	0.711453	0.0000127	QPCT	1.15945	5.06E-09

NOCT	0.514038	0.0001895	FAM43A	0.714085	0.0000161	PROCR	1.16068	2.84E-09
CD63	0.514105	3.14E-06	ZNF662	0.71593	5.02E-07	GCNT2P	1.17533	0.0000345
PIFO	0.514615	7.51E-06	IL1RAP	0.717086	0.0001	INHBA	1.20556	0.0001755
TM9SF2	0.515817	1.10E-07	EDEM2	0.717797	1.77E-07	NPHP3-ACAD11	1.21333	8.06E-14
NEBL	0.516598	0.0000557	ANKRD24	0.731676	8.81E-06	MIR492	1.26117	6.72E-10
GDF5OS	0.519052	0.0001215	MBD2	0.7319	0.0000735	PAPPA2	1.26257	8.16E-09
RRS1	0.520254	0.0002188	DNAJB2	0.732097	2.63E-07	SPAG4	1.26354	3.10E-08
FANCE	0.520775	3.89E-06	LOC101929908	0.739831	1.60E-06	FLT1	1.28657	1.61E-06
LOC102723968	0.522372	0.0002177	BCAT2	0.740091	0.000034	SASH1	1.35482	4.44E-16
SPCS3	0.528789	0.0002077	GFOD2	0.744214	2.11E-08	LEP	1.4688	0.0000951
DEFB135	0.531872	0.0001045	CTSA	0.74514	0.000025	SDCBPP3	1.49087	0.0001302
SGPP2	0.535728	3.97E-06	YPEL4	0.746133	7.56E-10	WDR86-AS1	1.54523	0.0000343
EGLN1	0.53579	0.0001521	ABCA7	0.747146	3.58E-07	LINC00592	1.63924	0.0002131
FRMD4A	0.536223	0.0001194	ADAM12	0.747481	3.14E-06	SNHG14	1.66669	0.0001777
IFNLR1	0.536786	0.0000287	LOC102723566	0.748184	1.24E-06	LOC91450	1.67905	0.0001637
ATP8B3	0.536813	0.0002066	SREBF1	0.750927	0.0000773	LOC145474	1.69893	0.0001436
IL27RA	0.540013	0.0000497	NRIP1	0.752369	4.46E-06	LOC729867	1.73052	0.0001167
CLC	0.54007	2.48E-07	DERL3	0.754053	7.29E-08	LINC02288	1.81783	0.0000659
MTX1	0.540469	0.0000311	PKM	0.754388	4.06E-06	MIF-AS1	1.91604	2.83E-07
LMAN2L	0.542329	0.0000507	NCOR2	0.756617	0.0000401	LINC00421	1.9969	0.0000209
LINC01556	0.545877	0.0000944	BTNL9	0.761848	0.000189	LOC283038	2.03076	0.0000169
IL27	0.549857	0.0000484	GALNT2	0.766963	0.000117	SLC2A1-AS1	2.31501	3.04E-06
AMOTL2	0.551473	7.61E-06	C10ORF90	0.769457	0.0000532	LOC100268168	3.04849	6.92E-08

Πίνακας 4 Τα ΔΕΓ που προέκυψαν από τη μετα-ανάλυση για τη σοβαρή προεκλαμψία σε FDR 1%

Γονίδιο	Μέγεθος Επίδρασης	p-value	Γονίδιο	Μέγεθος Επίδρασης	p-value	Γονίδιο	Μέγεθος Επίδρασης	p-value
Υποεκφραζόμενα								
ANKRD20A2	-3.88279	6.39E-06	SLC20A1	-1.13755	7.39E-08	MMP1	-0.920641	7.37E-06
FKBP1C	-2.2648	3.83E-09	GHR	-1.13341	7.33E-08	NR2C1	-0.918004	7.97E-06
ZC2HC1A	-1.9499	3.69E-09	MORC4	-1.12676	8.62E-08	TMEM74B	-0.911566	8.95E-06
LCAL1	-1.76192	0.0000111	RASL11B	-1.11347	0.000053	BMP5	-0.909443	9.43E-06
UACA	-1.65872	0.0000271	AMY2B	-1.09831	4.27E-05	HSPB11	-0.905162	1.08E-05
HSD17B1	-1.64256	1.81E-07	ZNF700	-1.09228	6.61E-06	MS4A4A	-0.890251	1.47E-05
P3H2	-1.6253	6.55E-13	PDHA1	-1.08268	1.67E-05	VWA8	-0.885573	5.99E-05
TRAF3	-1.57865	3.67E-06	HYAL4	-1.07629	2.43E-07	CAPRIN2	-0.884246	1.45E-05
NBPF3	-1.577	0.0000617	BHLHE41	-1.07487	3.49E-05	MCCC1	-0.877046	1.69E-05
ZNF554	-1.57392	0.0000242	MYO1B	-1.06261	0.000112	CALML4	-0.870652	2.02E-05
KLHDC9	-1.55843	1.33E-06	ARL4C	-1.056	1.07E-05	MRPS30	-0.867027	2.18E-05
OR1J4	-1.53486	7.03E-06	ADCY7	-1.05322	5.02E-07	ALDH5A1	-0.856652	0.000028
TMEM139	-1.50507	1.72E-06	RCBTB1	-1.04511	1.61E-05	ZNF10	-0.851464	2.88E-05
SOD1	-1.50233	8.14E-06	BORCS5	-1.04245	0.000015	SLC23A2	-0.849488	5.76E-05
MTA3	-1.49381	2.30E-06	TMEM168	-1.04143	4.62E-05	RWDD3	-0.849207	0.000033
APMAP	-1.4634	9.15E-08	ABCG2	-1.03166	7.02E-07	SH3YL1	-0.848553	3.12E-05
TRAK1	-1.4018	1.07E-06	CDH1	-1.02974	7.82E-07	ENPP2	-0.844393	3.63E-05
SH3TC2	-1.38244	1.39E-06	F13A1	-1.0086	1.09E-06	LDHB	-0.844175	3.43E-05

SLC25A27	-1.33143	3.47E-07	MAD2L2	-1.00523	2.85E-05	NR2F1	-0.832106	4.51E-05
GPCPD1	-1.33057	1.52E-06	KLRD1	-1.00201	1.47E-06	TRAK2	-0.824508	5.23E-05
ALDH4A1	-1.29189	0.0000264	ENPP1	-0.984554	1.29E-05	TKT	-0.814696	6.33E-05
ACOXL	-1.27259	0.0000347	RAD54L2	-0.983582	2.04E-06	FERMT1	-0.809284	6.94E-05
GOT1	-1.26554	3.08E-09	FAM84B	-0.977598	4.31E-05	ATP5F1E	-0.804481	7.88E-05
ZNF555	-1.23539	1.98E-06	RANBP17	-0.975786	8.63E-06	FGF13	-0.793708	9.01E-05
GSTA4	-1.22135	1.13E-08	CHKA	-0.959587	3.31E-06	COL1A2	-0.790566	9.92E-05
FBN2	-1.16869	0.0000267	ZNF137P	-0.951748	3.90E-06	PCOLCE	-0.789812	0.0001
GNG10	-1.1675	3.11E-07	MKNK1	-0.949288	4.26E-06	VNN1	-0.788042	9.84E-05
TRPV6	-1.16583	2.22E-07	POPDC2	-0.946651	4.41E-06	CTSC	-0.783121	0.000113
KCTD3	-1.15601	6.51E-07	MKL2	-0.933768	1.33E-05	STIM1	-0.781533	0.000109
CLDN1	-1.15597	4.09E-08	PLAC1	-0.931947	5.78E-06	PC	-0.780516	0.000112
EPHB6	-1.15104	4.45E-08	GDA	-0.924729	9.58E-05	UGGT2	-0.779132	0.00012
PRIMPOL	-1.14811	2.32E-06	RWDD2A	-0.921537	7.86E-06	KMT5B	-0.778219	0.000122
CCDC174	-1.14307	2.83E-06						
Υπερεκφραζόμενα								
CYP4F12	0.777377	0.0001239	LYN	1.0345	6.62E-07	SFXN3	1.33978	0.000047
GUCA2A	0.777569	0.0001237	FMNL1	1.04374	4.62E-05	TBC1D22A	1.34299	5.46E-10
TFAP2A	0.780108	0.0001149	DNAJB2	1.04662	4.68E-07	NPNT	1.34907	6.59E-08
FRMPD4	0.781726	0.0001217	DNM2	1.04854	4.98E-07	NPEPPS	1.35219	1.04E-08
IER5	0.78848	0.0000973	GALNT2	1.06394	3.56E-07	SLC2A1	1.35511	4.02E-10
PAK3	0.792394	0.0000942	DDR1	1.06642	1.02E-05	BHLHE40	1.3567	4.03E-10
ZCCHC2	0.798803	0.0000789	FKBP2	1.06955	3.01E-07	OCRL	1.35893	4.04E-10
CLDN7	0.803687	0.0000757	LRRC10	1.0712	0.000061	SH3BP5	1.3613	8.27E-07

P2RY2	0.807197	0.0000673	HBG2	1.07363	6.08E-05	PPL	1.37166	1.61E-08
FOXJ2	0.812383	0.0000603	TGOLN2	1.07455	2.57E-07	TPP1	1.37234	2.82E-10
GALE	0.822633	0.0000552	ZNF331	1.0783	2.48E-07	SLCO2A1	1.37383	2.87E-10
HPCAL1	0.823381	0.0000533	MME	1.07838	2.47E-07	DLX5	1.38542	1.63E-08
TRIM14	0.825334	0.0000513	S100A6	1.07863	2.31E-05	CSNK2A2	1.38585	2.06E-10
BTD	0.845627	0.0000326	GFOD2	1.08195	2.29E-07	LHB	1.38733	3.13E-08
MTX1	0.848898	0.0000326	EFHD1	1.08481	2.39E-07	KRT19	1.38863	1.88E-10
PTBP3	0.853346	0.0000302	PAM	1.09268	1.70E-07	CCSAP	1.3888	3.07E-08
INSL4	0.856545	0.0000266	TGFB1	1.09268	9.75E-06	FSTL3	1.39706	1.59E-10
CEBPB	0.857849	0.0000254	SLC35F2	1.09676	1.62E-07	SCARB1	1.4001	1.69E-10
ZMYND8	0.861305	0.0000248	CYP11A1	1.09734	1.71E-07	NDRG1	1.40937	1.12E-10
ACBD4	0.866316	0.0000237	GPX3	1.10045	1.51E-07	EBI3	1.40977	3.96E-06
TM9SF1	0.879526	0.0000174	HK2	1.10088	1.51E-07	RAB25	1.41975	7.68E-11
HMGB3	0.883762	0.0000447	KDM3A	1.11138	1.37E-07	FAM120AOS	1.42251	2.04E-08
EGFR	0.88449	0.000062	BASP1	1.11222	1.16E-07	HERPUD2	1.4482	1.18E-08
PPP1R15A	0.884832	0.0000156	RUVBL2	1.12279	6.29E-06	QPCT	1.46044	8.63E-09
VEGFC	0.894969	0.000013	BCAT2	1.12416	8.85E-08	KIF2A	1.46309	8.39E-05
ELF3	0.899312	0.0000114	PREX1	1.12967	3.21E-06	CYP2J2	1.4683	2.15E-08
BCKDK	0.906685	9.89E-06	SOX13	1.13537	7.58E-08	PLA2G16	1.46992	1.05E-06
ABCA12	0.914703	8.59E-06	GPX6	1.14228	2.80E-06	JUP	1.47432	2.30E-11
RMDN3	0.917893	8.03E-06	RDH13	1.14267	2.83E-06	IGSF8	1.48622	5.94E-09
FBLIM1	0.918709	0.000112	CDK2AP2	1.14677	5.64E-08	PSMA8	1.51069	2.80E-07
GLA	0.921436	7.17E-06	CA4	1.14792	0.000061	CRIP2	1.51358	8.69E-12
TRIM58	0.923993	0.0000503	BNIP3L	1.14872	1.31E-05	TECR	1.51743	8.84E-12

TRAPPC1	0.92712	0.0000977	TMEM45A	1.15157	1.06E-06	C10ORF90	1.52035	8.52E-05
ZNFX1	0.928769	0.0001189	ITPRIP	1.15189	4.98E-06	HTRA1	1.54077	5.32E-12
KIT	0.92966	6.33E-06	SIPA1L1	1.16048	3.95E-08	HTRA4	1.55474	1.68E-09
SERF2	0.931622	0.0000836	BAD	1.16591	1.72E-05	SPAG4	1.56239	3.78E-12
STS	0.932973	5.67E-06	ACSF2	1.16609	3.49E-08	PROCR	1.56636	6.78E-06
PRKD3	0.93351	6.25E-06	HILPDA	1.17175	2.75E-08	TPBG	1.56696	2.42E-12
MXI1	0.93483	5.20E-06	SLC6A8	1.17926	2.45E-08	GBA	1.58232	9.09E-10
LDHA	0.936978	5.48E-06	NEURL1	1.17951	2.78E-08	MOB3A	1.59981	3.05E-06
SMARCA1	0.940451	4.88E-06	MLX	1.1867	2.14E-08	TPI1	1.61271	6.47E-05
LRAT	0.948625	0.0000643	LIMD1	1.19022	1.88E-08	BCL6	1.61636	3.39E-05
ADAM12	0.950261	0.0000652	SP5	1.19055	1.07E-06	LRRC1	1.62094	0.000015
PPP1R12C	0.957215	0.0000567	TUBA4A	1.19435	1.63E-08	SIGLEC6	1.62316	4.12E-09
NDUFB10	0.957519	0.0000564	SLC35A2	1.20829	4.00E-07	ARHGAP45	1.65249	1.39E-11
CARHSP1	0.959254	3.36E-06	GRN	1.22035	1.51E-08	NEK11	1.65269	2.06E-06
GAPDH	0.964566	3.08E-06	FZD10	1.2293	1.23E-05	ASTE1	1.67969	1.71E-05
MARCKS	0.964689	2.87E-06	SPTLC3	1.22984	2.88E-06	TMEM141	1.72172	6.95E-11
BACH1	0.971166	0.0000151	SH3PXD2A	1.23127	8.23E-09	INHA	1.72457	6.26E-14
P4HA1	0.972042	2.48E-06	OR1C1	1.2322	1.36E-05	CGB7	1.73467	8.41E-07
REEP4	0.977173	2.66E-06	JUNB	1.23545	6.72E-09	GBAP1	1.75144	8.08E-09
ELL	0.983841	2.88E-06	DUSP1	1.24425	6.67E-09	SASH1	1.79437	8.22E-15
SMAD6	0.994861	1.52E-06	GREM2	1.2522	5.49E-09	LRRC8A	1.82489	1.86E-08
CST6	0.996884	1.54E-06	TREM1	1.27274	2.98E-09	ENG	1.85083	2.22E-15
TAX1BP3	1.00052	0.0000278	VDR	1.28706	2.10E-09	CGB3	1.87106	5.39E-06
PSG11	1.00759	1.26E-06	MAN1C1	1.29468	1.75E-09	FLT1	1.93504	4.03E-13

RHBDD2	1.00768	0.0000245	ARID3A	1.29858	1.52E-09	INHBA	2.22546	3.32E-06
ITGA5	1.01079	1.09E-06	WDR60	1.30682	1.50E-09	RNU1-134P	2.24111	0.000121
PDLIM2	1.01318	9.98E-07	ERO1A	1.31273	1.64E-05	RAB6C	2.24154	1.30E-11
GADD45G	1.01678	9.22E-07	LIMCH1	1.32195	4.86E-06	PAPPA-AS1	2.28861	9.96E-05
C12ORF75	1.01747	0.0000208	PHYHIP	1.32601	8.87E-10	ST8SIA6	2.36675	7.29E-05
CEBPA	1.018	9.41E-07	SLCO4A1	1.3266	1.18E-07	LEP	2.72592	3.32E-05
AKAP12	1.02111	0.000085	PUSL1	1.33727	7.87E-08	HSBP1L1	2.88056	7.38E-05
HEY1	1.02284	8.66E-07						

Πίνακας 5 Τα ΔΕΓ που προέκυψαν από τη μετα-ανάλυση για την ήπια προεκλαμψία σε FDR 1%

Γονίδιο	Μέγεθος Επίδρασης	p-value	Γονίδιο	Μέγεθος Επίδρασης	p-value	Γονίδιο	Μέγεθος Επίδρασης	p-value
Υποεκφραζόμενα								
ABCA7	0.781878	5.89E-07	OR51B2	-0.678692	0.000045	ARSI	-0.527592	2.17E-06
TUG8	-1.98133	0.0000231	RNA5SP502	-0.676033	3.21E-05	UBE2E3	-0.52562	2.39E-06
TUG3	-1.93585	0.0000308	PARP15	-0.675874	1.83E-09	DBNDD2	-0.525548	0.000122
NPSA	-1.90261	0.0000381	RPS15P9	-0.673746	0.000068	NUDT9P1	-0.524913	0.000127
TUG6	-1.8361	0.0000585	VTN	-0.670662	3.52E-07	TMEM86A	-0.52203	4.59E-06
UBE2Q2L	-1.80618	0.0000711	COL6A1	-0.670393	2.53E-09	FAM20A	-0.521986	2.79E-06
LOC200772	-1.77806	0.0000854	RN7SKP174	-0.669517	3.79E-05	NR2F1	-0.520807	8.02E-05
LOC100132977	-1.50769	0.0001128	RNU6-590P	-0.667536	3.99E-05	FCGRT	-0.520145	5.09E-05
RN7SKP258	-1.08675	1.43E-10	GSTA3	-0.66732	2.89E-09	ZNF717	-0.520012	2.95E-06
HSD17B1P1	-1.08558	1.49E-10	PER3	-0.665321	8.12E-05	RNF14	-0.518878	3.13E-06

RN7SKP20	-1.06325	3.15E-10	FMOD	-0.663447	3.52E-09	CD180	-0.517416	3.41E-06
RNU6-316P	-0.966947	7.28E-09	RN7SKP53	-0.661734	4.61E-05	CD37	-0.514498	1.57E-05
FAM74A7	-0.946206	1.40E-08	GAPDHP67	-0.658972	4.94E-05	CPT1B	-0.514243	3.85E-06
TCF21	-0.910874	5.69E-06	PDPN	-0.656446	7.07E-07	CD33	-0.512715	4.01E-06
RNU6-703P	-0.899494	5.98E-08	PRDM6	-0.651597	5.83E-08	GAS6	-0.508961	4.80E-06
CREB1	-0.899388	0.0001008	RNA5SP385	-0.647636	6.55E-05	PLEKHM3	-0.506423	5.26E-06
TLR8	-0.888435	9.33E-15	FCGR1A	-0.646259	1.21E-05	BAMBI	-0.502884	5.99E-06
RIPPLY1	-0.883686	4.07E-10	ZNF555	-0.646122	4.88E-05	PRKCB	-0.500883	3.27E-05
RNU6-1320P	-0.877762	1.16E-07	LINC01963	-0.643116	1.77E-05	ADCY3	-0.499636	6.75E-06
SOD1	-0.87628	1.01E-06	SORBS2	-0.640722	1.05E-08	MUM1	-0.496809	7.59E-06
CHURC1-FNTB	-0.874657	1.27E-07	HPGDS	-0.640127	5.58E-05	ZNF215	-0.491388	9.76E-06
CARD16	-0.863797	8.23E-10	SET	-0.639768	1.15E-08	TSHZ3	-0.485676	6.59E-05
RNU6-1229P	-0.861726	1.87E-07	MCTP2	-0.63841	3.08E-06	TRMT5	-0.483903	1.32E-05
RN7SKP153	-0.861136	1.90E-07	RNU6-1321P	-0.637474	8.41E-05	AOC3	-0.480916	1.49E-05
MORC4	-0.858198	3.07E-06	RNU4-53P	-0.635033	8.92E-05	SENP8	-0.479865	5.73E-05
RNU4-66P	-0.851694	2.52E-07	PAK2	-0.630802	5.59E-05	MS4A6A	-0.479751	1.56E-05
RNU4-44P	-0.849793	2.67E-07	FBLN5	-0.625627	1.23E-06	ADCK1	-0.479683	9.33E-05
BMP6	-0.849665	4.06E-11	CD53	-0.620714	8.81E-06	LIPG	-0.479618	1.55E-05
RNU2-47P	-0.823676	5.73E-07	RPL7AP28	-0.619654	0.000129	ZNF214	-0.479366	9.22E-05
TRGV7	-0.821291	6.81E-08	RAB5C	-0.613419	2.19E-05	PSMD2	-0.479273	0.000016
RNU6-1218P	-0.818427	6.67E-07	DELE1	-0.612528	4.59E-08	MRPL44	-0.477649	1.62E-05
RN7SL709P	-0.797796	1.20E-06	CYTL1	-0.610415	4.82E-08	COL6A6	-0.475533	1.87E-05
RUBCNL	-0.796172	4.60E-07	XPO4	-0.603994	6.90E-08	NUMA1	-0.473771	9.62E-05
RNU4-15P	-0.789702	1.51E-06	APMAP	-0.600038	3.74E-05	ZBTB20	-0.473756	1.93E-05

ZBTB39	-0.783477	3.83E-09	LOC642366	-0.599849	0.000105	ARL14EP	-0.473419	2.02E-05
RNU6-732P	-0.779887	2.00E-06	TLR6	-0.595094	1.68E-06	GJC1	-0.473098	3.69E-05
CTSC	-0.768136	1.30E-11	GRAMD1A	-0.591566	3.61E-07	JAZF1	-0.472799	0.000021
TMPRSS11F	-0.751517	5.56E-07	DLK1	-0.590326	1.31E-07	USP27X	-0.472007	0.000021
RN7SKP283	-0.750368	4.52E-06	ARID5B	-0.588994	8.74E-07	NAV3	-0.471865	2.17E-05
METTL7B	-0.749179	1.61E-09	MIR199A2	-0.584605	6.95E-05	PIP5K1B	-0.467616	2.46E-05
RNU6-1303P	-0.741206	5.80E-06	CCDC50	-0.584179	4.65E-05	ABL1	-0.463008	0.00003
CCND2	-0.73914	6.49E-11	DYRK4	-0.575487	7.26E-05	SLC16A7	-0.461469	3.11E-05
DPYSL3	-0.735496	0.0000144	FGF12	-0.573945	1.50E-06	IGF2BP3	-0.46068	3.28E-05
RNU6-1328P	-0.733194	7.20E-06	NDN	-0.569773	3.26E-07	SYT17	-0.457598	3.69E-05
RASL11B	-0.732499	9.26E-11	ZNF610	-0.569458	1.26E-05	ADGRD1	-0.452666	4.45E-05
PALLD	-0.729551	0.0000245	RIN2	-0.566483	2.26E-06	SPDYA	-0.452093	7.24E-05
RGS10	-0.713147	5.16E-09	MIR214	-0.561388	0.000126	GPR65	-0.451825	4.64E-05
LOC494127	-0.712088	0.0000126	LRRC41	-0.560086	1.67E-06	RASSF4	-0.45113	4.68E-05
AIF1	-0.709035	3.20E-10	MAP4K3	-0.556423	1.24E-05	NOX5	-0.449149	4.97E-05
RNU1-58P	-0.708377	0.0000139	PABPN1	-0.554679	7.76E-05	ZNF112	-0.445024	6.07E-05
RNU6-466P	-0.705415	0.000015	EPHX1	-0.551232	7.62E-07	C1QA	-0.443571	6.18E-05
ADIRF	-0.702761	0.0000344	RNASE1	-0.551209	7.78E-07	NDUFA9	-0.443392	6.22E-05
WDFY3	-0.69901	5.20E-10	ZNF780A	-0.549258	6.81E-05	DNAJA4	-0.438495	7.45E-05
MPEG1	-0.698431	4.72E-07	PDGFRA	-0.543557	3.14E-05	AGO2	-0.438273	7.64E-05
LOC102725104	-0.695346	0.0000196	ASPH	-0.540946	0.000047	RPL10	-0.436432	8.25E-05
DACH1	-0.694338	0.000081	SOCS2	-0.536846	1.43E-06	PTPRO	-0.435867	0.000108
ZSWIM6	-0.694249	7.36E-10	MSRB3	-0.532682	1.78E-06	CNOT9	-0.434576	8.49E-05
MIR222	-0.690272	3.13E-06	ITGAM	-0.531294	1.81E-06	IGFLR1	-0.432717	9.23E-05

MTCH2	-0.689123	1.01E-06	FILIP1L	-0.530974	3.48E-05	TM4SF18	-0.428903	0.000105
RNU6-1272P	-0.688867	0.0000231	NCF2	-0.528858	2.07E-06	ACSS2	-0.425706	0.000119
RAB3IL1	-0.686382	1.05E-09	RARB	-0.527696	2.16E-06	TMEM176A	-0.42414	0.000127
RSPO2	-0.682931	1.13E-06						
Υπερεκφραζόμενα								
ATP8B3	0.422983	0.000131	SMOX	0.556133	1.43E-05	ACSF2	0.765845	9.66E-05
XK	0.425468	0.0001224	GLRX	0.55647	5.81E-07	CA4	0.77116	1.02E-11
GIPR	0.426719	0.0001132	PFKL	0.557456	7.35E-05	CCDC183	0.776614	6.45E-05
DNASE2	0.42786	0.0001094	LSAMP	0.558481	1.11E-05	LOC102723566	0.778155	2.10E-06
EXOSC5	0.428837	0.0001056	GDF5OS	0.570515	0.000101	GLOD5	0.784546	1.98E-08
FGD4	0.429164	0.0001071	TMEM234	0.574177	2.76E-07	RNU4-26P	0.784977	1.73E-06
ERLIN1	0.430591	0.0001006	SSR4	0.577833	0.000105	B9D2	0.788003	3.80E-12
TMEM184B	0.434067	0.0000901	CSF2RB	0.579589	0.000094	KRT17P2	0.79949	1.15E-06
CTU1	0.436082	0.0000827	HELZ2	0.580383	0.000101	HTRA1	0.805363	1.56E-05
KCNJ6	0.438767	0.0000758	SEMA3G	0.580455	6.21E-05	RNU6-185P	0.808614	8.84E-07
AMELY	0.439224	0.0001094	COX6C	0.581804	6.39E-05	RN7SL159P	0.8147	7.43E-07
P4HA2	0.441359	0.0000687	LIMS4	0.584664	0.000089	RN7SKP296	0.814778	7.41E-07
SALL4	0.441843	0.0000647	MYCNOS	0.585475	2.13E-05	TMEM91	0.818062	2.77E-06
PLPPR4	0.442161	0.0000664	PHPT1	0.586034	7.06E-07	CT47A10	0.831782	4.53E-07
QPRT	0.442909	0.0000643	VSTM2B	0.595785	1.54E-05	GPR32P1	0.837381	3.84E-07
APC2	0.446018	0.0000555	DNAJC1	0.606843	0.000102	DNM2	0.840748	2.27E-07
SNX32	0.448405	0.000053	MIR375	0.608366	3.62E-05	BTD	0.844159	6.49E-07
KANSL1L	0.449017	0.0000515	GLB1L2	0.610549	1.38E-05	EDEM2	0.844787	4.80E-06
PIM2	0.44914	0.0000516	NTF4	0.611254	4.82E-08	RN7SKP106	0.844862	3.08E-07

KRTCAP3	0.449911	0.0000476	GLRXP3	0.614042	3.01E-05	RNU6-996P	0.887723	8.56E-08
RBM38	0.45222	0.0000443	ZNF525	0.614153	4.10E-08	GPER1	0.896113	0.00004
PEBP1	0.455779	0.0000397	RN7SL180P	0.619077	0.000131	OR10AC1	0.898656	6.13E-08
LRP2BP	0.455839	0.0000398	MAP3K8	0.622055	4.77E-05	RNA5SP227	0.912723	3.98E-08
CDK19	0.457093	0.0000372	KRT8P41	0.623189	3.11E-05	RNU6-1151P	0.920714	3.11E-08
C1ORF21	0.458479	0.0001052	RN7SKP63	0.62349	0.000118	RN7SKP267	0.933806	2.07E-08
C8ORF31	0.461259	0.0000319	SLC11A1	0.627831	9.96E-05	EPS8L1	0.944596	7.88E-06
TM9SF2	0.468011	0.0000246	HDAC10	0.628006	5.37E-06	PROCR	0.957324	6.16E-05
FBN3	0.468782	0.0000243	CGB3	0.629272	5.20E-06	RN7SL854P	0.967545	7.15E-09
C19ORF24	0.470429	0.0000229	RN7SKP77	0.630303	0.0001	DIO2	0.970516	4.51E-07
SH2D3A	0.473894	0.0000194	SGPP2	0.634682	1.49E-08	RN7SL265P	0.971506	6.30E-09
SPAG9	0.474255	0.0000196	LINGO3	0.635587	4.35E-06	RNU6-694P	0.980143	4.78E-09
RNF151	0.474634	0.0000193	FTH1P4	0.636939	8.52E-05	FTH1P18	0.98351	4.29E-09
MROH6	0.475766	0.0000184	PTPRF	0.639585	4.72E-07	CORO2A	0.990423	1.76E-08
FER1L4	0.477396	0.0000173	SNHG12	0.646892	2.78E-06	RNU6-1324P	1.00043	2.48E-09
ZC3H6	0.477636	0.0000169	TXNP4	0.656033	5.32E-05	HK2	1.00804	1.44E-14
C17ORF77	0.484794	0.0000122	KCNF1	0.659331	4.37E-09	RNU6-643P	1.02076	1.28E-09
PEG3	0.487661	0.0000115	RN7SL364P	0.662624	4.51E-05	CSNK1E	1.04601	6.78E-06
CACNA1A	0.488106	0.0000111	DERL3	0.664087	3.59E-09	RNU6-725P	1.06584	2.89E-10
SDC4	0.49125	0.0000874	RN7SL322P	0.667363	0.00004	RNU6-1150P	1.07593	2.06E-10
P4HA3	0.494098	8.49E-06	TPI1P1	0.667655	5.49E-05	LOC729870	1.08425	7.99E-05
RPS27L	0.495351	8.17E-06	ANKRD24	0.672993	6.46E-06	PAPPA2	1.08473	0.000017
PPP1R15A	0.495617	8.00E-06	WDR86	0.678402	2.25E-06	RN7SL57P	1.09055	1.26E-10
CEACAM19	0.496165	8.02E-06	RN7SL711P	0.685683	2.51E-05	SNORA18	1.09691	6.92E-05

CMYA5	0.498537	7.43E-06	RNU6-104P	0.686144	2.48E-05	SASH1	1.14841	8.05E-08
MXD1	0.499775	0.0000593	TNIP2	0.692538	7.83E-10	GCNT2P	1.17533	3.45E-05
RECQL4	0.502473	6.18E-06	WDR90	0.693245	6.52E-05	ERVK3-2	1.18791	4.44E-12
ONECUT1	0.504446	5.83E-06	RN7SKP130	0.694078	2.02E-05	NPHP3-	1.22082	1.40E-12
						ACAD11		
DPF1	0.504994	5.68E-06	PKM	0.701455	0.000057	MIR492	1.35538	1.13E-14
CD63	0.507798	0.0001057	RNU6-302P	0.701966	1.65E-05	RNU6-447P	1.37242	6.00E-15
GTPBP3	0.510467	0.0000273	TENM4	0.722984	1.99E-07	SDCBPP3	1.49087	0.00013
CEBPG	0.51455	0.0001289	RNU6-1003P	0.730972	7.64E-06	WDR86-AS1	1.54523	3.43E-05
NFATC2IP	0.518619	3.22E-06	SDF2L1	0.736923	7.76E-06	LOC729867	1.73052	0.000117
CCDC78	0.519206	2.97E-06	RNU6-957P	0.739225	6.12E-06	LINC02288	1.81783	6.59E-05
TTC39A	0.524794	0.0001262	ZNF662	0.74199	2.66E-05	LINC00421	1.9969	2.09E-05
MAPK13	0.527178	2.23E-06	PLBD1	0.746608	7.75E-05	LOC283038	2.03076	1.69E-05
ELL2	0.529681	1.96E-06	GREM2	0.74962	3.80E-09	LHFPL3-AS2	2.06498	1.37E-05
DEFB135	0.531872	0.0001045	LOC101929908	0.756302	3.84E-06	MIF-AS1	2.19269	6.27E-06
GJD3	0.539175	0.0000839	HMG3P1	0.760854	3.39E-06	SLC2A1-AS1	2.31501	3.04E-06
SDC3	0.541645	0.00011	BIN2P1	0.76301	3.19E-06	GABPB1-IT1	2.5777	7.01E-07
TVP23A	0.543314	0.0000745	YPEL4	0.765124	1.78E-05	LOC100268168	3.04849	6.92E-08
OCEL1	0.555248	6.37E-07						

Πίνακας 6 Τα ΔΕΓ που προέκυψαν από τη μετα-ανάλυση για τη συνολική προεκλαμψία σε FDR 5%

Γονίδιο	Μέγεθος Επίδρασης	p-value	Γονίδιο	Μέγεθος Επίδρασης	p-value	Γονίδιο	Μέγεθος Επίδρασης	p-value
---------	-------------------	---------	---------	-------------------	---------	---------	-------------------	---------

Υποεκφραζόμενα

TUG8	-1.98133	0.000023	MAD2L2	-0.570157	0.00089	SLFN12	-0.438443	0.000247
TUG3	-1.93585	0.000031	EID3	-0.568122	0.000163	TACC2	-0.438358	0.000186
TUG6	-1.8361	0.000059	DDX47	-0.564709	0.002063	GABRA4	-0.437945	0.000651
UBE2Q2L	-1.80618	0.000071	TTC28	-0.563664	0.001812	ACOT2	-0.437752	0.000085
LOC200772	-1.77806	0.000085	SLC44A5	-0.563078	0.000259	TMEM74B	-0.437733	0.000094
SEC24B-AS1	-1.616	0.000249	PIP5K1B	-0.562422	0.000201	TSHZ3	-0.437347	0.000012
TUG7	-1.59334	0.000289	XPO4	-0.558649	0.000016	TBC1D15	-0.436094	0.001229
NPIP12	-1.58911	0.000545	RNU6-285P	-0.558044	0.000536	VPS39	-0.434813	6.50E-06
DPY19L2P3	-1.56671	0.000345	YTHDC2	-0.555597	0.000037	PTRHD1	-0.434763	0.001485
LOC100132977	-1.50769	0.000113	CAP2	-0.554901	0.000991	NOTCH2	-0.434184	0.001329
TUG5	-1.50658	0.000514	NRCAM	-0.554226	0.000933	ZNF215	-0.433389	0.000037
TUG4	-1.46676	0.00067	CHKA	-0.553372	0.002373	SRXN1	-0.432483	0.002281
CRSP8P	-1.45113	0.000181	TBC1D14	-0.553007	0.00001	MAGEA9B	-0.432325	0.001849
FBXO26	-1.43972	0.000802	ZNF10	-0.552893	0.000112	FOXF1	-0.431746	8.00E-06
SOX1-OT	-1.43325	0.000837	FXVD3	-0.55272	0.000841	FADS2	-0.430803	0.000377
TUG10	-1.37275	0.001251	BMP5	-0.552268	0.000161	FEZ1	-0.430692	0.001991
SLIT2-IT1	-1.36156	0.001347	FRZB	-0.551714	0.000028	PTK2	-0.430213	0.001214
TUG9	-1.30173	0.001998	ARHGAP18	-0.549896	0.000026	NOX5	-0.430119	8.20E-06
FGF13-AS1	-1.27803	0.002333	GRAMD1A	-0.549393	0.00000016	GIMAP7	-0.429713	0.000016
MIR448	-1.25702	0.000867	PDGFD	-0.549192	0.000035	FBXO8	-0.428819	0.002243
NPIP2	-1.25595	0.000852	MS4A4A	-0.548903	0.000039	PTGFR	-0.42868	0.000245
NPM1P35	-1.16396	0.00193	CXCL14	-0.546196	0.0000021	C6ORF62	-0.428638	0.000238
MIR1265	-1.16094	0.001833	DYRK4	-0.54577	4.90E-08	MRPL44	-0.428513	0.002079

MIR324	-1.14342	0.002344	MSRB3	-0.545731	2.80E-07	ENOX1	-0.428488	0.001177
HSD17B1P1	-1.11226	4.20E-12	RNU2-26P	-0.545067	0.000713	CDKAL1	-0.427613	0.000065
RN7SKP258	-1.08675	1.40E-10	ASAH1	-0.543775	1.72E-03	IQGAP2	-0.426908	0.001831
LCAL1	-1.08381	0.002334	SETMAR	-0.543202	0.000971	SCAMP5	-0.426462	0.000598
SOD1	-1.06177	1.30E-10	RNU4-50P	-0.543032	0.000745	ZNF625	-0.426452	0.000102
ZC2HC1A	-0.979874	0.001638	RBMS3	-0.54303	3.90E-05	SIRT5	-0.425899	0.001077
HSD17B1	-0.976389	0.000311	F5	-0.54221	0.000771	CPQ	-0.425768	0.001246
CALHM4	-0.938845	4.50E-06	RNA5SP272	-0.541184	7.75E-04	SYNPO2	-0.425258	0.00002
TMEM136	-0.928731	0.002243	POPDC2	-0.540603	1.03E-03	CCL13	-0.423838	0.000805
MORC4	-0.928422	7.70E-13	ANTXR2	-0.539875	1.30E-04	MS4A6A	-0.423654	0.000012
RNU6-1320P	-0.92602	3.90E-09	FGF10	-0.539496	2.35E-03	CLMN	-0.423453	0.000012
RNU6-703P	-0.899494	6.00E-08	GLUD1	-0.539448	1.04E-03	ACBD6	-0.423182	0.001396
F13A1	-0.898481	1.10E-10	RNU6-519P	-0.538578	0.00082	ACSS2	-0.422836	0.000022
APMAP	-0.877938	4.40E-07	KEAP1	-0.538076	1.40E-05	ABCA8	-0.422599	0.000595
FAM106CP	-0.873983	0.001144	CPT1B	-0.537536	7.57E-04	KCTD17	-0.422356	0.000013
CHURC1-FNTB	-0.872453	2.30E-08	PTPRD	-0.537473	0.001128	GNPDA1	-0.420989	0.000013
ACOXL	-0.872424	3.40E-06	MYLK	-0.536695	1.00E-03	ABCA6	-0.420254	0.000015
SLAMF1	-0.868695	0.000088	RNA5SP465	-0.535784	8.71E-04	DELE1	-0.420227	0.001501
RIPPLY1	-0.868602	1.30E-10	MCCC1	-0.535433	0.0000093	ARHGAP9	-0.420124	0.001724
FBN2	-0.867898	9.00E-06	MCTP2	-0.535232	0.000024	UBE2E3	-0.419196	0.000015
RNU6-732P	-0.866327	0.000021	BLOC1S2	-0.535095	0.001487	SATB1	-0.418456	0.000015
P3H2	-0.865597	0.000087	CD53	-0.534695	1.30E-05	ZNF214	-0.417968	0.000024
GABRE	-0.865336	0.001716	C15ORF38-	-0.533673	0.000439	SPATA20	-0.417684	0.000107
			AP3S2					

RNU6-1229P	-0.861726	1.90E-07	SEC22C	-0.533109	0.000012	NDUFAF1	-0.417614	0.001515
TMEM139	-0.851867	0.000397	RN7SKP115	-0.531134	4.63E-04	RARB	-0.416732	0.000251
TMEM273	-0.851642	0.002111	TLR8	-0.530868	8.06E-04	RETREG1	-0.416515	0.000396
CLDN1	-0.846953	1.10E-08	CYTL1	-0.530127	5.00E-08	SRSF6	-0.416418	0.000763
RNU2-47P	-0.823676	5.70E-07	DACH1	-0.52948	0.000618	VSIG4	-0.415893	0.000952
CARD16	-0.823393	2.80E-11	ZSWIM5	-0.529456	0.0000062	MAP3K12	-0.415175	0.000017
BHLHE41	-0.823021	0.000018	ARID5B	-0.52938	0.0000041	PRRX1	-0.414927	0.000018
SRGAP2	-0.822746	0.000125	TLR6	-0.529258	5.30E-08	PTGER2	-0.413986	0.000953
TRAF3	-0.820576	0.00067	CYP4F35P	-0.528484	1.85E-04	CCDC85A	-0.41358	0.001819
RASL11B	-0.819225	8.00E-14	MIR105-1	-0.528464	0.000494	TBC1D19	-0.413475	0.000019
SLC23A2	-0.813707	6.60E-06	RNU6-534P	-0.528107	1.03E-03	ZNF738	-0.412973	0.000091
OLFML3	-0.809505	0.000093	C1QTNF7	-0.52804	1.99E-03	ASCC1	-0.41273	0.00002
GSTA3	-0.792495	2.00E-08	SELENOP	-0.527833	3.39E-04	VSNL1	-0.411858	0.000456
RNU4-15P	-0.789702	1.50E-06	CFI	-0.527372	0.0000068	FKBP7	-0.409783	0.000038
ZNF555	-0.786665	1.10E-07	BMX	-0.527125	1.32E-03	RASL12	-0.408944	0.000329
GOT1	-0.786162	0.000038	POSTN	-0.526249	1.90E-03	CTDSP2	-0.408739	0.000023
BMP6	-0.78094	1.40E-07	ANKRD10	-0.525348	2.25E-03	FZD4	-0.40785	0.000024
HMG1	-0.780088	1.00E-05	RIN2	-0.52524	6.60E-08	GCFC2	-0.407686	0.000024
RN7SKP283	-0.772073	6.00E-07	CMKLR1	-0.524899	0.000163	PRR14L	-0.406725	0.00089
CTSC	-0.771709	6.70E-15	VPS45	-0.52475	0.000127	CYB5B	-0.406288	0.000026
HS3ST3A1	-0.768802	0.000223	HAPLN1	-0.522719	0.000002	TNC	-0.405991	0.001864
TKT	-0.766917	0.00005	CPEB3	-0.522586	8.30E-04	APPBP2	-0.405716	0.000181
CATSPERB	-0.750329	0.000035	LRRN3	-0.522374	2.91E-04	VCAM1	-0.405598	0.001077
RNU6-466P	-0.747633	1.30E-06	PIP4P2	-0.521525	0.000105	LNX2	-0.405096	0.000552

CCND2	-0.743952	5.10E-14	NCMAP	-0.521177	4.00E-05	TMEM204	-0.403558	0.000304
COASY	-0.743829	0.000452	ARL14EP	-0.519574	0.00000022	NMNAT2	-0.402162	0.000049
HMGB1	-0.743207	7.20E-06	DHX57	-0.519431	1.50E-05	DAAM2	-0.40171	0.000381
RCAN1	-0.741787	0.000057	KANK2	-0.517908	1.30E-04	MSH5	-0.40168	0.000284
RNU6-1303P	-0.741206	5.80E-06	MID1IP1	-0.517024	1.30E-06	LINC00470	-0.401662	0.000409
ARPC5	-0.740788	6.40E-06	SYT17	-0.516979	2.90E-05	CEP76	-0.401423	0.00114
ENPP1	-0.73339	1.00E-06	ADCY3	-0.516625	0.000000097	PSMD2	-0.401412	0.000392
RNU6-1328P	-0.733194	7.20E-06	RUBCNL	-0.516338	0.002164	CDC40	-0.400806	0.002165
ALDH4A1	-0.730104	0.001099	ZBTB39	-0.515813	0.00031	TM4SF18	-0.39906	0.00006
CADM3	-0.728007	0.000058	RNU1-72P	-0.514586	1.37E-03	NCK2	-0.398413	0.000381
SELE	-0.727102	0.000253	RIMS3	-0.514522	4.30E-04	ZFYVE9	-0.398097	0.000297
RAP1GAP2	-0.72473	2.30E-06	MIR30B	-0.512719	0.000237	COL6A6	-0.397842	0.000668
MPEG1	-0.723699	4.40E-06	PAWR	-0.512559	0.002186	MMD	-0.395774	0.000042
GPATCH2	-0.723475	0.000397	ADCK1	-0.512512	3.20E-07	PNP	-0.395442	0.000461
SLC20A1	-0.722224	0.000433	RNF130	-0.512019	0.000148	PIAS3	-0.394639	0.000143
TRAK2	-0.71959	1.60E-06	HNMT	-0.509495	0.000017	PNO1	-0.394336	0.002396
RNA5SP385	-0.718357	0.000076	ZNF625-ZNF20	-0.50912	7.82E-04	TBC1D8B	-0.394162	0.000044
BTBD3	-0.714685	0.001833	ZMYND11	-0.508455	0.0000093	ANKAR	-0.393072	0.002174
METTL7B	-0.710229	2.30E-09	MTCH2	-0.508027	2.25E-04	GSTZ1	-0.393063	0.000047
SLC45A4	-0.709582	0.000014	ZNF501	-0.506622	0.001072	COL5A2	-0.393039	0.001587
DPYSL3	-0.703506	5.40E-11	C6	-0.506386	3.13E-04	GAS6	-0.391351	0.000848
LCP1	-0.700656	0.000385	ZNF471	-0.506371	0.00084	NRXN3	-0.390723	0.000755
AIF1	-0.700113	1.10E-12	ARL4AP1	-0.506279	0.000832	HPS5	-0.389351	0.00009
RNF138P1	-0.698555	0.00015	KLHL3	-0.505071	0.000406	SMIM11A	-0.386933	0.000101

TCF21	-0.696225	0.000022	DUS2	-0.504311	0.00098	AIG1	-0.386763	0.000411
CALHM5	-0.686522	0.000103	DLG4	-0.50429	0.000135	CD209	-0.385477	0.000063
PXK	-0.683367	0.000917	LOC100996517	-0.50316	8.95E-04	NTS	-0.385208	0.000081
SH3TC2	-0.682152	0.000407	RNU6-329P	-0.503139	0.001732	SLIT3	-0.384985	0.000066
SYNPO2L	-0.682006	4.60E-06	RFX5	-0.502695	0.00000023	IGF2BP1	-0.382174	0.000123
ACSS1	-0.681698	0.000849	TMSB4X	-0.502591	0.000074	CEP19	-0.381985	0.000123
ZNF185	-0.681512	0.000113	NR1D2	-0.502398	0.000635	VIM	-0.381826	0.000077
RMC1	-0.680204	0.000036	ADGRD1	-0.501731	8.50E-07	OR2B6	-0.38136	0.000177
MKNK1	-0.679445	0.000516	TMEM131	-0.500704	0.000055	FZD2	-0.381271	0.001303
LOC102725104	-0.677667	9.90E-06	RN7SKP225	-0.500145	1.84E-03	SENP6	-0.380565	0.000427
FGF7P2	-0.677667	0.001146	BIN1	-0.499222	0.001993	ANKRD36B	-0.380518	0.000079
RGS10	-0.675043	6.40E-08	RNU6-616P	-0.49919	0.001878	LPIN1	-0.379845	0.000106
RPS15P9	-0.673746	0.000068	NUDT9P1	-0.498731	0.000087	KCTD1	-0.378595	0.000139
PLPP3	-0.671342	0.000089	NFIX	-0.49865	2.47E-04	CR1	-0.377077	0.00147
MRPS30	-0.669875	0.001236	ZFH3	-0.498214	0.00000029	ARHGDI3	-0.376184	0.000095
PER3	-0.669829	4.30E-07	NDUFAF5	-0.497864	9.63E-04	CES2	-0.375497	0.001279
RN7SKP174	-0.669517	0.000038	HSD3BP5	-0.496644	0.000356	CD163L1	-0.37415	0.001912
CLDN16	-0.669176	6.00E-07	RNA5SP104	-0.496167	1.05E-03	PKNOX2	-0.373675	0.001951
ABCA11P	-0.669162	0.000435	RN7SKP191	-0.495783	2.01E-03	CCDC68	-0.373117	0.000109
GHR	-0.668687	9.40E-06	ZNF717	-0.495551	0.001978	ZNF660	-0.372148	0.000289
ATXN7L3B	-0.665767	5.60E-07	FAM20A	-0.495176	7.60E-07	C9ORF72	-0.370721	0.000189
ATP8B4	-0.664724	0.000164	NME2	-0.495047	0.000896	ZNF136	-0.370373	0.000524
ZSWIM6	-0.664706	1.20E-09	WDSUB1	-0.494685	0.001305	ZDHHC1	-0.368031	0.002092
KL	-0.664588	0.000705	TXK	-0.494421	0.001865	MPHOSPH6	-0.367318	0.00014

PARP6	-0.664482	0.00002	PARM1	-0.494393	0.00000061	MFSD12	-0.367196	0.00014
MOB3C	-0.66425	0.001044	PLAC1	-0.49184	0.000178	DPY19L2	-0.366967	0.00161
MAF	-0.663429	0.000343	NDN	-0.491201	0.000038	EXOC5P1	-0.366143	0.00102
ADCY7	-0.662373	0.000072	TRGV9	-0.49104	0.001177	INPP1	-0.36605	0.000969
PDZRN3	-0.659427	0.000132	RN7SL181P	-0.489938	0.002266	LOXL1	-0.365852	0.000145
RPL7AP28	-0.656951	0.000018	RNASE1	-0.489863	0.0000015	RDH11	-0.365842	0.001406
CABLES1	-0.656947	0.000131	FAM172A	-0.489538	1.08E-04	CUBN	-0.36554	0.00015
LRIG3	-0.654652	0.000236	CASP1	-0.488698	0.000044	DNAJC6	-0.362249	0.002199
PDPN	-0.654089	2.70E-09	WFDC1	-0.488678	0.000169	ACAT2	-0.361928	0.000177
MYO1B	-0.653043	0.000222	OTUB2	-0.488552	0.001983	SNRPC	-0.360098	0.000189
EID1	-0.652538	0.001534	ZNF780A	-0.488147	0.00018	CDADC1	-0.360095	0.001029
CA10	-0.6507	0.001222	IFT122	-0.487854	4.39E-04	URGCP	-0.35991	0.00019
FAM20B	-0.650315	0.001953	SENP8	-0.487404	1.10E-06	TMEM232	-0.359315	0.000615
HPGDS	-0.649324	5.20E-09	PLEKHM3	-0.486714	0.0015	SOAT1	-0.359241	0.001583
C10ORF25	-0.648595	1.40E-07	AGTR1	-0.486378	0.00000054	LAPTM5	-0.357885	0.000204
TRPV6	-0.645116	0.002256	RAB5C	-0.484687	0.000109	PCDH17	-0.357656	0.000206
RAD54L2	-0.642809	0.000047	ZNF610	-0.483814	0.000092	GZF1	-0.356931	0.00172
PALLD	-0.63764	0.000113	SNX24	-0.482684	0.0000062	TSPAN14	-0.355074	0.002369
RNU6-1321P	-0.637474	0.000084	VRK2	-0.481336	0.001808	CLEC3B	-0.355016	0.00034
NPL	-0.635723	0.000024	DGLUCY	-0.481161	0.000885	LHFPL2	-0.354242	0.001524
RNU4-53P	-0.635033	0.000089	FES	-0.479087	8.62E-04	ZNF546	-0.353981	0.00051
RNF114	-0.634965	0.001546	FADS3	-0.478946	2.00E-05	KIAA1551	-0.353888	0.002237
ADIRF	-0.634317	7.40E-06	SCMH1	-0.478231	0.000739	PSMD3	-0.353839	0.001684
PRDM6	-0.634231	6.20E-09	AFF3	-0.477314	0.000012	ZBED8	-0.353796	0.000239

TMPRSS11F	-0.633674	0.000084	LANCL1	-0.477205	0.000185	CSGALNACT2	-0.352935	0.001889
LOC644584	-0.633449	0.000168	USP27X	-0.477015	4.60E-06	ACOX2	-0.352505	0.00154
TMEM168	-0.631614	0.000786	QDPR	-0.476963	2.12E-03	HMGNI1P38	-0.351143	0.001641
CHN1	-0.631238	1.30E-08	NOMO2	-0.47653	0.00106	KCNJ8	-0.349304	0.000693
WDFY3	-0.628632	1.30E-10	CD180	-0.476163	9.10E-07	IRAK3	-0.348115	0.000303
FPGT	-0.625529	0.001352	MIR30A	-0.475791	0.000607	ZNF740	-0.347999	0.002209
PARP15	-0.624671	5.10E-09	DPYD	-0.475624	0.0000039	TLR1	-0.34738	0.000306
ANK3	-0.623125	0.001909	APIP	-0.475413	0.002398	CCL2	-0.346571	0.000965
TMEM176A	-0.622181	0.000115	MMP1	-0.474952	8.90E-06	ZNF562	-0.34642	0.00032
MTFMT	-0.621754	0.000184	FAM111A	-0.474782	2.03E-03	IFI16	-0.345144	0.000996
MIR222	-0.621633	9.70E-06	TMEM110-	-0.47216	0.00179	PLPP2	-0.344505	0.001556
			MUSTN1					
HIGD1B	-0.618748	3.10E-06	GABRB1	-0.471772	0.000276	NDUFA9	-0.342973	0.001492
MIR199A2	-0.616842	0.000011	FERMT1	-0.470954	2.14E-04	THY1	-0.342662	0.000738
MEST	-0.615807	0.000695	AGO2	-0.4708	8.40E-06	CACNB4	-0.341492	0.000396
FBLN5	-0.613367	3.40E-10	LIPG	-0.469584	0.0000012	PM20D2	-0.338762	0.000653
IPO13	-0.613066	1.70E-08	FASTKD1	-0.468036	0.000316	ZNF197	-0.335	0.000484
VWA5A	-0.612692	0.000239	SYTL2	-0.466943	9.40E-05	PLCL1	-0.334645	0.001017
DSN1	-0.612185	0.001625	MUM1	-0.466713	1.40E-06	MRPS14	-0.333581	0.000534
RSPO2	-0.609081	7.00E-08	CCDC149	-0.466065	0.000063	ARL2BP	-0.333555	0.000534
NR2F1	-0.605999	1.40E-08	BDH2	-0.46459	0.00063	STARD13	-0.331925	0.001249
GUCY1A1	-0.605131	0.00068	PLPP6	-0.464357	4.96E-04	CENPBD1	-0.330965	0.000981
DISP1	-0.60312	0.000914	ERV3-1	-0.464334	0.000005	RBAK	-0.328807	0.000904
COL6A1	-0.602973	6.90E-10	LRRC37A4P	-0.46402	0.001386	CASP10	-0.327847	0.000647

SNX27	-0.602738	0.001112	PDGFRA	-0.4639	0.0000017	LY96	-0.327183	0.000673
RN7SKP56	-0.601659	0.00008	LINC00597	-0.463536	1.70E-06	TMEM97	-0.326993	0.000655
LOC642366	-0.599849	0.000105	PRICKLE2	-0.462999	0.0000035	OR52H1	-0.326141	0.001232
LACC1	-0.598805	4.50E-06	PLCG2	-0.461374	1.41E-03	COX4I2	-0.326128	0.00102
SPX	-0.597637	0.000073	CCDC3	-0.461022	0.000854	FSTL5	-0.32469	0.001409
FGF12	-0.594438	1.50E-08	SPATA7	-0.460208	2.90E-06	RCAN3	-0.324154	0.000756
RNU6-842P	-0.594362	0.000098	DLK1	-0.459388	0.000221	NDUFV1	-0.323966	0.000766
MAP4K3	-0.594242	1.20E-09	AMACR	-0.458007	0.000556	SELENOH	-0.323716	0.001101
SCN9A	-0.593762	0.001665	CCDC77	-0.457833	7.59E-04	APOL2	-0.323695	0.00077
NPC1	-0.593623	0.000324	IDI1	-0.4576	0.000807	STK39	-0.320493	0.000873
ALDH1A1	-0.591606	1.40E-09	ZNF280C	-0.457482	0.00062	COQ7	-0.318735	0.002316
ASB3	-0.59141	0.000453	DBNDD2	-0.457212	9.30E-05	LOXL3	-0.318339	0.001314
METTL21A	-0.591407	0.002266	AOC3	-0.457096	2.30E-06	EFEMP2	-0.318107	0.001528
FMOD	-0.590378	1.40E-09	MAP1LC3C	-0.457032	0.000074	WDR48	-0.317653	0.000931
RGL1	-0.59031	1.50E-09	BCKDHB	-0.455933	0.000097	EHD4	-0.317617	0.000958
RNU6-1189P	-0.588547	0.000269	SH2B3	-0.455429	0.000906	FAM129A	-0.316985	0.000965
MIR214	-0.588373	0.000026	SOCS2	-0.455402	0.0000025	SYK	-0.312189	0.002329
IGF1	-0.587852	0.00051	TMEM37	-0.455055	0.000017	F3	-0.311897	0.001179
MRPS31P5	-0.587717	0.000115	PAMR1	-0.454996	0.0000025	PHTF1	-0.311173	0.0019
H2AFY	-0.587335	0.000501	CYB5R4	-0.454182	2.56E-04	FAM234B	-0.309265	0.002306
SAMD5	-0.583518	0.000789	ZNF568	-0.453889	0.002088	FYB1	-0.307732	0.00137
SORBS2	-0.583106	2.10E-09	FCGRT	-0.452948	0.000338	AAMDC	-0.305682	0.001496
RN7SL138P	-0.581959	0.000312	ZDHHC2	-0.452721	5.80E-06	LGALS1	-0.305657	0.001485
LARP1	-0.580238	0.000013	WLS	-0.451735	1.97E-03	ALDH3A2	-0.303069	0.001612

NCF2	-0.577983	3.20E-09	CDH11	-0.448799	0.000894	SGPL1	-0.302525	0.001658
ZBTB20	-0.576815	6.40E-06	THAP12	-0.447858	0.000072	APOBEC3F	-0.300905	0.00176
C1QTNF3	-0.576357	0.000826	PCDH18	-0.447431	2.21E-04	ARHGAP29	-0.300592	0.001769
OR51B2	-0.576111	0.000105	PARP1	-0.44722	0.001183	CAVIN1	-0.300301	0.001765
ASPH	-0.576097	4.60E-06	ZNF583	-0.444367	0.001171	FCER1G	-0.298406	0.001913
SRPX	-0.575869	0.000605	PIP4K2B	-0.443465	4.70E-06	ASMTL	-0.297976	0.001943
CCDC50	-0.575571	0.001507	RASSF4	-0.443033	0.0000045	SNRNP25	-0.297669	0.001969
PAK2	-0.574952	2.00E-08	RALYL	-0.442674	2.41E-03	NIT2	-0.296358	0.002058
AP5M1	-0.573686	0.00061	CAV1	-0.442544	0.000293	AMBRA1	-0.295188	0.002134
ISM1	-0.572429	9.90E-06	GJC1	-0.441533	0.0000052	SLF2	-0.29468	0.002143
SLC27A2	-0.572121	0.000016	BAMBI	-0.439314	0.0000055	EVI2B	-0.293942	0.002227
UTP6	-0.570981	0.001507	MEIS1	-0.439103	5.80E-06	GABRB3	-0.293742	0.002227
RNU6-984P	-0.570733	0.000175	KCNK3	-0.438522	0.001192	PEX5	-0.293596	0.002231
Υπερεκφραζόμενα								
CDH16	0.292335	0.002375	OTOA	0.468289	0.000679	JUNB	0.657868	0.0013
PHOX2A	0.293012	0.002312	MIR425	0.468321	0.001613	DNAJC1	0.661683	4.50E-08
OAS1	0.294541	0.002191	LYPD3	0.469189	0.000242	MIEN1	0.664138	0.000716
ICAM5	0.295669	0.002119	RN7SL521P	0.470096	0.001881	MICAL3	0.666737	0.000623
RRP12	0.296144	0.002085	RN7SL681P	0.470457	0.001853	GADD45G	0.668721	0.000046
SMUG1	0.297707	0.001937	ANKRD61	0.470772	0.00184	OCEL1	0.669677	2.10E-07
GRM6	0.298345	0.001919	ATP5MFP3	0.471625	0.001814	BMPR2	0.671513	0.002236
CPNE6	0.29891	0.001846	RN7SL555P	0.472475	0.001783	EGFR	0.675995	7.70E-06
DUOX1	0.298961	0.001852	EPS15P1	0.473036	0.002042	SDF2L1	0.677915	4.20E-07
EHF	0.300186	0.00179	RHBDD2	0.473449	0.000779	TENM4	0.67821	3.50E-08

PFKFB1	0.300481	0.001766	PTPRF	0.474342	0.000132	FAM184A	0.679194	0.001602
POU3F1	0.30133	0.001672	RN7SL149P	0.474996	0.001678	COMTD1	0.679511	0.000029
ADAM8	0.301806	0.001685	CARD10	0.475554	0.000138	CYP11A1	0.681159	0.000247
PRG2	0.301926	0.001671	POLR2F	0.475659	0.001046	CRIP2	0.681255	0.001944
WDR24	0.302686	0.00226	CEBPB	0.475664	0.000671	ARNT2	0.682055	6.30E-07
TINAGL1	0.303259	0.001589	REEP4	0.477084	0.000299	SLC35F2	0.683141	0.000269
KLF7	0.304187	0.001565	TMEM184B	0.477775	3.00E-06	PUSL1	0.684733	0.000821
PPP1R13L	0.304942	0.001515	LRP2BP	0.478376	8.20E-07	HERPUD1	0.685157	0.000059
NRGN	0.305038	0.001526	LFNG	0.478656	0.001571	TUBA4A	0.685475	1.50E-06
SCCPDH	0.305114	0.001997	RN7SL299P	0.478687	0.001551	RNU6-104P	0.686144	0.000025
STAT2	0.30716	0.001777	RN7SL553P	0.478904	0.00154	ZNF765	0.686205	0.001725
CNPPD1	0.307607	0.001384	SDC4	0.478957	7.90E-07	ZNF114	0.688117	0.001255
RGS1	0.308088	0.002322	WHAMMP2	0.479629	0.0018	PAM	0.69086	0.000187
CELSR2	0.308765	0.00152	OBSL1	0.481917	0.000071	LOC100130449	0.696807	0.000494
SHISAL2A	0.309022	0.001854	HS6ST1	0.483378	0.000025	TPI1P1	0.701412	9.10E-07
DNMT3L	0.309355	0.001307	RN7SL62P	0.483838	0.001381	ZCCHC2	0.703635	0.001536
PDE4D	0.309746	0.001295	FGFR4	0.484178	0.002254	RN7SKP130	0.704568	4.60E-06
CEP72	0.310027	0.00128	RN7SL42P	0.484555	0.001376	LHB	0.706062	0.000662
TLR9	0.311822	0.001667	FMNL1	0.484771	0.001273	BASP1	0.707199	0.000466
ACOT7	0.312146	0.001187	RN7SL643P	0.485082	0.001351	SP5	0.708328	0.000332
SLC6A12	0.313332	0.001135	RNU6-482P	0.48511	0.001347	MAN1C1	0.708912	0.000752
DLEU1	0.313592	0.00111	TMEM239	0.485158	0.000588	STBD1	0.709227	0.000023
GCAT	0.313718	0.001969	SCD	0.485568	0.000551	RN7SL159P	0.710792	0.00078
ZG16	0.314132	0.001103	TVP23A	0.486435	0.000207	KIT	0.711453	0.000013

SOST	0.315518	0.00146	VEGFA	0.486943	0.001863	RAB17	0.712127	0.000325
PSORS1C1	0.315959	0.001019	LINC01530	0.487338	0.000556	ADORA2A	0.712595	0.000237
CTDP1	0.316389	0.001001	LAMC2	0.488317	0.000018	FAM43A	0.714085	0.000016
SYNE1	0.31722	0.000986	PIP4K2A	0.489779	0.000024	ZNF662	0.71593	5.00E-07
ZNF696	0.317321	0.000954	PSG11	0.490761	0.002079	IL1RAP	0.717086	0.0001
SRARP	0.31742	0.00138	NT5E	0.491285	4.10E-07	EDEM2	0.717797	1.80E-07
TRPV4	0.31764	0.000966	SNAPC1	0.493052	1.10E-06	SLCO4A1	0.717831	0.001606
SH3BP4	0.318026	0.000955	RNU1-148P	0.493407	0.001121	NNAT	0.722086	0.000655
C3ORF36	0.318151	0.000944	CXXC1	0.494715	0.000407	PICSA	0.725268	0.002239
SSUH2	0.319419	0.001779	P4HA2	0.495621	7.40E-07	C12ORF75	0.726732	0.000827
SETD1B	0.319938	0.001939	SH2D3A	0.495798	0.000052	PPP1R14B	0.727469	0.001069
PARS2	0.325729	0.001025	GATA3	0.495933	0.001552	ANKRD24	0.731676	8.80E-06
PAGE1	0.326698	0.001669	FBXO22	0.496587	0.000027	MBD2	0.7319	0.000073
FAM83E	0.326811	0.000682	DVL2	0.496636	0.001926	DNAJB2	0.732097	2.60E-07
CD70	0.329002	0.000625	RNF185-AS1	0.496767	0.000454	CCDC120	0.737004	0.000447
RFPL4B	0.329753	0.002231	DDIT4L	0.49736	0.000619	ARHGEF4	0.739309	0.000399
SIK1	0.33017	0.0006	RN7SL359P	0.498971	0.000989	LOC101929908	0.739831	1.60E-06
NKX2-6	0.331697	0.001979	GSE1	0.49934	0.001879	BCAT2	0.740091	0.000034
ASCL5	0.332529	0.00226	RN7SKP47	0.499776	0.000977	SPTLC3	0.742842	0.000518
DDX53	0.333449	0.000795	RNA5SP415	0.50012	0.000956	GFOD2	0.744214	2.10E-08
CCR2	0.33384	0.000524	IFI44L	0.500178	2.60E-07	CTSA	0.74514	0.000025
TBC1D26	0.336071	0.001387	CRTC1	0.500217	0.001163	YPEL4	0.746133	7.60E-10
FOXB1	0.337086	0.000471	SMOX	0.501302	0.000261	ABCA7	0.747146	3.60E-07
XK	0.337416	0.001197	PFKP	0.50311	0.000228	ADAM12	0.747481	3.10E-06

COG6	0.337588	0.000668	PCSK9	0.504483	0.000031	LOC102723566	0.748184	1.20E-06
RHOBTB1	0.338434	0.000436	IGSF3	0.506315	0.000688	SREBF1	0.750927	0.000077
NSUN5	0.339835	0.000636	GNPTG	0.506396	4.30E-07	NRIP1	0.752369	4.50E-06
KLC3	0.339886	0.000616	RNU6-389P	0.506719	0.000824	DERL3	0.754053	7.30E-08
MT1G	0.339903	0.000398	MIR191	0.507187	0.000262	PKM	0.754388	4.10E-06
GAST	0.340698	0.000403	DEGS1	0.509491	0.00003	NCOR2	0.756617	0.00004
SPRR4	0.341328	0.001802	RNU4-76P	0.509701	0.001512	LDHA	0.75764	0.000732
LINC01270	0.342628	0.001401	RN7SL368P	0.509876	0.00078	DUSP1	0.761482	0.001591
CATSPERD	0.344038	0.000522	RPL6P20	0.50992	0.002329	BTNL9	0.761848	0.000189
JSRP1	0.344602	0.000523	RABAC1	0.510139	8.50E-06	EFNB1	0.763631	0.000482
C1QTNF8	0.344608	0.000656	TRIM71	0.510146	0.000112	CGB5	0.766241	0.000472
C11ORF96	0.344805	0.001275	LGALS8	0.510188	0.001308	GALNT2	0.766963	0.000117
IRF7	0.34532	0.001759	ASIC4	0.510451	0.001701	C10ORF90	0.769457	0.000053
LIFR	0.346589	0.000322	RN7SL864P	0.510698	0.000751	CNPY2	0.769836	4.50E-06
HSD17B14	0.347161	0.000393	DAXX	0.511863	0.001487	TECR	0.770765	0.000164
ATP2C2	0.3478	0.000299	RN7SL309P	0.512404	0.000729	WDR60	0.773368	0.000395
ABRACL	0.347907	0.00092	KRTCAP3	0.513477	2.90E-07	HEXB	0.774843	0.002064
OVOL2	0.348888	0.000291	NOCT	0.514038	0.00019	ILVBL	0.774879	0.000029
FGF9	0.349529	0.001817	CD63	0.514105	3.10E-06	GLOD5	0.778398	5.90E-09
SLC6A11	0.349574	0.000283	SNX9	0.514605	0.001706	BAD	0.778673	0.000196
ASB18	0.35091	0.000526	PIFO	0.514615	7.50E-06	ARRDC3	0.782862	0.000057
DPF1	0.351054	0.001735	SHANK2	0.514803	0.001415	GPR32P1	0.783211	4.30E-07
PODXL	0.351531	0.000266	CD300LG	0.51509	0.001188	UBE2Q2	0.784336	1.50E-07
FER1L6-AS2	0.35166	0.0004	RN7SKP76	0.515672	0.000668	PPL	0.784752	0.0001

RASGEF1A	0.352223	0.000801	TM9SF2	0.515817	1.10E-07	EAF1	0.784776	0.000084
ZBP1	0.353486	0.00024	HCAR3	0.516236	0.002116	RNU4-26P	0.784977	1.70E-06
SULF2	0.35531	0.000607	NEBL	0.516598	0.000056	CSNK2A2	0.787687	0.000721
ABCG1	0.35664	0.000215	BAK1P2	0.517672	0.000638	CRH	0.787944	0.001449
CTSD	0.356842	0.000914	LOC107987235	0.518126	0.000382	FKBP2	0.790693	0.001452
ZNF710	0.3571	0.000208	GDF5OS	0.519052	0.000122	PLIN2	0.791447	0.000136
BIRC7	0.358036	0.002128	RN7SL748P	0.519898	0.000609	RN7SKP296	0.792505	3.10E-07
CEBPA-AS1	0.358114	0.00042	PFKFB4	0.520211	0.000709	MXI1	0.794555	0.000572
FOXJ1	0.358154	0.000201	RRS1	0.520254	0.000219	GAPDH	0.794812	0.000019
C17ORF47	0.358603	0.000649	FANCE	0.520775	3.90E-06	GPX3	0.79647	2.20E-06
NDEL1	0.359279	0.001595	LOC102723968	0.522372	0.000218	CEBPA	0.797692	0.000151
FGF20	0.359557	0.000193	RN7SL712P	0.523106	0.001067	HMGB3P1	0.802927	0.000131
SLC25A4	0.359857	0.000555	HELZ2	0.523557	0.00033	JUP	0.805305	0.000051
PIGZ	0.360512	0.000184	RN7SL34P	0.5242	0.000546	ARID3A	0.807129	0.00003
PTTG1IP	0.361169	0.001315	TFPI	0.525128	0.001923	RNU6-185P	0.808614	8.80E-07
GIPR	0.361213	0.000356	LTF	0.526097	0.001806	TMEM141	0.811172	0.000642
IGFBP5	0.362608	0.000167	RNU6-882P	0.526276	0.000521	ERGIC1	0.813718	0.000019
SLC39A4	0.362657	0.000169	RN7SL419P	0.527712	0.000503	LIMD1	0.815345	0.001922
NVL	0.362897	0.002109	ARMH4	0.527726	0.000421	DLX5	0.818182	0.000017
TRHDE-AS1	0.363488	0.001884	SPCS3	0.528789	0.000208	CCSAP	0.819301	0.000117
ZIC2	0.364366	0.000245	NCOA4	0.530486	0.001139	KRT17P2	0.822043	1.20E-07
OR2W3	0.364427	0.001799	MARCKS	0.530964	0.000464	PHYHIP	0.824171	4.60E-07
LINC00982	0.36449	0.000879	FAM90A6P	0.531325	0.001544	LRRC8A	0.826857	0.001
MGAT1	0.36483	0.001832	DEFB135	0.531872	0.000105	LOC154761	0.827379	0.001989

CDKL2	0.365196	0.000208	ZNF331	0.53324	0.001454	CA4	0.83306	0
ERBB3	0.367622	0.002248	SGPP2	0.535728	4.00E-06	PRKD3	0.833505	0.000108
ATP2A1	0.367785	0.001	EGLN1	0.53579	0.000152	CCDC183	0.833706	1.60E-06
GRK3	0.367906	0.000137	FRMD4A	0.536223	0.000119	NEURL1	0.8362	4.80E-06
PAK6	0.368712	0.000129	PNPLA6	0.536473	0.000408	P4HA1	0.836603	0.000089
FAM3B	0.368934	0.000206	IFNLR1	0.536786	0.000029	ERRFI1	0.844814	0.002054
BASP1-AS1	0.371276	0.000185	ATP8B3	0.536813	0.000207	KDM3A	0.847723	1.10E-07
NECTIN1	0.371339	0.000491	MAML2	0.536845	0.000362	CT47A10	0.84913	5.00E-08
BRI3	0.371539	0.001396	GAPDHP29	0.537422	0.000399	CYP2J2	0.849664	0.000968
IL17C	0.372041	0.000219	TUSC1	0.538625	0.000525	SNORD119	0.850156	0.001505
LINC00575	0.373572	0.001272	ATP6V1D	0.53952	0.000284	PPP1R1C	0.854732	0.001486
RNF185	0.374443	0.002199	IL27RA	0.540013	0.00005	IGSF8	0.865063	0.000167
LOC100287792	0.374758	0.000477	CLC	0.54007	2.50E-07	GPER1	0.870534	1.20E-07
SLC41A2	0.375995	0.000152	MTX1	0.540469	0.000031	TMEM91	0.871012	1.80E-06
C2ORF71	0.376312	0.000458	FBLIM1	0.54169	0.002388	KIF2A	0.871121	0.000022
TPD52L1	0.377263	0.001083	HNRNPA1P7	0.542248	0.001239	SYDE1	0.874717	0.000057
AADAACL4	0.378147	0.001067	LMAN2L	0.542329	0.000051	BTD	0.881932	6.70E-16
ZBED6CL	0.378825	0.000139	RN7SKP286	0.543447	0.000738	SLC2A1	0.882441	1.80E-06
C17ORF77	0.379333	0.000186	ACER3	0.545089	0.000244	GREM2	0.882543	3.10E-12
ANO1	0.380329	0.000082	LINC01556	0.545877	0.000094	RNU6-996P	0.887723	8.60E-08
NIM1K	0.380512	0.000131	RUVBL2	0.547291	0.000446	ACSF2	0.888762	1.70E-09
HYAL2	0.380975	0.00008	RNU6-376P	0.547691	0.000673	MIR193BHG	0.890401	0.000953
DPP7	0.380983	0.000225	RAET1E-AS1	0.548142	0.000309	EPS8L1	0.891898	7.50E-06
FOS	0.380999	0.000079	RNU6-955P	0.548894	0.000655	FAM120AOS	0.896029	5.90E-07

KANSL1L	0.381788	0.000077	IL27	0.549857	0.000048	TMEM45A	0.898615	0.001147
BFAR	0.383524	0.000071	RNASEH2C	0.551025	0.002282	DIO2	0.901835	2.30E-10
PRX	0.384742	0.001938	AMOTL2	0.551473	7.60E-06	MME	0.905447	0.000137
VN2R1P	0.385781	0.000546	LRRC32	0.55259	5.00E-07	OCRL	0.906412	0.001475
ZNF614	0.386623	0.000061	RNA5SP360	0.554915	0.000261	RNA5SP227	0.912723	4.00E-08
SIRT6	0.387212	0.000061	CENPX	0.556455	0.000576	BHLHE40	0.915197	5.50E-06
TDGF1	0.388269	0.002211	PRR31	0.558097	0.000242	RN7SL854P	0.920112	4.60E-09
ELMOD1	0.389754	0.00009	CEBPG	0.560355	0.000513	RNU6-1151P	0.920714	3.10E-08
PPP2R5A	0.390662	0.000307	POR	0.560437	0.000024	NPNT	0.923281	0.000139
FABP4	0.391683	0.000679	RNA5SP210	0.560628	0.000226	SH3BP5	0.924079	0.000371
CNIH4	0.391978	0.00042	SEMA3G	0.564207	1.10E-06	DNM2	0.9262	0
CLDN3	0.391994	0.001181	RIMBP3	0.564225	0.000138	LRRC1	0.92898	0.000232
FGD4	0.39203	0.000083	VSTM2B	0.564551	0.000018	ASTE1	0.932076	0.000205
IGFBP4	0.396284	0.000365	CHPF	0.566734	0.000015	MIR320A	0.938815	0.000058
THEM6	0.396519	0.000661	SNORA14A	0.567292	0.000578	TREM1	0.941062	1.20E-06
ZER1	0.396834	0.000839	RNU2-63P	0.56738	0.000435	KRT19	0.942306	0.000038
ERLIN1	0.398212	0.000272	MIDN	0.567639	0.001242	BIN2P1	0.942937	0.001936
PPP4R4	0.398675	0.000037	ZNF525	0.568273	9.10E-08	HILPDA	0.944923	3.80E-06
SDF4	0.40033	0.000797	NMUR1	0.568521	0.000964	LIMCH1	0.946326	2.50E-06
FER1L4	0.40126	0.001584	RNU1-85P	0.568522	0.000186	SLCO2A1	0.953066	0.000071
OASL	0.402128	0.000631	RN7SL241P	0.569524	0.000183	PPP1R12C	0.961512	0.000015
NUDT1	0.405064	0.000027	SLC6A6	0.571715	0.000046	CORO2A	0.961659	3.70E-11
EDN1	0.405515	0.000929	PREX1	0.571816	0.000011	MIF	0.966552	6.00E-07
NUCB2	0.405807	0.000144	TNIP2	0.572038	0.00059	RN7SL265P	0.971506	6.30E-09

CDC42EP1	0.406313	0.000056	ELL	0.573209	0.000459	SCARNA18	0.974218	0.000315
SEC1P	0.40701	0.001789	MUC1	0.574069	7.60E-07	FTH1P18	0.975719	6.30E-10
SALL4	0.408115	0.000042	RARRES1	0.574188	0.000187	SLC6A10P	0.976456	2.10E-07
POLM	0.408241	0.00171	SNHG12	0.5751	6.90E-06	KCNIP4-IT1	0.977722	0.000315
MOB2	0.408571	0.000052	MAFF	0.575578	0.000322	RNU6-694P	0.980143	4.80E-09
C1QTNF12	0.408748	0.000105	RN7SL381P	0.577166	0.00015	SCARB1	0.98188	0.000069
SPRN	0.410153	0.001996	RNA5SP484	0.579303	0.000143	SPDYE10P	0.986322	0.000281
TMEM38B	0.410477	0.000022	GJD3	0.57932	0.000011	SCARNA10	0.997419	0.00162
FLJ34503	0.411297	0.001612	ELF3	0.579535	0.000025	SFXN3	1.00123	0.000031
DUSP5	0.411657	0.001067	RNU6-934P	0.580327	0.000324	NEK11	1.007	0.000223
KDM6B	0.411815	0.000877	MED12L	0.580647	0.001744	TBC1D22A	1.00755	2.30E-08
IGSF9B	0.414529	0.000063	SEMA7A	0.582736	0.000362	CSNK1E	1.01025	3.40E-08
CHRM2	0.41576	0.000144	FBXL16	0.583233	5.20E-06	RDH13	1.01499	0.000013
CCDC127	0.41645	0.001581	PPP1R15A	0.583988	2.20E-09	TPI1	1.02016	0.000291
PRR23A	0.417381	0.001591	LSR	0.584086	0.000015	TPBG	1.02291	0.000058
NR1I3	0.41929	0.000639	DVL1	0.587307	0.000037	SH3PXD2A	1.0271	5.50E-06
RPN1	0.419632	0.000575	PP7080	0.589523	0.000137	ENG	1.03026	0.000756
PEG3	0.420783	0.001474	RN7SL797P	0.590172	0.000108	HK2	1.04648	0
APOBEC2	0.420949	0.000082	SSR4	0.590387	5.90E-08	SNORD91A	1.04711	0.001579
HIST2H2AA4	0.421869	0.001232	YOD1	0.592243	0.000308	ERO1A	1.04907	0.000017
C9ORF47	0.426162	0.000053	LIMS4	0.592557	0.00003	ARHGAP45	1.04946	0.000208
SNORA70F	0.426569	0.002395	ANG	0.593251	0.00027	CGB3	1.05149	0.000034
GRAMD1B	0.4286	0.000392	PRSS40B	0.594931	0.000168	HTRA1	1.05296	4.60E-10
ESYT1	0.428714	0.000934	LINC00189	0.595312	0.000219	INHA	1.05391	0.00003

GCHFR	0.428986	9.30E-06	ARHGEF16	0.596126	0.000983	LOC729870	1.08425	0.00008
TOR1B	0.430111	0.000028	IGHV3-35	0.596557	0.000091	FSTL3	1.08441	0.000389
TTC39A	0.430959	0.000367	HPS3	0.596612	0.000087	RNU6-1150P	1.09086	9.30E-12
PHLDB3	0.431928	0.00005	QPRT	0.598298	0.000324	SNORA18	1.09691	0.000069
MDS2	0.432118	0.002105	SEC24A	0.599586	0.000033	C5ORF66-AS2	1.12448	0.000697
TNIP1	0.432203	0.000466	AKAP12	0.599974	0.00003	PLA2G16	1.12703	0.000014
NFATC2IP	0.433331	0.000607	RN7SL180P	0.601506	0.00008	BCL6	1.1438	0.000016
ASB5	0.433824	0.000024	SP8	0.60157	0.000264	KMT2E-AS1	1.14927	0.000164
NCAPD2	0.434537	0.000207	OR52K3P	0.601731	0.00004	NECTIN4	1.15054	0.000023
ODAPH	0.43534	0.001521	RN7SKP63	0.602319	0.000078	SIGLEC6	1.15527	2.70E-06
BRI3BP	0.435413	0.000013	MIR365A	0.604509	0.002172	QPCT	1.15945	5.10E-09
ADSSL1	0.435858	0.000749	PFKL	0.604889	3.80E-08	PROCR	1.16068	2.80E-09
SYNGR1	0.437966	5.90E-06	RNU6-504P	0.604939	0.000183	HTRA4	1.17151	0.000415
HCFC2	0.438399	0.000916	PRRG1	0.605065	0.000188	GCNT2P	1.17533	0.000035
CTU1	0.438482	0.000011	WDR86	0.605704	1.40E-06	AKR1B10P1	1.19024	0.001432
SCAND2P	0.438664	0.000841	DDA1	0.605729	0.000447	ST13P6	1.2005	0.001525
LOC100131496	0.438999	0.000768	C5ORF52	0.606706	0.001526	INHBA	1.20556	0.000175
RPS27L	0.439734	5.30E-06	TMEM107	0.607073	0.000678	NPHP3-ACAD11	1.21333	8.10E-14
ONECUT1	0.440231	5.40E-06	CLN3	0.607566	0.000117	MIR532	1.2247	0.001225
MIR4697HG	0.440749	0.000473	GLRX	0.607589	8.70E-09	MIR492	1.26117	6.70E-10
FGD6	0.441269	0.000029	ZNF175	0.609888	3.20E-06	PAPPA2	1.26257	8.20E-09
GLB1L2	0.441609	0.000486	MLPH	0.6108	1.40E-06	SPAG4	1.26354	3.10E-08
CEMP1	0.442706	0.000347	NARF	0.612134	0.000749	FLT1	1.28657	1.60E-06
TLCD2	0.44282	0.000711	RNA5SP292	0.612677	0.000153	LOC100507507	1.29082	0.002146

CIART	0.442883	0.000039	PMEL	0.614056	0.001372	GAPDHP33	1.30126	0.000641
BPIFB3	0.442958	0.000483	SSR1	0.614645	2.20E-06	HTT-AS	1.30339	0.001976
HEXIM1	0.444294	0.00163	PLBD1	0.617132	0.00006	AP1B1P1	1.30359	0.001974
MZT2A	0.444772	0.000037	CRIP1P4	0.617279	0.000053	LINC01056	1.32156	0.001753
H1F0	0.445445	0.000053	LOC399886	0.617938	0.000064	LOC100507547	1.32962	0.001663
C8G	0.445475	4.20E-06	TMEM234	0.618539	4.70E-08	CACNA1G-AS1	1.34742	0.001479
RECQL4	0.445748	4.10E-06	HIST2H2AA3	0.619928	0.000014	SASH1	1.35482	4.40E-16
RNF151	0.446598	0.000024	GALK1	0.621054	0.000077	DNAH21	1.35693	0.001389
ADGRG6	0.446877	3.80E-06	PNPLA2	0.622224	0.000884	LOC100506514	1.35827	0.001376
CPTP	0.447416	0.000216	B9D2	0.623695	0.000041	RAD51-AS1	1.35965	0.001364
RMDN3	0.448204	0.00051	CALM1	0.626369	0.000077	LOC400212	1.36196	0.001343
PPIB	0.448445	0.002247	DCP2	0.627289	0.000148	LINC00928	1.36461	0.00132
KRT86	0.448674	7.10E-06	CCNY	0.627827	0.000215	LINC00106	1.37583	0.001226
CAPN15	0.448728	0.000744	BACH1	0.628469	0.00025	CDRT3	1.37653	0.00122
AFTPH	0.450517	0.000572	CSF2RB	0.628613	2.20E-07	LINC01776	1.37907	0.0012
PLPPR4	0.45237	2.90E-06	MIR375	0.629078	7.60E-06	PICART1	1.38157	0.00118
RIMBP3C	0.452738	0.00109	FTH1P4	0.630633	0.000037	GAPDHP23	1.3965	0.000308
USP18	0.453067	0.001797	PHPT1	0.631647	0.000317	LINC01538	1.44221	0.000789
SLC25A37	0.453205	0.000636	PRKAG3	0.634212	0.000331	LINC02181	1.44463	0.000776
RNF38	0.454064	0.001301	NTF4	0.634296	2.20E-09	ALOX12-AS1	1.45427	0.000728
LIN7B	0.455838	0.000253	GLRXP3	0.635007	6.10E-06	BACE1-AS	1.45785	0.000711
ITPR3	0.456806	0.000449	ORMDL2	0.635489	0.000169	LEP	1.4688	0.000095
TLE2	0.457136	0.000166	TMEM151A	0.635513	0.001167	LOC100289333	1.48388	0.000598
KCNJ6	0.457341	2.30E-06	SMARCA1	0.635991	0.000189	SDCBPP3	1.49087	0.00013

P4HA3	0.45805	0.000015	CGB8	0.63702	0.00205	LINC01684	1.5054	0.000518
RN7SKP50	0.458877	0.002395	LINGO3	0.637139	1.50E-06	LOC100505585	1.50767	0.000511
ZNF865	0.459136	0.001118	SERPINA3	0.638321	0.000592	TUSC8	1.52483	0.000456
RN7SL217P	0.459269	0.002384	MTMR4	0.638811	0.000571	CFLAR-AS1	1.52611	0.000452
RN7SL87P	0.459686	0.002359	ACBD4	0.639035	0.000037	CEBPZOS	1.53007	0.00044
RN7SL708P	0.460139	0.002317	VWA1	0.639777	0.001028	WDR86-AS1	1.54523	0.000034
RN7SL735P	0.460954	0.002316	RN7SL364P	0.642038	0.000027	MORC1-AS1	1.60945	0.00026
HBM	0.460968	0.000285	ZSWIM4	0.643221	0.001961	LINC00592	1.63924	0.000213
MXRA5Y	0.46166	0.002242	BNIP3L	0.643525	0.000217	SNHG14	1.66669	0.000178
LOC100294362	0.46224	0.002216	WFS1	0.645453	0.000028	LOC91450	1.67905	0.000164
C1ORF21	0.462699	0.000236	SIPA1L1	0.645534	0.001178	LOC145474	1.69893	0.000144
BICD2	0.463556	1.70E-06	SLC11A1	0.646314	9.30E-06	LOC729867	1.73052	0.000117
MX1	0.463859	0.000115	P2RY2	0.650092	1.50E-08	LINC02288	1.81783	0.000066
MIR31	0.464665	0.000835	SLC11A2	0.651698	0.000024	MIF-AS1	1.91604	2.80E-07
DNASE2	0.464849	0.000015	MAP3K8	0.655373	2.50E-11	LINC00421	1.9969	0.000021
RN7SL263P	0.464924	0.002095	TXNP4	0.656033	0.000053	LOC283038	2.03076	0.000017
OLAH	0.465879	1.50E-06	NTRK2	0.65747	0.002144	SLC2A1-AS1	2.31501	3.00E-06
RN7SL258P	0.465968	0.002054	KCNF1	0.657517	1.90E-11	MIG7	2.42697	0.00031
PGBD2	0.466903	0.002107	IER2	0.657795	0.000379	LOC100268168	3.04849	6.90E-08
CDK19	0.468061	0.001106						

Πίνακας 7 Τα ΔΕΓ που προέκυψαν από την πολυμεταβλητή μετα-ανάλυση

Νόσος του Crohn		
Γονίδιο	Μέγεθος επίδρασης	p-value
MAPK14	0.739521	0.000000000014
GK	0.719576	0.000000000011
GBP2	0.721228	0.000000000013
MUC1	0.684937	0.000000000038
H2BFS	1.45281	0.000000000043
KREMEN1	0.780149	0.000000000024
GNAI3	0.652637	0.000000000029
OPLAH	0.649633	0.000000000032
USB1	0.648913	0.000000000039
DSC2	0.639449	0.0000000001
ALPL	0.614051	0.00000000027
CD55	0.610327	0.00000000034
SH2D1B	-0.70448	0.00000000044
PGS1	0.630272	0.00000000059
LIMK2	0.842782	0.00000000074
GALNS	0.591774	0.0000000013
MMP9	0.629649	0.0000000013
STAT1	0.593578	0.0000000018
PHTF1	0.588195	0.0000000022
KIF1B	0.580611	0.0000000024
WDR33	-0.66913	0.0000000024
PFKFB3	0.574893	0.0000000026
IFNAR1	0.602539	0.0000000027
CACNA1E	0.571418	0.0000000029
TMEM8A	0.570592	0.0000000031
ID3	-0.65562	0.0000000033
GALNT14	0.56786	0.0000000044
PLPP3	0.560853	0.0000000051
CR1	0.571683	0.0000000054
SRPK1	0.560963	0.0000000055

SERTAD2	-0.56247	0.00000057
ACSL1	0.608012	0.00000062
SOCS3	0.750625	0.00000067
CRACR2B	0.643708	0.00000084
TMEM185B	0.55284	0.00000085
CEACAM1	0.606386	0.00000011
FTH1P5	1.13012	0.00000011
UBE2J1	0.56489	0.00000013
N4BP2L2-IT2	-0.84645	0.00000019
CDCA5	0.617236	0.00000025
PARP9	0.731591	0.0000003
SCARF1	0.620118	0.0000003
AVL9	0.539909	0.00000034
PPP1R15A	0.648524	0.00000034
WDFY3	0.539776	0.00000036
EXOC6	0.607712	0.00000039
TMEM243	-0.55711	0.00000044
EIF5	-0.54493	0.00000045
PSTPIP2	0.519978	0.00000048
VIPAS39	0.524379	0.0000005
DYSF	0.645547	0.00000059
KCNJ15	0.787891	0.00000061
IL1R1	0.617605	0.00000066
TTC26	0.510377	0.00000067
SLC26A8	0.806335	0.00000071
TRIB1	0.86291	0.00000072
NTNG2	0.59614	0.00000074
PHF21A	0.514691	0.00000078
PGD	0.508543	0.00000082
RPL31P17	-1.04729	0.00000093
DLGAP5	0.503322	0.00000095
TYW1	-0.56028	0.00000095
STAM2	0.540224	0.00000098
ENTPD1	0.653973	0.00000099

ADGRG3	0.732174	0.0000011
NACA	-0.56882	0.0000011
FFAR2	0.550776	0.0000012
FAM53C	0.502336	0.0000014
LINC00963	0.781503	0.0000014
OSM	0.648819	0.0000015
PLIN3	0.501632	0.0000015
ARG1	0.531585	0.0000016
CD274	0.823752	0.0000016
PDK3	0.657807	0.0000016
LINC00528	0.648732	0.0000017
PLPP2	0.523466	0.0000017
RPL18	-0.53628	0.000002
ADM	0.564599	0.0000021
F5	0.555233	0.0000021
FUT7	0.48748	0.0000021
CHAD	0.4928	0.0000022
SBNO2	0.565324	0.0000022
TRAC	-0.76251	0.0000023
VNN2	0.486764	0.0000023
XPO6	0.896072	0.0000023
GTPBP2	0.536433	0.0000025
GYG1	0.671906	0.0000025
DOK4	0.49002	0.0000026
FAM43A	-0.63668	0.0000026
NCF4	0.636907	0.0000026
CEACAM3	0.672241	0.0000029
LINC02363	0.556629	0.000003
TUBA1A	0.696393	0.000003
BMX	0.480629	0.0000031
PRICKLE4	-0.98894	0.0000032
TMUB2	0.526106	0.0000032
IL1R2	0.49247	0.0000034
SLC2A3	0.615852	0.0000036

CDK14	0.473765	0.000042
PYGL	0.526888	0.000044
HMMR	0.467598	0.000051
EXO1	0.530942	0.000053
ST3GAL2	0.481299	0.000055
DSTYK	-0.47693	0.000058
TM9SF1	0.524247	0.000058
KIR2DS1	-0.5342	0.00006
SIPA1L1	0.568179	0.00006
ITGB4	0.463563	0.000061
CDC47	0.562718	0.000064
RPL32	-0.50155	0.000087
RARA-AS1	1.18959	0.000089
TUBA1B	0.611746	0.000090
RAB20	0.46402	0.000093
TM2D3	-0.45429	0.000094
SIGLEC5	0.538219	0.000095
C1RL	0.454192	0.00001
TIFA	0.524833	0.00001
HMGB3	0.520606	0.000011
LOC100507507	1.17383	0.000011
SH2B2	0.464457	0.000011
TECPR2	0.593009	0.000011
OSGEP	-0.53131	0.000012
TRPM6	0.481356	0.000012
CDC45	0.446406	0.000013
NLRX1	0.676895	0.000013
SLC2A3P2	0.633853	0.000013
SQOR	0.692725	0.000013
CSRNP1	0.630791	0.000014
FEZ1	-0.8278	0.000014
GPR27	0.44825	0.000014
PADI2	0.462035	0.000014
SOX1	0.561619	0.000014

CDCA2	0.61667	0.000015
FCGR2C	0.51045	0.000015
RPL23AP43	-0.58293	0.000015
CASP7	0.457318	0.000016
GMP	0.440057	0.000018
MTHFD2	0.455302	0.000018
NFIL3	0.642663	0.000018
RXYLT1	-0.44017	0.000018
TRIP6	0.829395	0.000018
LINC01094	1.13959	0.000019
PPP1R3B	0.519209	0.000019
TP53I11	0.5308	0.000019
DNAJC5	0.708636	0.00002
PMS2P8	-0.88711	0.00002
SEMA4A	0.526893	0.00002
GSEC	1.13452	0.000021
KAZN	0.435285	0.000021
MCTP1	0.470965	0.000021
SSBP2	-0.43502	0.000021
TESMIN	0.44477	0.000021
C5AR2	0.434861	0.000022
ADAM9	0.487862	0.000023
KNL1	0.442612	0.000024
AURKA	0.431496	0.000025
PHF1	-0.44935	0.000025
PML	0.581823	0.000025
UBTD1	0.440106	0.000026
SAMD9L	0.508396	0.000027
SH3GLB1	0.437285	0.000027
ANXA3	0.687837	0.000028
VAPA	0.478214	0.000028
RUBCN	0.454689	0.000029
SAMSN1	0.584079	0.000029
VNN1	0.428253	0.000029

PSMB8-AS1	1.11025	0.000030
TBX21	-0.59452	0.000030
KIF2C	0.42643	0.000031
KLRC3	-0.42558	0.000033
ACSL4	0.735211	0.000034
CD177	0.524755	0.000034
NUCB1	0.4561	0.000034
C16ORF86	0.5723	0.000035
BUB1	0.422067	0.000037
GRB10	0.516793	0.000037
RPL36AP8	-0.86652	0.000039
ETV7	0.450516	0.00004
CRISP2	0.419812	0.000041
GBP5	0.499501	0.000041
IGFBP3	-0.42766	0.000041
MMP24OS	0.706207	0.000041
S1PR5	-0.52354	0.000041
LOC100505549	-1.0878	0.000042
GBA	0.607512	0.000044
KLHL2	0.544317	0.000044
LOC100507557	-1.08265	0.000045
POR	0.584273	0.000045
C1ORF174	-0.92692	0.000046
LRG1	0.501172	0.000046
ATF6	0.416488	0.000047
BCL6	0.679295	0.000047
PTCH1	-0.42286	0.000047
RPS10	-0.46631	0.000047
CATIP-AS1	1.07854	0.000048
MTMR9	-0.41618	0.000048
PADI4	0.455006	0.00005
MSRB1	0.628873	0.000052
CXCR1	0.46377	0.000053
CYB5R4	0.413701	0.000053

GBAT2	1.07247	0.000053
ETV6	0.511121	0.000054
LOC644265	-0.84869	0.000054
ZYX	0.755996	0.000054
HAUS4	0.659976	0.000056
ITPKC	0.424714	0.000056
B3GNT8	0.548852	0.000057
KIFC1	0.411741	0.000057
CAV1	0.432761	0.000058
LILRA5	0.411693	0.000059
AQP9	0.752544	0.000062
ERO1A	0.472862	0.000062
LILRB3	0.564311	0.000062
ARMC5	0.644413	0.000063
CLTCL1	0.415494	0.000063
LRPAP1	0.518233	0.000064
PTGDS	-0.79695	0.000064
TFEB	0.4091	0.000064
KIF20A	0.432481	0.000065
SERPINA5	0.408436	0.000065
NSUN7	0.412352	0.000066
RPL19P12	-0.5371	0.000066
THOC5	0.810202	0.000066
JUNB	0.407811	0.000067
TNPO2	-0.62982	0.000067
OARD1	-0.42817	0.000068
CCNF	0.487583	0.000069
HPGD	0.504797	0.00007
TAP1	0.678987	0.00007
MTSS1	-0.46546	0.000074
BCL3	0.486476	0.000081
RPS2P48	-0.53267	0.000082
SMARCD3	0.617211	0.000082
SNX13	0.520788	0.000083

ATG14	-0.74427	0.000084
FOSL2	0.506278	0.000085
FLOT2	0.560649	0.000086
CCDC15	-0.40148	0.000087
FBXL13	0.49015	0.000087
RPL15P5	-0.82379	0.000087
RPS8	-0.44526	0.000088
SIGLEC9	0.506705	0.000089
SNX20	0.582277	0.00009
SLC16A3	0.611566	0.000093
MED16	0.399116	0.000095
STX11	0.414715	0.000095
TLR6	0.644738	0.000095
RAD52	-0.40136	0.000096
FAM90A1	0.530193	0.000097
Ελκώδης κολίτιδα		
Γονίδιο	Μέγεθος επίδρασης	p-value
ALPL	0.626957	0.00000019
ARHGAP27P1	0.574747	0.0000052
BUB1	0.513134	0.0000038
C1ORF174	-0.68161	0.00000026
CA12	0.492095	0.0000098
CAMKK1	0.552749	0.0000097
CD177	0.52584	0.0000093
CD55	0.559302	0.0000051
CDCA5	0.58506	0.0000029
CR1	0.529502	0.0000062
DLGAP5	0.526772	0.0000021
FBXO17	0.48357	0.000013
GLDC	0.545772	0.000001
GPRC5D	0.61921	0.0000051
IFNAR1	0.533976	0.000014
IGHG1	1.22505	0.0000024

IGHV1-69	0.622751	0.000002
IGKC	1.24049	0.00000014
IGLC1	1.42208	0.000000069
IGLJ3	1.24119	0.0000018
KIF4A	0.50259	0.0000087
LINC00963	0.898353	0.000003
LMNB1	0.552499	0.0000029
LOC100505549	-1.54161	0.00000068
LOC101928896	-1.50055	0.0000012
MAPK13	0.551263	0.000011
MKI67	0.514777	0.0000049
MMP9	0.574016	0.0000015
NDUFAF1	0.560827	0.000001
PFKFB3	0.493626	0.000011
PPP1R15A	0.50431	0.0000064
RARA-AS1	1.4197	0.0000039
RRM2	0.606491	0.0000016
SLC2A3	0.61759	0.000000047
SRPK1	0.546177	0.000013
TM2D3	-0.52705	0.0000021
TYW1	-0.49907	0.0000075
XPO6	0.679153	0.0000053

Συνολικός Έλεγχος

Γονίδιο	Μέγεθος	Μέγεθος	<i>p</i> -value του
	επίδρασης UC	επίδρασης CD	συνολικού ελέγχου
MAPK14	0.739521	0.575865	0.0000000000056
GK	0.719576	0.45143	0.0000000000087
GBP2	0.721228	0.311567	0.000000000011
H2BFS	1.45281	0.682675	0.0000000003
MUC1	0.684937	0.415682	0.00000000032
ALPL	0.614051	0.626957	0.0000000006
MMP9	0.629649	0.574016	0.0000000016

GNAI3	0.652637	0.473805	0.000000017
KREMEN1	0.780149	0.5373	0.000000017
LIMK2	0.842782	0.71536	0.000000018
USB1	0.648913	0.546312	0.000000019
SOCS3	0.750625	0.470511	0.000000021
OPLAH	0.649633	0.401009	0.000000026
DSC2	0.639449	0.480293	0.000000033
CD55	0.610327	0.559302	0.000000037
XPO6	0.896072	0.679153	0.000000047
SH2D1B	-0.70448	-0.23413	0.000000049
SLC2A3	0.615852	0.61759	0.000000006
TRIP6	0.829395	0.316846	0.00000002
PGS1	0.630272	0.469819	0.000000024
ACSL1	0.608012	0.491126	0.000000025
GALNS	0.591774	0.568692	0.000000025
IGKC	0.446977	1.24049	0.000000029
RRM2	0.51786	0.606491	0.000000032
CR1	0.571683	0.529502	0.000000034
PFKFB3	0.574893	0.493626	0.000000056
SRPK1	0.560963	0.546177	0.000000089
PARP9	0.731591	0.486971	0.000000096
IFNAR1	0.602539	0.533976	0.00000011
IL1R1	0.617605	0.465519	0.00000011
KIF1B	0.580611	0.581477	0.00000011
C1ORF174	-0.92692	-0.68161	0.00000012
TMEM8A	0.570592	0.46177	0.00000012
STAT1	0.593578	0.232914	0.00000013
CEACAM1	0.606386	0.404044	0.00000014
PHTF1	0.588195	0.490476	0.00000014
WDR33	-0.66913	-0.3718	0.00000016
PPP1R15A	0.648524	0.50431	0.00000019
CACNA1E	0.571418	0.343004	0.00000021
ID3	-0.65562	-0.55672	0.00000021
SERTAD2	-0.56247	-0.50389	0.00000021

KCNJ15	0.787891	0.718377	0.0000022
TOR1A	0.490314	0.572101	0.0000022
VIPAS39	0.524379	0.609459	0.0000023
CDCA5	0.617236	0.58506	0.0000024
IGHG1	-0.09159	1.22505	0.0000026
SECTM1	0.727088	0.367424	0.0000027
PLPP3	0.560853	0.407948	0.0000029
GALNT14	0.56786	0.402958	0.0000031
DLGAP5	0.503322	0.526772	0.0000038
TYW1	-0.56028	-0.49907	0.0000039
LMNB1	0.494763	0.552499	0.0000041
IGLC1	0.438281	1.42208	0.0000042
CD274	0.823752	0.435603	0.0000049
LINC00963	0.781503	0.898353	0.0000051
TMEM185B	0.55284	0.320336	0.0000055
CRACR2B	0.643708	0.448354	0.0000056
SCARF1	0.620118	0.400496	0.0000065
FTH1P5	1.13012	0.583443	0.0000072
UBE2J1	0.56489	0.427476	0.0000008
WDFY3	0.539776	0.402825	0.0000082
TMEM243	-0.55711	-0.5877	0.0000009
ENTPD1	0.653973	0.576965	0.0000096
LINC02363	0.556629	0.097231	0.0000099
OSM	0.648819	0.519242	0.000001
SLC26A8	0.806335	0.717226	0.0000011
EXO1	0.530942	0.451671	0.0000012
N4BP2L2-IT2	-0.84645	-0.57058	0.0000013
PLPP2	0.523466	0.431266	0.0000013
PSTPIP2	0.519978	0.204334	0.0000014
RPL31P17	-1.04729	-0.8609	0.0000015
TM2D3	-0.45429	-0.52705	0.0000015
EXOC6	0.607712	0.4843	0.0000016
FGF9	-0.45969	-0.52287	0.0000016
NACA	-0.56882	-0.33732	0.0000016

AVL9	0.539909	0.35746	0.0000017
LOC100505549	-1.0878	-1.54161	0.0000023
CEACAM3	0.672241	0.511129	0.0000026
PGD	0.508543	0.581812	0.0000027
NDUFAF1	0.356186	0.560827	0.0000028
ACSL4	0.735211	0.541766	0.0000029
DYSF	0.645547	0.705784	0.0000029
EIF5	-0.54493	-0.65275	0.0000029
CASP5	0.625171	0.318752	0.0000031
MGAM	0.448001	0.509984	0.0000031
PDK3	0.657807	0.64045	0.0000031
PHF21A	0.514691	0.480755	0.0000031
MATN4	-0.03239	0.440044	0.0000032
CHAD	0.4928	0.453545	0.0000034
NTNG2	0.59614	0.435293	0.0000034
ARG1	0.531585	0.355944	0.0000037
TRIB1	0.86291	0.748032	0.0000037
RARA-AS1	1.18959	1.4197	0.0000039
CDCA7	0.562718	0.518926	0.000004
PCLAF	0.522322	0.492664	0.000004
ADGRG3	0.732174	0.672204	0.0000042
TTC26	0.510377	0.338647	0.0000042
ARHGAP27P1	0.614616	0.574747	0.0000043
ZNF438	0.841968	0.655647	0.0000045
BUB1	0.422067	0.513134	0.0000046
IGLJ3	0.213705	1.24119	0.0000046
POLB	-0.17316	-0.57468	0.0000046
NCF4	0.636907	0.597949	0.0000047
VNN2	0.486764	0.452222	0.0000047
CDCA2	0.61667	0.404171	0.0000051
FFAR2	0.550776	0.383156	0.0000058
FAM53C	0.502336	0.503766	0.000006
STAM2	0.540224	0.476626	0.0000062
GLDC	0.304978	0.545772	0.0000063

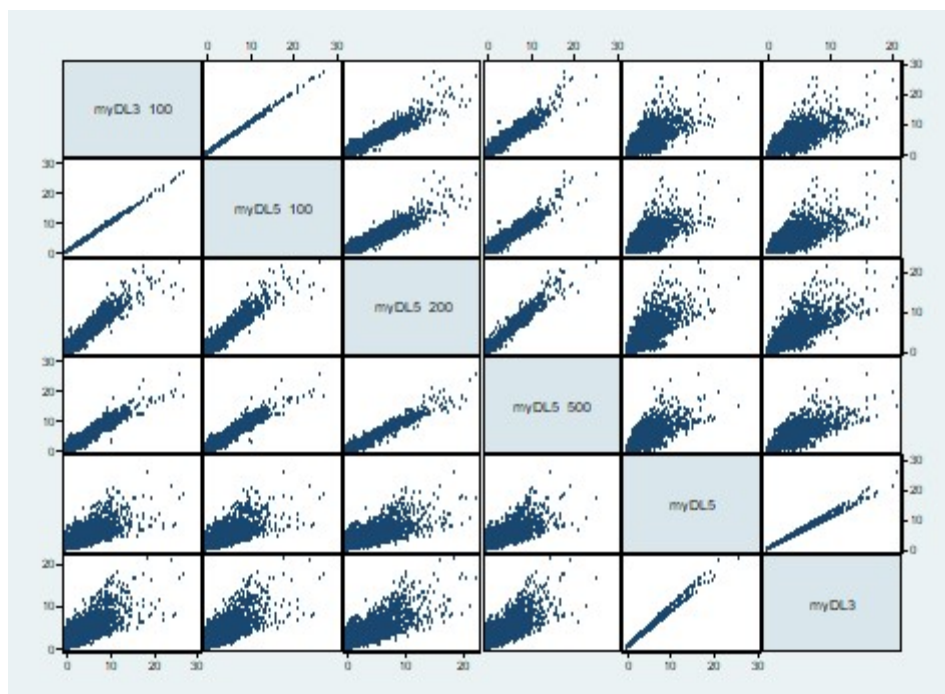
CDC45	0.446406	0.478378	0.0000065
FUT7	0.48748	0.408297	0.0000074
HMMR	0.467598	0.442037	0.0000074
LOC101928896	-0.82397	-1.50055	0.0000077
SIPA1L1	0.568179	0.394597	0.0000077
BMX	0.480629	0.455628	0.000008
PLIN3	0.501632	0.341614	0.0000085
IGHV1-69	0.236199	0.622751	0.0000091
GPRC5D	0.402476	0.61921	0.0000096
CDK14	0.473765	0.413715	0.0000099
CAV1	0.432761	0.473465	0.000011
LINC00528	0.648732	0.275383	0.000011
SBNO2	0.565324	0.547846	0.000011
INSL3	0.54483	0.337243	0.000012
PBLD	0.427891	0.483134	0.000012
RPL18	-0.53628	-0.37268	0.000012
TECPR2	0.593009	0.436226	0.000012
ADM	0.564599	0.567227	0.000013
CCNF	0.487583	0.440277	0.000013
CD177	0.524755	0.52584	0.000013
DOK4	0.49002	0.438447	0.000013
F5	0.555233	0.66357	0.000013
TRAC	-0.76251	-0.4147	0.000013
CSRNP1	0.630791	0.571205	0.000014
GTPBP2	0.536433	0.610079	0.000014
TMUB2	0.526106	0.361552	0.000014
TUBA1A	0.696393	0.67806	0.000014
DNAJC5	0.708636	0.783554	0.000015
GYG1	0.671906	0.546341	0.000015
HMGB3	0.520606	0.373058	0.000015
AURKA	0.431496	0.457211	0.000016
F8	0.512283	0.69257	0.000016
FAM43A	-0.63668	-0.59543	0.000016
TIFA	0.524833	0.159799	0.000016

CA12	0.413801	0.492095	0.000017
ITGB4	0.463563	0.396095	0.000017
ZYX	0.755996	0.711724	0.000017
KIF20A	0.432481	0.451908	0.000018
KNL1	0.442612	0.454941	0.000019
IL1R2	0.49247	0.367656	0.00002
KIR2DS1	-0.5342	-0.06595	0.00002
PRICKLE4	-0.98894	-0.49165	0.00002
BCAT1	0.427352	0.449007	0.000023
CATIP-AS1	1.07854	1.31091	0.000023
NECTIN2	0.581962	0.383348	0.000023
PPP1R3B	0.519209	0.522801	0.000023
ALKAL2	-0.58869	-0.49532	0.000024
GBA	0.607512	0.514796	0.000024
WARS	0.545412	0.38566	0.000024
CDKN1C	-0.46261	-0.15394	0.000025
ECE1	0.454549	0.412631	0.000025
GMIP	0.440057	0.424219	0.000025
LINC00294	-0.59667	-0.32959	0.000025
KAZN	0.435285	0.427647	0.000026
PYGL	0.526888	0.541827	0.000026
TM9SF1	0.524247	0.672235	0.000026
ADAM18	0.408722	0.465864	0.000027
POLG2	-0.40597	-0.46605	0.000027
PADI2	0.462035	0.365507	0.000028
SLC2A3P2	0.633853	0.440074	0.000028
TACR3	-0.15511	0.300311	0.000028
MKI67	0.337138	0.514777	0.000029
GBP5	0.499501	0.043644	0.00003
KIR2DL1	-0.47424	-0.09382	0.000031
ST3GAL2	0.481299	0.357597	0.000031
DSTYK	-0.47693	-0.3786	0.000033
FCF1	-0.24085	-0.63531	0.000034
MMP24OS	0.706207	0.65742	0.000034

FBXO17	0.358049	0.48357	0.000035
RPL32	-0.50155	-0.37417	0.000036
SELL	0.332779	0.553661	0.000036
GSEC	1.13452	1.1692	0.000038
ZWINT	0.313079	0.506609	0.000038
GZMB	-0.8212	-0.27677	0.000039
NNT-AS1	-1.0033	-1.30375	0.000041
CKAP4	0.422788	0.457591	0.000042
NLRX1	0.676895	0.711226	0.000045
LINC01094	1.13959	1.12577	0.000046
NSUN7	0.412352	0.43265	0.000046
MAPK13	0.440285	0.551263	0.000047
C1RL	0.454192	0.383566	0.000049
CWH43	0.423502	0.432248	0.000049
KIF4A	0.359538	0.50259	0.000049
PITPNM3	0.36167	0.474852	0.00005
BCL3	0.486476	0.556956	0.000051
CAMKK1	0.35216	0.552749	0.000051
DRAM1	0.592815	0.501623	0.000051
RUBCN	0.454689	0.344533	0.000051
TRPM6	0.481356	0.380289	0.000052
TUBA1B	0.611746	0.468969	0.000052
LOC100507507	1.17383	0.977536	0.000053
SIGLEC5	0.538219	0.584509	0.000053
RAB20	0.46402	0.303667	0.000054
ATP7B	0.513541	0.321202	0.000055
MTHFD2	0.455302	0.319547	0.000055
NFATC4	0.343018	0.482904	0.000055
ZNF552	0.472987	0.373704	0.000055
KIFC1	0.411741	0.421898	0.000056
FCGR2C	0.51045	0.232898	0.000057
LOC100507557	-1.08265	-1.18607	0.000057
TP53I11	0.5308	0.669395	0.000059
TESMIN	0.44477	0.354131	0.00006

C16ORF86	0.5723	0.280045	0.000061
RPS16	-0.39228	-0.45203	0.000061
TAP1	0.678987	0.366085	0.000061
ANXA3	0.687837	0.563711	0.000062
SOX1	0.561619	0.369224	0.000062
CDCA3	0.393407	0.451145	0.000063
KIF11	0.393452	0.450132	0.000063
SH2B2	0.464457	0.358585	0.000063
OSGEP	-0.53131	-0.56307	0.000064
CDK1	0.445096	0.438112	0.000065
CCDC115	-0.50523	-0.19195	0.000067
JUNB	0.407811	0.41865	0.000067
RPL23AP43	-0.58293	-0.26833	0.000067
RXYLT1	-0.44017	-0.34774	0.000069
SQOR	0.692725	0.521545	0.00007
VAPA	0.478214	0.162042	0.00007
FEZ1	-0.8278	-0.49571	0.000073
NCAPG	0.44973	0.441652	0.000073
SLC16A3	0.611566	0.489196	0.000073
ETV7	0.450516	0.072082	0.000075
KLHL2	0.544317	0.298276	0.000076
FAS	0.525229	0.324106	0.000077
GPR27	0.44825	0.340341	0.000078
LOC644265	-0.84869	-0.73968	0.000078
ADAM9	0.487862	0.32153	0.000079
VNN1	0.428253	0.371626	0.000079
CXCR1	0.46377	0.453467	0.00008
GAPDH	0.422623	0.401675	0.00008
VPS9D1	0.765517	0.409128	0.00008
LILRB3	0.564311	0.321661	0.000082
MTMR9	-0.41618	-0.3947	0.000082
NUP43	-0.43274	-0.4805	0.000082
KIF2C	0.42643	0.372325	0.000083
SLC22A18	0.430757	0.111471	0.000084

TOP2A	0.314762	0.47982	0.000084
MSRB1	0.628873	0.591158	0.00009
CASP7	0.457318	0.302406	0.000091
LOC101927069	-0.79684	-1.30468	0.000094
UVRAG-DT	0.991873	1.2153	0.000094
PDE11A	0.376048	0.435861	0.000096
PML	0.581823	0.306068	0.000097
CRISP2	0.419812	0.140547	0.000098
NFIL3	0.642663	0.639744	0.000098
ERO1A	0.472862	0.336945	0.000099
DLG1	-0.53017	-0.70781	0.0001
GIGYF2	-0.72008	-0.39449	0.0001
PSMB8-AS1	1.11025	1.02354	0.000101
SSBP2	-0.43502	-0.21226	0.000101
ZNF254	0.46865	0.419122	0.000101



Εικόνα 1 Διαγράμματα του αρνητικού λογάριθμου των p -values των διαφόρων ελέγχων ανά ζεύγη σε μικρό σύνολο δεδομένων

ado file με τον κώδικα της μονομεταβλητής μετα-ανάλυσης για το Stata

```
program define myDL5, rclass
    version 10.1
    syntax , x(varlist numeric max=1) Type(varlist numeric max=1)
Study(varlist numeric max=1) [ Effect(string ) ]
preserve
tempvar last
qui bysort `study': gen `last'=_n==_N
qui sum `study'
local max=r(max)
tempvar x0 x1 sd0 sd1 n0 n1 d sed m gmfm gmfm_1 J g seg
qui gen double `x0'=.
qui gen double `x1'=.
qui gen double `sd0'=.
qui gen double `sd1'=.
qui gen double `n0'=.
qui gen double `n1'=.
qui gen double `d'=.
qui gen double `sed'=.
qui gen double `m'=.
qui gen double `gmfm'=.
qui gen double `gmfm_1'=.
qui gen double `J'=.
qui gen double `g'=.
qui gen double `seg'=.
forvalues i=1(1)`max' {
    qui sum `x' if `study'==`i' & `type'==0
    qui replace `x0'=r(mean) if `study'==`i'
    qui replace `sd0'=r(sd) if `study'==`i'
    qui replace `n0'=r(N) if `study'==`i'
    qui sum `x' if `study'==`i' & `type'==1
    qui replace `x1'=r(mean) if `study'==`i'
    qui replace `sd1'=r(sd) if `study'==`i'
    qui replace `n1'=r(N) if `study'==`i'
    qui replace `m'= (`n1'+ `n0'-2) if `study'==`i'
    qui replace `d'=(`x1' - `x0')/(sqrt( ( (`n1'- 1)*(`sd1')^2 +(`n0'-
1)*(`sd0')^2)/(`n1' + `n0' -2 ))) if `study'==`i'
```

```

    qui replace `sed`= sqrt( (`n1' + `n0')/(`n1'*`n0')
+(`d')^2/(2*(`n1'+`n0'))) if `study'==`i'
    qui replace `gmfm' = exp((lngamma(`m'/2))) if `study'==`i'
    qui replace `gmfm_1'= exp((lngamma((`m'-1)/2))) if `study'==`i'
    qui replace `J' = `gmfm'/((sqrt(`m'/2))*`gmfm_1') if `study'==`i'
    qui replace `g' = `J'*`d' if `study'==`i'
    qui replace `seg'=(`J')^2*`sed' if `study'==`i'
** var(g)=J*var(d)

}

qui keep if `last'
qui keep if `x' !=.
metan `g' `seg' , `effect' nograph
**metan `n1' `x1' `sd1' `n0' `x0' `sd0', `effect' nograph
**metan `n1' `x1' `sd1' `n0' `x0' `sd0',fixed nograph hedges
**metan `d' `sed' , random nograph
**list
Restore
return scalar g=r(ES)
return scalar seg=r(seES)

end

```

Κώδικας Stata για πραγματοποίηση μονομεταβλητής μετα-ανάλυσης

```

set more off
file open meta using resultmyDL5.txt, write append
file write meta "gene"
file write meta ","
file write meta "g"
file write meta ","
file write meta "seg" _n

foreach var of varlist albg - vtrna3_1p {
    di ""
    di ""
    di in ye "Performing meta-analysis for `var' gene"
    di ""
}

```

```

myDL5,x(`var') type(type) study(study) effect(random)

file write meta "`var'"
file write meta ","
file write meta "`r(g)'"
file write meta ","
file write meta "`r(seg)'"_n
}
file close meta

```

Κώδικας Stata για πραγματοποίηση μονομεταβλητής μετα-ανάλυσης με bootstrap

```

set more off
file open meta using resultsbootstrap5.txt, write append
file write meta "gene"
file write meta ","
file write meta "g"
file write meta ","
file write meta "se" _n

foreach var of varlist albg-vtrna3_1p{
    di ""
    di ""
    di in ye "Performing meta-analysis for `var' gene"
    di ""
    bootstrap g=r(g), nowarn nohead reps(200) strata(study type):
myDL5,x(`var') type(type) study(study) effect(random)

    mat g=e(b)
    local g=g[1,1]
    mat se=e(se)
    local se=se[1,1]

    file write meta "`var'"
    file write meta ","
    file write meta "`g'"
    file write meta ","
    file write meta "`se'"_n
}

```

```
}  
file close meta
```

Κώδικας Stata για πραγματοποίηση πολυμεταβλητής μετα-ανάλυσης

```
set maxvar 32767  
  
use "test13.dta",clear  
  
set more off  
  
file open meta using mvmeta.txt, write append  
  
file open metan using single_study.txt, write append  
  
file write meta "gene"  
file write meta " "  
file write meta "b1"  
file write meta " "  
file write meta "se_b1"  
file write meta " "  
file write meta "p_b1"  
file write meta " "  
file write meta "b2"  
file write meta " "  
file write meta "se_b2"  
file write meta " "  
file write meta "p_b2"  
file write meta " "  
file write meta "global1"  
file write meta " "  
file write meta "global2"_n  
  
file write metan "gene"  
file write metan " "  
file write metan "b1"  
file write metan " "  
file write metan "se_b1"
```

```

file write metan " "
file write metan "z_b1"
file write metan " "
file write metan "p_b1"
file write metan " "
file write metan "df_b1"
file write metan " "
file write metan "b2"
file write metan " "
file write metan "se_b2"
file write metan " "
file write metan "z_b2"
file write metan " "
file write metan "p_b2"
file write metan " "
file write metan "df_b2"
file write metan " "
file write metan "w1"
file write metan " "
file write metan "global1"
file write metan " "
file write metan "w2"
file write metan " "
file write metan "global2" _n

qui sum study
local x=r(max)

foreach var of varlist alcf - znf888 {
preserve

    qui gen last=0
    qui gen x0=0
    qui gen x1=0
    qui gen x2=0
    qui gen n0=0
    qui gen n1=0
    qui gen n2=0

```

```

qui gen sd0=0
qui gen sd1=0
qui gen sd2=0
qui gen d1=0
qui gen d2=0
qui gen v11=0
qui gen v22=0
qui gen v12=0
qui gen n=0
qui gen m=0
    qui gen double gmfm=0
qui gen double gmfm_1=0
qui gen J=0
qui gen b1=0
    qui gen b2=0
qui gen V11=0
qui gen V22=0
qui gen V12=0
qui gen se1=0
qui gen se2=0

forvalues i=1/`x' {

    qui sum `var' if study==`i' & status==0
    qui replace x0=r(mean) if study==`i'
    qui replace n0=r(N) if study==`i'
    qui replace sd0=r(sd) if study==`i'

    qui sum `var' if study==`i' & status==1
    qui replace x1=r(mean) if study==`i'
    qui replace n1=r(N) if study==`i'
    qui replace sd1=r(sd) if study==`i'

    qui sum `var' if study==`i' & status==2
    qui replace x2=r(mean) if study==`i'
    qui replace n2=r(N) if study==`i'
    qui replace sd2=r(sd) if study==`i'

```

```

    qui replace n= (n1+ n2+ n0) if study=='i'
    qui replace m= (n1+ n2+ n0 -3) if study=='i'
    qui replace d1=(x1 - x0)/(sqrt( ( (n1 - 1)*(sd1^2)) +((n2 -
1)*(sd2^2)) + ((n0 - 1)*(sd0^2)))/(m)) if study=='i'
    qui replace d2=(x2 - x0)/(sqrt( ( (n1 - 1)*(sd1^2)) +((n2 -
1)*(sd2^2)) + ((n0 - 1)*(sd0^2)))/(m)) if study=='i'
    qui replace v11= (1/n1) + (1/n0) + ((d1^2)/(2*(n))) if
study=='i'
    qui replace v22= (1/n2) + (1/n0) + ((d2^2)/(2*(n))) if
study=='i'
    qui replace v12= (1/n0) +((d1*d2)/(2*(n))) if study=='i'
    qui replace gmfm = exp((lngamma(m/2))) if study=='i'
    qui replace gmfm_1 = exp((lngamma((m-1)/2))) if study=='i'
    qui replace J = gmfm/((sqrt(m/2))*gmfm_1) if study=='i'
    qui replace b1 = (J*(d1)) if study=='i'
    qui replace b2 = (J*(d2)) if study=='i'
    qui replace V11 = ((J^2)*(v11)) if study=='i'
    qui replace V22 = ((J^2)*(v22)) if study=='i'
    qui replace V12 = ((J^2)*(v12)) if study=='i'
    qui replace se1=sqrt(V11)
    qui replace se2=sqrt(V22)

}

bysort study : replace last=_n==_N
    keep if last==1

    qui summarize x0
    qui return list

if (r(N)==1) {

    keep if x0!=.

    qui metan b1 se1
    local var_es1= r(ES)

```



```

local var_sees1= r(seES)
local var_z1= r(z)
local var_p1= r(p_z)
local var_df1= r(df)

    qui metan b2 se2
local var_es2= r(ES)
local var_sees2= r(seES)
local var_z2= r(z)
local var_p2= r(p_z)
local var_df2= r(df)

qui mat d=(b1[1],b2[1])
qui mat list d
qui mat g=(b1[1]\b2[1])
qui mat list g
qui mat v=(V11[1],V12[1]\V12[1],V22[1])
qui mat list v
qui mat V=invsym(v)
qui mat list V
qui mat W=d*V*g
qui mat list W
local var_w1=W[1,1]
local var_global1=1-chi2(2, W[1,1])

qui mat w2=(b1[1]-b2[1])/ (sqrt( (V11[1]+V22[1])-(2*V12[1])))
qui mat list w2
local var_w2=w2[1,1]
local var_global2=2*(1-normal(abs( w2[1,1] )))

dis `var_es1', `var_sees1', `var_z1', `var_p1', `var_df1', `var_es2',
`var_sees2', `var_z2', `var_p2', `var_df2', `var_w1', `var_w2',
`var_global1', `var_global2'

    file write metan "`var'"
    file write metan " "
    file write metan "`var_es1'"
    file write metan " "

```

```

file write metan "`var_sees1'"
file write metan " "
file write metan "`var_z1'"
file write metan " "
file write metan "`var_p1'"
file write metan " "
file write metan "`var_df1'"
file write metan " "
file write metan "`var_es2'"
file write metan " "
file write metan "`var_sees2'"
file write metan " "
file write metan "`var_z2'"
file write metan " "
file write metan "`var_p2'"
file write metan " "
file write metan "`var_df2'"
file write metan " "
file write metan "`var_w1'"
file write metan " "
file write metan "`var_global1'"
file write metan " "
file write metan "`var_w2'"
file write metan " "
file write metan "`var_global2'"_n
}

```

```

else {
  qui mvmeta b V, vars(b1 b2) mm
  qui matrix v=e(V)

  qui matrix a=e(b)

  qui matrix c=e(V)

  ** coef b1
  local var_b1=a[1,1]
}

```

```

** coef b2
    local var_b2=a[1,2]

** std.err sto tetragono b1
    local var_st_b1=c[1,1]

** std.err b1
local se_b1=sqrt(c[1,1])

** std.err sto tetragono b2
    local var_st_b2=c[2,2]

** std.err b2
    local se_b2=sqrt(c[1,1])

**calculation of p-value b1
    local var_p_b1=2*(1-normal(abs(a[1,1]/(sqrt(c[1,1])))))

**calculation of p-value b2
    local var_p_b2=2*(1-normal(abs(a[1,2]/(sqrt(c[2,2])))))

qui test b1 b2
qui return list
local var_g1= r(p)

qui test b1=b2
qui return list
local var_g2= r(p)

dis `var_b1', `se_b1', `var_p_b1', `var_b2', `se_b2', `var_p_b2',
`var_g1', `var_g2'

file write meta "`var'"

```

```
file write meta " "  
file write meta "`var_b1'"  
file write meta " "  
file write meta "`se_b1'"  
file write meta " "  
file write meta "`var_p_b1'"  
file write meta " "  
file write meta "`var_b2'"  
file write meta " "  
file write meta "`se_b2'"  
file write meta " "  
file write meta "`var_p_b2'"  
file write meta " "  
file write meta "`var_g1'"  
file write meta " "  
file write meta "`var_g2'"_n  
}  
restore  
}  
  
file close metan  
  
file close meta
```