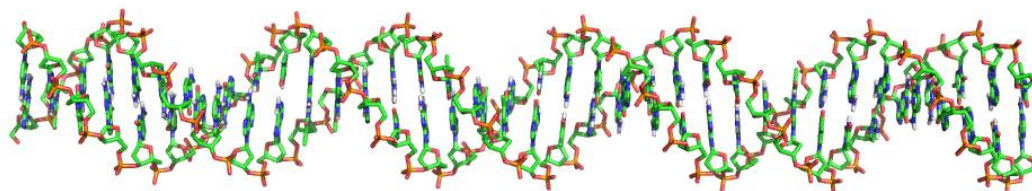




ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

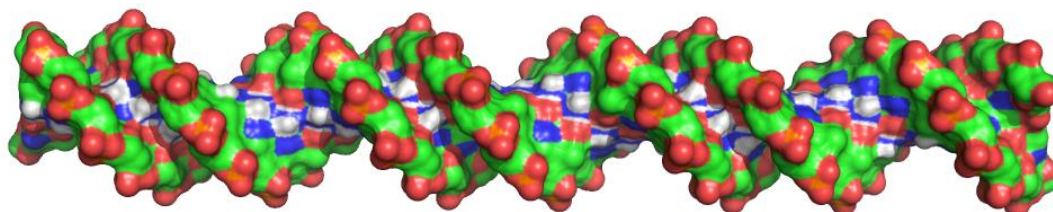
**Ανίχνευση βιοδεικτών μετά από έκθεση σε ιοντίζουσα ακτινοβολία
μέσω της βιοπληροφορικής ανάλυσης δεδομένων μικροσυστοιχιών
DNA και μετα-ανάλυσης**



ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΒΑΣΙΛΕΙΟΥ ΧΡΙΣΤΙΝΑ



Επιβλέπων : Αλέξανδρος Γεωργακίλας
Αναπληρωτής Καθηγητής Ε.Μ.Π.

Σε συνεργασία με τον: Δρ. Ιωάννη Μιχαλόπουλο, Ε.Λ.Ε. Β' Ι.Ι.Β.Ε.Α.Α.

Αθήνα, Ιούλιος 2017



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ
ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

Ανίχνευση βιοδεικτών μετά από έκθεση σε ιοντίζουσα ακτινοβολία μέσω της βιοπληροφορικής ανάλυσης δεδομένων μικροσυστοιχιών DNA και μετα-ανάλυσης.

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

ΒΑΣΙΛΕΙΟΥ ΧΡΙΣΤΙΝΑ

Επιβλέπων : Αλέξανδρος Γεωργακίλας
Αναπληρωτής Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την:

**Μπάγκος
Παντελής**
Αν. Καθηγητής
Παν/μιο Θεσσαλίας
(Υπογραφή)

.....

**Γεωργακίλας
Αλέξανδρος**
Αν. Καθηγητής
ΕΜΠ
(Υπογραφή)

.....

**Μακροπούλου-Λουκογιαννάκη
Μυρσίνη**
Αν. Καθηγήτρια
ΕΜΠ
(Υπογραφή)

.....

Αθήνα, Ιούλιος 2017

© 2017 – All rights reserved

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα αρχικά να ευχαριστήσω όλους όσους συνέβαλαν στην επιτυχή εκπόνηση αυτής της διπλωματικής εργασίας. Ειδικότερα, ευχαριστώ θερμά τον καθηγητή μου κ. Αλέξανδρο Γεωργακίλα Αναπληρωτή Καθηγητή του ΕΜΠ για την ανάθεση και επίβλεψη αυτής της διπλωματικής εργασίας, για την πολύτιμη βοήθειά του καθώς και για την εμπιστοσύνη που μου έδειξε κατά τη διάρκεια της εκπόνησής της. Στη συνέχεια, ευχαριστώ ιδιαίτερα τον ερευνητή του Ιδρύματος Ιατροβιολογικών Ερευνών της Ακαδημίας Αθηνών, Δρ. Ιωάννη Μιχαλόπουλο για την συνεργασία που είχαμε και τη βοήθεια που μου προσέφερε, παρέχοντας μου τα απαραίτητα γνωστικά εργαλεία και τη σωστή καθοδήγηση. Θα ήθελα ακόμη να ευχαριστήσω τον κ. Παντελή Μπάγκο αναπληρωτή Καθηγητή του πανεπιστημίου Θεσσαλίας και την ερευνητική ομάδα του στο Πανεπιστήμιο Θεσσαλίας για την συνεισφορά τους και την καθοδήγηση τους σε θέματα μετα-ανάλυσης. Ευχαριστώ εκ νέου τον Αναπ. Καθηγητη. κ. Π. Μπάγκο αλλά και την Ανάπ.Καθηγήτρια κα Μ.Μακροπούλου που δέχθηκαν να συμμετέχουν στην τριμελή εξεταστική επιτροπή μου.

Περίληψη

Εισαγωγή

Η αυξανόμενη χρήση ιοντιζουσών ακτινοβολιών σε πλήθος βιοιατρικών εφαρμογών ακτινοθεραπείας και ακτινοδιάγνωσης δημιουργεί την ανάγκη μελέτης των επιπτώσεων της. Επιπτώσεις και τρόποι αλληλεπίδρασης που συνδέονται με υψηλής δόσης ακτινοβολίες είναι γνωστές ως ένα βαθμό. Σκοπός της παρούσας μελέτης είναι η διερεύνηση των κινδύνων στις χαμηλότερες δόσεις με μετα-ανάλυση σε συνδυασμό μελετών γονιδιακής έκφρασης μικροσυστοιχιών DNA.

Μέθοδο

Επιλέχθηκαν δείγματα γονιδίων από ανθρώπινο ιστό ακτινοβολημένα με χαμηλή δόση ακτινοβολίας από μικροσυστοιχίες DNA και ελέγχθηκε σε αυτά η διαφορική έκφραση τους. Τα πρωτογενή δεδομένα από τις μελέτες λήφθηκαν από το GEO και ανήκουν στις σειρές GSE8917, GSE12435, GSE16935, GSE23901, GSE29344, GSE52918, GSE59861, υποβλήθηκαν σε διόρθωση υποβάθρου, κανονικοποίηση ποσοστημορίου και λογαρίθμιση με την βοήθεια του πακέτου της Limma από την σουίτα Bioconductor. Στη συνέχεια με δεδομένα από το Biomart σχεδιάστηκε η βάση δεδομένων MySQL "Microarrays" με σκοπό την αποθήκευση των δεδομένων μας και την διεξαγωγή μετα-ανάλυσης για την επιλογή των στατιστικά σημαντικών διαφορικά εκφρασμένων γονιδίων. Χρησιμοποιώντας την STRING.db και το WebGestalt, καταφέραμε να σχεδιάσουμε γονιδιακά δίκτυα και να καθορίσουμε τις βιολογικές διεργασίες που σχετίζονται με την απόκριση στις χαμηλές δόσεις ιοντίζουσας ακτινοβολίας.

Αποτελέσματα

Στα αποτελέσματα της ανάλυσης αυτής για την επίδραση των υψηλών δόσεων ιοντίζουσας ακτινοβολίας βρέθηκαν γονίδια που συνδέονται με την εμφάνιση καρκίνου, ενεργοποίηση του αναστολέα p21, ενεργοποίηση μεταγραφικών παραγόντων, όπως ο CUX1, ο AP1 και ρύθμιση μέσω miRNA. Όσον αφορά την επίδραση των χαμηλών δόσεων βρέθηκαν γονίδια που σχετίζονται με την ενεργοποίηση μηχανισμών επιδιόρθωσης DNA, τη διακοπή του κυτταρικού κύκλου, το μεταβολισμό πυριμιδινών με χαμηλό ωστόσο επίπεδο εμπιστοσύνης.

Συζήτηση

Τα αποτελέσματα των υψηλών δόσεων όπως αναμένεται διακρίνονται έντονα. Ωστόσο ο συσχετισμός της απόκρισης στις χαμηλές δόσεις δεν είναι ξεκάθαρος. Διαφαίνεται η ενεργοποίηση miRNA που συμμετέχουν στην αποσιώπηση γονιδίων καθώς και λειτουργιών όπως ρύθμιση της μετάβασης από G0 σε G1, αρνητική ρύθμιση της διαφοροποίησης των νευρώνων, μονοπάτια για το μεταβολισμό πυριμιδινών. Ωστόσο λόγω του χαμηλού επιπέδου στατιστικής σημαντικότητας των αποτελεσμάτων ενδείκνυται η διενέργεια περαιτέρω αναλύσεων και ο σχεδιασμός νέων μελετών

Λέξεις-κλειδιά: ιοντίζουσες ακτινοβολίες, χαμηλές δόσεις, υψηλές δόσεις, διαφορική έκφραση γονιδίων, βιοπληροφορική, ανάλυση μικροσυστοιχιών, βάσεις δεδομένων, μετα-ανάλυση, βιοδείκτες, βιολογικά δίκτυα

Abstract

Introduction

The increasing use of ionizing radiation in a number of biomedical applications such as radiotherapy and radiodiagnosis creates the need to study its effects. Effects and modes of interaction associated with high-dose irradiation are known. The purpose of this study was to discover through the use of bioinformatics and DNA microarray analysis proper gene biomarkers in the case of exposure to low doses of ionizing radiation. By using meta-analysis, we aimed to strengthen our results and reveal a set of genes to be used as possible low dose exposure markers.

Method

Human genes were selected from human tissue irradiated with low dose irradiation from microarrays and their differential expression was analyzed. The raw data were taken from GEO and belong to the series GSE8917, GSE12435, GSE16935, GSE23901, GSE29344, GSE52918, GSE59861, which underwent background correction, quantile normalization and log transformation with the help of the Limma package from the Bioconductor Suite. Then, with data from Biomart, the MySQL database 'Microarrays' was designed to store our data and perform meta-analysis to select statistically significant differentially expressed genes. Using STRING.db and WebGestalt, we were able to design gene networks and determine the biological processes associated with the response to low doses of ionizing radiation.

Results

In the results of this analysis for the effect of high doses of ionizing radiation, genes associated with the onset of cancer, activation of the p21 inhibitor, activation of transcription factors like CUX1, AP1 and miRNA regulation were found. Regarding the effect of low doses, genes related to the activation of DNA repair mechanisms, cell cycle arrest, pyrimidine metabolism were found though with a low level of confidence.

Discussion

The results of high doses as expected are clearly distinguishable. However, the correlation of the low dose response is not clear. MiRNA activation appears to be involved in gene silencing as well as functions such as regulation of the transition from G0 to G1, negative regulation of neuron differentiation, pathways for pyrimidine metabolism. However, due to the low level of statistical significance of the results, it is advisable to carry out further analyzes and to design new studies

Key words: ionizing radiation, low doses, high doses, differential gene expression, bioinformatics, microarray analysis, databases, meta-analysis, biomarkers, biological networks

There are no facts, only interpretations.

Frederick Nietzsche (1844-1900)

Στους γονείς μου και στο σκύλο μου, Πέρλα.

ΠΕΡΙΕΧΟΜΕΝΑ

Εισαγωγή.....	14
Ιοντιζουσα ακτινοβολια.....	14
Προέλευση ακτινοβολίας	17
Εφαρμογές ιοντιζουσας ακτινοβολίας.....	18
Αλληλεπίδραση ακτινοβολίας με την Υλη	22
Επιδράσεις της ιοντιζουσας ακτινοβολίας στα κύτταρα	25
Βλάβες και επιδιόρθωση DNA.....	32
Τρόποι μελέτης επίδρασης ‘χαμηλών’ δόσεων (<0.5Gy).....	39
Μικροσυστοιχίες.....	41
Είδη μικροσυστοιχιών	45
Βιολογικές Βάσεις Δεδομένων.....	58
G.E.O (Gene Expression Omnibus).....	59
ArrayExpress.....	59
MIAME.....	60
Βάσεις Δεδομένων.....	61
Μετα-ανάλυση.....	64
Μέθοδος.....	68
Αποτελέσματα βιοπληροφορικής αναζήτησης	68
Διόρθωση υποβάθρου	72
Λογαρίθμηση.....	72
Κανονικοποίηση.....	73
GSE52918	76
GSE29344	83
GSE59861	89
Βάση Δεδομένων Microarrays.....	98

Μετα-ανάλυση.....	133
Βιβλιογραφική αναζήτηση	134
Αποτελέσματα.....	135
Αποτελέσματα μετα-ανάλυσης για υψηλές και χαμηλές δόσεις Ιοντίζουσας Ακτινοβολίας	136
Ανάλυση GO	158
Ανάλυση KEGG.....	161
Ανάλυση Transcription Factor Target.....	167
Ανάλυση MicroRNA Target.....	171
Απεικόνιση Βιολογικών Δικτύων.....	177
Γονίδια με απόκριση σε ιοντίζουσα ακτινοβολία βάσει βιβλιογραφίας	181
Κοινός Γεωμετρικός Τόπος Αποτελεσμάτων	185
Συζήτηση-Προοπτικές.....	188
Βιβλιογραφία.....	195

Εικόνα 1: Διάσπαση-α.	14
Εικόνα 2: Ανασχετική ισχύς για θετικά μόνια σε χαλκό.....	23
Εικόνα 3: Αριστερά: Ο σχηματισμός της αιχμής Bragg. Δεξιά: Θεραπεία πρωτονίων και η κορυφή Bragg συγκριτικά με την καμπύλη εναπόθεσης δόσης μέσω ακτινοβολίας X.....	24
Εικόνα 4: Επιδράσεις ακτινοβολίας χαμηλού/υψηλού LET.....	26
Εικόνα 5: Σχηματική απεικόνιση πιθανών αλληλεπιδράσεων λόγω ραδιόλυσης ύδατος.....	28
Εικόνα 6: Οι τέσσερις φάσεις ενός τυπικού κυτταρικού κύκλου ευκαρυωτικού κυττάρου. Οι φάσεις G1, S, G2 αποτελούν τη μεσόφαση, εφόσον το κύτταρο αυξάνεται συνεχώς. Η κυτταρική διαίρεση συμβαίνει στη φάση M. Η σύνθεση του DNA περιορίζεται στη φάση S. Η φάση G1 είναι η περίοδος ανάμεσα στη φάση M και S, ενώ η G2 ανάμεσα στη φάση S και M. Η διάρκεια του κυτταρικού κύκλου 24 ώρες είναι ενδεικτική. Υπάρχουν κύτταρα με κυτταρικούς κύκλους μεγαλύτερης και μικρότερης διάρκειας.	31
Εικόνα 7: Βλάβες ενός ή δύο κλώνων, εκτομή νουκλεοτιδίου και ομαδοποιημένες βλάβες.....	33
Εικόνα 8: Οι δύο γενικές κατηγορίες των βιολογικών αποτελεσμάτων των ακτινοβολιών: τα καθορισμένα (για την περιοχή των μεγάλων δόσεων) και τα στοχαστικά αποτελέσματα.....	36
Εικόνα 9: Η εξάρτηση από τη δόση της ακτινοβολίας της πιθανότητας εμφάνισης των στοχαστικών και μη αποτελεσμάτων.	38
Εικόνα 10: Τρόποι μελέτης επίδρασης χαμηλών δόσεων.....	40
Εικόνα 11: Human Genome Project (Πηγή: http://humangenomeproject.yolasite.com/background.php)	41
Εικόνα 12: Χρονική εξέλιξη γενωμικής ανάλυσης (Πηγή: (Lander et al., 2001)).....	42
Εικόνα 13: Συσκευές τύπωσης μικροσυστοιχιών. Αριστερά βρίσκεται η μέθοδος εξ' επαφής και δεξιά η μέθοδος εκτύπωσης με εκτόξευση μελάνης (Πηγή: http://www.slideshare.net/ruchibioinfo/microarray).....	46
Εικόνα 14: Βιολογική Επεξεργασία δείγματος για χρήση σε μικροσυστοιχία Illumina.....	47
Εικόνα 15: Διάταξη BeadArray (Fan et al., 2006).....	48
Εικόνα 16: Σχηματική απεικόνιση ενός ολιγομερούς που προσκολλάται στο σφαιρίδιο. Τα πραγματικά σφαιρίδια έχουν προσκολλημένα στην επιφάνεια τους εκατοντάδες χιλιάδες αντίγραφα της ίδιας ακολουθίας (source: http://www.ohsu.edu/).....	48
Εικόνα 17: Τύποι μικροσυστοιχιών Illumina.....	49
Εικόνα 18: Οι δύο τύποι μικροσυστοιχιών της Illumina. Αριστερά η SAM και δεξιά η Sentrix BeadChip.....	50
Εικόνα 19: Αποκωδικοποίηση μικροσυστοιχιών Illumina.....	51

Εικόνα 20: Αντίστροφη μεταγραφή και δημιουργία cDNA βιβλιοθήκης.....	53
Εικόνα 21: Διαδικασία υβριδισμού.	53
Εικόνα 22: Φωτολιθογραφία με χρήση μασκών (τεχνολογία που εφαρμόζει η εταιρεία Affymetrix για την κατασκευή μικροσυστοιχιών DNA).....	54
Εικόνα 23: Ανιχνευτές PM/MM.	56
Εικόνα 24: Είδη αρχείων Affymetrix.....	57
Εικόνα 25: Ενημέρωση για νέες σειρές δεδομένων μέσω GEO.	70
Εικόνα 26: Κατανομή αναφοράς.....	73
Εικόνα 27: Αναδιάταξη τιμών με βάση τη δοκιμαστική κατανομή αναφοράς.....	74
Εικόνα 28: Κανονικοποίηση ποσοστημορίου τριών κατανομών.....	75
Εικόνα 29:Chip Illumina HumanHT-12 V4.0 expression beadchip, product information.....	76
Εικόνα 30:ΑντικείμενοEListRaw(47.221 παρατηρήσεις από 24 μεταβλητές).....	79
Εικόνα 31:Κανονικοποιημένα δεδομένα Illumina.....	80
Εικόνα 32:Δείγμα τιμών των κανονικοποιημένων δεδομένων από το Rstudio.....	81
Εικόνα 33:Boxplot δεδομένων σειράς GSE52918 πριν και μετά τη κανονικοποίηση με τη μέθοδο <code>negc()</code>	82
Εικόνα 34: HumanRef-8 Beadchip Illumina.....	83
Εικόνα 35: Boxplot δεδομένων σειράς GSE29344 πριν τη κανονικοποίηση.....	87
Εικόνα 36: Boxplot δεδομένων σειράς GSE29344 μετά τη κανονικοποίηση με τη μέθοδο <code>negc()</code>	88
Εικόνα 37: GeneChip® PrimeView™.	91
Εικόνα 38: Boxplot κανονικοποιημένων δεδομένων (17.593 παρατηρήσεις σε 22 μεταβλητές) της Affymetrix.....	93
Εικόνα 39: Box Plot Probe Cell Intensity πριν την επεξεργασία.....	94
Εικόνα 40: Relative Probe Cell Intensity.....	94
Εικόνα 41:Κανονικοποιημένα δεδομένα με RMA της GSE59861.....	95
Εικόνα 42: Box Plot Relative Signal της GSE59861. Δε διαπιστώνεται κάποιο ιδιαίτερα αποκλίνον δείγμα.....	96
Εικόνα 43: Ανάλυση κύριων συνιστωσών (PCA) σε δεδομένα έντασης κυψελών ανίχνευσης (CEL). Συστάδες γονιδίων παρουσιάζουν ισχυρή συσχέτιση σημάτων έκφρασης.....	97
Εικόνα 44: Ανάλυση κύριων συνιστωσών (PCA) σε δεδομένα .CHP.	97
Εικόνα 45: Διάγραμμα Συσχέτισης Οντοτήτων βάσης δεδομένων Microarrays.	98
Εικόνα 46: : Control-Low (16-24hrs),GO, adjP<0.1. Ρίζα της οντολογίας δεξιά είναι ο όρος: GO:0003674, molecular_function και αριστερά ο όρος: GO:0008150, biological_process	160

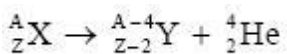
Εικόνα 47: Εντερικό ανοσοποιητικό δίκτυο για παραγωγή IgA, Control-Low,(16-24)hrs, adjP<0.1,KEGG.....	164
Εικόνα 48: Λοίμωξη από <i>Staphylococcus aureus</i> , Control-Low,(16-24)hrs, adjP<0.1,KEGG.....	165
Εικόνα 49: Μεταβολισμός Πυριμιδίνων, Control-Low,(16-24)hrs, adjP<0.1,KEGG.....	166
Εικόνα 50: Βιολογικό δίκτυο γονιδίων όπως προέκυψαν από το σύνολο των δειγμάτων που εκτέθηκαν σε υψηλή δόση ακτινοβολίας και συλλέχθηκαν στο διάστημα μεταξύ 16-24 ωρών.....	178
Εικόνα 51: Βιολογικά δίκτυα διαφορετικά εκφρασμένων γονιδίων, όπως προέκυψαν από το σύνολο των δειγμάτων που εκτέθηκαν σε χαμηλή δόση ακτινοβολίας και συλλέχθηκαν στο διάστημα μεταξύ 16 και 24 ωρών, ορίζοντας την ουδό του score αλληλεπιδράσεων σε: A) 0.7 (υψηλή σημαντικότητα) και B) 0.4 (μέση σημαντικότητα).	179
Εικόνα 52: Διάγραμμα Venn μεταξύ των διαφορετικά εκφρασμένων γονιδίων της παρούσας μελέτης, των γονιδίων από τη βιβλιογραφία και τέλος τη λίστα γονιδίων που σχετίζονται με καρκίνο.....	186

ΕΙΣΑΓΩΓΗ

ΙΟΝΤΙΖΟΥΣΑ ΑΚΤΙΝΟΒΟΛΙΑ

Η ιοντίζουσα ακτινοβολία είναι ακτινοβολία υψηλής ενέργειας που μπορεί να προκαλέσει ιονισμό, να φορτίσει δηλαδή την ύλη διώχνοντας ηλεκτρόνια από τα άτομα και δημιουργώντας ιόντα, καθώς και να σπάσει τους δεσμούς των χημικών ενώσεων. Θεωρείται επικίνδυνη για τους ζωντανούς οργανισμούς καθώς μπορεί να προκαλέσει καρκίνο με την αλλοίωση των μορίων DNA (United Nations. Scientific Committee on the Effects of Atomic Radiation., 1988). Ιοντίζουσες είναι οι ακτινοβολίες που μεταφέρουν ενέργεια ικανή να εισχωρήσει στην ύλη τέτοια ώστε να προκαλέσει ιονισμό των ατόμων, να διασπάσει βίαια χημικούς δεσμούς και να προκαλέσει βιολογικές βλάβες στον ανθρώπινο οργανισμό (Hüttermann et al., 1978). Ο ιοντισμός ενός ουδέτερου ατόμου είναι η βίαιη απομάκρυνση ενός ηλεκτρονίου από τις στοιβάδες του, λόγω εξωτερικού αιτίου, με αποτέλεσμα την παραγωγή δύο αντίθετα φορτισμένων ιόντων, του θετικού ατόμου (ιόντος) και του αρνητικού ηλεκτρονίου. Η πιο διεισδυτική ιοντίζουσα ακτινοβολία είναι η ηλεκτρομαγνητική και η ακτινοβολία νετρονίων, τα οποία έχουν ουδέτερο φορτίο και δεν αντιδρούν με τα άτομα της ύλης. Ιοντίζουσα ακτινοβολία είναι: η υπεριώδης ακτινοβολία, οι ακτίνες X, οι ακτίνες γ, η ακτινοβολία άλφα, η ακτινοβολία βήτα, ακτινοβολία νετρονίων που έχουν υψηλές ταχύτητες (ταχέα νετρόνια) (Celitans, 1973).

Ακτινοβολία α: Σωματιδιακή ακτινοβολία που αποτελείται από δύο πρωτόνια και δύο νετρόνια. Έχει μικρή διεισδυτικότητα και μπορεί να θωρακιστεί από ένα φύλλο χαρτί. Είναι δύσκολα ανιχνεύσιμη και αποτελεί σημαντικό κίνδυνο εσωτερικής έκθεσης. Προκύπτει από τη διάσπαση α (Εικόνα 1) όπου το σωματίο α είναι ένας πυρήνας ηλίου. Η μάζα του είναι 4,002777 amu και το φορτίο του +2e. Όταν συμβαίνει εκπομπή σωματίων α από ένα βαρύ πυρήνα, που λέγεται μητρικός, ο μαζικός αριθμός μειώνεται κατά 4 και ο νέος πυρήνας, που λέγεται θυγατρικός, είναι σταθερότερος (μεταστοιχείωση) (Beyer, 1949).

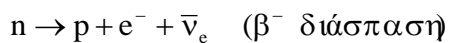


Εικόνα 1: Διάσπαση-α.

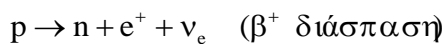
Η ακτινοβολία α είναι χαρακτηριστικό των βαρέων ($A > 140$) πυρήνων. Θεωρείται σαν αποτέλεσμα της αστάθειας που προκαλείται από περίσσειμα μάζας στον πυρήνα και της επικράτησης απωστικών δυνάμεων μεταξύ των πρωτονίων του. Η δέσμη των σωματιδίων α εκπέμπεται από τους πυρήνες με ταχύτητα (10.000-20.000) km/s. Προκαλούν τον ιονισμό των αερίων και εκτρέπονται από το μαγνητικό και ηλεκτρικό πεδίο. Η ακτινοβολία α απορροφάται εύκολα από τα συνήθη μέσα του περιβάλλοντος (προστατευτική στολή, γάντια κλπ). Χρειάζεται όμως ιδιαίτερη προσοχή, γιατί είναι δυνατόν να μεταφερθεί δια μέσου της σκόνης στο εσωτερικό του σώματος με την αναπνοή και να προκαλέσει σοβαρότατες βλάβες. (Urbach and American Society for Photobiology., 1992)

Ακτινοβολία β: Σαν ακτινοβολία β χαρακτηρίζεται η εκπομπή β-σωματιδίων είτε ηλεκτρονίων είτε ποζιτρονίων από ασταθείς ραδιενεργούς πυρήνες. Τα σωματίδια αυτά εμφανίζονται κατά τις β-πυρηνικές διασπάσεις ή δημιουργούνται από την αλληλεπίδραση, ηλεκτρομαγνητικής ακτινοβολίας, με το πεδίο των πυρήνων (δίδυμη γένεση). Κατά τη διάσπαση β ο μαζικός αριθμός παραμένει αμετάβλητος ενώ αλλάζει ο ατομικός αριθμός του στοιχείου (μεταστοιχείωση).

Με την εκπομπή ενός ηλεκτρονίου ο ατομικός αριθμός του θυγατρικού πυρήνα είναι μεγαλύτερος του μητρικού κατά 1. Σε αυτή την περίπτωση ένα νετρόνιο του μητρικού πυρήνα μετατρέπεται σε πρωτόνιο με ταυτόχρονη εκπομπή ενός ηλεκτρονίου και ενός αντινετρίνου. (Beyer, 1949)



Με την εκπομπή ενός ποζιτρονίου ο ατομικός αριθμός του θυγατρικού πυρήνα είναι μικρότερος του μητρικού κατά 1. Σε αυτή την περίπτωση ένα πρωτόνιο του μητρικού πυρήνα μετατρέπεται σε νετρόνιο με ταυτόχρονη εκπομπή ενός ποζιτρονίου και ενός νετρίνου.



Η μελέτη της εκτροπής των σωματιδίων β από το μαγνητικό πεδίο φανερώνει ότι για ένα δεδομένο ραδιενεργό στοιχείο, οι αρχικές τους ταχύτητες διαφέρουν. Σε ορισμένες περιπτώσεις πλησιάζουν την ταχύτητα του φωτός. Η εμβέλεια τους στον αέρα φτάνει τα μερικά μέτρα, ενώ μέσα σε στερεά σώματα είναι πολύ μικρότερη. Προκαλούν τον ιονισμό των αερίων αλλά σε μικρότερο βάθος από τις ακτίνες α. Το ενεργειακό φάσμα των ακτίνων β είναι συνεχές σε αντίθεση με το κβαντισμένο των ακτίνων α. Αυτό οφείλεται στην παρουσία του νετρίνου. Η έκθεση του ανθρώπινου ιστού σε ακτινοβολία β είναι πολύ επικίνδυνη. (Lehnert, 2008) Μπορεί να προκαλέσει εγκαύματα λόγω της εύκολης απορρόφησής τους και αν είναι ακτινοβολία μεγάλης έντασης, τότε εισχωρεί σε μικρό βάθος μέσα στους ιστούς προξενώντας μεγάλες βλάβες στα κύτταρα. Είναι πιο διεισδυτική από την ακτινοβολία α, αλλά μπορεί να αποκοπεί από φύλλα πλαστικού ή λεπτά μεταλλικά φύλλα. Αποτελεί κίνδυνο εξωτερικής έκθεσης ματιών και δέρματος και κίνδυνο εσωτερικής έκθεσης. Ο βαθμός ανίχνευσης εξαρτάται από την ενέργεια των β σωματιδίων.

Ηλεκτρομαγνητική ακτινοβολία

Ακτινοβολία X: Ακτίνες X ή ακτίνες Ρέντγκεν (Röntgen) αποκαλείται ένα τμήμα του ηλεκτρομαγνητικού φάσματος με περιοχή μήκους κύματος μεταξύ 10 nm με 10 pm, που αντιστοιχεί σε περιοχή ενέργειας 120 eV-120 keV. Αυτό το τμήμα του ηλεκτρομαγνητικού φάσματος βρίσκεται μεταξύ των τμημάτων της υπεριώδους ακτινοβολίας και των ακτίνων γ. Οι ακτίνες X πρωταρχικά χρησιμοποιήθηκαν από την Ιατρική ως διαγνωστικό εργαλείο με τη μορφή της ακτινογραφίας και από τη Φυσική και τη Χημεία με τη μορφή της κρυσταλλογραφίας (Agarwal, 1991). Όμως, οι ακτίνες X ανήκουν στις ιονίζουσες ακτινοβολίες, αφού η

ενέργειά τους είναι ικανή να προκαλέσει τον ιονισμό ατόμων και μορίων από εσωτερικά τους ηλεκτρόνια. Οι ακτίνες X διαχωρίζονται σε 2 υποπεριοχές μήκους κύματος, συχνότητας και ενέργειας:

«Μαλακές ακτίνες X»: 10nm-100pm,120eV-12keV.

«Σκληρές ακτίνες X»: 100-10pm,12-120keV.

Για θωράκιση χρειάζεται να χρησιμοποιηθεί χάλυβας, μόλυβδος ή σκυρόδεμα (ή ένας συνδυασμός). Είναι πολύ πιο διεισδυτική από την β ακτινοβολία και αποτελεί κίνδυνο εξωτερικής και εσωτερικής έκθεσης (Allewell et al., 2013).

Ακτινοβολία γ: Η ακτινοβολία γ, δηλαδή φωτόνια υψηλής ενέργειας, συχνά συνοδεύει τη ραδιενέργεια α και β. Οι ακτίνες γ είναι ηλεκτρομαγνητική ακτινοβολία, όπως οι ακτίνες-X, αλλά έχουν μεγαλύτερη διαπεραστική ικανότητα γιατί έχουν υψηλότερη ενέργεια (Celitans, 1973).

ΠΡΟΕΛΕΥΣΗ ΑΚΤΙΝΟΒΟΛΙΑΣ

Στο φυσικό περιβάλλον στο οποίο ζούμε η ακτινοβολία προέρχεται από το **έδαφος** και την ατμόσφαιρα. Ειδικότερα, τα πετρώματα, το νερό και ο αέρας περιλαμβάνουν φυσικά ραδιενεργά στοιχεία, όπως είναι το κάλιο, το ράδιο, το ουράνιο και το ραδόνιο. Το ραδόνιο είναι ευγενές αέριο που εκλύεται από το έδαφος και τα οικοδομικά υλικά και στο οποίο αποδίδεται το μεγαλύτερο μέρος της ετήσιας δόσης ακτινοβολίας που λαμβάνει ο οργανισμός μας. **Η ακτινοβολία του εδάφους** σε δεδομένη θέση εξαρτάται άμεσα από τη γεωλογική σύσταση των πετρωμάτων της περιοχής (Bond and Duval, 1998). Η επιφάνεια της γης δέχεται συνεχώς και **κοσμική ακτινοβολία**, η οποία προέρχεται από το εξωτερικό διάστημα, δηλαδή από τον ήλιο, καθώς και άλλες άγνωστες ακόμη αστρικές πηγές. Αύξηση της κοσμικής ακτινοβολίας έχουμε κατά τις εξάρσεις της ηλιακής δραστηριότητας. Η κοσμική ακτινοβολία κατά τη διέλευσή της μέσα από τα στρώματα της γήινης ατμόσφαιρας απορροφάται μερικώς και η έντασή της μειώνεται σταδιακά με αποτέλεσμα να είναι σχετικά εξασθενημένη στο επίπεδο της επιφάνειας της θάλασσας (Swenberg et al., 1993). **Η τροφική αλυσίδα** αποτελεί ακόμη μια φυσική πηγή πρόσληψης ραδιενεργών στοιχείων. Χαρακτηριστικό είναι το παράδειγμα του καλίου (^{40}K), ενός μετάλλου απαραίτητου σε κάθε οργανισμό. Τέλος ένα μέρος της ιοντίζουσας ακτινοβολίας που προσλαμβάνουμε οφείλεται σε **ακτινοδιαγνωστικές-ακτινοθεραπευτικές** διαδικασίες που αναλύονται παρακάτω.

Πέρα από τις φυσικές πηγές ιοντίζουσας ακτινοβολίας ο άνθρωπος ανακάλυψε τις τεχνητές πηγές παραγωγής ακτινοβολιών κατά τα τέλη του 19ου αιώνα. Έκτοτε η συστηματική έρευνα οδήγησε τόσο στην εκτεταμένη χρήση τους όσο και στη λήψη μέτρων για την προστασία από τις ενδεχόμενες βλαβερές επιπτώσεις τους. (Bacq, 1965; Ferguson et al., 2003) Οι ιοντίζουσες ακτινοβολίες παράγονται και χρησιμοποιούνται καθημερινά στην ιατρική διάγνωση και θεραπεία: διαγνωστική ακτινολογία, θεραπευτική ακτινολογία (Niu et al., 2010), στη βιομηχανία: ραδιογραφίες, ακτινοβολητές για αποστείρωση υλικών, συσκευές για έλεγχο ποιοτικών παραμέτρων κ.ά. (Gaughran et al., 1974), στην παραγωγή ενέργειας: πυρηνικά εργοστάσια στη γεωργία, (Agora, 2008), στην έρευνα και στην εκπαίδευση. Η χρήση της ιοντίζουσας ακτινοβολίας στην ιατρική είναι διαδεδομένη σε όλο τον κόσμο (Cincu et al., 2008). Στην Ελλάδα οι ιατρικές εφαρμογές των ακτινοβολιών καλύπτουν περίπου το 90% του συνολικών εφαρμογών τους. Οι διαγνωστικές διαδικασίες, και ιδιαίτερα η χρήση των ακτίνων X, είναι η πιο κοινή εφαρμογή της ακτινοβολίας στην ιατρική. Οι διαγνωστικές εκθέσεις χαρακτηρίζονται από αρκετά χαμηλές δόσεις στους ασθενείς που είναι επαρκείς για να παρέχουν τις απαραίτητες κλινικές πληροφορίες (WHO Expert Committee on the Use of Ionizing Radiation and Radionuclides on Human Beings for Medical Research Training, 1977). Αντίθετα, οι θεραπευτικές εκθέσεις είναι λιγότερο συχνές και περιλαμβάνουν πολύ υψηλότερες δόσεις για τη θεραπεία της ασθένειας, κυρίως καρκίνος, ή για να ανακουφίσουν τα συμπτώματα. Η ακτινοθεραπεία χρησιμοποιείται κυρίως για θεραπεία του καρκίνου, με σκοπό ο κακοήθης ιστός να λάβει μια μεγάλη δόση ακτινοβολίας μέσα σε μια καλά καθορισμένη περιοχή (όγκος-στόχος), αποφεύγοντας την ακτινοβόληση υγιών ιστών (Schottenfeld and Fraumeni, 2006). Θα αναλύσουμε τις εφαρμογές στην Ιατρική που περιλαμβάνουν το μεγαλύτερο ποσοστό και ειδικότερα στην ακτινοδιάγνωση που χρησιμοποιεί επί το πλείστον χαμηλές δόσεις.

Ακτινοδιαγνωστική

Η χρήση της ιοντίζουσας ακτινοβολίας γίνεται όλο και πιο διαδεδομένη σε διάφορους τομείς. Οι ιοντίζουσες ακτινοβολίες βρίσκουν σημαντικές διαγνωστικές και θεραπευτικές εφαρμογές στην ιατρική. Η έγκαιρη και έγκυρη διάγνωση παθήσεων μέσω απεικονιστικών τεχνικών και οι εξελιγμένες θεραπείες συμβάλλουν στην επιμήκυνση του προσδόκιμου ζωής και στη βελτίωση της ποιότητας ζωής του ανθρώπου. Στη διαγνωστική ακτινολογία χρησιμοποιούνται απλά, αλλά και πολύ σύνθετα απεικονιστικά συστήματα, παρέχοντας εικόνες δύο και τριών διαστάσεων και εικόνες σε πραγματικό χρόνο, αποτυπώνοντας την κινητική των οργάνων. Γενικά ένα διαγνωστικό ακτινολογικό σύστημα αποτελείται από την ακτινολογική λυχνία παραγωγής της ακτινοβολίας (ακτίνες-X) και το μηχανισμό απεικόνισης (π.χ. φιλμ, διατάξεις ανιχνευτών, ενισχυτής εικόνας). Ο εξεταζόμενος τοποθετείται κατάλληλα ανάμεσα στη λυχνία ακτίνων-X και στο σύστημα απεικόνισης. (Simon, 1967) Η δέσμη ακτίνων-X επικεντρώνεται στην περιοχή του σώματος που πρέπει να απεικονιστεί. Κατά τη διέλευση της ακτινοβολίας από το σώμα η ακτινοβολία αλληλεπιδρά με τα διαφορετικά είδη ιστών που το αποτελούν. Οι ιστοί ανάλογα με τη πυκνότητά τους εξασθενούν σε διαφορετικό βαθμό στη

διερχόμενη ακτινοβολία. Αυτό έχει ως αποτέλεσμα, η εξασθένηση της δέσμης η οποία προκαλείται από τους ιστούς μικρής πυκνότητας (μαλακοί ιστοί) να είναι μικρότερη από την αντίστοιχη των ιστών μεγάλης πυκνότητας (οστά). Όσο μεγαλύτερη είναι η ένταση της εξερχόμενης από το σώμα ακτινοβολίας τόσο πιο έντονη είναι και η αμαύρωση του ακτινογραφικού φιλμ. Για αυτό το λόγο, τα οστά, απεικονίζονται με λευκό χρώμα στο ακτινογραφικό φιλμ, ενώ οι μαλακοί ιστοί με διαφορετικές αποχρώσεις του γκρι.

Ενδεικτικά να αναφέρουμε πως ανάμεσα στις πηγές έκθεσης ιοντίζουσας ακτινοβολίας που προαναφέρθηκαν η φυσιολογική ακτινοβολία στο περιβάλλον (στην επιφάνεια της γης) διαφέρει από τόπο σε τόπο αλλά είναι της τάξης των 2.4 mGy/έτος περίπου (Swenberg et al., 1993). Ακολουθεί η τυπική μέση δόση ακτινοβολίας σε διάφορες ακτινοδιαγνωστικές και ακτινοθεραπευτικές μεθόδους

Ακτινογραφία: Η ακτινογραφία είναι η πιο συνηθισμένη και η πιο απλή διαγνωστική εξέταση. Οι συνηθέστερες ακτινογραφικές εξετάσεις περιλαμβάνουν αυτές του θώρακα, των άκρων, της αυχενικής, θωρακικής και οσφυϊκής μοίρας της σπονδυλικής στήλης, της λεκάνης - ισχίων, του κρανίου, των ιγμορείων, κ.ά. Η ακτινοβόληση διαρκεί μερικά κλάσματα του δευτερολέπτου. Η απορρόφηση από μία ακτινογραφία είναι της τάξης των 1.4 mGy (De la Matta Martin et al., 2011).

Μαστογραφία: Μια ιδιαίτερη περίπτωση ακτινογράφησης είναι η μαστογραφία που χρησιμοποιείται για την απεικόνιση της μορφολογίας, της δομής και των παθολογικών αλλοιώσεων του μαστού. Η μαστογραφία παίζει πρωταρχικό ρόλο στην έγκαιρη ανίχνευση καρκίνου του μαστού, καθώς μπορεί να αναδείξει αλλαγές στη μορφολογία του μαστού σε προ-κλινικό στάδιο, όταν το μέγεθος του όγκου είναι πολύ μικρό και δεν είναι δυνατή η ψηλάφησή του. Κατά τη διάρκεια της μαστογραφίας, ο τεχνολόγος χειριστής τοποθετεί τον κάθε μαστό χωριστά πάνω στη βάση στήριξης και ακολούθως τον συμπιέζει προσεκτικά με το ειδικό σύστημα συμπίεσης, προκειμένου το πάχος του μαστού να παραμένει σταθερό και όσο το δυνατό μικρότερο σε όλο το ακτινοβολούμενο πεδίο, ώστε κατά τη διέλευση των ακτίνων-X να μην επισκιάζονται από τον υπερκείμενο μαστικό ιστό πιθανές αλλοιώσεις. Κατά την ακτινοβόληση, που διαρκεί κλάσματα του δευτερολέπτου, ο μαστός εκτίθεται σε μικρή δόση ακτινοβολίας χαμηλής ενέργειας, προκειμένου να παραχθεί μια εικόνα της εσωτερικής του δομής. Η μέση ενεργός δόση από 2 λήψεις μαστογραφίας είναι 0.56 mSv από αναλογική με φιλμ ή 0.44 mSv από ψηφιακή μαστογραφία και ισοδυναμεί περίπου με 2 μήνες ακτινοβολία περιβάλλοντος. (Johnson, 2017)

Αξονική Τομογραφία: Η ακτινοσκόπηση είναι η τεχνική λήψης εικόνων της μορφολογίας και φυσιολογίας του ανθρώπινου σώματος σε πραγματικό χρόνο. Συνήθεις ακτινοσκοπικές εξετάσεις είναι αυτές του πεπτικού συστήματος (π.χ. με βαριούχο γεύμα, βαριούχος υποκλυσμός). Κατά την εξέταση, ο εξεταζόμενος προσλαμβάνει ένα σκιαγραφικό (π.χ. βαριούχο), δηλαδή υλικό το οποίο απεικονίζεται έντονα από το ακτινολογικό σύστημα. Η πορεία του υλικού μέσα στο σώμα παρακολουθείται με χρήση ακτίνων-X, οι οποίες μετατρέπονται σε εικόνα. Ο εξεταζόμενος δέχεται ακτινοβολία ακτίνων-X για αρκετό χρόνο (1-2 λεπτά), ενώ παράλληλα λαμβάνονται ακτινογραφίες. Για τους λόγους αυτούς, η δόση ακτινοβολίας από αυτή την εξέταση

είναι υψηλότερη σε σύγκριση με τις απλές ακτινογραφίες. Η αξονική τομογραφία αποτελεί την πιο σύγχρονη μορφή εξέτασης με ακτίνες-Χ. Οι συνηθέστερες εξετάσεις είναι αυτές του κρανίου, της αυχενικής, θωρακικής, οσφυϊκής μοίρας, της σπονδυλικής στήλης, του θώρακα, της κοιλιάς κ.α. Τα βασικά πλεονεκτήματα της αξονικής τομογραφίας, σε σχέση με άλλες διαγνωστικές τεχνικές, είναι η δυνατότητα λήψης πολλών δισδιάστατων ή και τρισδιάστατων εικόνων του ανθρωπίνου σώματος, μεγάλης ευκρίνειας, σε μικρό χρονικό διάστημα, οι οποίες παρέχουν σημαντικές μορφολογικές και λειτουργικές πληροφορίες. (Seeram, 1997) Ο εξεταζόμενος τοποθετείται στην τράπεζα (κρεβάτι) του αξονικού τομογράφου η οποία μετατοπίζεται με σταθερό βήμα μέσα στο «άνοιγμα» - δακτύλιο του συστήματος. Η λυχνία ακτίνων-Χ του συστήματος και το σύστημα απεικόνισης (ανιχνευτές) που βρίσκεται αντιδιαμετρικά της λυχνίας περιστρέφονται γύρω από τον εξεταζόμενο. Σε κάθε περιστροφή παράγεται μια λεπτή δέσμη ακτίνων-Χ (πλάτους 1 mm - 10 mm) σε μορφή «βεντάλιας», η οποία ακτινοβολεί εγκάρσια τον εξεταζόμενο στη συγκεκριμένη τομή. Ανάλογα με την εξέταση, λαμβάνεται ένας αριθμός τομών (π.χ. 20-30), ώστε να «σαρώνεται» όλη η περιοχή ενδιαφέροντος. Η ακτινοβολία σε κάθε τομή διαρκεί λιγότερο από δευτερόλεπτο, ενώ ο συνολικός χρόνος της εξέτασης (δηλαδή λήψη όλων των τομών) διαρκεί λίγα λεπτά ανάλογα με το σύστημα και την εξέταση. Η δημιουργία των εικόνων αξονικής τομογραφίας επιτυγχάνεται με την κατάλληλη επεξεργασία του σήματος των ανιχνευτών. Η δόση ακτινοβολίας από μία εξέταση αξονικής τομογραφίας είναι αρκετά μεγαλύτερη από τις υπόλοιπες συμβατικές ακτινολογικές εξετάσεις. Για το λόγο αυτό η εξέταση πρέπει να γίνεται αφού προηγουμένως έχει τεκμηριωθεί και αιτιολογηθεί η αναγκαιότητά της και έχουν εξαντληθεί οι επιλογές άλλων συμβατικών τεχνικών μικρότερης ή μηδενικής δόσης (π.χ. ακτινογράφιση και υπέρηχοι (Pitman, 1997)).

Αγγειογραφία: Μια άλλη χρήση της ακτινογραφίας είναι στην αγγειογραφία (Abrams, 1983). Αγγειογραφία είναι μια εξέταση κατά την οποία χορηγείται μέσω ένεσης υγρής μορφής χρωστική ουσία, η οποία εγχύεται στις αρτηρίες, με αποτέλεσμα να είναι εύκολα ορατές στις ακτινογραφίες. Η αγγειογραφία μπορεί να χρησιμοποιηθεί εάν ο ιατρός εξετάζει την περίπτωση χειρουργικής επέμβασης, καθώς δίνει μια σαφή εικόνα για την κατάσταση των αιμοφόρων αγγείων, μπορεί να αποκαλύψει ανευρύσματα (εξογκώματα που οφείλονται σε αποδυνάμωση των τοιχωμάτων ενός αιμοφόρου αγγείου), να δώσει μια καλή εικόνα της καρωτιδικής αρτηρίας και των υποσυστημάτων της στο λαιμό και το κεφάλι, μπορεί να δείξει τον βαθμό στένωσης των αρτηριών της καρδιάς κλπ (Bradač and Oberson, 1983). Η αγγειογραφία γίνεται ως εξής : πριν από την λήψη μιας ακτινογραφίας, υγρής μορφής χρωστική ουσία εκχύεται στα αιμοφόρα αγγεία μέσω ένεσης. Εάν η εξέταση αφορά τις αρτηρίες της καρδιάς, την καρωτιδική αρτηρία ή τις μεγάλες αρτηρίες που προέρχονται από την αορτή, ο καθετήρας εισάγεται στην βουβωνική χώρα ή στο χέρι. Πριν ο καθετήρας μπορέσει να εισχωρήσει σε μια αρτηρία, η γύρω περιοχή πρέπει να μουδιάσει με τοπικό αναισθητικό. Με τη χρήση βελόνας, ένα κοντό, λεπτό σύρμα με στρογγυλευμένο άκρο παρεμβάλλεται προσεκτικά γύρω από την αρτηρία. Η καθοδήγηση γίνεται με τη βοήθεια ακτινοσκόπησης (X-ray εικόνες) στο σημείο όπου είναι απαραίτητη η χρωστική ουσία. Κατόπιν, η βελόνα αφαιρείται και ένα αγγειακό περίβλημα παρεμβάλλεται

γύρω από το σύρμα. Ο καθετήρας μπορεί να προστεθεί στη συνέχεια, κατά μήκος του σύρματος-οδηγού. Όταν ο καθετήρας βρίσκεται στη σωστή θέση, το σύρμα αποτραβιέται και η χρωστική ουσία εκχύνεται μέσω του καθετήρα. Ο ασθενής μπορεί να εμφανίσει μια αίσθηση ζεστασιάς στην περιοχή, αλλά αυτό θα εξαφανιστεί μετά από μερικά δευτερόλεπτα. Τα αιμοφόρα αγγεία μπορούν να ελεγχθούν πλέον σε μια οθόνη, ή σε μια σειρά από ταχύτατα καταγραφόμενες ακτινογραφίες. Έχουμε έτσι εικόνες υψηλής ποιότητας, χρησιμοποιώντας την ελάχιστη δόση ακτινοβολίας.

Ακτινοθεραπεία

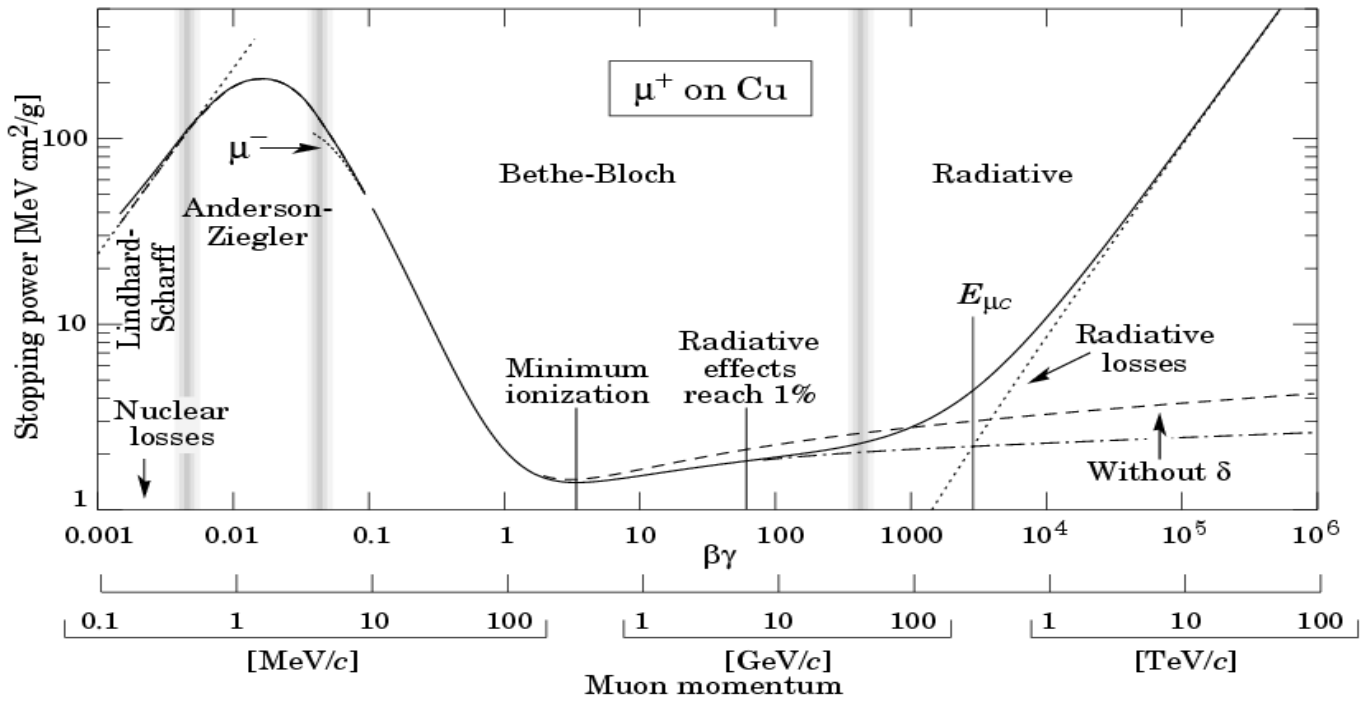
Στην ακτινοθεραπεία, η ακτινοβολία χρησιμοποιείται για να καταστρέψει τα νεοπλασματικά κύτταρα (καρκινικούς όγκους). Η θεραπεία σχεδιάζεται προσεκτικά για να εξασφαλιστεί ότι σωστή δόση ακτινοβολίας προσδίδεται στη σωστή περιοχή του όγκου, ενώ ταυτόχρονα η δόση στους φυσιολογικούς ιστούς ελαχιστοποιείται, ώστε η βλάβη στα υγιή όργανα να είναι ελάχιστη. Στην ακτινοθεραπεία, η δόση ακτινοβολίας στον όγκο είναι αρκετές χιλιάδες ή και εκατομμύρια φορές μεγαλύτερη από τη δόση σε σύγκριση με διαγνωστικές εξετάσεις με ακτίνες X ή πυρηνικής ιατρικής. Η ακτινοθεραπεία γίνεται με διάφορους τρόπους, τεχνικές και τύπους ακτινοβολίας (Hendee et al., 2005).

ΑΛΛΗΛΕΠΙΔΡΑΣΗ ΑΚΤΙΝΟΒΟΛΙΑΣ ΜΕ ΤΗΝ ΥΛΗ

Τα φορτισμένα σωματίδια προκαλούν *απευθείας ιονισμούς* σε αντίθεση με τα φωτόνια (ακτίνες X, ακτίνες γ κλπ) που προκαλούν κυρίως *έμμεσους ιονισμούς*. Όταν ένα φορτισμένο σωματίδιο κινείται μέσα στην ύλη αλληλεπιδρά H/M με τα αρνητικά e^- και τους θετικούς πυρήνες ανταλλάσσοντας φωτόνια. Το αποτέλεσμα αυτών των αλληλεπιδράσεων για το φορτισμένο σωματίδιο είναι: να χάσει ενέργεια, να αλλάξει κατεύθυνση η τροχιά του, και τελικά να σταματήσει και να απορροφηθεί διανύοντας συνολικά μια απόσταση που ονομάζεται διάστημα εμβέλειας (range). Οι μηχανισμοί μέσω των οποίων χάνει ενέργεια το σωματίδιο είναι: αλληλεπίδραση Coulomb με τα e^- και τους πυρήνες, ατομικές διεγέρσεις, ιονισμό ατόμων, H/M ακτινοβολία πέδησης (ακτινοβολείται όταν το σωματίδιο επιβραδύνεται σ' ένα πεδίο Coulomb), πυρηνικές αλληλεπιδράσεις, ακτινοβολία Cherenkov (όταν ξεπεράσει ένα κατώφλι και αν τα υλικά είναι διαφανές), H/M ακτινοβολία μετάπτωσης (transition radiation) (ακτινοβολείται όταν το σωματίδιο κινείται σε υλικό με ασυνεχή διηλεκτρική σταθερά). Κατά τη διέλευση μέσα από ύλη έχουμε απώλεια ενέργειας κυρίως λόγω ατομικών σκεδάσεων και λόγω απόκλισης του σωματιδίου από την αρχική του διεύθυνση. Το μέγεθος που έχει οριστεί για την μέση απώλεια ενέργειας λόγω αλληλεπιδράσεων Coulomb ανά μονάδα μήκους είναι γνωστό και ως ανασχετική ισχύς (Εικόνα 2) (Stopping Power) ή απλά dE/dx και έχει υπολογιστεί πρώτα από τον Bohr (Κλασικός Υπολογισμός) και αργότερα από τους Bethe-Bloch (Κβαντομηχανική). Η βασική έκφραση που χρησιμοποιείται είναι αυτή των Bethe-Bloch, η οποία δίνεται από τον παρακάτω τύπο:

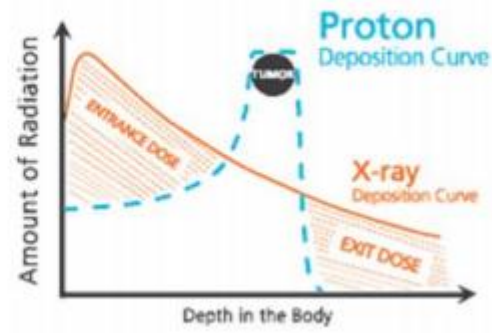
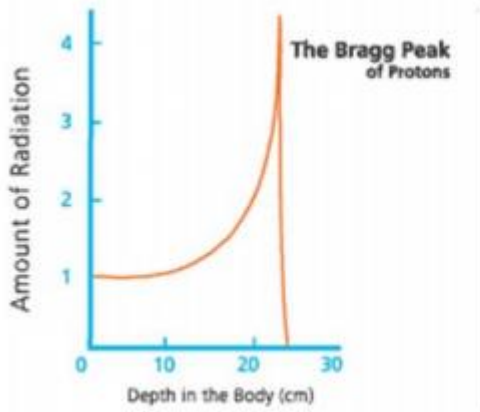
$$-\frac{dE}{dx} = \left(\frac{1}{4\pi\epsilon_0}\right)^2 \frac{4\pi z^2 e^4}{m u^2} \frac{Z}{A} \ln\left(\frac{m u^2}{I}\right)$$

Εξίσωση 1: Γενική μορφή της εξίσωσης Bethe-Bloch, όπου $(1/4\pi\epsilon_0)$ σταθερή ποσότητα από τον Νόμο Coulomb, m , z και u καθορίζονται από την ακτινοβολία (μάζα, ατομικός αριθμός, π.χ. $z=1$ για πρωτόνια, ταχύτητα), e το φορτίο του ηλεκτρονίου και I το δυναμικό ιοντισμού..



Εικόνα 2: Ανασχετική ισχύς για θετικά μόνια σε χαλκό.

Όσον αφορά τα φορτισμένα σωματίδια η διαφορά των πρωτονίων από τα ηλεκτρόνια ή φωτόνια είναι ότι εναποθέτουν την ψηλότερη δόση ενέργειας όταν σταματούν στο στόχο τους στο σώμα (Bragg Peak)(Εικόνα 3). Όταν εισέρχονται πχ στο σώμα, στο σημείο εισδοχής η ενέργεια είναι χαμηλή και στο σημείο εξόδου τους από το σώμα η ενέργειά τους είναι πλέον πολύ μικρή έως μηδενική. Η βασική αυτή διαφορά επιτρέπει την ασφαλέστερη χορήγηση υψηλότερων δόσεων ενέργειας, για την καταστροφή των κακοηθών όγκων και καλύτερη ανεκτικότητα από τους ασθενείς. Οι επιπλοκές λόγω ακτινοθεραπείας με πρωτόνια, όπως η ανορεξία, η διάρροια και οι πονοκέφαλοι, που παρουσιάζουν οι ασθενείς, είναι λιγότερες σε σύγκριση με τις άλλες κλασικές μορφές ακτινοθεραπείας. Οι υγιείς ιστοί που βρίσκονται στη διαδρομή της δέσμης της ιοντίζουσας ακτινοβολίας στην ακτινοθεραπεία με πρωτόνια και αυτοί που γειτνιάζουν με τον κακοήθη όγκο, λαμβάνουν πολύ λιγότερη ακτινοβολία σε σύγκριση με άλλες μεθόδους ακτινοθεραπείας (Hendee et al., 2005).



Εικόνα 3: Αριστερά: Ο σχηματισμός της αιχμής Bragg. Δεξιά: Θεραπεία πρωτονίων και η κορυφή Bragg συγκριτικά με την καμπύλη εναπόθεσης δόσης μέσω ακτινοβολίας Χ.

ΕΠΙΔΡΑΣΕΙΣ ΤΗΣ ΙΟΝΤΙΖΟΥΣΑΣ ΑΚΤΙΝΟΒΟΛΙΑΣ ΣΤΑ ΚΥΤΤΑΡΑ

Η ιοντίζουσα ακτινοβολία μπορεί να θεωρηθεί ένας ενεργειακός τοξικός παράγοντας του περιβάλλοντος. Ο ιοντισμός της οργανικής ύλης προκαλεί μεταβολές (μόνιμες ή παροδικές) στα άτομα και στα μόρια των χημικών ενώσεων. Εάν η βλάβη δεν επιδιορθωθεί, παρεμποδίζεται η αναπαραγωγή ή η επιβίωση του κυττάρου. (Hüttermann et al., 1978)

Γενικά, το βιολογικό αποτέλεσμα της επίδρασης της ιοντίζουσας ακτινοβολίας (ενός συγκεκριμένου είδους ακτινοβολίας και συγκεκριμένης ενέργειας) στην οργανική ύλη, εξαρτάται από τους κάτωθι παράγοντες:

1. Απορροφούμενη Δόση: Το ζητούμενο είναι ένα όριο δόσης, όπου επίπεδα δόσεων κάτω από αυτό δεν προκαλούν βιολογικές αντιδράσεις και δεν υφίσταται πρόβλημα ακτινοπροστασίας. Η μέση τιμή της ενέργειας που μεταφέρεται στο βιολογικό υλικό είναι η απορροφούμενη δόση.

2. Ρυθμός δόσης: Το ίδιο ποσό δόσης όταν αποδίδεται σε ένα μικρό χρονικό διάστημα (π.χ. 1 sec) προκαλεί σοβαρότερο βιολογικό αποτέλεσμα σε σχέση με την ακτινοβόληση για μεγαλύτερο χρονικό διάστημα (π.χ. 1 εβδομάδα).

3. Όγκο του υπό ακτινοβόληση ιστού και τις ιδιότητες αυτού.

Μονάδα μέτρησης της δόσης είναι:

$$1 \text{ rad} = 100 \text{ erg/gr} = 10^{-2} \text{ J/kg}$$

$$\text{SI: } 1 \text{ Gy} = 1 \text{ J/Kg } 1 \text{ Gy} = 100 \text{ rad}$$

Οι ως άνω παράγοντες επηρεάζονται άμεσα ή έμμεσα από μία σειρά από ακτινοβιολογικούς παράγοντες όπως η ευαισθησία των κυττάρων, η φάση του κυτταρικού κύκλου, η ευαισθησία των ιστών και οργάνων, κ.α.

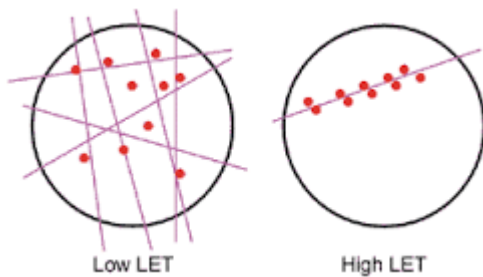
Η αποτελεσματικότητα μίας ιοντίζουσας ακτινοβολίας εξαρτάται σημαντικά από το είδος (π.χ. φωτόνιο ή σωματίδιο) και την ενέργειά της. Συνεπώς, αυτές οι διαφορές προκύπτουν βασικά από το γεγονός ότι ακτινοβολίες διαφορετικής ποιότητας και είδους, παράγουν διαφορετικά μονοπάτια (patterns) ιοντισμού (Lehnert, 2008).

Η ιοντίζουσα ακτινοβολία αλληλεπιδρά με διάφορους μηχανισμούς. Διακρίνουμε το φυσικό, το βιολογικό και το χημικό στάδιο:

Φυσικό στάδιο

Οι φυσικοί παράγοντες περιλαμβάνουν το είδος της ακτινοβολίας, την κατάτμηση της δόσης δεδομένης ακτινοβολίας, και τον ρυθμό δόσης.

Στο φυσικό στάδιο η μεταφορά ενέργειας πραγματοποιείται σε χρόνο μικρότερο των 10-18 s αλλά η κατανομή της στον χώρο εξαρτάται από το είδος και την ενέργεια της ιοντίζουσας ακτινοβολίας. Με το είδος ή ποιότητα της ακτινοβολίας αναφερόμαστε κυρίως στο LET (Linear Energy Transfer-LET = dE/dx) της ακτινοβολίας, μέγεθος ενδεικτικό της απορρόφησης ενέργειας, ή ισοδύναμα της πυκνότητας ιοντισμών, ανά μονάδα διαδρομής των σωματιδίων της ακτινοβολίας σε ένα υλικό. Ακτινοβολία με μεγάλο LET (σωματίδια α, πρωτόνια ή βαρέα ιόντα) έχει μεγαλύτερη πιθανότητα να πλήξει τον απειροστών διαστάσεων στόχο του DNA και να προκαλέσει DSB. Για χαμηλό LET τότε η ακτινοβολία (πχ φωτόνια ή ηλεκτρόνια) προκαλεί μικρό αριθμό ιοντισμών και διεγέρσεων ανά μονάδα μήκους (Εικόνα 4).



Εικόνα 4: Επιδράσεις ακτινοβολίας χαμηλού/υψηλού LET.

Είναι εμφανές από την Εικόνα 4 ότι και στα δύο παραδείγματα ακτινοβολήσης προκαλείται ο ίδιος αριθμός ιοντισμών άρα πρόκειται για την ίδια δόση και διαφαίνεται ο διαφορετικός τρόπος δράσης ανάλογα το LET.

Για τη σύγκριση του αποτελέσματος διαφορετικών ειδών ακτινοβολίας έχει οριστεί η σχετική βιολογική αποτελεσματικότητα (Relative Biological Effectiveness-RBE) ως ο λόγος της δόσης μιας ακτινοβολίας αναφοράς (συνήθως αυτής δέσμης ακτίνων X 250keV) προς της δόση της ακτινοβολίας υπό δοκιμή που θα επάγει το ίδιο βιολογικό αποτέλεσμα.

Η σχετική βιολογική αποτελεσματικότητα για ακτινοβολία τύπου R σε ιστό τύπου T ορίζεται ως ο λόγος:

$$RBE = D_X / D_R$$

Εξίσωση 2: Τρόπος υπολογισμού σχετικής βιολογικής αποτελεσματικότητας ακτινοβολίας.

Όπου το D_X είναι η δόση αναφοράς ακτινοβολίας (X, γ) που απορροφά ένας τύπος X και D_R είναι η απορροφούμενη δόση ακτινοβολίας τύπου R που προκαλεί την ίδια ποσότητα βιολογικής βλάβης. Και οι δύο δόσεις υπολογίζονται από την ποσότητα ενέργειας που απορροφάται από τα κύτταρα (Smith, 2010).

Λόγω της σημασίας της ποιότητας της ακτινοβολίας, εκτός από την απορροφούμενη δόση (D με μονάδες Gy) έχει οριστεί το μέγεθος της ισοδύναμης δόσης (H με μονάδες Sievert-Sv) που συναντούμε σε θέματα σχετικά με την ακτινοπροστασία η οποία ορίζεται ως η απορροφούμενη δόση στον ιστό ή το όργανο (D_T) σταθμισμένη για το είδος και την ποιότητα της ακτινοβολίας (W_R)

$$H_T = W_R D_T$$

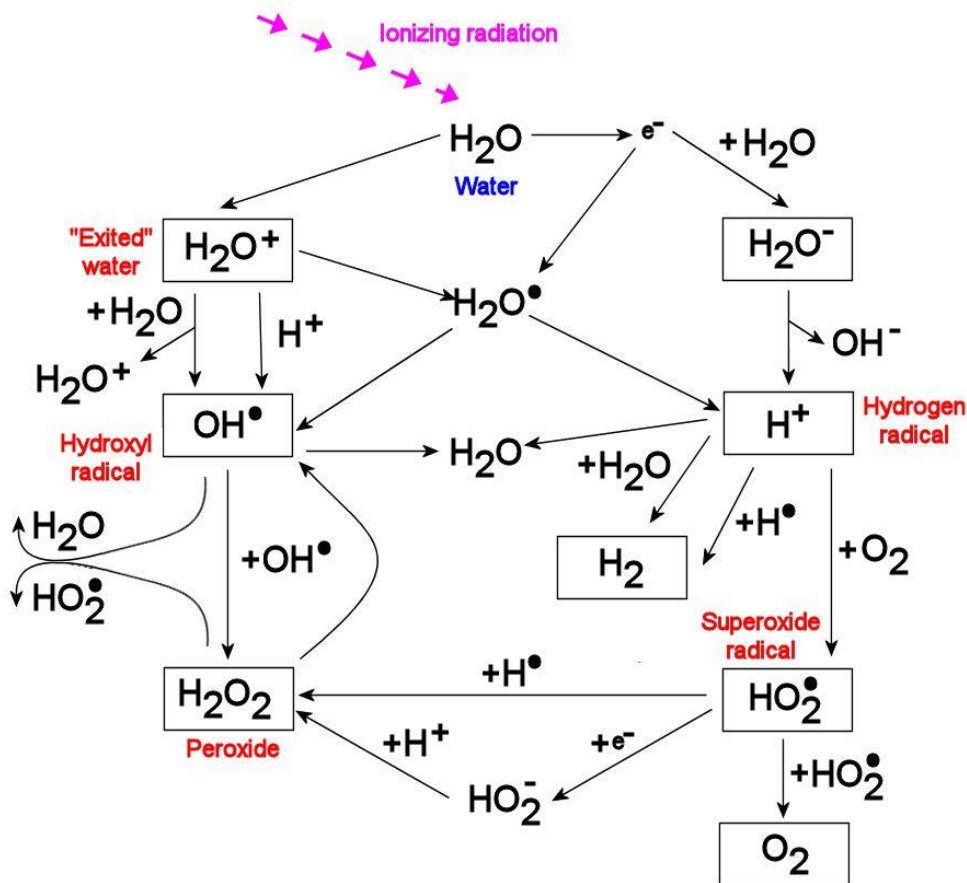
Μονάδες: 1 Sievert (Sv) = 1 Joule / kilogram

1 Sievert (Sv) = 100 rem

Σημαντικό ρόλο διαδραματίζει επίσης από απόψεως βιολογικού αποτελέσματος ο τρόπος χορήγησης της δόσης δεδομένης ακτινοβολίας. Το γεγονός αυτό σχετίζεται με τους επανορθωτικούς μηχανισμούς των κυττάρων. Η κλασματική χορήγηση μιας δόσης έχει μικρότερο βιολογικό αποτέλεσμα από την εφάπαξ χορήγηση της ίδιας δόσης καθώς στην πρώτη περίπτωση δίνεται χρόνος να αναλάβουν δράση οι επανορθωτικοί μηχανισμοί. Η αύξηση του ρυθμού δόσης σχετίζεται επίσης με τον περιορισμό της αποτελεσματικότητας των κυτταρικών επιδιορθωτικών μηχανισμών και κυρίως με αύξηση της πιθανότητας δημιουργίας DSB λόγω ιονισμών κατά μήκος της τροχιάς δύο διαφορετικών σωματιδίων ακτινοβολίας, και οδηγεί σε αύξηση του βιολογικού αποτελέσματος (και άρα μείωση της κυτταρικής επιβίωσης) (Brooks et al., 2016).

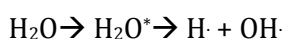
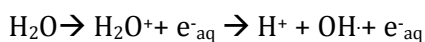
Χημικό στάδιο

Πλήθος χημικών παραγόντων εμφανίζει ενισχυτικό ή κατασταλτικό ρόλο στη βιολογική επίδραση της ιοντίζουσας ακτινοβολίας. Οι κυριότεροι τροποποιητικοί χημικοί παράγοντες περιλαμβάνουν το μοριακό οξυγόνο και διάφορους δεσμευτές των δραστικών ελευθέρων ριζών. Ο όρος ελεύθερη ρίζα αναφέρεται σε ένα ελεύθερο άτομο, μόριο ή συγκρότημα ατόμων, το οποίο φέρει ένα ασύζευκτο ηλεκτρόνιο και συνεπώς χαρακτηρίζεται από αυξημένη δραστηριότητα ανεξαρτήτως εάν είναι ηλεκτρικά ουδέτερο ή φορτισμένο. Στο στάδιο αυτό αναφερόμαστε στην πρόκληση βιολογικής βλάβης λόγω της χημικής αντίδρασης σημαντικών οργανικών μορίων του κυττάρου με δραστικές ελεύθερες ρίζες που σχηματίζονται κατά τη ραδιόλυση του νερού (Εικόνα 5), την αλληλεπίδραση δηλαδή της ακτινοβολίας με άτομα των μορίων του νερού που αφθονεί στα βιολογικά συστήματα (Muroya et al., 2006).



Εικόνα 5: Σχηματική απεικόνιση πιθανών αλληλεπιδράσεων λόγω ραδιόλυσης ύδατος.

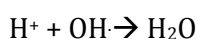
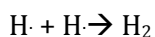
Από την αλληλεπίδραση ιοντίζουσας ακτινοβολίας με τα μόρια του νερού παράγονται ιονισμένα (H_2O^+) (Εικόνα 5 πάνω αριστερά) και διεγερμένα (H_2O^*) μόρια νερού καθώς και ελεύθερα ηλεκτρόνια (e^-). Τα ελεύθερα ηλεκτρόνια έλκουν μόρια νερού λόγω της πολικότητας των τελευταίων (e^-_{aq} , καλούμενα και «ενυδατωμένα» ηλεκτρόνια). Τα ιονισμένα μόρια του νερού διασπώνται άμεσα παράγοντας ένα κατιόν υδρογόνου (H^+ , ένα πρωτόνιο δηλαδή) και μια ελεύθερη ρίζα υδροξυλίου ($\text{OH}\cdot$) ενώ τα διεγερμένα μόρια (H_2O^*) μπορεί αντί να αποδιεγερθούν να διασπαστούν σε μια ελεύθερη ρίζα υδρογόνου ($\text{H}\cdot$) και μια ελεύθερη ρίζα υδροξυλίου ($\text{OH}\cdot$) (Muroya et al., 2006). Η αλληλεπίδραση ακτινοβολίας με τα άτομα των μορίων του νερού έχει τα εξής, αρχικά, αποτελέσματα:



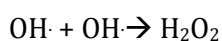
Οι ελεύθερες ρίζες υδρογόνου και υδροξυλίου, τα κατιόντα υδρογόνου και τα ενυδατωμένα ηλεκτρόνια που παράγονται, μπορούν ακολούθως μέσω της διάχυσής τους να συμμετάσχουν σε διάφορες αντιδράσεις μεταξύ τους ή με άλλα μόρια του συστήματος. Για την ανάλυση των αντιδράσεων αυτών θα πρέπει να λάβουμε

υπόψη τη σταθερά διάχυσης κάθε αντιδρώντος, την ακτίνα αντίδρασης κάθε ζεύγους αντιδρώντων καθώς και τη σταθερά του ρυθμού κάθε αντίδρασης. Αναφέρουμε ενδεικτικά κάποιες από αυτές όπως :

- η «αδρανοποίηση» ελευθέρων ριζών κατά την αντίδραση ριζών υδρογόνου μεταξύ τους, ή την αντίδραση ρίζας υδρογόνου και ρίζας υδροξυλίου, ή την αντίδραση κατιόντος υδρογόνου και ρίζας υδροξυλίου:



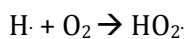
- η αντίδραση δύο ριζών υδροξυλίου κατά την οποία παράγεται υπεροξείδιο του υδρογόνου:



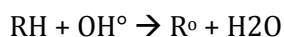
, το οποίο αποτελεί ισχυρό οξειδωτικό παράγοντα και μπορεί να αντιδράσει με ρίζα υδροξυλίου οδηγώντας στο σχηματισμό μιας νέας δραστικής ελεύθερης ρίζας υπεροξυλίου:



- και ο σχηματισμός ρίζας υπεροξυλίου από την αντίδραση ρίζας υδρογόνου με μοριακό οξυγόνο:



Οι ελεύθερες ρίζες OH° , είναι ασταθείς, εξαιρετικά δραστικές και η εμβέλεια τους είναι μικρότερη από 100 Angstrom. Εάν ένα οργανικό μόριο, RH (π.χ. DNA), βρεθεί μέσα στην τροχιά των ελευθέρων ριζών, θα συμβούν αντιδράσεις της μορφής:



θα παραχθούν δηλαδή οργανικές ρίζες R° , που ως ασταθείς και με υψηλό ενεργειακό περιεχόμενο, θα μετασχηματιστούν γρήγορα προκαλώντας μεταβολή-βλάβη στη δομή του οργανικού μορίου (Belon et al., 2016).

Η ρίζα υδρογόνου ($H\cdot$) έχει αναγωγικό χαρακτήρα και η ρίζα υδροξυλίου ($OH\cdot$) έχει οξειδωτικό χαρακτήρα. Παρ' όλα αυτά, ενώ η αντίδραση ριζών υδρογόνου οδηγεί σε σχηματισμό μοριακού υδρογόνου, το υπεροξείδιο του υδρογόνου που παράγεται από την αντίδραση ριζών υδροξυλίου έχει οξειδωτικό χαρακτήρα, όπως εξάλλου και η ρίζα υπεροξυλίου (Muroya et al., 2006). Έτσι υπερισχύει ο οξειδωτικός χαρακτήρας και το αποτέλεσμα της ραδιόλυσης του νερού από την έκθεση του κυτταροπλάσματος σε ιοντίζουσα ακτινοβολία είναι η οξείδωση διαφόρων μικρομορίων και μακρομορίων. Οι οξειδώσεις ανόργανων ουσιών που βρίσκονται διαλυμένες υπό μορφή ιόντων στο κυτταρόπλασμα είναι άνευ βιολογικής σημασίας. Η οξείδωση όμως

οργανικών ενώσεων είναι σημαντική καθώς, κατ' αντιστοιχία με την άμεση δράση της ακτινοβολίας, οδηγεί σε πολυμερισμούς ή από-πολυμερισμούς που συνοδεύονται από μεταβολή των φυσικών και χημικών ιδιοτήτων. Οι ελεύθερες ρίζες συμπεραίνουμε ότι έχουν την δυνατότητα να είναι εξαιρετικά χημικώς δραστικές.

Η αυξημένη συγκέντρωση του μοριακού οξυγόνου στα κύτταρα οδηγεί σε ενίσχυση του βιολογικού αποτελέσματος που εκφράζεται με το λόγο ενίσχυσης λόγω του οξυγόνου (Oxygen Enhancement Ratio-OER) που ορίζεται ως ο λόγος της δόσης ακτινοβολίας που απαιτείται σε ανοξικές συνθήκες αναφοράς για την πρόκληση δεδομένου βιολογικού αποτελέσματος προς την δόση της ίδιας ακτινοβολίας που θα επάγει το ίδιο βιολογικό αποτέλεσμα σε συνθήκες οξυγόνωσης (Cercek et al., 1968; Kal and Barendsen, 1976).

Βιολογικό στάδιο

Αμέσως μετά το χημικό στάδιο αρχίζει η εμφάνιση ενζυμικών μηχανισμών που δρουν επιδιορθωτικά στους μοριακούς σχηματισμούς που υπέστησαν βλάβη. Η διαδικασία της επιδιόρθωσης διαρκεί κατά το κύριο μέρος της από 15min-3h και ολοκληρώνεται σε (24-48) περίπου ώρες. Στην περίπτωση όπου η βλάβη είναι αδύνατον να επιδιορθωθεί, τα κύτταρα οδηγούνται είτε στη δημιουργία μεταλλάξεων ή/και σε κυτταρικό θάνατο μέσω απόπτωσης (apoptosis). Οι διαδικασίες αυτές εξαρτώνται από την φάση του κυτταρικού κύκλου αλλά και το είδος των κυττάρων.

Αν και το κύτταρο είναι ευπρόσβλητο από την ιοντίζουσα ακτινοβολία σε όλες τις φάσεις του κύκλου του, η ευαισθησία δεδομένου τύπου κυττάρων ποικίλλει από φάση σε φάση. Ο κυτταρικός κύκλος χωρίζεται σε δύο κύριες φάσεις: α) την μεσόφαση και β) την μίτωση ή μιτωτική διαίρεση. Η μεσόφαση χωρίζεται στις εξής:

Φάση G1 : Κύτταρο με ένα μόριο DNA / χρωματίδα στον πυρήνα.

Φάση S : Αντιγραφή του DNA : Διπλασιασμός χρωματίδας και ένωση των αδελφών χρωματίδων στο κεντρομερίδιο.

Φάση G2 : Διπλοειδές κύτταρο με διπλασιασμένο DNA . Προκύπτει ένα χρωμόσωμα με 2 αδελφές χρωματίδες, ενωμένες στο κεντρομερίδιο.

Και η μίτωση στις εξής :

ΠΡΟΦΑΣΗ : Ένωση κεντρομεριδίων χρωμοσωμάτων με τα ινίδια της ατράκτου.

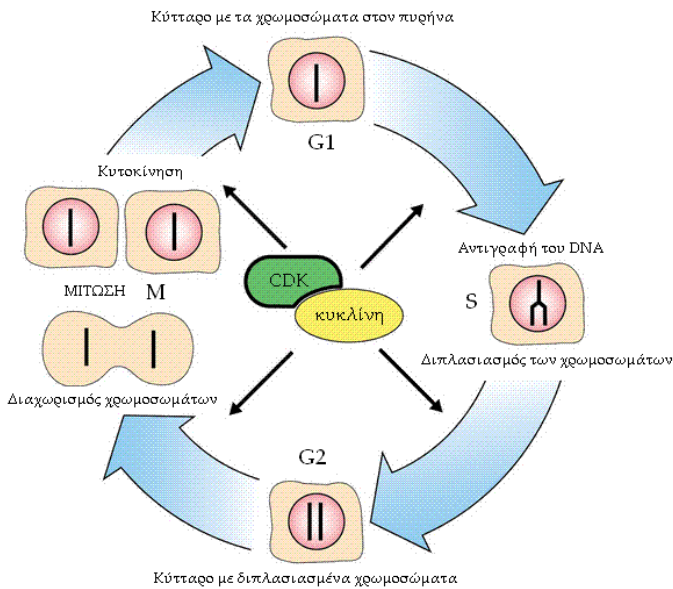
ΜΕΤΑΦΑΣΗ : Μέγιστη συσπείρωση χρωμοσωμάτων.

ΑΝΑΦΑΣΗ : Διαχωρισμός αδελφών χρωματίδων (διάσπαση κεντρομεριδίου).

ΤΕΛΟΦΑΣΗ : Σχηματισμός θυγατρικού πυρήνα και αποδιάταξη ινιδίων της ατράκτου.

ΚΥΤΟΚΙΝΗΣΗ : Διαχωρισμός σε δυο πανομοιότυπα απλοειδή θυγατρικά κύτταρα με το ίδιο γενετικό υλικό.

Γενικά, η φάση της μεγαλύτερης ευαισθησίας είναι η G2 και η μίτωση ενώ τα κύτταρα είναι, σχετικά, λιγότερο ευαίσθητα στη φάση S (Εικόνα 6). Αντίστοιχα, οι επιδιορθωτικοί μηχανισμοί της κυτταρικής βλάβης λειτουργούν σε όλες τις φάσεις του κυτταρικού κύκλου αλλά είναι αποδοτικότεροι στη φάση S.



Εικόνα 6.: Οι τέσσερις φάσεις ενός τυπικού κυτταρικού κύκλου ευκαρυωτικού κυττάρου. Οι φάσεις G1, S, G2 αποτελούν τη μεσόφαση, εφόσον το κύτταρο αυξάνεται συνεχώς. Η κυτταρική διαίρεση συμβαίνει στη φάση M. Η σύνθεση του DNA περιορίζεται στη φάση S. Η φάση G1 είναι η περίοδος ανάμεσα στη φάση M και S, ενώ η G2 ανάμεσα στη φάση S και M. Η διάρκεια του κυτταρικού κύκλου 24 ώρες είναι ενδεικτική. Υπάρχουν κύτταρα με κυτταρικούς κύκλους μεγαλύτερης και μικρότερης διάρκειας.

Το είδος των κυττάρων καθορίζει επίσης σε μεγάλο βαθμό και τη σχετική τους ευαισθησία στην ακτινοβολία. Η ακτινοευαισθησία του κυττάρου είναι ανάλογη της μιτωτικής του δραστηριότητας και αντιστρόφως ανάλογη του βαθμού διαφοροποίησής του.

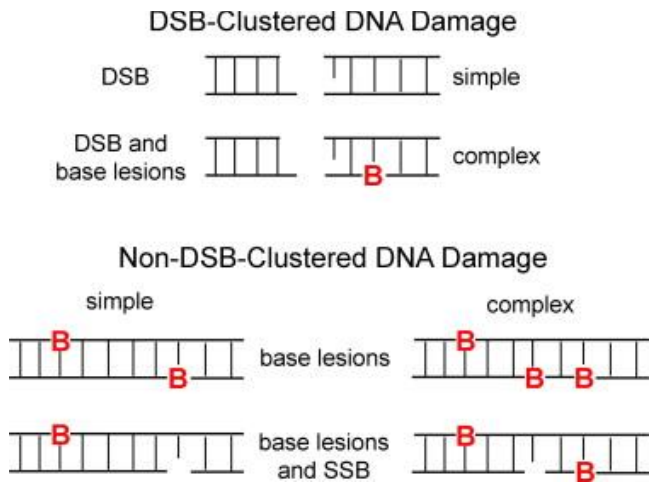
Σύμφωνα με τους μηχανισμούς δράσης της ακτινοβολίας που σκιαγραφήθηκαν στα προηγούμενα, η έκθεση σε ιοντίζουσα ακτινοβολία μπορεί να προκαλέσει στο DNA διάφορα είδη βλάβης όπως: αποσύνθεση, θραύση (συνήθως στο φωσφο-διεστερικό δεσμό) του ενός ή και των δύο κλώνων (Single Strand Break-SSB και Double Strand Break-DSB)(Nikitaki et al., 2016). Τέτοιου είδους θραύσεις και ακολουθούμενες με λανθασμένη αντικατάσταση βάσεων μπορούν να οδηγήσουν σε αλλαγές που μεταφέρονται στους απογόνους (αν γίνονται σε γενετικά κύτταρα) ή προκαλούν καρκινογένεση στο ίδιο το άτομο (αν γίνονται σε σωματικά κύτταρα). Επίσης συναντούμε και άλλα είδη βλάβης όπως βλάβη βάσεων, αλλοίωση μιας βάσης (οξειδώσεις, μεθυλιώσεις, αποπουρινώσεις, απαμινώσεις), απώλεια μιας βάσης (αβασικά σημεία-AP sites), σπάσιμο δεσμού υδρογόνου μεταξύ των δύο αλυσίδων, εγκάρσιες συνδέσεις μεταξύ των ελικών (διαδεσμοί-crosslinks), ενδο-κλωνική, δια-κλωνική ή και δια-μοριακή σύνδεση (με άλλα μόρια DNA ή πρωτεϊνών) (Taleei et al., 2015).

Το κύτταρο διαθέτει αποτελεσματικούς μηχανισμούς επιδιόρθωσης της βλάβης (σύστημα επιδιόρθωσης των αταίριαστων βάσεων - mismatch repair, μηχανισμός εκτομής βάσεων/νουκλεοτιδίων-base/nucleotide excision repair, ομόλογος ανασυνδυασμός-homologous recombination, μη-ομόλογη σύνδεση άκρων non homologous end joining). Στις χαμηλές δόσεις ακτινοβολίας κύριος μηχανισμός για ιοντίζουσες ακτινοβολίες (<1 Gy) είναι το BER (Rahmanian et al., 2014). Έτσι πρόδρομο των διαφόρων βιολογικών αποτελεσμάτων αποτελεί το, εξαιρετικά μικρό, ποσοστό μη επιδιορθωμένων ή ανεπιτυχώς επιδιορθωμένων βλαβών του DNA, και γνωρίζουμε ότι το είδος της βλάβης που συσχετίζεται καλύτερα με τα βιολογικά αποτελέσματα της ιοντίζουσας ακτινοβολίας, και ειδικότερα τον κυτταρικό θάνατο, είναι οι DSB .

Μπορεί όμως όλες οι DSB να μην είναι εξίσου σημαντικές. Ακόμα και για ακτινοβολίες χαμηλού LET, η αύξηση της απώλειας ενέργειας των δευτερογενών φορτισμένων ηλεκτρονίων στο τέλος της τροχιάς τους μπορεί να οδηγήσει σε σύνθετες συστοιχίες βλάβης του DNA που να περιλαμβάνουν μια ή περισσότερες DSB καθώς και αρκετές SSB, βλάβες βάσεων κτλ. (Georgakilas, 2008; Georgakilas et al., 2013). Τέτοιες εστίες σύνθετης συστοιχίας βλαβών αντιπροσωπεύουν το 60% με 90% της συνολικής βλάβης του DNA κατά την ακτινοβολία με χαμηλού και υψηλού LET ακτινοβολία, αντίστοιχα. Το γεγονός αυτό υποδεικνύει επίσης μια πιθανώς σημαντική διαφορά μεταξύ της βλάβης στο DNA από ιοντίζουσα ακτινοβολία και άλλους παράγοντες (αυθόρμητες μεταλλάξεις, χημικοί και βιολογικοί παράγοντες).

Οι βλάβες βάσεων και οι SSB επιδιορθώνονται στο μεγαλύτερο ποσοστό τους (Nikitaki et al., 2016). Σε περίπτωση ανεπιτυχούς επιδιόρθωσης όμως μπορούν να οδηγήσουν σε αλλαγή στην αλληλουχία των βάσεων του γενετικού υλικού και συνεπώς μετάλλαξη (λόγω αντικατάστασης, ελλείμματος ή ένθεσης βάσεων). Οι DSB (ή οι σύνθετες συστοιχίες βλάβης του DNA που να περιλαμβάνουν μια ή περισσότερες DSB) επιδιορθώνονται δυσκολότερα (Hada and Georgakilas, 2008). Μη επιδιορθωμένες DSB, εκτός από

χρωμοσωματικές ανωμαλίες (Aten et al., 2004), μπορούν να οδηγήσουν επίσης σε μετάλλαξη λόγω ελλείμματος βάσεων εάν τμήμα του DNA δεν επανασυνδεθεί και παραμένει ελεύθερο (Wang et al., 2005).



Εικόνα 7: Βλάβες ενός ή δύο κλώνων, εκτομή νουκλεοτιδίου και ομαδοποιημένες βλάβες.

Το αποτέλεσμα των μεταλλάξεων ως γνωστόν ποικίλει και μπορεί να οδηγήσει στην έκφραση μιας τροποποιημένης πρωτεΐνης και την αύξηση ή μείωση των επιπέδων μιας φυσιολογικής πρωτεΐνης. Τα στοχαστικά αποτελέσματα της ακτινοβολίας που αναλύονται παρακάτω (συμπεριλαμβανομένου του καρκίνου) οφείλονται σε μεταλλάξεις. Η συχνότητα εμφάνισης των μεταλλάξεων αυξάνει με τη δόση ακτινοβολίας αν και σε μεγάλες δόσεις κυριαρχεί ο κυτταρικός θάνατος οπότε ο αριθμός των μεταλλαγμένων κυττάρων μειώνεται. Σημειώνεται ότι ο ανθρώπινος μεταβολισμός οδηγεί σε 240.000 μεταλλάξεις σε κάθε ανθρώπινο κύτταρο καθημερινά, και η απορρόφηση δόσης ιοντίζουσας ακτινοβολίας προσθέτει σε αυτές μόλις 2000 μεταλλάξεις ανά Gy (Vladimirov, 1989).

Πέραν των μεταλλάξεων, μη επιτυχώς επιδιορθωμένες DSB λόγω της έκθεσης σε ιοντίζουσα ακτινοβολία (η πιθανότερα σύνθετες συστοιχίες βλάβης του DNA που να περιλαμβάνουν μια ή περισσότερες DSB) μπορεί να οδηγήσουν σε χρωμοσωματικές ανωμαλίες που θα καταστούν εμφανείς στην επόμενη διαίρεση του κυττάρου, στη φάση της μετάφασης. Οι χρωμοσωματικές ανωμαλίες περιλαμβάνουν μια ποικιλία ανταλλαγών των θραυσμάτων του DNA μεταξύ διαφορετικών χρωμοσωμάτων, μεταξύ των χρωματίδων του ίδιου χρωμοσώματος ή ακόμα και μεταξύ περιοχών της ίδιας χρωματίδας. Διακρίνονται σε χρωμοσωματικές ανωμαλίες (π.χ. δακτύλιοι και δικεντρικά χρωμοσώματα, και χρωματιδικές ανωμαλίες (π.χ. χρωματιδικές θραύσεις) ανάλογα με το εάν η βλάβη συνέβη κατά τη G1 φάση ή τη G2 φάση, αντίστοιχα (βλάβη κατά την S φάση ή τη μίτωση μπορεί να οδηγήσει και στα δύο είδη). Η εμφάνιση άτυπων χρωμοσωμάτων σε κύτταρα που πολλαπλασιάζονται μπορεί να οδηγήσει στον κυτταρικό θάνατο.

Η έκταση της βλάβης των χρωμοσωμάτων που θα εκδηλωθεί κατά την πρώτη μίτωση μετά την ακτινοβόληση είναι ο παράγοντας που συσχετίζεται καλύτερα από κάθε άλλο με τον κυτταρικό θάνατο (Georgakilas et al., 2014; Singleton et al., 2002).

Μετά από την εφάπαξ απορρόφηση σημαντικής δόσης ιοντίζουσας ακτινοβολίας που θα προκαλέσει εκτεταμένη βλάβη στο DNA (καθώς και σε πρωτεΐνες, ένζυμα και την κυτταρική μεμβράνη) προκαλείται άμεση νέκρωση του κυττάρου (μη ελεγχόμενος μηχανισμός κυτταρικού θανάτου που δεν συνοδεύεται από βιοχημικούς δείκτες και χαρακτηρίζεται από αύξηση του όγκου του κυττάρου και τελικώς ρήξη της κυτταρικής μεμβράνης) (Lehnert, 2008).

Χαμηλότερες τιμές δόσης που θα οδηγήσουν σε σχετικά μικρότερη, αλλά μη επιδιορθώσιμη βλάβη του DNA, θα προκαλέσουν άμεση απόπτωση (Georgakilas, 2008; Ray et al., 2012). Η απόπτωση είναι σημαντικός ελεγχόμενος μηχανισμός κυτταρικού θανάτου που θεωρείται μέσο των πολυκύτταρων οργανισμών για την απόρριψη ανεπιθύμητων ή κατεστραμμένων κυττάρων και ορίζεται βάσει αυστηρών μορφολογικών κριτηρίων που περιλαμβάνουν συμπύκνωση της χρωματίνης και συρρίκνωση του πυρήνα, πτύχωση της κυτταρικής μεμβράνης, συρρίκνωση του κυττάρου κ.α. (Narbutt et al., 2009).

Εάν η δόση της ακτινοβολίας είναι μικρή και οδηγήσει σε περιορισμένη αρχική βλάβη του DNA, δραστηριοποιούνται μηχανισμοί προαγωγής της κυτταρικής επιβίωσης που περιλαμβάνουν ανακοπή του κυτταρικού κύκλου για την αποτελεσματικότερη δράση των επιδιορθωτικών μηχανισμών. Σε περίπτωση επιτυχούς επιδιόρθωσης της βλάβης τα κύτταρα συνεχίζουν να πολλαπλασιάζονται. Σε περίπτωση ανεπιτυχούς επιδιόρθωσης τα κύτταρα μπορεί να συνεχίσουν να πολλαπλασιάζονται ωστόσο επέλθει μιτωτική καταστροφή (όρος που περιλαμβάνει κυτταρικούς θανάτους που εκδηλώνονται κατά τη φάση της μίτωσης είτε οφείλονται σε χρωμοσωματικές ανωμαλίες είτε όχι) που επιτελείται με καθυστερημένη νέκρωση ή καθυστερημένη απόπτωση (Georgakilas, 2008).

Σύμφωνα με τα παραπάνω συνοψίζοντας, η αλληλεπίδραση της ακτινοβολίας με τα κρίσιμα οργανικά μόρια ενός κυττάρου, που αποτελεί στοχαστικό φαινόμενο που υπόκειται στους νόμους της στατιστικής όπως γενικά η αλληλεπίδραση της ακτινοβολίας με την ύλη, πιθανώς να οδηγήσει σε βλάβη (Belon et al., 2016). Η βλάβη αυτή πιθανώς να μην επιδιορθωθεί επιτυχώς οπότε ίσως οδηγήσει σε κάποιο από τα ακόλουθα αποτελέσματα:

- τον άμεσο κυτταρικό θάνατο
- τον έμμεσο κυτταρικό θάνατο λόγω δυσλειτουργίας του κυττάρου ή της αδυναμίας του να πολλαπλασιαστεί ή
- τη μεταβολή, ή μετάλλαξη, του DNA του κυττάρου, το οποίο όμως παραμένει βιώσιμο και ικανό να πολλαπλασιαστεί.

Σε χαμηλές δόσεις ακτινοβολίας, η άμεση ή έμμεση θανάτωση ενός κυττάρου ή και ακόμη ενός μικρού αριθμού κυττάρων σε ένα ιστό ή όργανο του σώματος, δεν αναμένεται να έχει δυσμενή επίπτωση στη λειτουργία του οργάνου. Αυτό οφείλεται στο γεγονός ότι, η προκαλούμενη έκπτωση της λειτουργίας

αντισταθμίζεται πρακτικά από τον οργανισμό, ενώ τα νεκρά κύτταρα σταδιακά αντικαθίστανται. Εξάλλου χιλιάδες κύτταρα του ιστού ή του οργάνου πεθαίνουν και ανανεώνονται καθημερινά.

Οι βλάβες όμως που οδηγούν σε τροποποίηση του γενετικού υλικού ενός οποιουδήποτε μεμονωμένου σωματικού κυττάρου και δεν επιδιορθώνονται πλήρως, είναι πιθανό να έχουν μακροπρόθεσμα σοβαρή επίπτωση στην υγεία του εκτιθέμενου ατόμου. Πράγματι, στην περίπτωση που η επιδιόρθωση της βλάβης από τους διαθέσιμους κυτταρικούς μηχανισμούς δεν είναι τέλεια, τότε κατά τον πολλαπλασιασμό του κυττάρου, η μερικώς επιδιορθωμένη βλάβη στο DNA αποτελεί μετάλλαξη που μπορεί να μεταβιβαστεί σε θυγατρικά του κύτταρα και να αποτελέσει την απαρχή για την εμφάνιση μιας πιθανής κακοήθειας στο απώτερο μέλλον. Στην περίπτωση που η βλάβη συντελεστεί ειδικά σε αναπαραγωγικό κύτταρο του οργανισμού, τότε αυτή πιθανόν να οδηγήσει σε τροποποίηση του γενετικού κώδικα, και να μεταβιβαστεί στους απογόνους του εκτιθέμενου ατόμου. Οι βλάβες που οδηγούν σε τροποποίηση του γενετικού κώδικα καλούνται κληρονομήσιμες βλάβες.

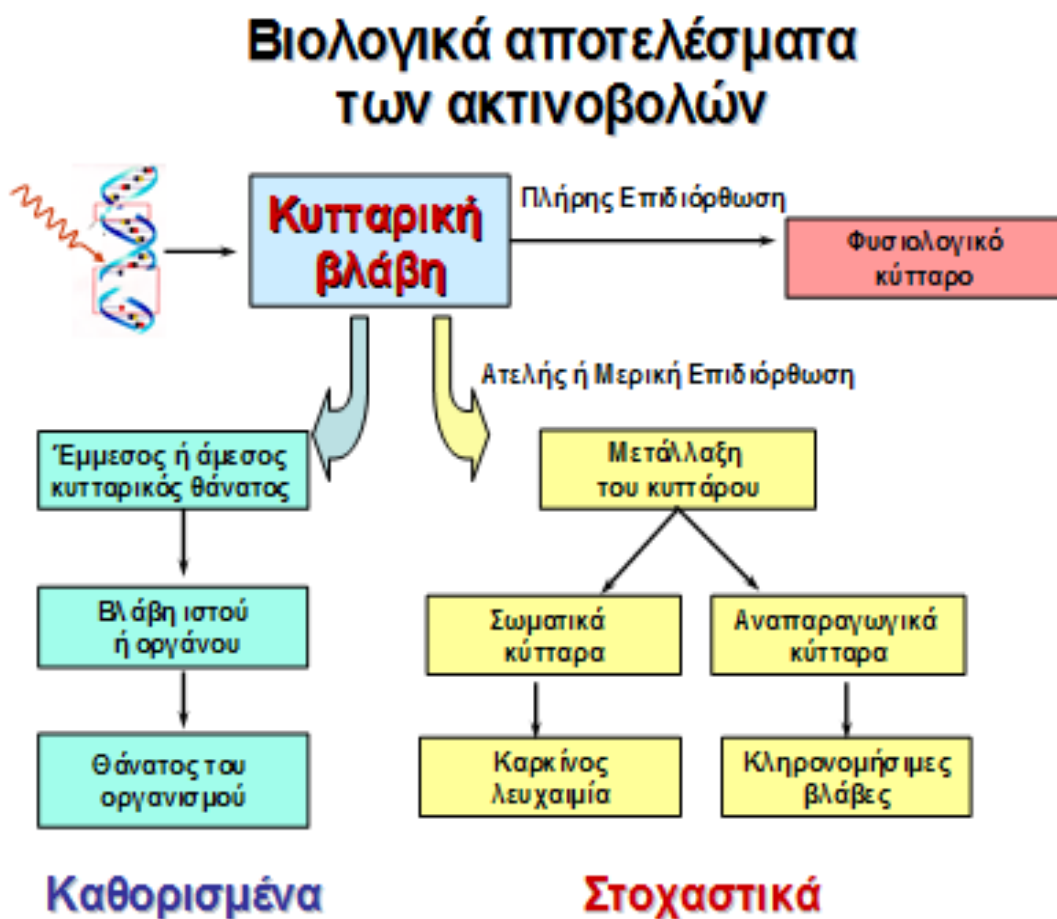
Λαμβάνοντας υπόψη ότι, τόσο η δημιουργία μιας κυτταρικής βλάβης, όσο και η μη πλήρης επιδιόρθωσή της είναι τυχαία γεγονότα, μπορούμε να υποθέσουμε ότι ακόμη και για πολύ μικρές δόσεις ακτινοβολίας, υπάρχει πάντοτε κάποια μικρή, αλλά πεπερασμένη πιθανότητα εμφάνισης μακροπρόθεσμων δυσμενών επιπτώσεων.

Σε πολύ μεγάλες δόσεις και μεγάλους ρυθμούς δόσεων ακτινοβολίας και σε αντίθεση με τα όσα προαναφέρθηκαν για τις αντίστοιχες χαμηλές δόσεις, τόσο η ποιότητα όσο και η έκταση της βλάβης αυξάνει με αποτέλεσμα τη βαριά δυσλειτουργία και τον άμεσο θάνατο μεγάλου πληθυσμού κυττάρων ενός ιστού ή οργάνου (Georgakilas, 2011). Οι βλάβες αυτές, μορφολογικές και λειτουργικές, είναι συνήθως εκτεταμένες και προφανώς μη αντιμετωπίσιμες από τους διαθέσιμους κυτταρικούς επιδιορθωτικούς μηχανισμούς και τους αντισταθμιστικούς μηχανισμούς του ιστού, με αποτέλεσμα την άμεση ή βραχυπρόθεσμη εμφάνιση σοβαρών κλινικών συμπτωμάτων (Brenner et al., 2003).

Με βάση τα παραπάνω, μπορούμε να διακρίνουμε τα δυσμενή αποτελέσματα στην υγεία ενός ατόμου που εκτέθηκε σε ιοντίζουσα ακτινοβολία στις ακόλουθες δύο γενικές κατηγορίες (Εικόνα 8)

- α. **Καθορισμένα αποτελέσματα (deterministic effects)**, είναι κλινικά αποτελέσματα, ανιχνεύσιμα μετά από έκθεση σε μεγάλες καταστροφικές δόσεις ακτινοβολίας, ενίοτε πολύ σοβαρά, που οφείλονται στη θανάτωση ή τη βαρεία δυσλειτουργία μεγάλων πληθυσμών κυττάρων ενός ιστού, και
- β. **Στοχαστικά αποτελέσματα (stochastic effects)**, στα οποία ανήκουν η λευχαιμία, ο καρκίνος και οι κληρονομήσιμες βλάβες. Οφείλουν την ονομασία τους στο γεγονός ότι η πιθανότητα εμφάνισής τους, ακολουθεί στατιστικούς νόμους. Προκύπτουν σαν αποτέλεσμα του πολλαπλασιασμού ενός μεταλλαγμένου από ακτινοβολία, σωματικού κυττάρου.

Η ανεπιτυχής ή η μερική επιδιόρθωση της κυτταρικής βλάβης, μπορεί να οδηγήσει σε μία από τις δύο γενικές κατηγορίες των βιολογικών αποτελεσμάτων των ακτινοβολιών: τα καθορισμένα (για την περιοχή των μεγάλων δόσεων) και τα στοχαστικά αποτελέσματα



Εικόνα 8: Οι δύο γενικές κατηγορίες των βιολογικών αποτελεσμάτων των ακτινοβολιών: τα καθορισμένα (για την περιοχή των μεγάλων δόσεων) και τα στοχαστικά αποτελέσματα.

Τα καθορισμένα αποτελέσματα (deterministic effects) εμφανίζονται μετά από έκθεση σε δόσεις ακτινοβολίας, μεγαλύτερες των 1Gy . Η σφοδρότητά τους αυξάνει με την αύξηση δόσης (Εικόνα 9). Από το διάγραμμα δεξιά, προκύπτει ότι για την εμφάνιση των καθορισμένων αποτελεσμάτων, απαιτείται απαραίτητα η υπέρβαση ενός κατωφλίου δόσης. Η πιθανότητα της εμφάνισής τους για δόσεις μικρότερες του κατωφλίου - δόσης είναι μηδενική, ενώ με την υπέρβαση του κατωφλίου, η τιμή της πιθανότητας εμφάνισης της βλάβης με την αύξηση δόσης, εκτινάσσεται στο 100%.. Στο αριστερό διάγραμμα της Εικόνα 9b, φαίνεται η βαρύτητα της βλάβης στον ιστό, η οποία μετά το κατώφλι κλιμακώνεται ταχύτατα με την αύξηση της δόσης. Αυτό είναι αναμενόμενο καθώς αυξάνει μεν η καταστροφικότητα της ακτινοβολίας, αλλά παράλληλα

εκπίπτει και η ικανότητα του ιστού να ανανήψει, λόγω του απεμπλουτισμού του σε ανανεώσιμους πληθυσμούς παρεγχυματικών κύτταρων που θα αντικαταστήσουν αυτούς που έχουν υποστεί τροποποιήσεις.

Οι *πρώιμες αντιδράσεις* ενός ιστού σε δόση ακτινοβολίας μεγαλύτερης από το κατώφλι δόσης, είναι αυτές που εμφανίζονται τις πρώτες ώρες, ημέρες ή και εβδομάδες μετά την ακτινοβολήση. Αυτές μπορεί να είναι φλεγμονώδεις (π.χ. το ερύθημα), λόγω της απελευθέρωσης διαφόρων κυτταρικών παραγόντων, ή να οφείλονται σε καταστροφή του πλακώδους επιθηλίου ή σε μεγάλη απώλεια κυττάρων. Οι φλεγμονώδεις αντιδράσεις οφείλονται κυρίως σε μεταβολές της διαπερατότητας των κυττάρων και της απελευθέρωσης ισταμίνης. Οι *καθυστερημένες αντιδράσεις* ενός ιστού, εμφανίζονται μερικούς μήνες ή και έτη μετά την ακτινοβολήση. Αυτές διακρίνονται σε καθυστερημένες αντιδράσεις γενικού τύπου (generic type) και σε επακόλουθου τύπου (consequential type) αντίστοιχα.

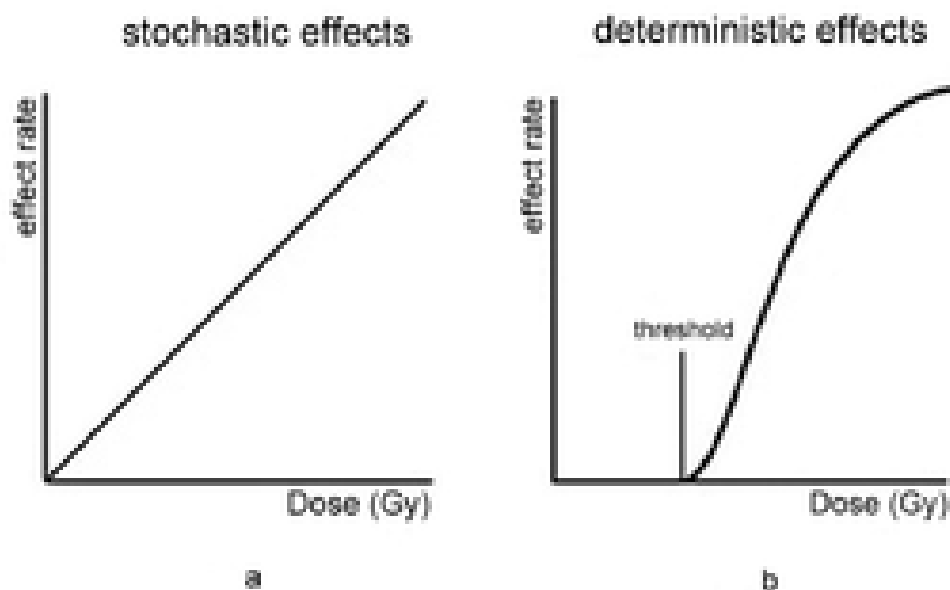
Ως γενικού τύπου καθυστερημένες αντιδράσεις, θεωρούνται αυτές που προκύπτουν σαν άμεσο αποτέλεσμα της μορφολογικής και λειτουργικής βλάβης του ίδιου του ιστού. Γενικού τύπου είναι π.χ. οι αρτηριακές αποφράξεις που οδηγούν σε εν τω βάθει ιστικές νεκρώσεις.

Ως επακόλουθου τύπου καθυστερημένες αντιδράσεις, θεωρούνται αυτές που προκύπτουν σαν αποτέλεσμα σοβαρής βλάβης που επηρέασε τον ιστό κατά την διάρκεια της φάσης των πρώιμων αντιδράσεων. Επακόλουθου τύπου αντιδράσεις είναι π.χ. η νέκρωση του δέρματος σαν αποτέλεσμα της βαρείας απογύμνωσης της επιδερμίδας και της χρόνιας φλεγμονής καθώς και οι εντερικές στενώσεις λόγω βαρείας εξέλκωσης του βλεννογόνου.

Σε περιπτώσεις καθορισμένων βιολογικών αποτελεσμάτων μετά από εφάπαξ, ολόσωμη έκθεση σε σημαντικές δόσεις ακτινοβολίας μεγαλύτερες του 1Gy, εμφανίζεται μια ακολουθία κλινικών συμπτωμάτων και κάνουμε λόγο για οξύ ακτινικό σύνδρομο.

Για την περιοχή απορροφούμενης δόσης έως 100 mGy, η πιθανότητα οποιοσδήποτε ιστός ή όργανο να παρουσιάσει κλινικά εκδηλούμενη αντίδραση είναι μηδενική. Τούτο ισχύει τόσο για χαμηλές ακαριαίες εφάπαξ δόσεις, όσο και χρόνιες χαμηλές δόσεις, εκφρασμένες ως επαναλαμβανόμενες ετήσιες δόσεις.

Η εξάρτηση της πιθανότητας εμφάνισης του στοχαστικού αποτελέσματος από τη δόση ακτινοβολίας αποδίδεται με γραμμικό χωρίς κατώφλι - δόσης μαθηματικό πρότυπο (Linear Non Threshold model – LNT). Αντιθέτως η βαρύτητα του στοχαστικού αποτελέσματος είναι ανεξάρτητη της δόσης της ακτινοβολίας (Εικόνα 10).



Εικόνα 9: Η εξάρτηση από τη δόση της ακτινοβολίας της πιθανότητας εμφάνισης των στοχαστικών και μη αποτελεσμάτων.

Όσον αφορά στη σχέση πιθανότητας εμφάνισης στοχαστικών αποτελεσμάτων συναρτήσει της δόσης για χαμηλές δόσεις και χαμηλούς ρυθμούς δόσεων, η επικρατούσα σήμερα άποψη, υποστηρίζει την υπόθεση του μαθηματικού γραμμικού πρότυπου πιθανότητας - δόσης, χωρίς κατώφλι δόσης (LNT). Αποδεχόμαστε εν ολίγοις, ότι η γραμμική συνάρτηση που συνδέει την πιθανότητα εμφάνισης θανατηφόρου καρκίνου με τη δόση των μεγάλων ακαριαίων δόσεων, μπορεί να προεκταθεί και στην αντίστοιχη περιοχή των χαμηλών δόσεων. Για τη μεταφορά όμως της συνάρτησης αυτής στην περιοχή των χαμηλών δόσεων, χρησιμοποιείται προσεγγιστικά, ο συντελεστής μείωσης της πιθανότητας εμφάνισης ανά μονάδα δόσης, γνωστός ως DDREF (DDREF - Dose and Dose Rate Effectiveness Factor - Συντελεστής Αποτελεσματικότητας Δόσης και Ρυθμού Δόσης). Ο συντελεστής DDREF συνδέει το αναμενόμενο ραδιοβιολογικό αποτέλεσμα των χαμηλών δόσεων και χαμηλών ρυθμών δόσης με το αντίστοιχο των υψηλών δόσεων και υψηλών ρυθμών δόσης. Η μέση τιμή για τον συντελεστή μείωσης είναι ίση με δύο, (DDREF = 2). Η περιοχή των χαμηλών δόσεων και χαμηλών ρυθμών δόσεων ακτινοβολίας (<100mSv) (Hamada and Fujimichi, 2014), παρουσιάζει ιδιαίτερο ενδιαφέρον για την ιατρική όπως αναφέραμε. Η ενεργός δόση ακτινοβολίας που δέχεται ένα άτομο που υποβάλλεται σε συνήθεις διαγνωστικές ακτινολογικές πρακτικές, κυμαίνεται ανά εξέταση από μερικά μSv (π.χ. απλή ακτινογραφία θώρακος) μέχρι μερικές δεκάδες mSv (π.χ. υπολογιστική τομογραφία θώρακος ~ 10 -15 mSv).

Το διεθνές σύστημα ακτινοπροστασίας (ICRP 103), στηρίζεται στο μαθηματικό γραμμικό πρότυπο LNT. Σύμφωνα με αυτό, για λόγους προάσπισης της υγείας δεχόμαστε ότι, οποιαδήποτε δόση ακτινοβολίας, οσοδήποτε μικρή και αν είναι, ενέχει κίνδυνο πρόκλησης καρκίνου, το μέτρο του οποίου είναι ανάλογο του μεγέθους της εν λόγω δόσης.

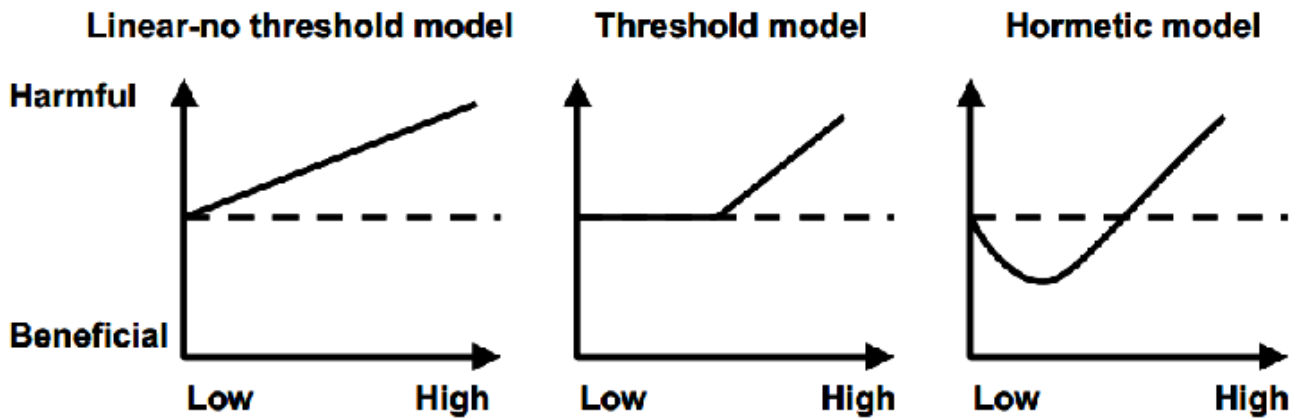
ΤΡΟΠΟΙ ΜΕΛΕΤΗΣ ΕΠΙΔΡΑΣΗΣ 'ΧΑΜΗΛΩΝ' ΔΟΣΕΩΝ (<0.5GY)

Οι βιολογικές επιδράσεις της ακτινοβολίας χαμηλής δόσης D ($D < 0.1$ Gy) και σχετικά χαμηλής δόσης D ($0.1 < D < 0.5$ Gy) υπολογίζονται επί του παρόντος συνάγοντας αποτελέσματα από τις βιολογικές επιδράσεις της ακτινοβολίας υψηλής δόσης (Bonner, 2003). Αυτά τα δεδομένα υπόκεινται σε πολλές αβεβαιότητες, συμπεριλαμβανομένων των ανεξέλεγκτων συνθηκών και της ανεπαρκούς δοσιμετρίας. Αυτή η μελέτη ενσωματώνεται στο γραμμικό χωρίς-κατώφλι δόσης μαθηματικό μοντέλο (linear non-threshold model), το οποίο δηλώνει ότι η ακτινοβολία χαμηλής δόσης είναι εξίσου επιβλαβής ανά Gy με την ακτινοβολία υψηλής δόσης (Krisiuk, 1989). Έτσι οποιαδήποτε δόση δεν έχει σημασία πόσο μικρή είναι δυνητικά επιβλαβής (Trott and Rosemann, 2000).

Ένα άλλο μοντέλο προτεινόμενο για την επεξήγηση της δράσης των χαμηλών δόσεων είναι το μοντέλο κατωφλίου (threshold model) που υποθέτει ότι η χαμηλή δόση ακτινοβολίας είναι ακίνδυνη κάτω από ένα ορισμένο επίπεδο. Η ανάλυση των επιδημιολογικών δεδομένων, κυρίως της μελέτης των επιζώντων ατομικής βόμβας, φαίνεται να δείχνει σε ορισμένες περιπτώσεις γραμμικότητα μεταξύ δόσης και κινδύνου σε χαμηλές δόσεις αλλά επίσης δεν μπορεί να αποκλείσει ένα όριο στα 60 mSv (Little and Muirhead, 2000) (Bonner, 2003).

Ένα άλλο μοντέλο είναι το μοντέλο προσαρμογής-απόκρισης, το οποίο δηλώνει ότι ορισμένες δόσεις ακτινοβολίας χαμηλής δόσης μπορεί να είναι ακόμη ευεργετικές. Το συγκεκριμένο προτείνει ότι χαμηλές δόσεις ακτινοβολίας επάγουν συνθήκες στο κύτταρο που μειώνουν τις ακολουθούμενες υψηλότερες δόσεις (Georgakilas et al., 1999; Kim et al., 2011; Olivieri et al., 1984).

Ακόμη ένα άλλο μοντέλο είναι το μοντέλο της μεταβίβασης σημάτων μεταξύ κυττάρων μετά την ακτινοβόληση (bystander-effect), το οποίο υποθέτει ότι η ακτινοβολία χαμηλής δόσης μπορεί να είναι ακόμη πιο καταστροφική από εκείνη που προβλέπεται από το γραμμικό χωρίς κατώφλι δόσης μοντέλο (Bonner, 2003; Georgakilas, 2015). Σύμφωνα με αυτό το μοντέλο της συστημικής δράσης της ακτινοβολίας, τα ακτινοβολημένα κύτταρα μπορεί να σηματοδοτήσουν την ανησυχία τους σε άλλα κύτταρα, ίσως με άμεση αλληλεπίδραση κυττάρου προς κύτταρο ή με μόρια που εκκρίνονται στο μέσο. Το φαινόμενο οφείλεται στην απόκριση σε σήματα που μεταβιβάζονται από κύτταρα που έχουν πληγεί από την ακτινοβολία μέσω μορίων που είτε εισέρχονται μέσω χασμοσυνδέσεων (gap junctions) στην κυτταρική μεμβράνη γειτονικών κυττάρων είτε διαχέονται στο υλικό καλλιέργειας των κυττάρων (Mothersill and Seymour, 1998).



Εικόνα 10: Τρόποι μελέτης επίδρασης χαμηλών δόσεων.

Τέλος υπάρχουν ενδείξεις ότι πολλές από τις μεθόδους που χρησιμοποιούνται για τη μέτρηση της DSB repair προκαλούν κατά λάθος την εν λόγω μορφή βλάβης ως μέρος της μεθοδολογίας και ως εκ τούτου περιορίζονται στην ευαισθησία τους (Rothkamm and Lobrich, 2003). Υπάρχουν στοιχεία ότι οι εστίες γ -H2AX φωσφορυλιωμένη ιστόνη που ανιχνεύονται με ανοσοφθορισμό είναι ποσοτικά οι ίδιες με αυτές των DSB και είναι ικανές να ποσοτικοποιήσουν την επιδιόρθωση μεμονωμένων DSB. Αυτό το εύρημα επιτρέπει τη διερεύνηση της DSB repair μετά από δόσεις ακτινοβολίας τόσο χαμηλές όσο 1 mGy. Παραδόξως, οι DSB που επάγονται σε καλλιέργειες μη διαχωρισμένων πρωτογενών ανθρώπινων ινοβλαστών με πολύ χαμηλές δόσεις ακτινοβολίας (≈ 1 mGy) παραμένουν αδιόρθωτες για πολλές ημέρες, σε έντονη αντίθεση με την αποτελεσματική αποκατάσταση DSB που παρατηρείται σε υψηλότερες δόσεις. Ωστόσο, το επίπεδο των DSB σε ακτινοβολημένες καλλιέργειες μειώνεται σε εκείνο των μη ακτινοβολημένων κυτταρικών καλλιεργειών εάν τα κύτταρα αφεθούν να πολλαπλασιαστούν μετά την ακτινοβολία και αυτό το αποτέλεσμα μπορεί να προκληθεί από την απομάκρυνση των κυττάρων που φέρουν ακέραια DSBs. Τα αποτελέσματα που παρουσιάζονται είναι σε αντίθεση με τα τρέχοντα μοντέλα εκτίμησης κινδύνου που υποθέτουν ότι οι κυτταρικές αποκρίσεις είναι εξίσου αποτελεσματικές σε χαμηλές και υψηλές δόσεις και παρέχουν τη δυνατότητα χρήσης του σχηματισμού εστιών γ -H2AX ως άμεσου βιοδείκτη για την έκθεση του ανθρώπου σε χαμηλές ποσότητες ιοντίζουσας ακτινοβολίας (Alessio et al., 2015; Bazyka et al., 2014; Rothkamm and Lobrich, 2003)

ΜΙΚΡΟΣΥΣΤΟΙΧΙΕΣ

Χαρτογράφηση ανθρώπινου γονιδιώματος.

Δυο πρόσφατες συμπληρωματικές επιστημονικές μέθοδοι διευκολύνουν εξαιρετικά τη μελέτη της γονιδιακής έκφρασης και την ανακάλυψη των ρόλων που παίζουν συγκεκριμένα γονίδια σε διάφορες βιολογικές λειτουργίες ή στην εξέλιξη μιας ασθένειας.

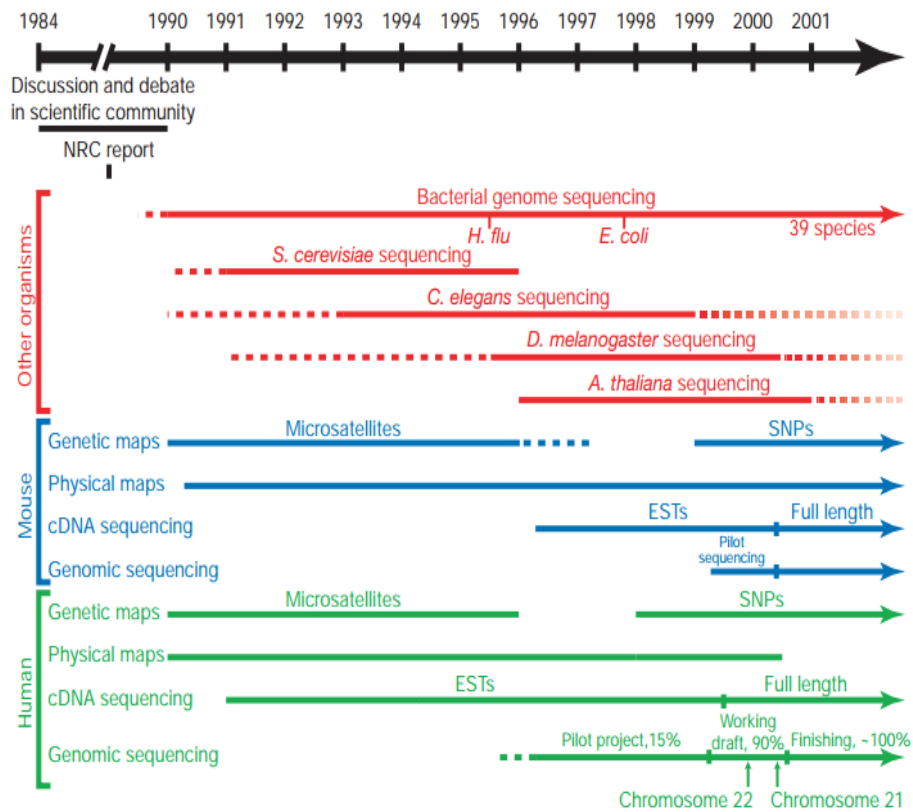
Η πρώτη είναι αυτή της χαρτογράφησης του ανθρώπινου γονιδιώματος (Human Genome Project, HGP) (Εικόνα 11) με την οποία έγινε διαθέσιμος ένας τεράστιος όγκος πληροφοριών πάνω στην αλληλουχία DNA του ανθρώπινου γονιδιώματος (Lander et al., 2001).



Εικόνα 11: Human Genome Project (Πηγή: <http://humangenomeproject.yolasite.com/background.php>)

Ταυτοποιήθηκε στη συνέχεια ένας μεγάλος αριθμός νέων γονιδίων μέσα σε αλληλουχίες άγνωστες μέχρι πρότινος. Αυτός ο όγκος της πληροφορίας δημιούργησε την ανάγκη να βρεθεί ένας τρόπος να οργανωθεί και να κατηγοριοποιηθεί έτσι ώστε να είναι δυνατόν να εξαχθούν χρήσιμα συμπεράσματα για τις λειτουργίες των

νέων αυτών γονιδίων. Στην Εικόνα 12 φαίνεται ένα διάγραμμα χρονικής εξέλιξης της χαρτογράφησης του ανθρώπινου γονιδιώματος σε μερικά ασπόνδυλα (κόκκινο χρώμα), στον άνθρωπο (πράσινο χρώμα) και τέλος στο ποντίκι (μπλε χρώμα) (Lander et al., 2001)



Εικόνα 12: Χρονική εξέλιξη γενωμικής ανάλυσης (Πηγή: (Lander et al., 2001))

Τεχνολογία μικροσυστοιχιών

Η δεύτερη μέθοδος απαντά σε αυτό το ερώτημα και με πολύ κομψό τρόπο μπορεί να διευκολύνει την ταυτοποίηση και την ταξινόμηση αυτής της πληροφορίας πάνω στην αλληλουχία του DNA και την αντιστοίχιση των λειτουργιών τους σε αυτά τα νέα γονίδια. Αναφερόμαστε φυσικά στην τεχνολογία των μικροσυστοιχιών DNA. Οι μικροσυστοιχίες έχουν το πλεονέκτημα ότι περιέχουν έναν πολύ μεγάλο αριθμό γονιδίων σε πολύ μικρό μέγεθος. Μπορούν να χρησιμοποιηθούν προκειμένου να γίνει ανάλυση της γονιδιακής έκφρασης με ένα μόνο δείγμα ή προκειμένου να συγκριθεί η γονιδιακή έκφραση σε δυο διαφορετικούς κυτταρικούς τύπους ή σε διαφορετικά είδη ιστών, όπως σε έναν υγιή και έναν παθολογικό ιστό. Η μικροσυστοιχία μπορεί να χρησιμοποιηθεί για να εξετάσει την έκφραση σε εκατοντάδες ή και σε χιλιάδες γονίδια με ένα μόνο πείραμα. Η τεχνολογία των μικροσυστοιχιών, όπως και οι άλλες μέθοδοι υψηλής απόδοσης, κρίνονται πλέον απαραίτητες για την κατανόηση των βιολογικών διεργασιών που λαμβάνουν χώρα στα βιολογικά συστήματα και συμπληρώνουν τις κοινές μεθόδους μελέτης (Lou et al., 2001).

Οι μικροσυστοιχίες γονιδίων είναι δισδιάστατες διατάξεις κηλίδων στις οποίες είναι ακινητοποιημένες αλληλουχίες χιλιάδων διαφορετικών γονιδίων. Είναι τοποθετημένες σε ένα στερεό υπόστρωμα όπως γυαλί, πυρίτιο, κ.α. (Smyth and Speed, 2003). Το DNA τυπώνεται, τοποθετείται ή στην πράξη συντίθεται άμεσα πάνω στο υπόστρωμα.

Η διαδικασία βασίζεται στην ανίχνευση μέσω υβριδισμού, μια τεχνική που χρησιμοποιεί μόρια νουκλεϊνικού οξέος σεσημασμένα με φθορίζουσα χρωστική ως «κινητούς ανιχνευτές» προκειμένου να ταυτοποιήσουν συμπληρωματικά μόρια, δηλαδή αλληλουχίες που μπορούν να ζευγαρώσουν με άλλες. Όταν δυο συμπληρωματικές ακολουθίες συναντώνται, όπως ο ακινητοποιημένος στόχος DNA και ο κινητός ανιχνευτής cDNA ή crRNA, τότε υβριδίζουν μεταξύ τους.

Υπάρχουν διάφοροι τύποι δειγμάτων που μπορούν να χρησιμοποιηθούν για να κατασκευαστούν μικροσυστοιχίες DNA όπως γενωμικοί, μεταγραφωμικοί, πρωτεωμικοί, κλπ. Αυτό που διαφοροποιεί το καθένα είναι το είδος του ακινητοποιημένου DNA που χρησιμοποιείται για να παραχθεί η συστοιχία και τέλος, το είδος της πληροφορίας που εξάγεται από το τσιπ.

Ανάλογα με το εκάστοτε βιολογικό ερώτημα, διαθέτουμε διαφόρους τύπους μικροσυστοιχιών. Ενδεικτικά αναφέρουμε ότι υπάρχουν μικροσυστοιχίες πρωτεϊνών, ιστών, cDNA, πεπτιδίων, αντισωμάτων (οι πρώτες που δημιουργήθηκαν), κυτταρικές, κ.α.

Οι μικροσυστοιχίες επιτρέπουν ποσοτικές ή ποιοτικές μετρήσεις εκμεταλλευόμενες την εκλεκτική φύση της αρχής της συμπληρωματικότητας μεταξύ νουκλεϊκών οξέων DNA-DNA ή DNA-RNA ή και μεταξύ των αμινοξέων των πρωτεϊνών, υπό αυστηρά ελεγχόμενες συνθήκες. Μια μικροσυστοιχία λειτουργεί εκμεταλλευόμενη την ιδιότητα ενός δοσμένου μορίου mRNA, να συνδέεται εξειδικευμένα, δηλαδή να υβριδίζει, με το μητρικό DNA από το οποίο προήλθε. Χρησιμοποιώντας μια συστοιχία, η οποία περιέχει πολλά δείγματα DNA, οι επιστήμονες μπορούν με ένα μόνο πείραμα να καθορίσουν τα επίπεδα έκφρασης χιλιάδων γονιδίων μέσα σε ένα κύτταρο μετρώντας μόνο το ποσό του υβριδισμένου mRNA σε κάθε σημείο πάνω στη συστοιχία. Με τη βοήθεια ενός υπολογιστή μπορεί να καταγραφεί με ακρίβεια η ποσότητα του mRNA που συνδέθηκε με τα επιτόπια (spots) της μικροσυστοιχίας, δημιουργώντας έτσι ένα προφίλ γονιδιακής έκφρασης μέσα στο κύτταρο. Οι μικροσυστοιχίες γονιδίων χρησιμοποιούνται σήμερα κατά κόρον για την εξέταση της γονιδιακής έκφρασης υπό ειδικές συνθήκες και για την ανίχνευση νουκλεϊκών οξέων παθογόνων. Η μέθοδος μοριακής βιολογίας αυτή, έχει το πλεονέκτημα ότι δύναται να εξετάζει ταυτόχρονα την έκφραση χιλιάδων γονιδίων, και ενδείκνυται για συγκριτικές μελέτες γονιδιωμάτων (Lou et al., 2001). Μπορεί κανείς να εξάγει χρήσιμες πληροφορίες για τη βιολογική λειτουργία ενός οργανισμού, βρίσκοντας ποια γονίδια επάγονται ή καταστέλλονται σε κάποια φάση του κυτταρικού κύκλου, σε κάποια αναπτυξιακή στιγμή ή σε απόκριση σε ερεθίσματα του περιβάλλοντος, όπως η απόκριση σε ορμόνες ή σε υψηλή θερμοκρασία ή σε ιοντίζουσα ακτινοβολία. Ομάδες γονιδίων των οποίων η έκφραση αυξάνεται ή μειώνεται υπό τις ίδιες συνθήκες, είναι πιθανό να έχουν συσχετιζόμενη βιολογική λειτουργία και ίσως, κοινή σχέση

ρύθμισης (Pollack et al., 1999). Θα μπορούσαν π.χ. να έχουν παρόμοιες αλληλουχίες υποκινητών για ίδιους μεταγραφικούς παράγοντες. Επιπρόσθετα, αν μπορούσε κανείς να αποκομίσει ένα πρότυπο έκφρασης για συγκεκριμένες συνθήκες, θα είχε ένα χρήσιμο εργαλείο στη διάθεσή του, ώστε να είναι σε θέση να χαρακτηρίσει παρόμοιες άγνωστες καταστάσεις. Η γονιδιακή έκφραση είναι άμεσα συσχετιζόμενη με τις βιολογικές λειτουργίες και οι μικροσυστοιχίες υπόσχονται τεράστιο μέγεθος δεδομένων πάνω σε ανθρώπινες ασθένειες, γήρανση, φαρμακευτική δράση, ορμονική δράση, διανοητικές ασθένειες, μεταβολισμό και σε αρκετά ακόμα κλινικά θέματα. Ανοίγουν ένα νέο δρόμο σε μεθόδους διάγνωσης καθώς με την εξέλιξή τους γίνονται όλο και πιο διαθέσιμα σε χρήση σε εργαστήρια και χώρους διάγνωσης (Schena et al., 1996).

ΕΙΔΗ ΜΙΚΡΟΣΥΣΤΟΙΧΙΩΝ

Ανάλογα με τον τρόπο κατασκευής της μικροσυστοιχίας, μια γενική διάκριση που μπορούμε να κάνουμε είναι οι τυπωμένες μικροσυστοιχίες, μικροσυστοιχίες με χρήση σφαιριδίων και μικροσυστοιχίες «επιτόπιας σύνθεσης».

Είναι αναγκαίο να αναφέρουμε μερικά στοιχεία για τη προετοιμασία του δείγματος πριν την εκτύπωση.

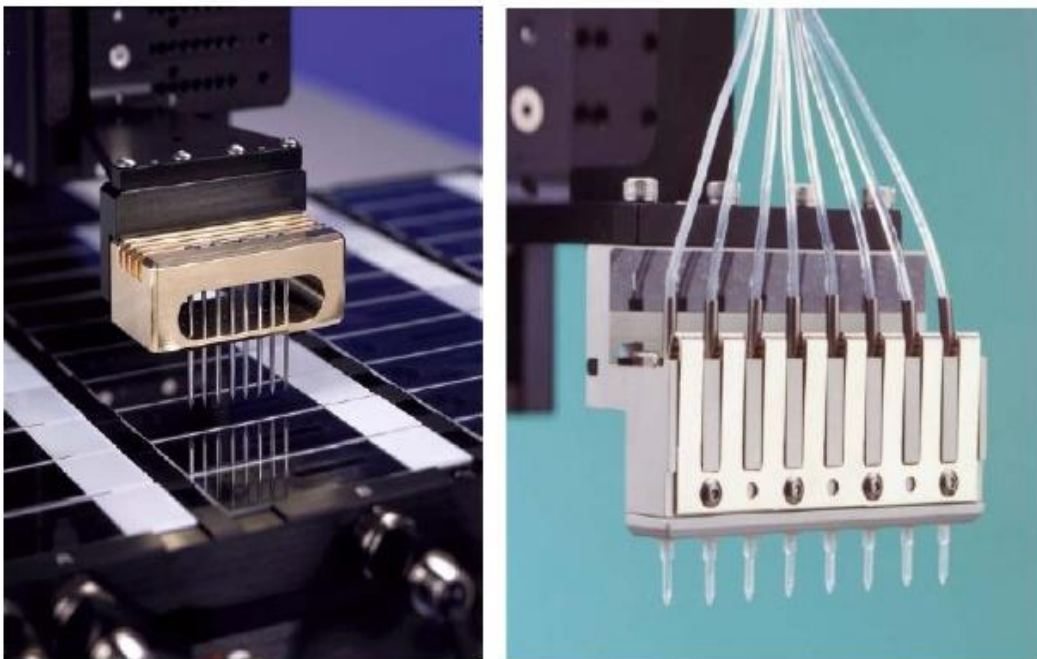
Πρώτον, τα δείγματα των ανιχνευτών βρίσκονται σε κατάλληλο διάλυμα προστασίας. Τα διαλύματα αυτά επιτελούν διάφορες λειτουργίες, μερικές από αυτές είναι να αυξήσουν την αποτελεσματικότητα της εκτύπωσης, να μειώσουν τη διασπορά του δείγματος, να εξασφαλίσουν ομοιόμορφη μεταφορά, να σταθεροποιήσουν τα δείγματα των ανιχνευτών, να απενεργοποιήσουν τους ανιχνευτές, να μειώσουν τη ξήρανση του υλικού και να αυξήσουν την ορατότητα των σημείων. Δεύτερον η επιφάνεια εκτύπωσης επίσης έχει υποστεί ειδική χημική επεξεργασία ώστε οι ανιχνευτές να τοποθετηθούν με το κατάλληλο προσανατολισμό στη μικροσυστοιχία. Απαιτείται επιφάνεια υψηλής ποιότητας γιατί παίζει σημαντικό ρόλο στην προσκόλληση των μορίων. Από την επιφάνεια επίσης εξαρτάται η απόδοση των αντιδράσεων καθώς και η ακρίβεια κατά την ανίχνευση και τέλος η ποιότητα των δεδομένων που προκύπτουν. Μερικά από τα χαρακτηριστικά της επιφάνειας τα οποία πρέπει να ελέγχονται και να διαμορφώνονται ανάλογα είναι οι διαστάσεις αυτής και το ότι πρέπει να είναι λεία και μεγάλης αντοχής. Το πιο κοινό υλικό είναι το γυαλί. Το γυαλί αποτελείται από διοξείδιο του πυριτίου. Σε αυτό πολλές φορές προστίθενται οξείδια μετάλλων προκειμένου να αυξηθεί το σημείο τήξης του γυαλιού. Το γυαλί επικαλύπτεται με αμινικές και αλδεϋδικές ομάδες που επιτρέπουν την πιο σταθερή προσκόλληση των νουκλεοτιδίων. Ο δεσμός προσκόλλησης στην επιφάνεια μπορεί να είναι ομοιοπολικός ή μη.

I. Τυπωμένες μικροσυστοιχίες.

Η ανάπτυξη της νανοτεχνολογίας έδωσε τη δυνατότητα να τυπώνουμε σε επίπεδο μορίου. Μπορούμε να διακρίνουμε ανάλογα τη τεχνολογία εκτύπωσης την εκτύπωση επαφής (contact printing) και την εκτύπωση χωρίς επαφή (non-contact printing) (Εικόνα 13). Οι τυπωμένες μικροσυστοιχίες ανήκουν στην πρώτη κατηγορία, του contact printing, που περιλαμβάνει απευθείας επαφή του εκτυπωτικού μηχανήματος ή του δείγματος που περιέχεται στο εκτυπωτικό όργανο με την επιφάνεια και είναι οι πιο ευρέως διαδεδομένες. Όταν οι ανιχνευτές δε συντίθεται απευθείας πάνω στο πλακίδιο αλλά εκτυπώνονται στην επιφάνεια μετά την ετοιμασία τους τότε κάνουμε λόγω για τυπωμένες μικροσυστοιχίες. Ποσότητες του ανιχνευτή τοποθετούνται πάνω στην γυάλινη επιφάνεια με τη βοήθεια μηχανημάτων που επιτρέπουν την ανάλογη ακρίβεια. Οι ανιχνευτές μπορεί να αποτελούνται από cDNA, αποπροϊόν PCR ή από ολιγονουκλεοτίδια, που προέρχονται από το ολικό mRNA των υπό εξέταση βιολογικών δειγμάτων. Καθένας είναι συμπληρωματικός ως προς ένα μοναδικό γονίδιο και μπορεί να τοποθετηθεί στην επιφάνεια με διάφορους τρόπους.

Οι συσκευές εκτύπωσης συμπεριλαμβάνουν μεγάλη οικογένεια οργάνων μεταφοράς όπως μικροτοποθετικές ακίδες (microspotting pins), τσιμπίδες και ακίδες σχισμών, τριχοειδή και άλλες παρόμοιες συσκευές. Κατά τη διάρκεια εκτύπωσης επαφής το δείγμα βρίσκεται στο εσωτερικό της ακίδας και συγκρατείται από δυνάμεις συνάφειας και επιφανειακή τάση. Οι μοριακές αυτές δυνάμεις είναι ελκτικές και αναπτύσσονται σε μικρή ακτίνα δράσης. Ασκούνται στο δείγμα από το υλικό της ακίδας που είναι συνήθως ανοξειδωτο ατσάλι. Η επιφανειακή τάση αφορά δυνάμεις που ασκούνται από το ίδιο το υγρό υλικό στα μόρια της επιφάνειας. Όταν η σταγόνα αγγίζει την επιφάνεια εκτύπωσης η επιφάνεια ασκεί μια δύναμη αντίθετη στις προαναφερθείσες η οποία είναι η αιτία που μένει η σταγόνα στην επιφάνεια (Lausted et al., 2006).

Περίληπτικά αξίζει να αναφέρουμε ότι όταν γίνεται λόγος για εκτύπωση χωρίς επαφή πρόκειται για την τεχνολογία που επιτρέπει εκτύπωση των ανιχνευτών χωρίς άμεση επαφή του εκτυπωτικού οργάνου με την επιφάνεια της μικροσυστοιχίας. Ειδικότερα μιλάμε για συσκευές εκτόξευσης μελάνης με χρήση πιεζοηλεκτρικών κρυστάλλων και μικροσωληνοειδών. Κάθε σταγόνα αντιστοιχεί σε 100 picolitres δείγματος Είναι σαφώς γρηγορότερη τεχνική τύπωσης αλλά η χρήση της παραμένει περιορισμένη λόγω κυρίως υψηλού κόστους.

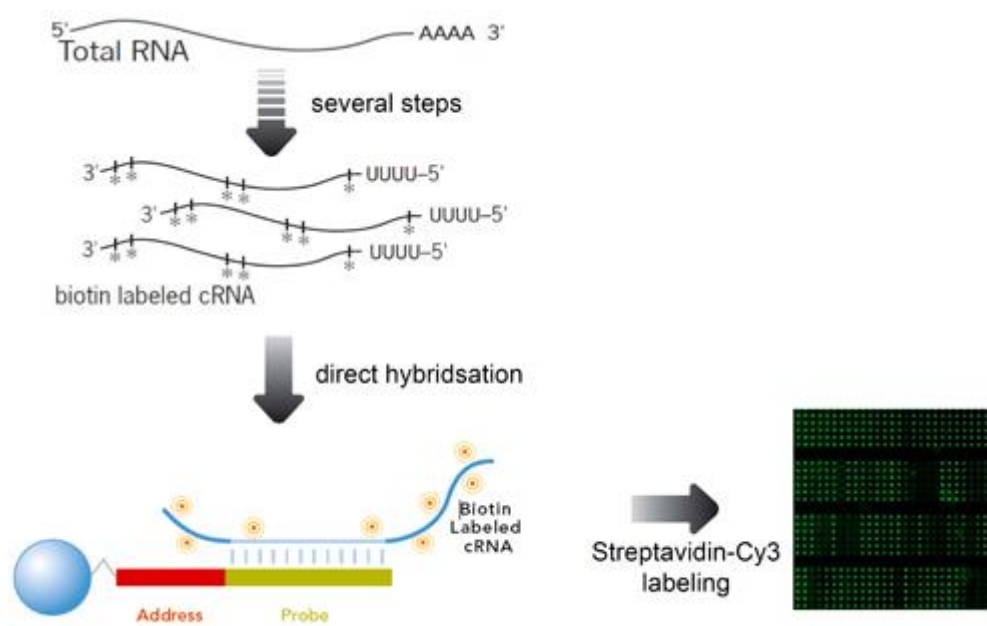


Εικόνα 13: Συσκευές τύπωσης μικροσυστοιχιών. Αριστερά βρίσκεται η μέθοδος εξ' επαφής και δεξιά η μέθοδος εκτύπωσης με εκτόξευση μελάνης (Πηγή: <http://www.slideshare.net/ruchibioinfo/microarray>).

II. Μικροσυστοιχίες με χρήση σφαιριδίων.

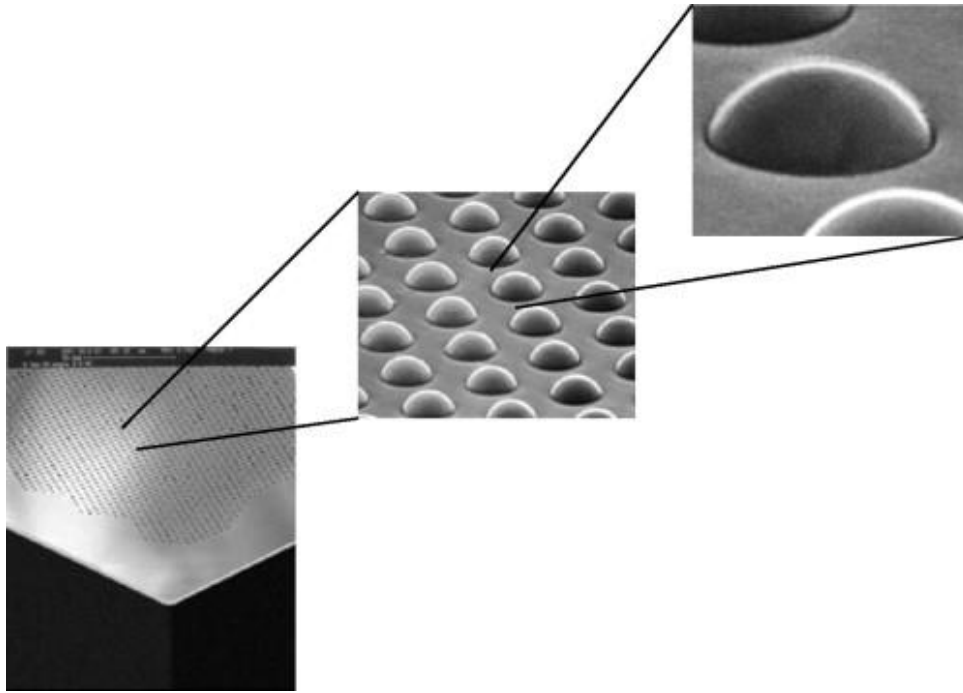
Οι μικροσυστοιχίες κατασκευάζονται εν γένει με τη τοποθέτηση ιχνηθετών σε δισδιάστατο υπόστρωμα γνωρίζοντας εκ των προτέρων τη θέση του καθενός. Η τεχνολογία της Illumina (Illumina, Inc., USA) βασίζεται στην τυχαία εναπόθεση σφαιριδίων (BeadArray®) πάνω σε ένα υπόστρωμα διαμορφωμένο ανάλογα, συνήθως γυαλί. Καταλήγουμε λόγω αυτής της κατασκευαστικής ιδιομορφίας σε αντίθεση με άλλα είδη μικροσυστοιχιών στην ανάγκη εύρεσης ενός τρόπου αποκωδικοποίησης προκειμένου να χαρτογραφήσουμε την ακριβή θέση των σφαιριδίων στο πλέγμα (Fan et al., 2006). Βρίσκουν εφαρμογή σε διάφορα βιολογικά προβλήματα και μπορεί να είναι ενός ή και δύο χρωμάτων. Οι μικροσυστοιχίες έκφρασης ωστόσο είναι πάντα ενός χρώματος.

Η τεχνολογία της Illumina όσον αφορά τη βιολογική επεξεργασία παρουσιάζεται συνοπτικά στην Εικόνα 14.



Εικόνα 14:Βιολογική Επεξεργασία δείγματος για χρήση σε μικροσυστοιχία Illumina.

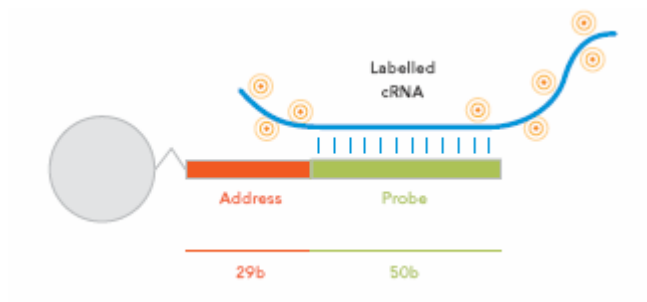
Η τεχνολογία Illumina BeadArray αφορά μικροσυστοιχίες ολιγονουκλεοτιδίων ακινητοποιημένων σε μαγνητικά σφαιρίδια πυριτίου (silica beads) τα οποία είναι τοποθετημένα τυχαία σε 'πηγάδια' διαμέτρου της τάξης μερικών μm στην επιφάνεια ενός υποστρώματος με δεδομένη διάταξη. Ειδικότερα τα σφαιρίδια πυριτίου μεγέθους 3μm, εναποτίθενται τυχαία στα πηγάδια μεγέθους 5μm. Τα πηγάδια είναι διαμορφωμένα ώστε να χωράνε ένα σφαιρίδιο το καθένα (Εικόνα 15).



Εικόνα 15: Διάταξη BeadArray (Fan et al., 2006).

Οι δυνατοί υπό μελέτη οργανισμοί στην Illumina είναι ο άνθρωπος και το ποντίκι.

Τα ολιγονουκλεοτίδια που χρησιμοποιούνται για τη μικροσυστοιχία γονιδιακής έκφρασης είναι μεγέθους 79 βάσεων (Εικόνα 16).



Εικόνα 16: Σχηματική απεικόνιση ενός ολιγομερούς που προσκολλάται στο σφαιρίδιο. Τα πραγματικά σφαιρίδια έχουν προσκολλημένα στην επιφάνεια τους εκατοντάδες χιλιάδες αντίγραφα της ίδιας ακολουθίας (source: <http://www.ohsu.edu/>).

Αποτελούνται από δύο μέρη:

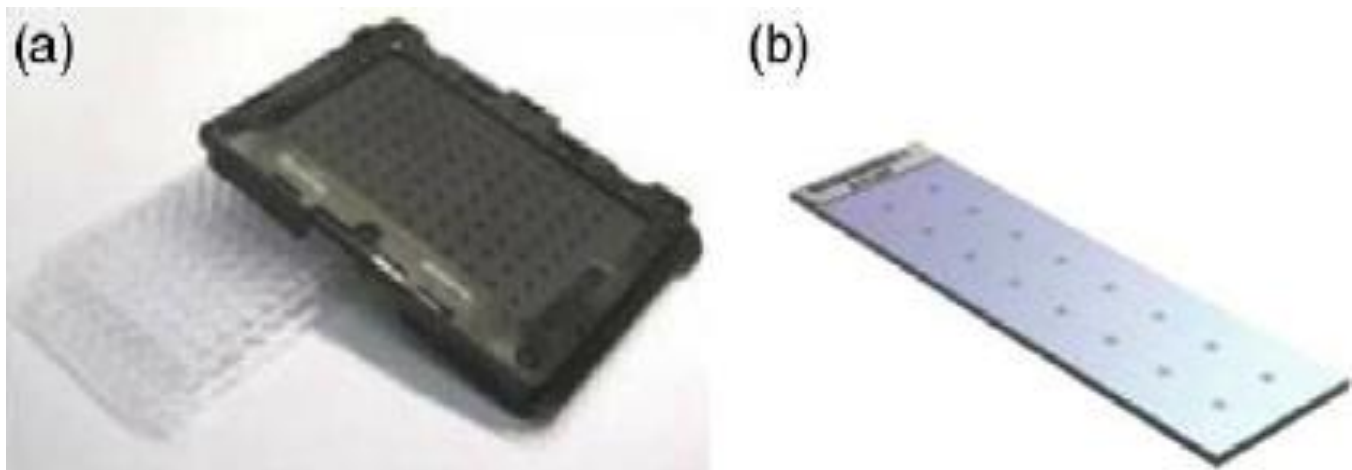
PROBE

Το ένα μέρος είναι οι ιχνηθέτες (probes) μεγέθους 50 βάσεων. Βρίσκεται στην 5' θέση και έχει σχεδιαστεί για να υβριδίζει ακολουθίες γονιδίων που είναι γνωστές οι αλληλουχίες τους από τα δημόσια αποθετήρια δεδομένων Αυτό το κομμάτι του ολιγονουκλεοτιδίου προορίζεται ώστε να υβριδιστεί με τον επισημασμένο

στόχο που έχει εξαχθεί από την polyA tail που είναι μέρος του ολικού RNA. Περίπου 30 αντίγραφα του κάθε ιχνηθέτη διατίθενται για τις συστοιχίες 6 και 8 θέσεων όσον αφορά τα ποντίκια και 15 αντίγραφα του κάθε ιχνηθέτη για τη συστοιχία 12-θέσεων που αφορά το ανθρώπινο γονιδίωμα. Οι ιχνηθέτες αυτοί εναποτίθενται τυχαία κατά μήκος της συστοιχίας. Η μικροσυστοιχία 12-θέσεων για τον άνθρωπο περιέχει περίπου 40.000 ιχνηθέτες.

ADDRESS

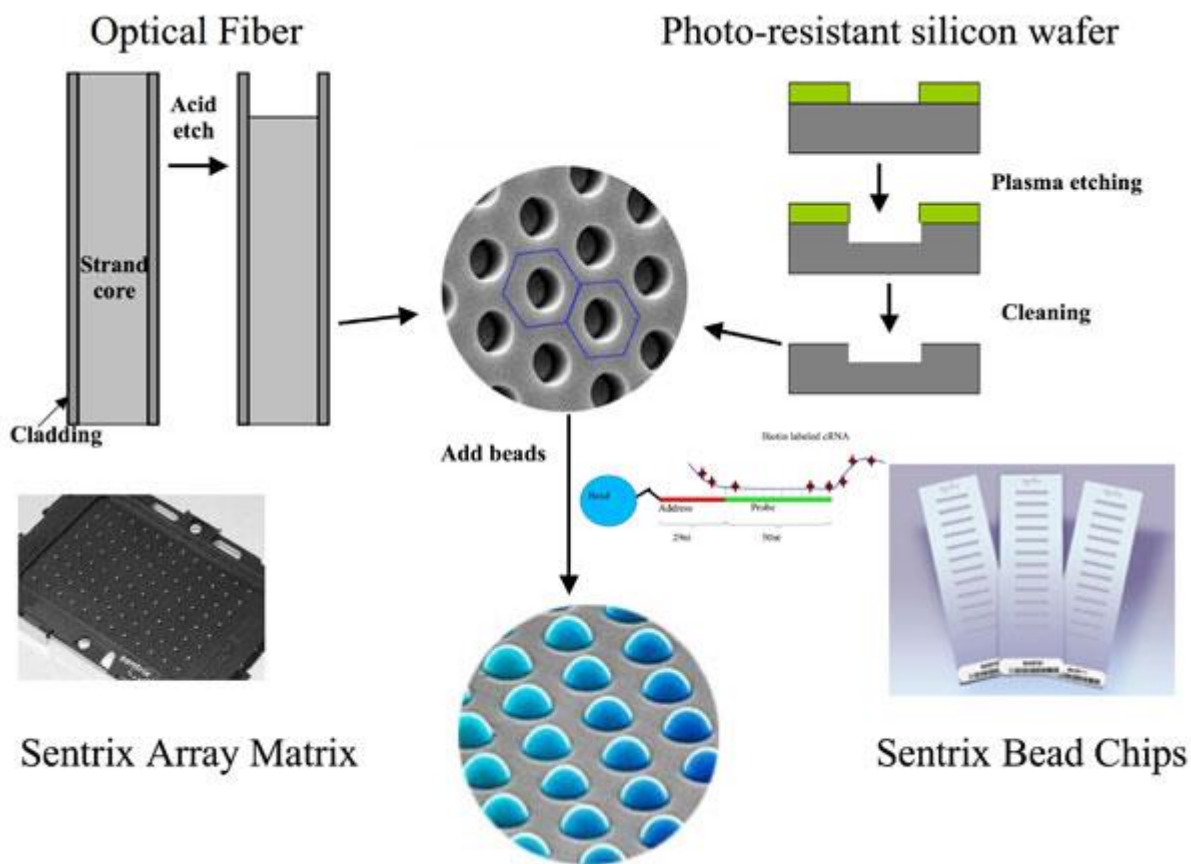
Το δεύτερο μέρος που βρίσκεται στο 3' άκρο αποτελείται από τις υπόλοιπες 29 βάσεις. Ονομάζεται Address και είναι μια μοναδική ακολουθία που δημιουργείται από την Illumina ώστε να επιτρέπει την μοναδική ταυτοποίηση του ολιγονουκλεοτιδίου αφού έχει αποθεθεί στη συστοιχία. Οι αλληλουχίες του Address σχεδιάζονται βιοπληροφορικά έτσι ώστε να μην έχουν κάποια ομολογία με γενωμικές ακολουθίες γονιδίων για τον υπό μελέτη οργανισμό. Ονομάζονται IllumiCodes και αποτελούν το βασικό εργαλείο για την αποκωδικοποίηση της συστοιχίας.



Εικόνα 17: Τύποι μικροσυστοιχιών Illumina.

Στο εμπόριο είναι διαθέσιμα δύο είδη μικροσυστοιχιών Illumina η SentrixArrayMatrix (SAM) και η SentrixBeadChip. Η SAM(Εικόνα 17,a) είναι μια διάταξη 8x12 θέσεων που αποτελείται από 96 οπτικές ίνες δέσμης με περίπου 50.000 φωτοαγωγίμες σειρές ανά δέσμη. Έπεται ότι στη SAM μπορούμε να τρέξουμε 96 διαφορετικά δείγματα ταυτόχρονα (Shen et al., 2005). Η SentrixBeadChip (Εικόνα 17,b) από την άλλη έχει διαστάσεις 2.5 x 8.25 cm σε υπόστρωμα πυριτίου. Ο αριθμός των δειγμάτων που μπορούμε να επεξεργαστούμε είναι 1 έως 16 δείγματα τη φορά και διαμορφώνεται ανάλογα το πείραμα. Στη παρούσα εργασία σε μια από τις σειρές δεδομένων υπό επεξεργασία γίνεται χρήση της Human-6 BeadChip πλατφόρμας, η οποία κάνει ανάλυση έκφρασης γονιδίων σε όλο το γονιδίωμα με τη δυνατότητα χρήσης 6 δειγμάτων τη φορά (Shen et al., 2005).

Και στους δύο προαναφερθέντες τύπους μετά την τυχαία εναπόθεση των σφαιριδίων στα εκάστοτε χαραγμένα υποστρώματα αυτά τοποθετούνται σε ομοιόμορφη απόσταση περίπου 5.7 μm και «πακετάρονται» με πυκνότητα που φτάνει τα 40.000 στοιχεία ανά τετραγωνικό χιλιοστό. Αυτό επιτρέπει στις μικροσυστοιχίες τύπου Illumina να περιέχουν περίπου 400 φορές περισσότερη πληροφορία από τις τυπωμένες συστοιχίες. Κάθε σφαιρίδιο καλύπτεται από εκατοντάδες χιλιάδες αντίγραφα ενός συγκεκριμένου ολιγονουκλεοτιδίου που λειτουργεί σαν την αλληλουχία σύλληψης σε κάθε πείραμα αντίστοιχα. Κάθε σφαιρίδιο έτσι έχει μια συγκεκριμένη αλληλουχία αλλά αυτή εμφανίζεται περίπου 30 φορές σε διάφορες θέσεις της συστοιχίας.



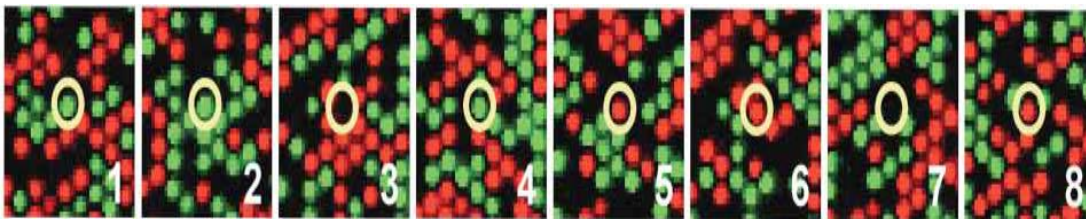
Εικόνα 18: Οι δύο τύποι μικροσυστοιχιών της Illumina. Αριστερά η SAM και δεξιά η Sentrix BeadChip.

Ένα πείραμα με χρήση μικροσυστοιχίας Illumina περιλαμβάνει τα εξής βήματα (Εικόνα 14):

- Διατύπωση βιολογικού ερωτήματος.
- Εκχύλιση ολικού mRNA.
- Υβριδισμός ενός ολιγονουκλεοτιδίου oligo-dT προς το κλάσμα polyA του ολικού RNA. Το ολιγονουκλεοτίδιο φέρει επίσης την ακολουθία ενός υποκινητή της ιικής πολυμεράσης RNA T7.

- Επέκταση του cDNA και στη συνέχεια, σύνθεση της συμπληρωματικής αλυσίδας για την παραγωγή δίκλωνου cDNA.
- Πρόσθεση πολυμεράσης RNA T7 και νουκλεοτιδίων για να ενισχυθεί γραμμικά το RNA. Το νεοπαραγόμενο cRNA ενσωματώνει UTP σεσημασμένο με βιοτίνη.
- Αντίστροφη μεταγραφή σε cDNA και σήμανση με μια φθορίζουσα ουσία
- Επίστρωση στο πλακίδιο και υβριδοποίηση
- Οπτική σάρωση με laser και μέτρηση των εντάσεων

Ένα τελευταίο σημείο που θα πρέπει να αναφέρουμε όσον αφορά τις μικροσυστοιχίες Illumina είναι αυτό της αποκωδικοποίησης της συστοιχίας (Εικόνα 19)



Εικόνα 19: Αποκωδικοποίηση μικροσυστοιχιών Illumina.

Μόλις τα σφαιρίδια αποτεθούν στην επιφάνεια του πλακιδίου, είναι απαραίτητο να ταυτοποιηθεί ποιος ανιχνευτής μεταγράφου είναι σε ποιο φρεάτιο. Αυτό γίνεται χρησιμοποιώντας το τμήμα διεύθυνσης του ολιγονουκλεοτιδίου. Η αποκωδικοποίηση των μικροσυστοιχιών περιλαμβάνει διαδοχική υβριδοποίηση διαφορετικά σεσημασμένων ανιχνευτών. Η διαφορεική σήμανση χρησιμοποιεί τρεις καταστάσεις: Σήμανση με καρβοξυφλουορεσκίνη (FAM) (πράσινο), κυανίνη 3 (Cy3) (κόκκινο), και μη σήμανση (κενό). Κατά τη διάρκεια κάθε κύκλου της διαδικασίας, ένα σφαιρίδιο είναι πράσινο, κόκκινο, ή κενό. Τα αποκωδικοποιώντας σεσημασμένα ολιγονουκλεοτίδια υβριδοποιούνται στις μικροσυστοιχίες σε υψηλές συγκεντρώσεις, το οποίο επιτρέπει ταχεία υβριδοποίηση, ακολουθούμενη από έκπλυση για την απομάκρυνση μη-ειδικού σήματος και του υποβάθρου. Αν ένας αριθμός αντιστοιχεί σε κάθε κατάσταση 0 στο κενό, 1 πράσινο, και 2 στο κόκκινο, τότε κάθε κύκλος της διαδικασίας δημιουργεί τριαδικό ψηφίο. Αν κοιτάξουμε ένα υποθετικό ανιχνευτή, τότε ο πρώτος γύρος μπορεί να είναι κόκκινο (πρώτο ψηφίο 2). Ο δεύτερος γύρος είναι πράσινο (ψηφίο 1), έτσι ώστε ο αριθμός είναι τώρα 21. Ο τρίτος γύρος κόκκινο, έτσι ώστε ο αριθμός είναι τώρα 212. Κάθε γύρος προσθέτει απλά ένα νέο ψηφίο στον αριθμό. Αυτό συνεχίζεται μέχρις ότου υπάρξουν επαρκή ψηφία για τον μοναδικό προσδιορισμό κάθε ανιχνευτή. Με ένα ψηφίο μπορεί να ταυτοποιήσουμε τρεις ανιχνευτές (0, 1, ή 2), με δύο ψηφία μπορούμε να προσδιορίσουμε 9 ανιχνευτές (00, 01, 02, 10, 11, 12, 21, 22, 23), με ένα τριψήφιο μπορούμε να εντοπίσουμε 27, κοκ. Η επανυβριδοποίηση συνεχίζεται έως ότου υπάρξουν επαρκή δεδομένα για να προσδιοριστεί σαφώς η ταυτότητα κάθε σφαιριδίου.

Ενδεικτικά παραθέτουμε και ένα πίνακα που συγκρίνει τα δύο είδη τεχνολογίας Illumina που χρησιμοποιήθηκαν στη παρούσα εργασία :

HumanHT-12 v4 Expression BeadChip Kit	<ul style="list-style-type: none"> • HumanHT-12 v4 BeadChips • Wash E1 BC Buffer (20 ml (4 × 5 ml)) • BeadChip Tweezer (1) • Wash Trays (14) • Wash Tray Lids (7) • High-Temperature Wash Buffer (325 ml) • Blocking E1 Buffer (40 ml) • Hybridization E1 Buffer (1.7 ml) • Humidity Control Buffer (2.8 ml) 	<ul style="list-style-type: none"> • BD-103-0204 <ul style="list-style-type: none"> — 24 Samples — 2 BeadChips • BD-103-0604 <ul style="list-style-type: none"> — 72 Samples — 6 BeadChips
HumanRef-8 v3.0 Expression BeadChip Kit	<ul style="list-style-type: none"> • HumanRef-8 v3.0 BeadChips • 8 microarrays per BeadChip • Wash E1 BC Buffer (20 ml (4 × 5 ml)) • BeadChip Tweezer (1) • Wash Trays (14) • Wash Tray Lids (7) • High-Temperature Wash Buffer (325 ml) • Blocking E1 Buffer (40 ml) • Hybridization E1 Buffer (1.7 ml) • Humidity Control Buffer (2.8 ml) 	<ul style="list-style-type: none"> • BD-102-0203 <ul style="list-style-type: none"> — 16 Samples — 2 BeadChips • BD-102-0603 <ul style="list-style-type: none"> — 48 Samples — 6 BeadChips

Πίνακας 1: Συγκριτικά χαρακτηριστικά των beadchip HumanHT-12 v4/HumanRef-8 v3.0.

Τεχνολογία επιτόπιας σύνθεσης μικροσυστοιχιών.

Η δημιουργία αυτού του είδους μικροσυστοιχιών πραγματοποιείται με την επιτόπια σύνθεσή τους πάνω σε ένα στερεό φορέα. Η σύνθεση αυτή δημιουργείται μέσω ομοιοπολικού δεσμού μεταξύ του υδροξυλίου του 5' άνθρακα του σακχάρου του τελευταίου συνδεδεμένου νουκλεοτιδίου και της φωσφορικής ομάδας του επόμενου νουκλεοτιδίου. Σε κάθε νουκλεοτίδιο που προστίθεται έχει προσδεθεί μία προστατευτική ομάδα διαδικασία γνωστή ως προστασία στο 5' άκρο του, για την αποφυγή σύνδεσης περισσότερων της μίας βάσης σε κάθε κύκλο σύνθεσης. Εν συνεχεία και προτού ξεκινήσει ο επόμενος κύκλος σύνθεσης, η προστατευτική ομάδα μετατρέπεται σε ομάδα υδροξυλίου διαδικασία γνωστή ως αποπροστασία είτε μέσω κάποιου οξέως (χημική αποπροστασία), είτε μέσω φωτός (φωτολιθογραφία). Αναλόγως της μεθόδου αποπροστασίας που ακολουθείται, μπορούμε να διακρίνουμε 3 βασικές τεχνολογίες για τις επιτόπιας σύνθεσης μικροσυστοιχίες DNA:

(α) φωτολιθογραφία με χρήση масκών (τεχνολογία που εφαρμόζει η εταιρεία Affymetrix για την κατασκευή μικροσυστοιχιών DNA)(Yap, 2002),

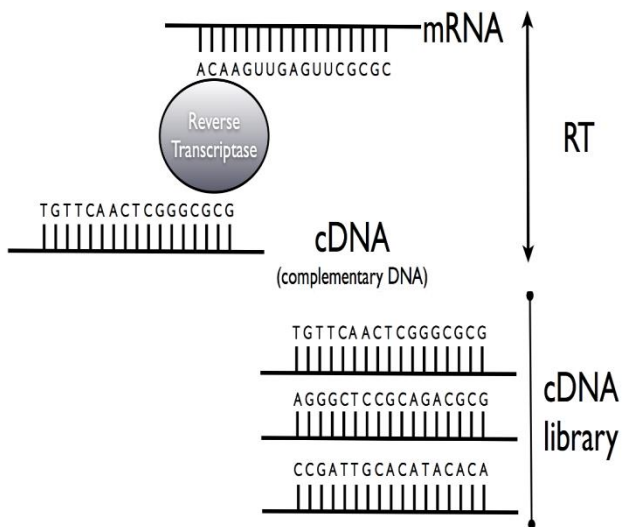
(β) φωτολιθογραφία χωρίς τη χρήση масκών (τεχνολογία που χρησιμοποιεί η εταιρεία Roche NimbleGen, Inc., USA),

(γ) χημική αποπροστασία με χρήση τεχνολογίας εκτύπωσης ψεκασμού (τεχνολογία που χρησιμοποιεί η εταιρεία Agilent, Inc., USA) (Stekel et al., 2005).

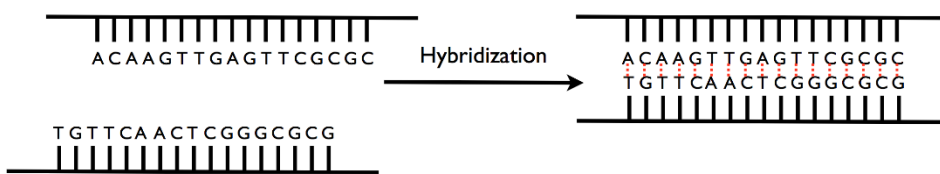
Στη παρούσα εργασία έγινε επίσης χρήση δεδομένων πλατφόρμας με τεχνολογία επιτόπιας σύνθεσης, της εταιρείας Affymetrix. Οι μικροσυστοιχίες Affymetrix αποτελούν το πιο διαδεδομένο τύπο συστοιχίας που χρησιμοποιείται σήμερα (Lipshutz et al., 1999).

Η διαδικασία κατασκευής της συστοιχίας περιλαμβάνει τα εξής βήματα (Yap, 2002):

Απομόνωση του ολικού mRNA και έπειτα με την βοήθεια του ενζύμου αντίστροφη μεταγραφάση (διαδικασία γνωστή από τη δράση των ρετροϊών) παίρνουμε το DNA. Το συμπληρωματικό DNA (Εικόνα 20) υβριδίζεται σε γνωστούς DNA στόχους (γονίδια) για να μπορέσουμε να υπολογίσουμε την έκφραση τους



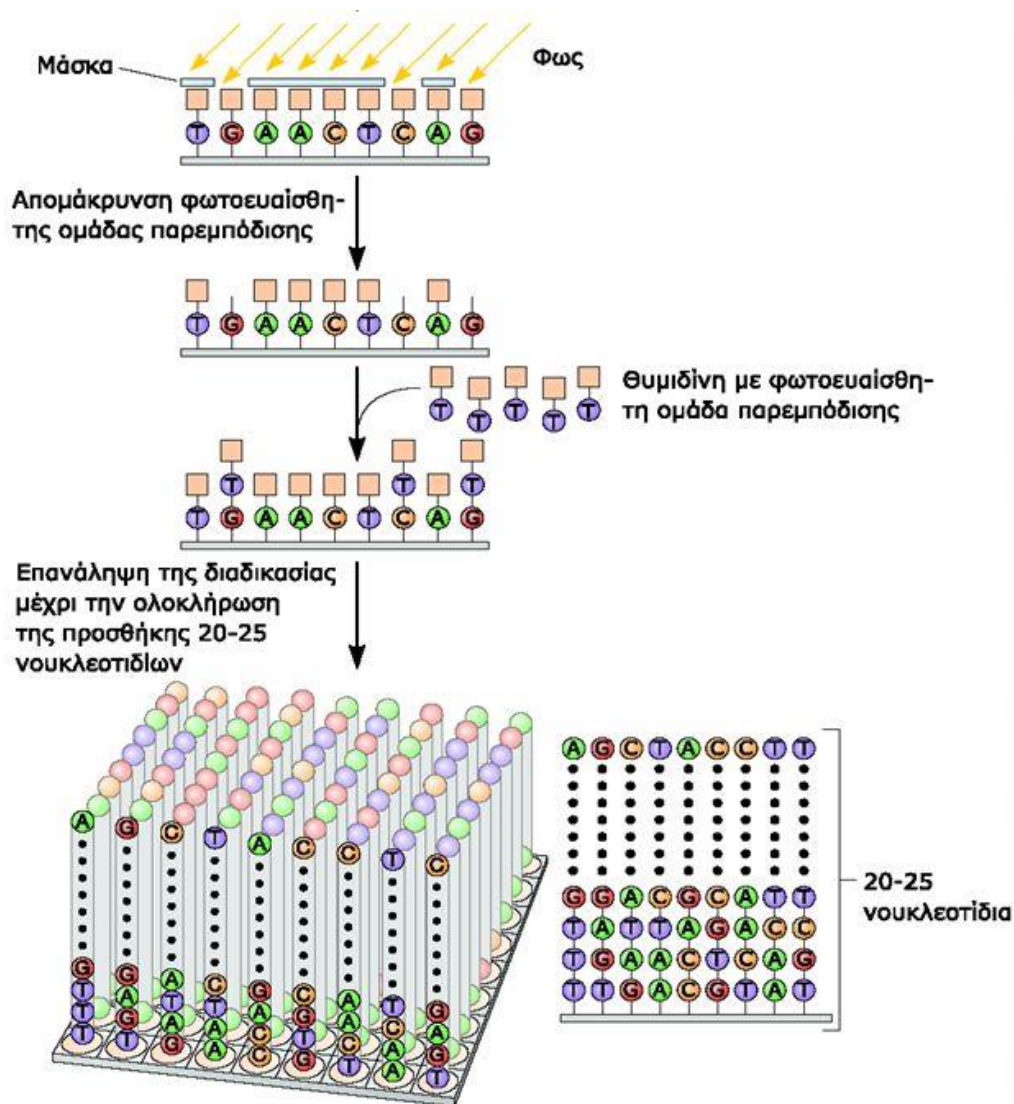
Εικόνα 20: Αντίστροφη μεταγραφή και δημιουργία cDNA βιβλιοθήκης.



Εικόνα 21: Διαδικασία υβριδισμού.

Οι μικροσυστοιχίες GeneChip® της εταιρείας Affymetrix αποτελούνται από μονόκλινα 25μερή ολιγονουκλεοτίδια-ανιχνευτές, τα οποία συντίθενται στη στερεή επιφάνεια της μικροσυστοιχίας με τη μέθοδο της φωτολιθογραφίας (Εικόνα 22)(Yap, 2002). Η διαδικασία ξεκινάει με το γυάλινο πλακίδιο (wafer) που αποτελεί την στερεή επιφάνεια της μικροσυστοιχίας. Το πλακίδιο εμβαπτίζεται σε σιλάνιο (SiH₄) και τα μόρια του σιλανίου συνδυάζονται με το γυαλί. Ένα μόριο συνδέτης μαζί με ένα φωτοευαίσθητο μόριο προστίθενται σε κάθε μόριο σιλανίου και το μόριο συνδέτης είναι το σημείο έναρξης πρόσδεσης του πρώτου

δεσοξυνουκλεοτιδίου. Κάθε νουκλεοτίδιο είναι φωτοχημικά τροποποιημένο αφού φέρει μια προστατευτική ομάδα η οποία απομακρύνεται μετά από επίδραση υπεριώδους ακτινοβολίας και αποτελεί το υπόστρωμα, όπου θα προσδεθεί το επόμενο νουκλεοτίδιο. Η διαδικασία αυτή επαναλαμβάνεται μέχρι να σχηματιστούν συγκεκριμένα 25μερή ολιγονουκλεοτίδια σε κάθε ανιχνευτικό σημείο (Lipshutz et al., 1999).

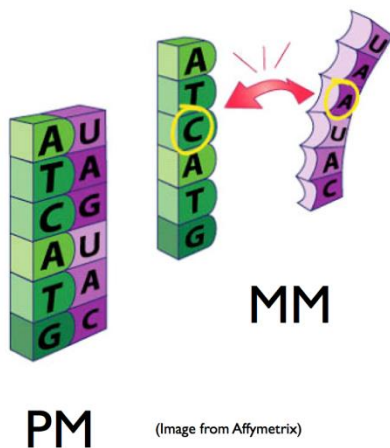


Εικόνα 22: Φωτολιθογραφία με χρήση μασκών (τεχνολογία που εφαρμόζει η εταιρεία Affymetrix για την κατασκευή μικροσυστοιχιών DNA)

Η διαδικασία της φωτολιθογραφίας αφορά υπεριώδη ακτινοβολία που διέρχεται μέσω της λιθογραφικής μάσκας που δρα ως φίλτρο είτε για να μεταδώσει είτε να μπλοκάρει την ακτινοβολία από τη χημικά προστατευμένη επιφάνεια της μικροσυστοιχίας. Η διαδοχική εφαρμογή των ειδικών λιθογραφικών μασκών προσδιορίζει τη σειρά της σύνθεσης του ολιγονουκλεοτιδικού ανιχνευτή. Η υπεριώδης ακτινοβολία αφαιρεί

τις προστατευτικές ομάδες, επιτρέποντας την προσθήκη ενός μόνο φωτοχημικά προστατευμένου νουκλεοτιδίου. Διαδοχικοί κύκλοι αποπροστασίας με ακτινοβολία, αλλαγής στο μοτίβο φιλτραρίσματος των λιθογραφικών μασκών, και προσθήκη ενός είδους μόνο νουκλεοτιδίων σχηματίζουν μικροσυστοιχίες με συγκεκριμένα 25μερή ολιγονουκλεοτίδια ανιχνευτές.

Κάθε γονίδιο ή νουκλεοτιδική αλληλουχία προς εξέταση αντιπροσωπεύεται από 11 έως 20 μοναδικούς ανιχνευτές που δημιουργούνται μέσω υπολογιστή και είναι διάσπαρτοι στη μικροσυστοιχία, ώστε να αποφευχθεί η λανθασμένη εκτίμηση της ποσοτικοποίησης της έκφρασης λόγω της θέσης τους. Οι ανιχνευτές χρησιμεύουν ως ευαίσθητοι, μοναδικοί, ειδικών αλληλουχιών αισθητήρες. Συνήθως, οι ανιχνευτές υβριδοποιούνται σε ανεξάρτητες περιοχές της αλληλουχίας, όμως ορισμένες φορές μπορεί να έχουν μικρή αλληλοεπικάλυψη (εφόσον αυτό κρίνεται απαραίτητο). Η ομάδα των ανιχνευτών που αφορά συγκεκριμένο γονίδιο ή ομάδα παρόμοιων γονιδίων, είναι γνωστή σαν σύνολο ανιχνευτών (probe set) το οποίο παρέχει, με υψηλή ακρίβεια, τον υπολογισμό της έκφραση του γονιδίου στόχου. Οι ολιγονουκλεοτιδικοί ανιχνευτές αναγνωρίζουν τμήματα του 3' άκρου του γονιδίου και καλούνται τέλειου ταιριάσματος (Perfect Match ή PM). Η μεγάλη ποσότητα των ανιχνευτών για διαφορετικές περιοχές του ίδιου RNA βελτιώνει σημαντικά το λόγο της έντασης σήματος ως προς το θόρυβο (λόγω του υπολογισμού της μέσης τιμής των εντάσεων των πολλαπλών ανιχνευτικών σημείων), παρέχει ακρίβεια στην ποσοτικοποίηση του RNA, ενώ αποτρέπει φαινόμενα διασταυρούμενης υβριδοποίησης (cross hybridization effect) και μειώνει δραστικά τα ψευδώς θετικά σήματα (Lipshutz et al., 1999). Επιπλέον έλεγχος γίνεται με την χρήση των ανιχνευτών ατελούς ταιριάσματος (Mismatch ή MM). Οι ανιχνευτές MM έχουν ακριβώς την ίδια νουκλεοτιδική ακολουθία με τα αντίστοιχα PM με τη διαφορά ότι στη 13η βάση υπάρχει η συμπληρωματική βάση (Εικόνα 23). Οι ανιχνευτές MM ενεργούν ως εξειδικευμένοι έλεγχοι που επιτρέπουν την άμεση αφαίρεση του θορύβου υποβάθρου και διασταυρούμενης υβριδοποίησης, και επιτρέπουν τη διάκριση μεταξύ πραγματικών σημάτων και εκείνων που οφείλονται σε μη ειδικό ή μερικό υβριδισμό. Η υβριδοποίηση των μορίων στόχων RNA στα PM παρέχει υψηλότερο σήμα από αυτό στα MM, αποδίδοντας ένα πρότυπο το οποίο δε γίνεται τυχαία. Ακόμη και σε χαμηλές συγκεντρώσεις RNA, η υβριδοποίηση στα PM/MM αποδίδει αναγνωρίσιμα πρότυπα που μπορούν να ποσοτικοποιηθούν. Κάθε ανιχνευτής MM βρίσκεται στη διπλανή θέση από αυτή του αντίστοιχου PM για να αποκλεισθεί οποιαδήποτε επίδραση λόγω θέσης. Για τους ευκαρυωτικούς οργανισμούς, οι ανιχνευτές επιλέγονται τυπικά από το 3' άκρο του γονιδίου ή μεταγράφου (κοντά στην ουρά polyA), για να μειωθούν τα προβλήματα που μπορεί να προκύψουν από τη χρήση του μερικώς αποδομημένου mRNA. Η χρήση της διαφοράς των PM από τα MM μειώνει σημαντικά το θόρυβο υποβάθρου και της διασταυρούμενης υβριδοποίησης, και αυξάνει την ποσοτική ακρίβεια και επαναληψιμότητα των μετρήσεων.

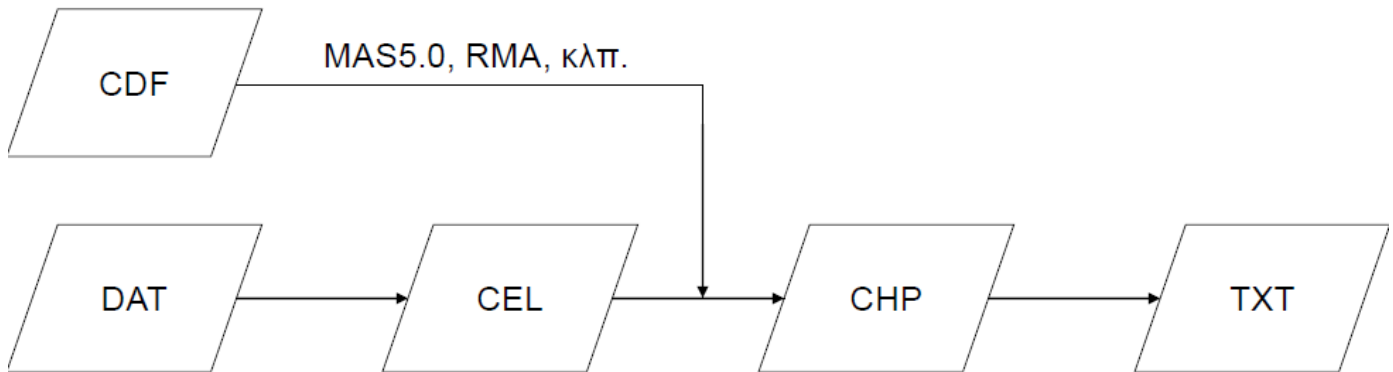


Εικόνα 23: Ανιχνευτές PM/MM.

Το λογισμικό της Affymetrix για την επεξεργασία των δεδομένων από τις μικροσυστοιχίες GeneChip χρησιμοποιεί πολλά αρχεία δεδομένων με διάφορες μορφές (Εικόνα 24) (Cao et al., 2004; Irizarry et al., 2003). Τα πιο κοινά είναι:

- Αρχείο DAT: Περιέχει τις τιμές έντασης των εικονοστοιχείων που συλλέγονται από ένα σαρωτή Affymetrix.
- Αρχείο CDF (Chip Description File): Περιγράφει τη διάταξη των ανιχνευτών σε μια μικροσυστοιχία Affymetrix GeneChip. Μια μικροσυστοιχία μπορεί να περιέχει εκφράσεις, γονοτυποποίηση και μαρκαρισμένα σύνολα ανιχνευτών. Όλα τα ονόματα των συνόλων ανιχνευτών εντός μιας μικροσυστοιχίας είναι μοναδικά. Πολλαπλά αντίγραφα ενός συνόλου ανιχνευτών μπορεί να υπάρχουν σε μία μικροσυστοιχία εφόσον κάθε αντίγραφο έχει ένα μοναδικό όνομα.
- Αρχείο CEL: Αποθηκεύει τα αποτελέσματα των υπολογισμών της έντασης σχετικά με τις τιμές των εικονοστοιχείων του αρχείου DAT. Αυτά περιλαμβάνουν την τιμή της έντασης, την τυπική απόκλιση της έντασης και τον αριθμό των εικονοστοιχείων που χρησιμοποιούνται για τον υπολογισμό της έντασης. Το αρχείο αποθηκεύει τα δεδομένα που αναφέρθηκαν προηγουμένως για κάθε ανιχνευτή της μικροσυστοιχίας.
- Αρχείο CHP: Περιέχει τα αποτελέσματα των αναλύσεων των συνόλων ανιχνευτών που δημιουργούνται από το λογισμικό Affymetrix. Υπάρχουν διάφορες εκδόσεις του αρχείου CHP που δημιουργούνται από τον MAS5, το GCOS και άλλα λογισμικά της Affymetrix. Έχει δυαδική μορφή για χρήση σε διάφορα προγράμματα ανάλυσης των αποτελεσμάτων.
- Αρχείο TXT: Είναι το αρχείο CHP σε αναγνώσιμη μορφή (κείμενο).

Αρχεία Affymetrix



- DAT: Μη επεξεργασμένη οπτική εικόνα του υβριδοποιημένου πλακιδίου
- CDF: Διάταξη του πλακιδίου (από την Affymetrix ή άλλες πηγές)
- CEL: Επεξεργασμένο αρχείο DAT (τιμές έντασης-θέσης)
- CHP: Πειραματικά αποτελέσματα από το συνδυασμό των αρχείων CEL και CDF
- TXT: Τιμές έκφρασης κάθε συνόλου ιχνηθέτη (αρχείο CHP σε διαμόρφωση αρχείου)

Εικόνα 24: Είδη αρχείων Affymetrix.

Αρχικά παράγεται το αρχείο DAT από τον σαρωτή της Affymetrix όπου στη συνέχεια με το κατάλληλο λογισμικό δημιουργείται το αρχείο CEL. Στη συνέχεια χρησιμοποιώντας το αρχείο CDF στον αλγόριθμο προεπεξεργασίας της επιλογής μας (MAS5, RMA, κλπ) παράγεται το αρχείο CHP όπου είναι σε δυαδική μορφή και στην συνέχεια μπορεί να μετατραπεί στο αρχείο κειμένου TXT (Yap, 2002).

ΒΙΟΛΟΓΙΚΕΣ ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ

Κατά την αναζήτηση δεδομένων προς επεξεργασία για τη παρούσα εργασία ανατρέξαμε σε διάφορες βιολογικές βάσεις δεδομένων. Οι βιολογικές βάσεις δεδομένων, αποτελούν βασικό κομμάτι της βιοπληροφορικής. γενικά, μπορούν να διακριθούν σε 2 μεγάλες κατηγορίες:

Πρωτογενείς βάσεις δεδομένων, οι οποίες περιέχουν τα πρωτογενή δεδομένα όπως αυτά προσδιορίζονται από τους πειραματικούς ερευνητές, και αναλύονται κυρίως σε βάσεις δεδομένων ακολουθιών νουκλεοτιδικών ακολουθιών, βάσεις δεδομένων ακολουθιών πρωτεϊνικών ακολουθιών, βάσεις δεδομένων τρισδιάστατων βιολογικών δομών, βάσεις δεδομένων γονιδιακής έκφρασης, βάσεις δεδομένων γενετικής ποικιλομορφία και βάσεις δεδομένων βιβλιογραφίας. Παραδείγματα τέτοιων βάσεων αποτελούν η GENBANK, η EMBL-Bank για δεδομένα νουκλεοτιδικών ακολουθιών, η UniprotKB για δεδομένα πρωτεϊνικών ακολουθιών, η ProteinData Bank για δεδομένα τρισδιάστατων βιολογικών δομών, η GeneExpression Omnibus (GEO) και η Array Express για δεδομένα γονιδιακής έκφρασης, η dbSNP, η HarMap για δεδομένα γενετικής ποικιλομορφίας και τέλος η PubMed, το SCOPUS, το Web of Science για δεδομένα βιβλιογραφίας.

Δευτερογενείς βάσεις δεδομένων, στις οποίες υπάρχουν κυρίως ταξινομήσεις των πρωτογενών δεδομένων, χρήσιμες για αναλυτικούς σκοπούς, και διακρίνονται περαιτέρω σε βάσεις δεδομένων οικογενειών κυρίως πρωτεϊνών και εξειδικευμένες βάσεις δεδομένων όπως η PROSITE, η PFAM, η CATH, η TCDB (σχεσιακή βάση) και πολλές άλλες.

Όσον αφορά τα δικά μας δεδομένα πρόκειται για δεδομένα γονιδιακής έκφρασης συνεπώς αναζητήσαμε δεδομένα στο GEO και στο ArrayExpress.

G.E.O (GENE EXPRESSION OMNIBUS)

Προκειμένου να στηριχθεί η δημόσια χρήση και η διασπορά των δεδομένων σχετικών με τη γονιδιακή έκφραση, το NCBI δημιούργησε το Gene Expression Omnibus ή αλλιώς GEO. Το GEO αντιπροσωπεύει την προσπάθεια να χτιστεί ένα αποθετήριο δεδομένων και μια διαδικτυακή πηγή πληροφοριών για την αποθήκευση και την επανάκτηση δεδομένων σχετικά με τη γονιδιακή έκφραση κάθε οργανισμού. Πολλοί τύποι δεδομένων μικροσυστοιχιών αρχειοθετήθηκαν και συνεχίζουν ακόμη μέχρι σήμερα να προστίθενται καθημερινά καθώς και τύποι δεδομένων από αλληλούχιση (next generation sequencing). Είναι διαθέσιμη στην ιστοσελίδα <http://www.ncbi.nlm.nih.gov/geo/>, ενώ στην ίδια διεύθυνση υπάρχουν διαθέσιμα και κάποια διαδικτυακά εργαλεία που επιτρέπουν απλές αναλύσεις των δεδομένων της βάσης. Τα δεδομένα υπάρχουν τόσο σε ακατέργαστη (raw data) όσο και σε επεξεργασμένη μορφή (με κανονικοποιήσεις κ.ο.κ.). Η μορφοποίηση της βάσης είναι ως εξής: στο βασικό επίπεδο οργάνωσης του GEO, υπάρχουν τέσσερις βασικοί τύποι οντότητας. Οι πρώτοι τρεις (Δείγμα, Πλατφόρμα, και Σειρά) συμπληρώνονται και παρέχονται στο GEO από τους χρήστες. Το προσωπικό του GEO καταρτίζει και επιμελείται τον τέταρτο τύπο, το Σύνολο Δεδομένων, χρησιμοποιώντας τα δεδομένα που έχουν υποβληθεί από τους χρήστες. Ειδικότερα η βάση περιέχει δεδομένα από διαφορετικές πλατφόρμες έκφρασης (GPL), προερχόμενα από διάφορα "δείγματα" (GSM), δηλαδή άτομα (στα οποία όμως δεν περιέχονται μόνο άνθρωποι, μπορεί να υπάρχουν δεδομένα από ζώα, φυτά ή ακόμα και μικροοργανισμούς), ταξινομημένα σε "σειρές" (GSE) και "σύνολα δεδομένων" (GDS). Το ίδιο δείγμα μπορεί να περιέχεται σε διαφορετικές σειρές και η ίδια σειρά σε ένα ή περισσότερα σύνολα δεδομένων.

ARRAYEXPRESS

Η ArrayExpress είναι δημόσια βάση δεδομένων μικροσυστοιχιών η οποία διατηρείται στο Ευρωπαϊκό Ινστιτούτο Βιοπληροφορικής, EBI. Είναι διαθέσιμη στην ιστοσελίδα <http://www.ebi.ac.uk/arrayexpress/>. Περιέχει ως υποσύνολο της την GEO μετά από συνεργασία των ιδρυμάτων. Στην ιστοσελίδα υπάρχουν επίσης διαθέσιμα εργαλεία για ανάλυση.

MIAME

Κάθε πείραμα μικροσυστοιχιών πριν κατατεθεί στο GEO πρέπει να υπόκειται σε κάποιες προϋποθέσεις. Αυτές είναι γνωστές ως MIAME (Minimum Information About a Microarray Experiment)(Brazma, 2009) και περιλαμβάνουν συνοπτικά τα εξής:

- Τα πρωτογενή δεδομένα (raw data) (π.χ. CEL ή GPR αρχεία).
- Τα τελικώς επεξεργασμένα δεδομένα (normalized data) για κάθε σετ υβριδισμών στο πείραμα (μελέτη) (π.χ., τον πίνακα με τα δεδομένα γονιδιακής έκφρασης).
- Αρχείο σχολιασμού της πειραματικής σχεδίασης και των επιλεγμένων παραμέτρων.
- Τον πειραματικό σχεδιασμό περιλαμβάνοντας τις σχέσεις μεταξύ των δεδομένων των δειγμάτων.
- Επαρκή σχολιασμό της συστοιχίας (π.χ., αναγνωριστικά γονιδίου, συντεταγμένες γονιδιωματικής, αλληλουχίες ολιγονουκλεοτιδικού ανιχνευτή, συστοιχίες ή αριθμός αναφοράς εμπορικού καταλόγου της ακολουθίας).
- Τα απαραίτητα εργαστηριακά πρωτόκολλα καθώς και τα πρωτόκολλα επεξεργασίας δεδομένων (π.χ., τρόποι κανονικοποίησης)(Brazma, 2009).

Έχοντας συγκεντρώσει όλα τα δεδομένα κανονικοποιημένα θα χρειαστεί να τα να τα οργανώσουμε σε μια βάση δεδομένων για την περαιτέρω επεξεργασία της μετα-ανάλυσης προκειμένου να διαπιστώσουμε τη διαφορική τους έκφραση μετά από έκθεση σε ιοντίζουσα ακτινοβολία. Υπάρχουν τέσσερα μοντέλα βάσεων δεδομένων: Το ιεραρχικό, το δικτυακό, το σχεσιακό και το αντικειμενοστραφές. Σήμερα στις περισσότερες βάσεις κυριαρχεί το σχεσιακό μοντέλο με ιδιαίτερα απλή και κατανοητή δομή ακόμα και σε εφαρμογές υψηλών απαιτήσεων το οποίο και επιλέξαμε για τη δημιουργία της δικής μας βάσης δεδομένων. Οι βασικοί στόχοι του σχεσιακού μοντέλου είναι οι εξής: η υποστήριξη της ανεξαρτησίας των δεδομένων, έτσι ώστε αλλαγές στη φυσική δομή και οργάνωση της βάσης δεδομένων να μην απαιτούν αλλαγές στα προγράμματα εφαρμογής, η αποφυγή του πλεονασμού, που εμφανίζεται όταν τα ίδια δεδομένα αποθηκεύονται πολλές φορές σε διαφορετικές περιοχές της βάσης δεδομένων, η διατήρηση της ακεραιότητας και της συνέπειας των δεδομένων, η υποστήριξη της ανάπτυξης γλωσσών χειρισμού δεδομένων, οι οποίες διευκολύνουν τη διατύπωση ερωτημάτων προς το Σύστημα Διαχείρισης Βάσεων Δεδομένων.

Στις σχεσιακές βάσεις περιλαμβάνονται βάσεις δεδομένων με δυνατότητες ταυτόχρονου χειρισμού και σύνδεσης πολλών συλλογών από εγγραφές διαφορετικών τύπων, οργανωμένες σε πίνακες. Ο όρος «σχεσιακό» αναφέρεται σε μια συγκεκριμένη μέθοδο οργάνωσης των βάσεων δεδομένων, σύμφωνα με την οποία οι πίνακες της βάσης δεδομένων μπορούν να συσχετισθούν μεταξύ τους με αποτελέσματα οι πληροφορίες να κατανέμονται ομοιόμορφα σε όλη τη βάση. Για την αναπαράσταση δεδομένων, το σχεσιακό μοντέλο χρησιμοποιεί πίνακες. Ο κάθε πίνακας έχει ένα μοναδικό όνομα και προσδιορίζεται από ένα σύνολο γραμμών και ένα σύνολο στηλών. Οι γραμμές του πίνακα ορίζουν τα χαρακτηριστικά της κάθε εγγραφής. Κάθε στήλη του πίνακα αναπαριστά μια εγγραφή δεδομένων και ονομάζεται πλειάδα. Το πλήθος των χαρακτηριστικών της σχέσης καλείται βαθμός (degree), ενώ ο αριθμός των πλειάδων καλείται πληθικότητα. Για κάθε χαρακτηριστικό υπάρχει ένα σύνολο επιτρεπτών τιμών, το οποίο καλείται πεδίο ορισμού του χαρακτηριστικού. Οι τιμές που μπορεί να πάρει ένα χαρακτηριστικό προσδιορίζονται από το αντίστοιχο πεδίο ορισμού, ενώ επίσης είναι απαραίτητο να γνωρίζουμε τον τύπο δεδομένων (datatype) και τη μορφοποίηση (format). Οι Βασικότερες ιδιότητες των πινάκων είναι οι εξής:

- Κάθε πίνακας της βάσης δεδομένων έχει ένα μοναδικό όνομα.
- Η τιμή ενός χαρακτηριστικού για μία πλειάδα είναι ατομική.
- Το κάθε χαρακτηριστικό έχει μοναδικό όνομα μέσα στον πίνακα. Δύο χαρακτηριστικά που ανήκουν σε διαφορετικούς πίνακες επιτρέπεται να έχουν ίδιο όνομα.
- Όλες οι τιμές ενός χαρακτηριστικού πρέπει να ανήκουν στο πεδίο ορισμού του χαρακτηριστικού.

- Η σειρά δήλωσης των χαρακτηριστικών ενός πίνακα δεν παίζει κανένα ρόλο. • Δύο πλειάδες μίας σχέσης δεν επιτρέπεται να ταυτίζονται σε όλα τα χαρακτηριστικά.

- Στο σχεσιακό μοντέλο δεν μας ενδιαφέρει η σειρά των πλειάδων στον πίνακα. Ωστόσο, η σειρά αποθήκευσης των δεδομένων επηρεάζει το χρόνο επεξεργασίας και επομένως λαμβάνεται υπόψη.

Ένα χαρακτηριστικό που μπορεί να διακρίνει τις διάφορες στήλες του πίνακα, ονομάζεται πρωτεύον κλειδί. Για παράδειγμα το πρωτεύον κλειδί στον πίνακα Expression της Microarray (Εικόνα 45) είναι το FeatureID μια μεταβλητή τύπου VARCHAR που χρησιμοποιείται για να αποθηκεύσουμε String/Αλφαριθμητικά μεγέθους έως 45 χαρακτήρων. Όμοια ισχύει και για τις μεταβλητές GSM και Plattform καθώς και σε αυτές τις μεταβλητές οι καταγραφές που θα αποθηκευτούν είναι μοναδικές. Το Expression δεν μπορεί να θεωρηθεί πρωτεύον κλειδί, καθώς μπορεί να υπάρχουν περισσότερες από μια ίδιες τιμές για την έκφραση. Σε πολλές περιπτώσεις απαιτούνται περισσότερα από ένα χαρακτηριστικά για να συνθέσουν ένα κλειδί, οπότε το κλειδί καλείται σύνθετο. Τα υπόλοιπα χαρακτηριστικά του πίνακα λέγονται δευτερεύοντα κλειδιά. Η συνέπεια των δεδομένων μετά από εισαγωγές, διαγραφές και ενημερώσεις διατηρείται με τη χρήση περιορισμών ακεραιότητας. Οι σημαντικότεροι είναι: περιορισμός χρήσης κενών τιμών (υπάρχουν περιπτώσεις όπου δεν γνωρίζουμε την τιμή ενός χαρακτηριστικού ή δεν μπορούμε να την προσδιορίσουμε. Σε αυτές τις περιπτώσεις αποδίδουμε στο χαρακτηριστικό την κενή τιμή (NULL)), περιορισμοί ακεραιότητας οντοτήτων όπου κάθε στήλη πρέπει να προσδιορίζεται μοναδικά τουλάχιστον από το πρωτεύον κλειδί του πίνακα, περιορισμοί αναφορών όπως πχ αν το κλειδί k ενός πίνακα A εμφανίζεται ως χαρακτηριστικό σε έναν άλλο πίνακα τότε λέγεται ότι το k αποτελεί ξένο κλειδί για τον δεύτερο πίνακα. Η ακεραιότητα αναφορών επιβάλλει η τιμή ενός ξένου κλειδιού να έχει αντίστοιχη τιμή στον αρχικό πίνακα.

Επιλέξαμε τις σχεσιακές βάσεις δεδομένων MySQL και δημιουργήσαμε τη βάση δεδομένων Microarray που φαίνεται παρακάτω (Εικόνα 45).

Η MySQL είναι ένα Σύστημα Διαχείρισης Σχεσιακών Βάσεων Δεδομένων που μετρά περισσότερες από 11 εκατομμύρια εγκαταστάσεις. Το πρόγραμμα τρέχει έναν εξυπηρετητή (server) παρέχοντας πρόσβαση πολλών χρηστών σε ένα σύνολο βάσεων δεδομένων. Ανήκει και χρηματοδοτείται από την Oracle. Η MySQL είναι δημοφιλής βάση δεδομένων για διαδικτυακά προγράμματα και ιστοσελίδες. Χρησιμοποιείται σε κάποιες από τις πιο διαδεδομένες διαδικτυακές υπηρεσίες, όπως Flickr, YouTube, Wikipedia, Google, Facebook, Twitter κλπ.

Τα δεδομένα μας εισήχθησαν στη βάση Microarray και διαμορφώθηκαν οι πίνακες με κάποια scripts σε PHP που θα αναλυθούν παρακάτω.

Η PHP είναι μια γλώσσα προγραμματισμού για τη δημιουργία ιστοσελίδων με δυναμικό περιεχόμενο. Μια σελίδα PHP περνά από επεξεργασία από ένα συμβατό εξυπηρετητή του Παγκόσμιου Ιστού (π.χ. Apache), ώστε να παραχθεί σε πραγματικό χρόνο το τελικό περιεχόμενο, που θα σταλεί στο πρόγραμμα περιήγησης

των επισκεπτών σε μορφή κώδικα HTML. Ένα αρχείο με κώδικα PHP θα πρέπει να έχει την κατάλληλη επέκταση (π.χ. *.php, *.php4, *.phtml κ.ά.). Όταν ένα αρχείο έχει την επέκταση .php, θα πρέπει ο εξυπηρετητής να είναι ρυθμισμένος για να επεξεργάζεται και να μεταγλωττίζει τον κώδικα PHP σε HTML που καταλαβαίνει το πρόγραμμα πλοήγησης του κάθε χρήστη. Ο εξυπηρετητής Apache, που χρησιμοποιείται σήμερα ευρέως σε συστήματα με λειτουργικά συστήματα Unix/Linux , Microsoft Windows και MacOS X υποστηρίζει εξ ορισμού την εκτέλεση κώδικα PHP, είτε με την χρήση ενός πρόσθετου (mod_php) ή με την αποστολή του κώδικα προς εκτέλεση σε εξωτερική διεργασία CGI.

Το XAMPP είναι ένα πακέτο προγραμμάτων ελεύθερου λογισμικού, λογισμικού ανοικτού κώδικα και ανεξαρτήτου πλατφόρμας το οποίο περιέχει το εξυπηρετητή ιστοσελίδων Apache, την βάση δεδομένων MySQL και ένα διερμηνέα για κώδικα γραμμένο σε γλώσσες προγραμματισμού PHP και Perl. Το XAMPP είναι ακρωνύμιο και αναφέρεται στα παρακάτω αρχικά:

- X (αναφέρεται στο «cross-platform» που σημαίνει λογισμικό ανεξάρτητο πλατφόρμας)
- ApacheHTTP εξυπηρετητής
- MySQL
- PHP
- Perl

Το XAMPP είναι ένα ελεύθερο λογισμικό το οποίο περιέχει ένα εξυπηρετητή ιστοσελίδων το οποίο μπορεί να εξυπηρετεί και δυναμικές ιστοσελίδες τεχνολογίας PHP/MySQL. Είναι ανεξάρτητο πλατφόρμας και τρέχει σε Microsoft Windows, Linux, Solaris, and MacOS X και χρησιμοποιείται ως πλατφόρμα για την σχεδίαση και ανάπτυξη ιστοσελίδων με την τεχνολογίες όπως PHP, JSP και Server. Το XAMPP υποστηρίζει την δημιουργία και διαχείριση βάσεων δεδομένων τύπου MySQL και File.

Όταν το XAMPP εγκατασταθεί στον τοπικό υπολογιστή διαχειρίζεται τον local host ως ένα απομακρυσμένο κόμβο, ο οποίος συνδέεται με το πρωτόκολλο μεταφοράς αρχείων FTP. Η σύνδεση στον localhost μέσω του FTP μπορεί να γίνει με το όνομα χρήστη «newuser» και το κωδικό «wampp». Για την βάση δεδομένων MySQL υπάρχει ο χρήστης «root» χωρίς κωδικό πρόσβασης.

Τα συμπεράσματα από πολλαπλές μελέτες πάνω στο ίδιο αντικείμενο ή θέμα θα πρέπει να θεωρούνται ως ένα σύνθετο σύνολο δεδομένων που απαιτεί την ίδια μορφή λεπτομερούς στατιστικής ανάλυσης, όπως θα απαιτούσε οποιαδήποτε μελέτη που αναφέρεται σε ένα μεγάλο πλήθος παρατηρήσεων. Δεδομένου ότι πειράματα μικροσυστοιχιών δεν είναι εύκολα επαναλήψιμα η μετα-ανάλυση απαντάται συχνά σε έρευνες συνδυασμών αποτελεσμάτων προηγούμενων πειραμάτων. Σαν βασική συνέπεια της διαδικασίας αυτής έχουμε την αύξηση της στατιστικής ισχύος της μελέτης μας αφού πλέον υπολογίζεται ένα μέγεθος το οποίο εκτιμά καλύτερα την έκφραση των γονιδίων. Στη παρούσα εργασία υλοποιήθηκε η μετα-ανάλυση των γονιδίων στο STATA (Boston and Sumner, 2003) από όπου υπολογίσαμε το αριθμό των στατιστικά σημαντικών γονιδίων με χρήση μεθόδων διόρθωσης για πολλαπλές συγκρίσεις. Ο έλεγχος της στατιστικής σημαντικότητας γονιδίων γίνεται με πληθώρα μεθόδων όπως ο έλεγχος t-test (ταυτόχρονη σύγκριση δύο συνόλων με τη σύγκριση των μέσων όρων τους) , έλεγχος με μέτρηση αλλαγής fold change (που ελέγχει την υπέρ/υπό έκφραση ενός γονιδίου σε σχέση με μια τιμή αναφοράς) κλπ.

t-test

Ως μέγεθος επίδρασης χρησιμοποιείται η τυποποιημένη διαφορά των μέσων τιμών έκφρασης των δειγμάτων ελέγχου και αναφοράς

$$d_i = \frac{\bar{x}_{1i} - \bar{x}_{2i}}{sd_i}$$

(Εξίσωση 3)

με τυπική απόκλιση

$$sd_i = \sqrt{\frac{(n_{1i} - 1)sd_{1i}^2 + (n_{2i} - 1)sd_{2i}^2}{n_{1i} + n_{2i} - 2}}$$

(Εξίσωση 4)

Από τους ελέγχους αυτούς προκύπτει ένα p-value που εκφράζει τη στατιστική σημαντικότητα του γονιδίου. Ανάλογα τη τιμή του p-value προκύπτει και η εμφάνιση λανθασμένων θετικά γονιδίων που επιβάλλει τη χρήση μεθόδων διόρθωσης του p-value που αναλύουμε παρακάτω.

Διόρθωση Bonferroni

Σε αυτή τη μέθοδο (Bonferroni, 1936) το p-value κάθε γονιδίου πολλαπλασιάζεται με το συνολικό αριθμό γονιδίων της μικροσυστοιχίας n και αν βρίσκεται κάτω από το όριο σφάλματος τότε θεωρείται στατιστικά σημαντικό:

$$p_{cor(i)} = p_{(i)} * n$$

(Εξίσωση 5)

Διόρθωση Bonferroni Step-Down (Holm)

Σε αυτή τη μέθοδο (Holm, 1979) το p-value ταξινομείται κατά αύξουσα σειρά και έπειτα το πρώτο πολλαπλασιάζεται με το συνολικό αριθμό γονιδίων της μικροσυστοιχίας n και η διόρθωση στα υπόλοιπα γίνεται με την (Εξίσωση 6)

$$p_{cor(i)} = (n - i) * p_{(i)}$$

(Εξίσωση 6)

Διόρθωση Benjamini & Hochberg false discovery rate (fdr)

Εδώ πρόκειται για λιγότερο αυστηρή μέθοδο με καλή ωστόσο εφαρμογή σε δεδομένα μικροσυστοιχιών και πάλι λοιπόν ταξινομούμε κατά αύξουσα σειρά τα p-value το μεγαλύτερο παραμένει το ίδιο και η διόρθωση στα υπόλοιπα γίνεται με το παρακάτω τύπο (Benjamini and Hochberg, 1995):

$$p_{cor(i)} = \frac{n}{n - i} * p_{(i)}$$

(Εξίσωση 7)

όπου i : δείκτης ταξινόμησης, n : συνολικός αριθμός γονιδίων μικροσυστοιχίας

Διόρθωση Holland

Η μέθοδος αυτή (Holland, 1987) αποτελεί παραλλαγή της μεθόδου Holm όπου μετά την ταξινόμηση το πρώτο p-value πολλαπλασιάζεται με το συνολικό αριθμό γονιδίων της μικροσυστοιχίας n-1 και η διόρθωση στα υπόλοιπα γίνεται με την:

$$P_{cor(i)} = (n - i + 1) * P_{(i)}$$

(Εξίσωση 8)

Διόρθωση Šidák

Η διόρθωση του p-value γίνεται με το παρακάτω τύπο (Šidák, 1967):

$$P_{cor(i)} = 1 - (1 - P_{(i)})^{\frac{1}{n}}$$

(Εξίσωση 9)

Μετά την βιβλιογραφική αναζήτηση (Πίνακας 2) των διαθέσιμων μελετών με δεδομένα ακτινοβόλησης <0.5Gy τη διαλογή τους και την αντίστοιχη προεπεξεργασία τους έχουμε έτοιμο το σύνολο των δεδομένων μας για την μετα-ανάλυση. Θα πρέπει στη συνέχεια να καθοριστεί το μέγεθος επίδρασης το οποίο μπορεί να είναι η διαφορά μέσων τιμών, ο συντελεστής συσχέτισης και το OR (Odds Ratio) ο οποίος αντιπροσωπεύει την πιθανότητα να συμβεί ένα γεγονός προς την πιθανότητα να μη συμβεί.

Η αποτελεσματικότητα της μετα-ανάλυσης εξαρτάται από τον αριθμό των μελετών και τη μέθοδο συνδυασμών των μεμονωμένων εκτιμήσεων του μεγέθους επίδρασης οι οποίες προέρχονται από τις αρχικές μελέτες. Έχουμε δύο μοντέλα συνδυασμών των παραπάνω εκτιμήσεων : το μοντέλο σταθερών επιδράσεων (fixed effect model) και το μοντέλο τυχαίων επιδράσεων (random effect model)

Το μοντέλο των σταθερών επιδράσεων υποθέτει ότι τα δείγματα των μελετών προέρχονται από ενιαίο πληθυσμό με κοινό μέγεθος επίδρασης. Αντίστοιχα το μοντέλο τυχαίων επιδράσεων υποθέτει ότι τα δείγματα των μελετών μπορούν να προέλθουν από μια κατανομή πληθυσμών. Κάθε μελέτη έχει ένα διαφορετικό μέγεθος επίδρασης θ_i και διακύμανση s_i^2 . Στην περίπτωση αυτή κάθε μελέτη έχει διαφορετικό μέγεθος επίδρασης και διακύμανση ενώ κάθε δείγμα έχει μέγεθος επίδρασης κατανομημένο σύμφωνα με την ανάλογη μέση τιμή και διακύμανση. Σαν μέγεθος επίδρασης χρησιμοποιείται η διαφορά μέσων τιμών μεταξύ των δειγμάτων ελέγχου και αναφοράς. Η διαφορά των μέσων τιμών μπορεί να είναι τυποποιημένη ή μη (Kulinskaya et al., 2011).

Η τυποποιημένη διαφορά μέσω των τιμών (standardized mean difference) και η τυπική της απόκλιση δίνονται από τους παρακάτω τύπους

$$d_i = \frac{\bar{x}_{1i} - \bar{x}_{2i}}{sd_i} \quad (\alpha) \quad sd_i = \sqrt{\frac{(n_{1i} - 1)sd_{1i}^2 + (n_{2i} - 1)sd_{2i}^2}{n_{1i} + n_{2i} - 2}} \quad (\beta)$$

Εξίσωση 10: (α) η διαφορά μέσω των τιμών, (β) η τυπική απόκλιση.

Η μη τυποποιημένη διαφορά μέσω των τιμών (unstandardized mean difference) δίνεται από το παρακάτω τύπο

$$d_i = \bar{x}_{1i} - \bar{x}_{2i}$$

(Εξίσωση 11)

Τέλος ένα βασικό στάδιο στη διεξαγωγή της μετα-ανάλυσης είναι η εξέταση του βαθμού ετερογένειας μεταξύ των μεμονωμένων μελετών. Ο υπολογισμός της ετερογένειας στο μοντέλο τυχαίων επιδράσεων που μας αφορά υλοποιείται στο STATA με τους κάτωθι τύπους

$$\theta(\tau)_{MLE} = \frac{\sum_i W_i(\tau) Y_i}{\sum_k W_i(\tau)} \quad \text{με} \quad W_i(\tau) = \frac{1}{s_i^2 + \tau^2} \quad \text{και} \quad Y_i = \bar{x}_{1i} - \bar{x}_{2i}$$

(Εξίσωση 12)

ΜΕΘΟΔΟΣ

ΑΠΟΤΕΛΕΣΜΑΤΑ ΒΙΟΠΛΗΡΟΦΟΡΙΚΗΣ ΑΝΑΖΗΤΗΣΗΣ

Έπειτα από ενδεδειγμένη αναζήτηση στα αποθετήρια δεδομένων (GEO, ArrayExpress, Quertle) με βασικό άξονα αναζήτησης τις χαμηλές δόσεις ακτινοβολίας (<0.5Gy) με όρους: ["radiation response"[All Fields] OR "response to radiation"[All Fields]) AND "Homo sapiens"[Organism]] και με εγγραφή στο site του GEO ώστε να έχουμε ενημέρωση σε όλη την πορεία της πραγματοποίησης της εργασίας για νέες σειρές (Εικόνα 25) προέκυψε ο παρακάτω πίνακας με τα εξής σύνολα δεδομένων (datasets) (Πίνακας 2):

GSE#	Low Dose [Gy]	High Dose [Gy]	IR Type	Platform
GSE20629 (bin< 4hrs)	0.1	2.5	lowLET protons	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Probe Name version)
GSE16935 (bin< 4hrs)	0.1	2.5	lowLET protons	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Probe Name version)
GSE23901 (bin< 4hrs)	0.5	NULL	a-particles	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Feature Number version)
GSE29344 (bin< 4hrs)	0.1	1	X-rays	Illumina HumanRef-8 v3.0 expression beadchip
GSE22639 (no control)	0.010, 0.025, 0.050, 0.075, 0.100	NULL	unknown	[HG-U133A] Affymetrix Human Genome U133A Array
GSE12435 (bystander removed)	0.5	NULL	a-particles	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Probe Name version)
GSE8917	0.5	2, 5, 8	Cs-137 (gamma-rays)	Agilent-012391 Whole Human Genome Oligo Microarray G4112A (Feature Number version)
GSE6978 (bin< 4hrs)	0.05, 0.50	NULL	Co-60 (gamma-rays)	CEA - DSV/DRR/SGF : Human 25k oligo v1.0
GSE52918 (chronic removed)	0.1(acute)	2 (acute)	Cs-137 (acute) Co-57(chronic)	Illumina HumanHT-12 V4.0 expression beadchip
GSE43151 (bin< 4hrs)	0.005,0.010,0.025,0.05 0,0.100,0.500	NULL	Co-60 (gamma-rays)	Agilent-026652 Whole Human Genome Microarray 4x44K v2 (Probe Name version)
GSE59861 (bin< 4hrs)	0.05	2	X-rays	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)

Πίνακας 2: Σειρές δεδομένων μικροσυστοιχιών με χαμηλή δόσης ακτινοβόλησης (<0.5Gy).

← → ↻ 🏠 Ασφαλές | <https://mail.google.com/mail/u/0/#search/radiation+response/15d0592cfa3cb363>

Google

Gmail

ΣΥΝΤΑΞΗ

Εισερχόμενα (7.928)
Με αστέρι
Σημαντικά
Απεσταλμένα
Πρόχειρα (18)

Κατηγορίες

Κοινωνικά (1.141)
Προσφορές (4.944)
Ενημερώσεις (1.678)
Φόρουμ (87)

Christine +

What's new for '(radiation response) AND "Homo sapiens"[porgn]' in Gds

My NCBI <efback@ncbi.nlm.nih.gov> Κατόνισση ενσωματώσεως
προς Εμένα

3 Ιουλ (Πριν από 2 ημέρες)

Αγγλικά > Ελληνικά > Μετάφραση μηνύματος

Απενεργοποίηση για: Αγγλικά

This message contains My NCBI what's new results from the National Center for Biotechnology Information (NCBI) at the U.S. National Library of Medicine (NLM). Do not reply directly to this message.

Sender's message:

Sent on Sunday, 2017 July 02
Search: (radiation response) AND "Homo sapiens"[porgn]
[View](#) complete results in Gds (results may change over time).
[Edit](#) saved search settings, or [unsubscribe](#) from these e-mail updates.

GEO DataSets Results

Items 1 - 5 of 8

1. [Waterpipe Smoking Induces Epigenetic Changes in the Small Airway Epithelium \[array\]](#)
(Submitter supplied) Waterpipe (also called hookah, shisha, or narghile) smoking is a common form of tobacco use in the Middle East. Its use is becoming more prevalent in Western societies, especially among young adults as an alternative form of tobacco use to traditional cigarettes. While the risk to cigarette smoking is well documented, the risk to waterpipe smoking is not well defined with limited information on its health impact at the epidemiologic, clinical and biologic levels with respect to lung disease. [more...](#)
Organism: Homo sapiens
Type: Expression profiling by array; Methylation profiling by genome tiling array
Platforms: [GPL570 GPL16419 28 Samples](#)
Download data: [GEO \(CEL, FAIR\)](#)
Series
Accession: GSE84101
ID: 222204101

Πραγματοποίηση κλήσης

Σχετική σελίδα Google+

NIH National Institute...

Εικόνα 25: Ενημέρωση για νέες σειρές δεδομένων μέσω GEO.

Οι τελικές προς επεξεργασία μελέτες είναι οι GSE20629, GSE23901, GSE29344, GSE12435, GSE8917, GSE52918, GSE59861. Οι υπόλοιπες αφαιρέθηκαν είτε λόγω ελλιπών δεδομένων (πχ μη ύπαρξη control sample) είτε λόγω του ότι τα δεδομένα ήταν τα ίδια σε δυο ταυτόχρονα μελέτες. Οι σειρές αυτές αντιστοιχούν στα chip Agilent, Illumina, Affymetrix.

Για να μελετήσουμε την επίδραση των χαμηλών δόσεων ιοντίζουσας ακτινοβολίας χρησιμοποιήθηκαν τις τρεις σειρές δεδομένων που αφορούν την Illumina και την Affymetrix. Τα πρωτογενή δεδομένα από τις προαναφερθείσες έρευνες πάρθηκαν από τη δημόσια βάση δεδομένων μικροσυστοιχιών GEO. Στην παρούσα εργασία επεξεργάστηκαν τα δεδομένα από τις GEO σειρές, GSE29344, GSE59861 και GSE52918, με τις εξής διαδικασίες: διόρθωση υποβάθρου, \log_2 μετασχηματισμός, κανονικοποίηση ποσοστημορίου χρησιμοποιώντας πακέτα (όπως limma) από τη σουίτα Bioconductor (Okoniewski and Miller, 2008), βασιζόμενοι στη γλώσσα προγραμματισμού R. Χρησιμοποιήθηκαν στατιστικές μέθοδοι για την ταυτοποίηση των στατιστικά σημαντικών διαφορικά εκφραζόμενων γονιδίων μεταξύ των διαφόρων δειγμάτων και η επεξεργασία της μετα-ανάλυσης.

Η διαδικασία επεξεργασίας (προσαρμόζεται στη συνέχεια στη κάθε πλατφόρμα) σε βήματα είναι η παρακάτω :

ΔΙΟΡΘΩΣΗ ΥΠΟΒΑΘΡΟΥ

Γενικά το πρώτο βήμα είναι η διόρθωση υποβάθρου της έντασης του κάθε ανιχνευτή. Ο φθορισμός υποβάθρου μπορεί να προκύψει από πολλές πηγές, όπως η μη ειδική σύνδεση του σεσημασμένου δείγματος στην επιφάνεια της μικροσυστοιχίας, επιδράσεις της επεξεργασίας όπως εναποθέσεις που παρέμειναν μετά το στάδιο της πλύσης ή οπτικού θορύβου από το σαρωτή. Υπάρχει πάντα κάποιο επίπεδο θορύβου υποβάθρου, ακόμη και εάν μόνο αποστειρωμένο νερό είναι επισημασμένο και υβριδιστεί στη μικροσυστοιχία, κάποιο επίπεδο φθορισμού θα εντοπιστεί από το σαρωτή. Οι διάφοροι αλγόριθμοι χρησιμοποιούν διαφορετικές μεθόδους διόρθωσης υποβάθρου, για παράδειγμα, ο αλγόριθμος RMA που χρησιμοποιούμε στην πλατφόρμα GSE59861 χρησιμοποιεί τη συνέλιξη του σήματος και του θορύβου.

ΛΟΓΑΡΙΘΜΗΣΗ

Ο νόμος του Benford (ή αλλιώς νόμος του πρώτου ψηφίου) (Benford, 1938) λέει ότι όταν έχουμε αριθμητικά δεδομένα (στατιστικά ή μετρήσεις) συχνά συμβαίνει το πρώτο ψηφίο των αριθμών να ακολουθεί μη-ομοιόμορφη κατανομή.

$$P(d) = \log\left(1 + \frac{1}{d}\right)$$

Όπου d : το πρώτο ψηφίο

Τα δεδομένα μας ακολουθούν τις προϋποθέσεις του νόμου του Benford οπότε τα λογαριθμούμε. Οι προϋποθέσεις αυτές αναφέρονται πχ σε δεδομένα που το εύρος τους καταλαμβάνει πολλές τάξεις μεγέθους ή σε δεδομένα που ακολουθούν κάποιες γνωστές ακολουθίες ακεραίων και ικανοποιούν ακριβώς τον νόμο του Benford όπως πχ ακολουθία Fibonacci.

ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ

Το επόμενο στάδιο είναι η κανονικοποίηση. Ο σκοπός αυτού του βήματος είναι να προσαρμοστούν τα δεδομένα στις τεχνικές διακυμάνσεις, σε αντίθεση με τις βιολογικές διαφορές μεταξύ των δειγμάτων. Υπάρχουν πάντα μικρές αποκλίσεις μεταξύ των διαδικασιών υβριδισμού για κάθε μικροσυστοιχία και αυτές οι αποκλίσεις τείνουν να οδηγούν σε μεγάλες αποκλίσεις μεταξύ των συνολικών εντάσεων για διαφορετικές μικροσυστοιχίες. Για παράδειγμα, η ποσότητα του RNA σε ένα δείγμα, ο χρόνος που ένα δείγμα υβριδίζεται ή ο όγκος ενός δείγματος μπορεί να εισάγουν σημαντικές διακυμάνσεις. Ακόμη και ανεπαίσθητες υλικές διαφορές μεταξύ των μικροσυστοιχιών ή μεταξύ των σαρωτών που χρησιμοποιούνται για να σαρώσουν τις μικροσυστοιχίες μπορεί να έχουν επίδραση στα αποτελέσματα. Η κανονικοποίηση εξασφαλίζει ότι η σύγκριση επιπέδων έκφρασης διαφορετικών μικροσυστοιχιών, είναι όσο το δυνατόν, σύγκριση ομοειδών. Οι μέθοδοι κανονικοποίησης που χρησιμοποιούνται έχουν σημαντική επίδραση στα τελικά επίπεδα διαφορικής έκφρασης (Altman, 2005), γι' αυτό είναι ζωτικής σημασίας να επιλεγεί η κατάλληλη μέθοδος.

Η μη-παραμετρική μέθοδος κανονικοποίησης που βρίσκει ευρεία εφαρμογή στην ανάλυση δεδομένων γονιδιακής έκφρασης, είναι η κανονικοποίηση ποσοστημορίων (quantile normalization) (). Η συγκεκριμένη μέθοδος οδηγεί σε συγκρίσιμες κατανομές τιμών ακόμα και στην περίπτωση που οι αρχικές κατανομές δεν είναι κανονικές. Η διαδικασία βασίζεται ουσιαστικά στη σύγκριση της κατάταξης των τιμών και για το λόγο αυτό είναι ανεξάρτητη των ροπών (μέση τιμή, διασπορά κλπ).

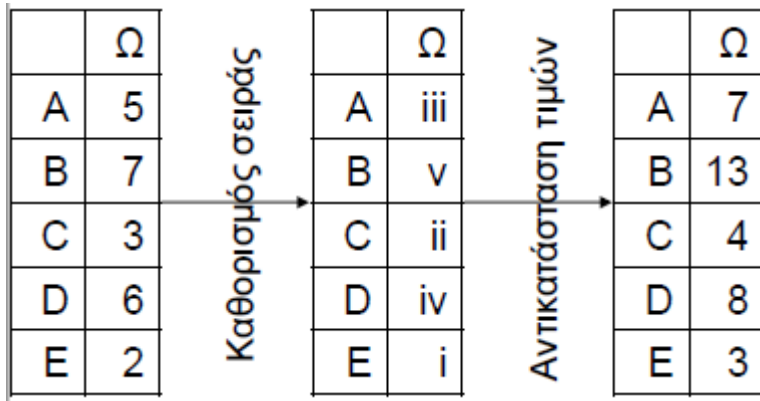
Για την κανονικοποίηση μιας κατανομής βάσει αυτής της μεθόδου, ορίζουμε μια κατανομή αναφοράς ίδιου μήκους και τις ταξινομούμε και τις δύο χωριστά. Η κατανομή αναφοράς φαίνεται στην Εικόνα 26

i	3
ii	4
iii	7
iv	8
v	13

Εικόνα 26: Κατανομή αναφοράς.

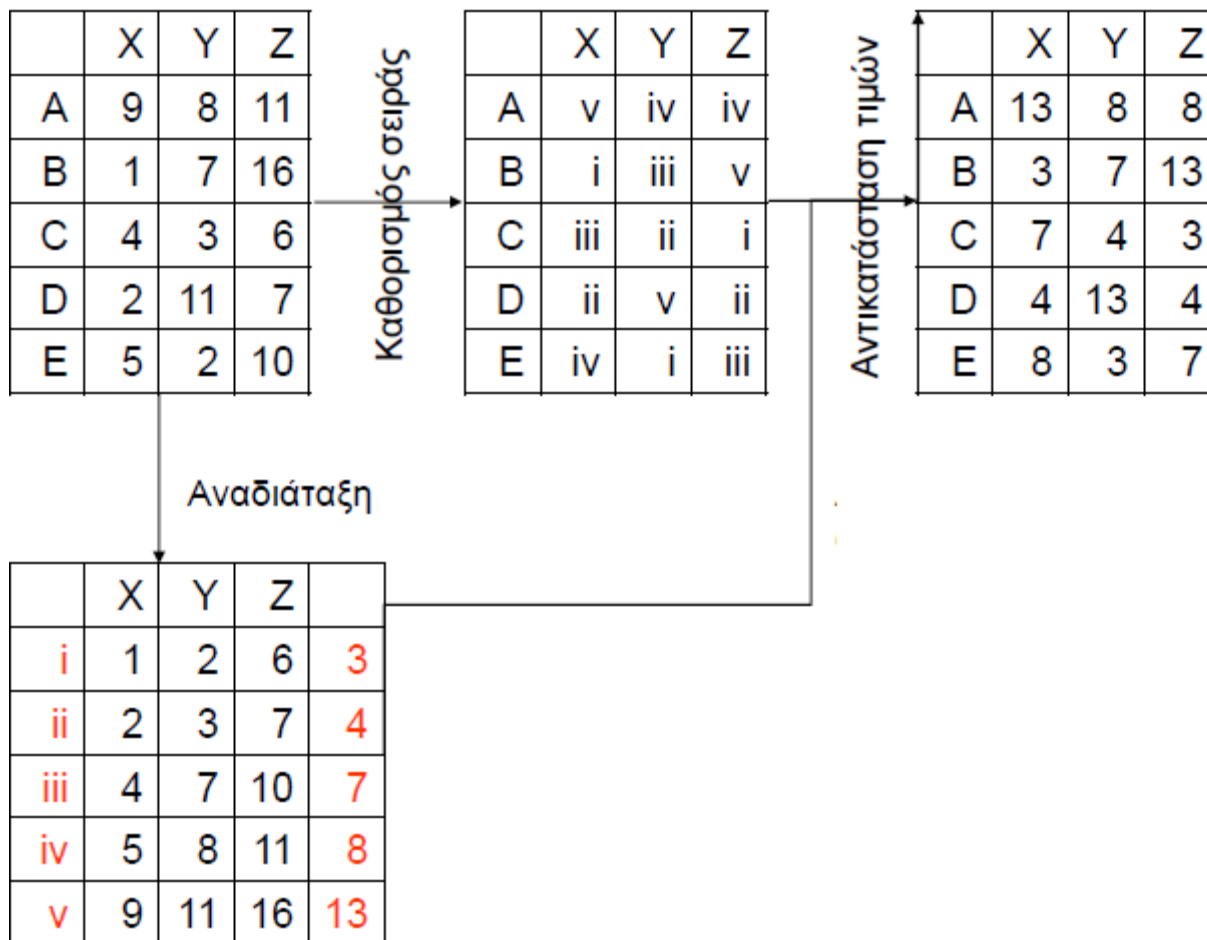
Για να κανονικοποιηθεί μια δοκιμαστική κατανομή σε κατανομή αναφοράς ίδιου μήκους, ταξινομούνται οι δύο κατανομές κατά αύξουσα σειρά. Η υψηλότερη τιμή της δοκιμαστικής κατανομής λαμβάνει στη συνέχεια την αξία της υψηλότερης τιμής της κατανομής αναφοράς, η επόμενη υψηλότερη τιμή της δοκιμαστικής

κατανομής την επόμενη υψηλότερη της κατανομής αναφοράς, κοκ, ώσπου η δοκιμαστική κατανομή είναι μια αναδιάταξη της κατανομής αναφοράς (Εικόνα 27).



Εικόνα 27: Αναδιάταξη τιμών με βάση τη δοκιμαστική κατανομή αναφοράς.

Για να κανονικοποιηθούν δύο ή περισσότερες κατανομές μεταξύ τους, χωρίς κατανομή αναφοράς, οι κατανομές ταξινομούνται κατά αύξουσα σειρά, και μετά ορίζεται ο μέσος όρος (συνήθως, ο αριθμητικός) των κατανομών. Έτσι, η υψηλότερη τιμή σε όλες τις περιπτώσεις γίνεται ο μέσος όρος των υψηλοτέρων τιμών, η δεύτερη υψηλότερη τιμή γίνεται ο μέσος όρος των τιμών της δεύτερη υψηλότερης τιμής, κοκ. (Εικόνα 28). Έτσι προκύπτουν οι νέες κανονικοποιημένες τιμές, οι οποίες ακολουθούν την ίδια κατανομή και μπορούν πλέον να συγκριθούν με ευκολία.



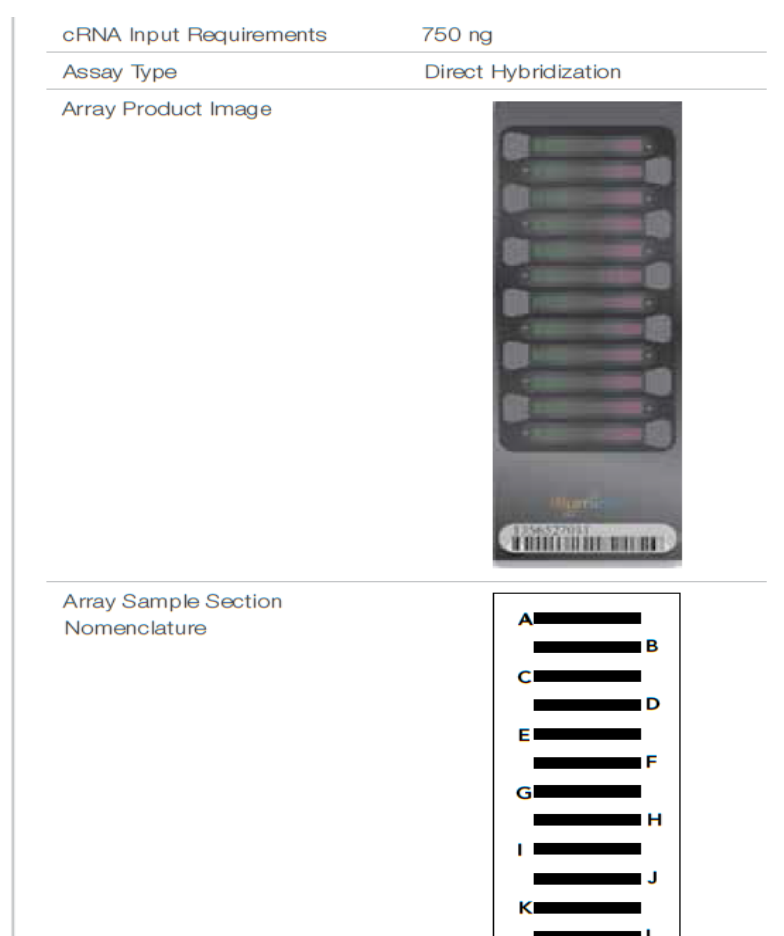
Εικόνα 28: Κανονικοποίηση ποσοστημορίου τριών κατανομών.

Πριν και μετά την κανονικοποίηση οπτικοποιούμε τα δεδομένα μας με θηκογράμματα για να παρουσιάσουμε τα κυριότερα χαρακτηριστικά της κατανομής μας. Στο boxplot διακρίνουμε την κατανομή των δεδομένων, τη διάμεσο, τα τεταρτημόρια, την ελάχιστη/μέγιστη τιμή των παρατηρήσεων του δείγματος καθώς και τυχόν ακραίες τιμές. Το θηκόγραμμα προσφέρεται ιδιαιτέρως για την ανίχνευση ακραίων τιμών και για την αναγνώριση της συμμετρίας ή του είδους της ασυμμετρίας της κατανομής. Αν το ευθύγραμμο τμήμα που αναπαριστά τη διάμεσο βρίσκεται στο μέσο του ορθογωνίου, η κατανομή είναι συμμετρική, αν βρίσκεται προς την κάτω πλευρά του ορθογωνίου, η κατανομή παρουσιάζει θετική ασυμμετρία και τέλος, αν βρίσκεται προς την πάνω πλευρά του ορθογωνίου, η κατανομή παρουσιάζει αρνητική ασυμμετρία. Επίσης, τα θηκογράμματα είναι εξαιρετικά χρήσιμα για τη σύγκριση των κατανομών δύο ή περισσότερων δειγμάτων όπως στην περίπτωση μας.

GSE52918

Η σειρά δεδομένων GSE52918 προέκυψε από τη μελέτη «Chronic low dose-rate versus acute ionizing radiation induced gene expression in primary derived prostate fibroblasts» (McDonald et al., 2014) πραγματοποιήθηκε στο Center of Cancer Systems Biology, GRI της Βοστώνης και δημοσιεύθηκε στις 20 Δεκεμβρίου 2013 στο περιοδικό International Journal of Radiation Biology.

Χρησιμοποιήθηκαν δεδομένα που αντλήθηκαν από την πλατφόρμα GPL10558. Η πλατφόρμα αυτή αντιστοιχεί στο chip Illumina HumanHT-12 V4.0 Expression beadchip.



Εικόνα 29: Chip Illumina HumanHT-12 V4.0 expression beadchip, product information.

Από τη σελίδα του κατασκευαστή (<http://support.illumina.com/>) αντλήσαμε πληροφορίες για την υπό μελέτη μικροσυστοιχία μέρος των οποίων εκθέτουμε παρακάτω (Εικόνα 29).

Πρόκειται για μία μικροσυστοιχία γονιδιακής έκφρασης. Ειδικότερα το chip HumanHT-12 v4 ExpressionBeadChip παρέχει τεχνολογία υψηλής απόδοσης με επεξεργασία 12 δειγμάτων (samples) ανά BeadChip. Κάθε συστοιχία (array) στο chip στοχοποιεί περισσότερα από 31,000 'σχολιασμένα' γονίδια με

περισσότερους από 47,000 ιχνηθέτες που αντλήθηκαν από το NCBI [National Center for Biotechnology Information Reference Sequence, RefSeqRelease 38 (November 7, 2009)] καθώς και από άλλες πηγές.

Στην σειρά δεδομένων GSE52918 συλλέχθηκαν δείγματα από ακτινοβόληση καλλιιεργειών ανθρώπινων ινοβλαστών του προστάτη δύο ασθενών. Η εν λόγω μελέτη περιλαμβάνει 32 συνολικά δείγματα στα οποία πραγματοποιήθηκε άμεση ακτινοβόληση με χαμηλή δόση ακτινοβολίας στα 10 cGy και 200 cGy καθώς και ακτινοβόληση με σταδιακή αύξηση της δόσης μεταξύ (1.0-2.45) cGy. Κρατήσαμε για τη μελέτη μας τα δεδομένα που αφορούν τις χαμηλές δόσεις ακτινοβολίας που αντιστοιχούν στα παρακάτω 24 δείγματα:

GSM1278163	p1541_0cGy_24hr_rep1
GSM1278164	p1541_0cGy_24hr_rep2
GSM1278165	p1541_0cGy_24hr_rep3
GSM1278167	p1617_0cGy_24hr_rep1
GSM1278168	p1617_0cGy_24hr_rep2
GSM1278169	p1617_0cGy_24hr_rep3
GSM1278170	p1617_0cGy_24hr_rep4
GSM1278171	p1541_10cGy_24hr_rep1
GSM1278172	p1541_10cGy_24hr_rep2
GSM1278173	p1541_10cGy_24hr_rep3
GSM1278174	p1541_10cGy_24hr_rep4
GSM1278175	p1617_10cGy_24hr_rep1
GSM1278176	p1617_10cGy_24hr_rep2
GSM1278177	p1617_10cGy_24hr_rep3

GSM1278178	p1617_10cGy_24hr_rep4
GSM1278179	p1541_200cGy_24hr_rep1
GSM1278180	p1541_200cGy_24hr_rep2
GSM1278181	p1541_200cGy_24hr_rep3
GSM1278182	p1541_200cGy_24hr_rep4
GSM1278183	p1617_200cGy_24hr_rep1
GSM1278184	p1617_200cGy_24hr_rep2
GSM1278185	p1617_200cGy_24hr_rep3
GSM1278186	p1617_200cGy_24hr_rep4

Πίνακας 3:Επιλεγμένα δείγματα από την σειρά δεδομένων GSE52918.

Η συλλογή των δειγμάτων έγινε μετά από 24 ώρες και τα δείγματα των καλλιιεργειών συλλέχθηκαν από 2 άτομα. Ανάμεσα στα υπό μελέτη δείγματα βρίσκονται και μη ακτινοβολημένα δείγματα απαραίτητα για να πραγματοποιηθεί σύγκριση της γονιδιακής έκφρασης.

Έχοντας προμηθευτεί τα πρωτογενή δεδομένα από τη σελίδα του GEO (Πίνακας 3) θα τα επεξεργαστούμε με ένα ειδικό πακέτο της R που ονομάζεται limma. Έχοντας εγκαταστήσει την τελευταία έκδοση της R χρησιμοποιούμε μια πιο φιλική προς το χρήστη διεπαφή για προγραμματισμό σε R που ονομάζεται RStudio και διανέμεται ελεύθερα.

Είναι σημαντικό το λογισμικό Rstudio να εκτελείται κάθε φορά με δικαιώματα διαχειριστή. Στο Rstudio βρίσκουμε το πακέτο Bioconductor το οποίο επί της ουσίας περιλαμβάνει πακέτα εντολών για την ανάλυση δεδομένων μικροσυστοιχιών.

```
>source("http://www.bioconductor.org/biocLite.R")
```

```
>biocLite()
```

Φορτώνουμε έπειτα τα πακέτα limma κι statmod του Bioconductor και φορτώνουμε της βιβλιοθήκες τους

```
>biocLite("limma")
```

```
>biocLite("statmod")
```

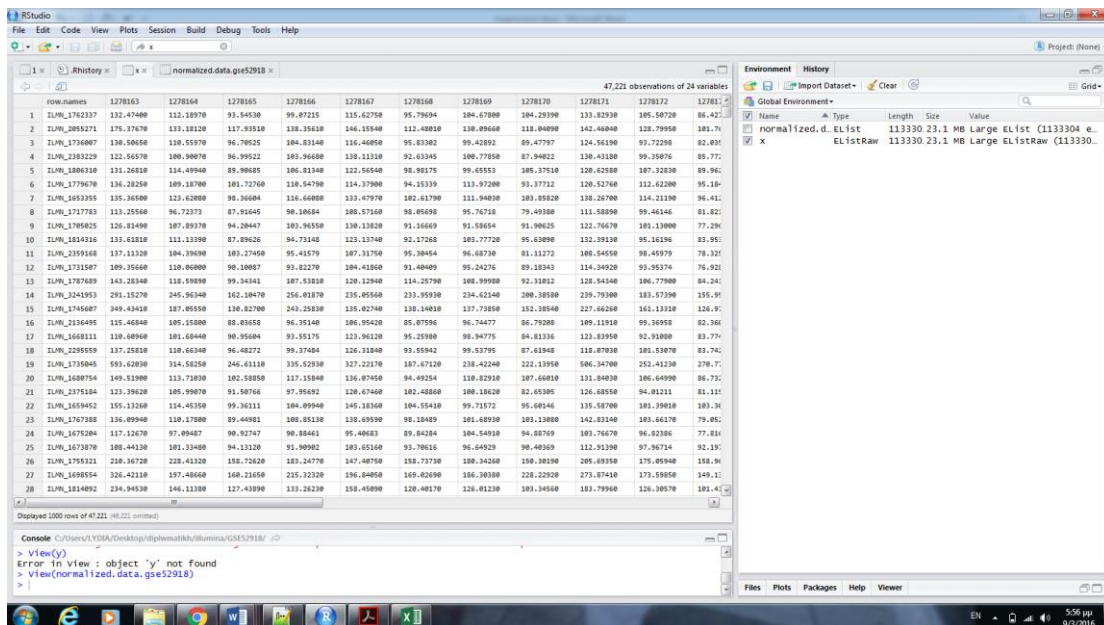
```
>library(limma)
```

Θα πρέπει να ορίσουμε το περιβάλλον εργασίας από το οποίο θα αντλούνται τα δεδομένα και αυτό γίνεται με την εντολή

```
>setwd("C:/Users/Illumina/GSE52918")
```

Η μέθοδος που θα χρησιμοποιήσουμε είναι η `negc()` η οποία πραγματοποιεί διόρθωση υποβάθρου, \log_2 μετασχηματισμό και κανονικοποίηση ποσοστημορίου. Για διόρθωση υποβάθρου υλοποιεί την `normexp.signal()` χρησιμοποιώντας τα αρνητικά control probes Σε περίπτωση που δεν υπάρχουν καλείται η μέθοδος `normexp.fit.detection.p()` η οποία συνάγει τις εντάσεις των αρνητικών control probes από τις τιμές των p-values που δίνονται από τα δεδομένα (detection p-values) συνδυασμένες με τα κανονικά probes. Για την κανονικοποίηση υλοποιεί την `.normexp.fit.control()`

Μετά την κανονικοποίηση εφαρμόζει \log_2 μετασχηματισμό. Επόμενο βήμα είναι η δημιουργία ενός αντικειμένου "EListRaw" ή αλλιώς RawExpressionList (Εικόνα 30). Η κλάση αυτή χρησιμοποιείται για να αποθηκεύσει πρωτογενή δεδομένα έντασης, μονοκαναλικής προέλευσης πριν αυτά υποστούν κάποιου είδους κανονικοποίηση αλλά αφού έχουν υποστεί διόρθωση υποβάθρου. Αντικείμενα αυτής της κλάσης περιέχουν μια γραμμή για κάθε probe και μια στήλη για κάθε δείγμα (Εικόνα 30). Αντικείμενα EListRaw δημιουργεί η μέθοδος `read.ilmm()`.



Εικόνα 30:ΑντικείμενοEListRaw(47.221 παρατηρήσεις από 24 μεταβλητές).


```
> head(normalized.data.gse52918$E)
```

```
1278163 1278164 1278165 1278166 1278167 1278168 1278169 1278170
ILMN_1762337 4.533312 4.555145 4.513908 4.450152 4.466452 4.505968 4.608671 4.988172
ILMN_2055271 5.266706 5.147371 5.552051 5.676969 5.300959 5.124925 5.516276 5.567485
ILMN_1736007 4.511253 4.518688 4.614745 4.587310 4.475635 4.506990 4.472128 4.454525
ILMN_2383229 4.430351 4.349025 4.623914 4.564458 4.971624 4.431634 4.503757 4.420682
ILMN_1806310 4.519551 4.611403 4.420373 4.642776 4.558663 4.595136 4.477375 5.038025
ILMN_1779670 4.580608 4.489085 4.813162 4.756317 4.453430 4.465882 4.925086 4.557753
1278171 1278172 1278173 1278174 1278175 1278176 1278177 1278178
ILMN_1762337 4.644341 4.627297 4.518807 4.579492 4.904804 4.414691 4.752954 4.596701
ILMN_2055271 4.829096 5.452661 5.178734 5.669661 4.900282 5.350108 5.333594 5.677365
ILMN_1736007 4.506316 4.354437 4.405791 4.399547 4.615653 4.542526 4.487237 4.370222
ILMN_2383229 4.587964 4.460575 4.500541 4.494216 4.737280 4.410104 4.756559 4.391546
ILMN_1806310 4.460432 4.687179 4.638241 4.709672 4.407988 4.672686 4.434281 4.630985
ILMN_1779670 4.459214 4.877312 4.860838 4.632696 4.479249 4.841386 4.719460 4.845799
1278179 1278180 1278181 1278182 1278183 1278184 1278185 1278186
ILMN_1762337 4.565776 4.407239 4.436835 4.407938 4.515374 4.501326 4.491686 4.595284
ILMN_2055271 5.531560 5.634481 5.304346 5.465553 5.370238 5.362296 5.221187 5.338386
ILMN_1736007 4.485482 4.585728 4.800458 4.597954 4.402373 4.574147 4.773721 4.407427
ILMN_2383229 4.724353 4.614488 4.370779 4.485877 5.029614 4.508379 4.415529 4.605269
ILMN_1806310 4.576398 4.510095 4.483834 4.585421 4.649169 4.293428 4.498529 4.423268
ILMN_1779670 4.666451 4.726693 4.954930 4.771613 4.484148 4.668722 5.075991 4.627991
```

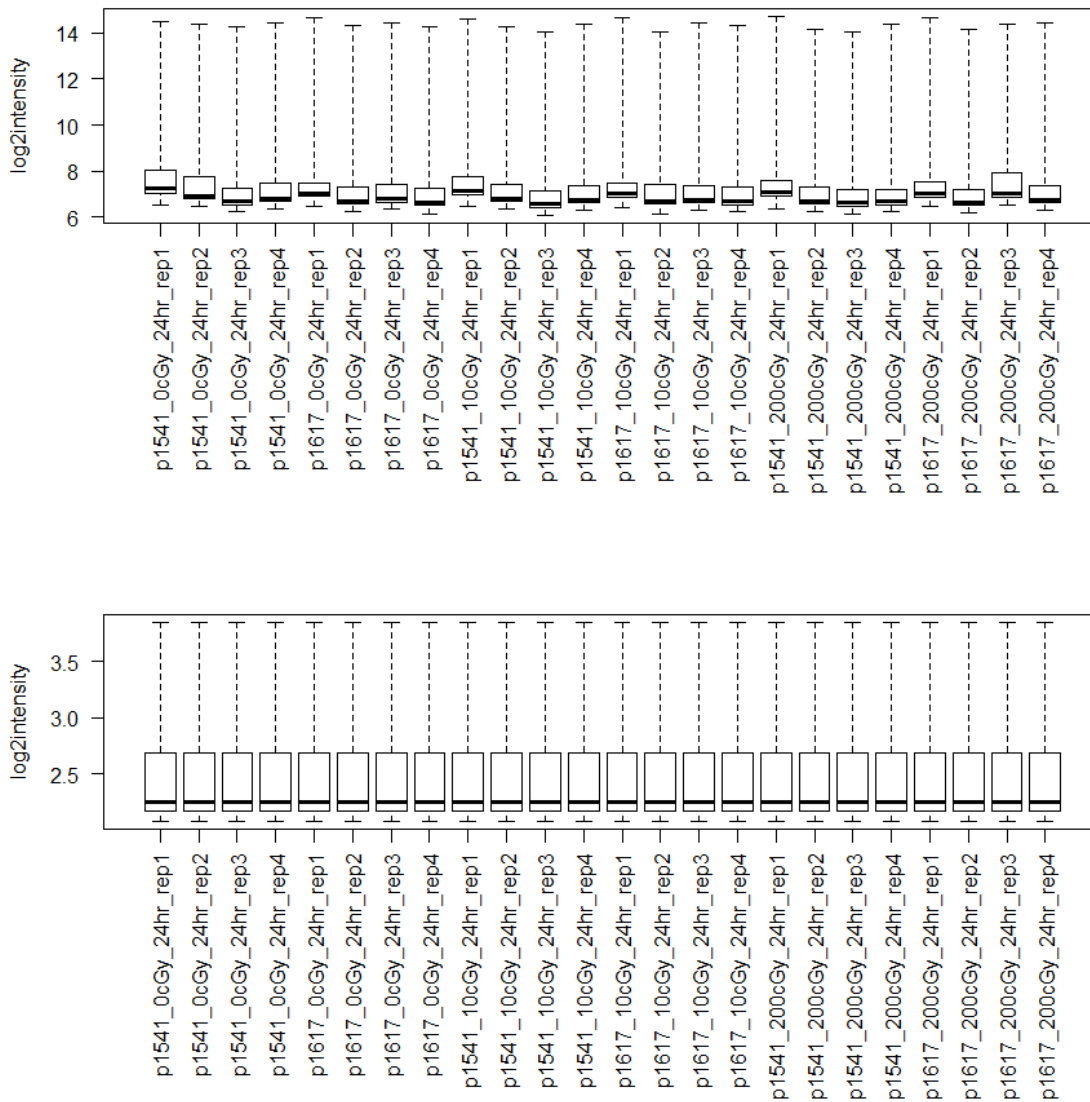
Εικόνα 32: Δείγμα τιμών των κανονικοποιημένων δεδομένων από το Rstudio.

Η συνιστώσα E περιέχει την τιμή έκφρασης του κάθε probe. Στην εικόνα βλέπουμε ένα boxplot των δεδομένων πριν και μετά την κανονικοποίηση που δημιουργήθηκαν με τις εντολές (Εικόνα 33):

```
>boxplot(log2(x$E), range=0, ylab="log2intensity", las=2, names=c("p1541_0cGy_24hr_rep1", "p1541_0cGy_24hr_rep2", "p1541_0cGy_24hr_rep3", "p1541_0cGy_24hr_rep4", "p1617_0cGy_24hr_rep1", "p1617_0cGy_24hr_rep2", "p1617_0cGy_24hr_rep3", "p1617_0cGy_24hr_rep4", "p1541_10cGy_24hr_rep1", "p1541_10cGy_24hr_rep2", "p1541_10cGy_24hr_rep3", "p1541_10cGy_24hr_rep4", "p1617_10cGy_24hr_rep1", "p1617_10cGy_24hr_rep2", "p1617_10cGy_24hr_rep3", "p1617_10cGy_24hr_rep4", "p1541_200cGy_24hr_rep1", "p1541_200cGy_24hr_rep2", "p1541_200cGy_24hr_rep3", "p1541_200cGy_24hr_rep4", "p1617_200cGy_24hr_rep1", "p1617_200cGy_24hr_rep2", "p1617_200cGy_24hr_rep3", "p1617_200cGy_24hr_rep4"))
```

```
>boxplot(log2(normalized.data.gse52918$E), range=0, ylab="log2intensity", las=2, names=c("p1541_0cGy_24hr_rep1", "p1541_0cGy_24hr_rep2", "p1541_0cGy_24hr_rep3", "p1541_0cGy_24hr_rep4", "p1617_0cGy_24hr_rep1", "p1617_0cGy_24hr_rep2", "p1617_0cGy_24hr_rep3", "p1617_0cGy_24hr_rep4", "p1541_10cGy_24hr_rep1", "p1541_10cGy_24hr_rep2", "p1541_10cGy_24hr_rep3", "p1541_10cGy_24hr_rep4", "p1617_10cGy_24hr_rep1", "p1617_10cGy_24hr_rep2", "p1617_10cGy_24hr_rep3", "p1617_10cGy_24hr_rep4"))
```

p1617_10cGy_24hr_rep2", "p1617_10cGy_24hr_rep3", "p1617_10cGy_24hr_rep4", "p1541_200cGy_24hr_rep1", "p1541_200cGy_24hr_rep2", "p1541_200cGy_24hr_rep3", "p1541_200cGy_24hr_rep4", "p1617_200cGy_24hr_rep1", "p1617_200cGy_24hr_rep2", "p1617_200cGy_24hr_rep3", "p1617_200cGy_24hr_rep4"))



Εικόνα 33:Boxplot δεδομένων σειράς GSE52918 πριν και μετά τη κανονικοποίηση με τη μέθοδο `pcqc()`

Η σειρά δεδομένων GSE29344 προέκυψε από τη μελέτη «Analysis of gene co-expression networks in skin cells exposed to different doses of ionizing radiation at different time points» των Ray M, Yunis R, Chen X, Rocke DM και πραγματοποιήθηκε από τον οργανισμό UC Davis στο τμήμα Biostatistics/Public Health των Ηνωμένων Πολιτειών και δημοσιεύθηκε 01 Φεβρουαρίου 2012 στο περιοδικό BMC Genomics.

Χρησιμοποιήθηκαν δεδομένα που αντλήθηκαν από την πλατφόρμα GPL6883. Η πλατφόρμα αυτή αντιστοιχεί στο chip Illumina HumanRef-8 v3.0 expression beadchip.

Από τη σελίδα του κατασκευαστή αντλήσαμε πληροφορίες για την υπό μελέτη μικροσυστοιχία μέρος των οποίων εκθέτουμε παρακάτω.

Πρόκειται για μία μικροσυστοιχία γονιδιακής έκφρασης (Εικόνα 34). Ειδικότερα το chip Illumina HumanRef-8 v3.0 expression beadchip παρέχει τεχνολογία υψηλής απόδοσης με επεξεργασία 8 δειγμάτων (samples) ανά BeadChip.



Εικόνα 34: HumanRef-8 Beadchip Illumina

Η εν λόγω μελέτη στοχεύει στη μελέτη των επιπτώσεων των χαμηλών δόσεων ιοντίζουσας ακτινοβολίας στον άνθρωπο. Ποσοτικοποιήθηκε η γενωμική απόκριση σε ακτίνες X 10cGy και 100cGy σε μοντέλο ιστού δέρματος λαμβάνοντας αποτελέσματα ανά διάφορα χρονικά διαστήματα. Το μοντέλο είναι ανακατασκευασμένος ιστός δέρματος, ο οποίος αποτελείται από τα κερατινοκύτταρα που απαρτίζουν το επιδερμικό στρώμα, και ινοβλάστες που απαρτίζουν το δερματικό στρώμα του δέρματος. Οι ιστοί ακτινοβολήθηκαν με 0, 10 και 100 cGy και δείγματα συλλέχθηκαν 0, 3, 8, και 24 ώρες μετά την ακτινοβολήση.

Προκειμένου να έχουμε την απαραίτητη βιολογική απόκριση τα δεδομένα που λήφθηκαν 3 ώρες μετά την ακτινοβολήση απορρίφθηκαν. Συνολικά η μελέτη περιλαμβάνει 24 δείγματα εκ των οποίων κρατήσαμε τα παρακάτω 18 δείγματα.

GSM725372	dose 0 cGy (control), 0 hrs (control), replicate 1
GSM725373	dose 0 cGy (control), 0 hrs (control), replicate 2
GSM725376	dose 0 cGy (control), 8 hrs, replicate 1
GSM725377	dose 0 cGy (control), 8 hrs, replicate 2
GSM725378	dose 0 cGy (control), 24 hrs , replicate 1
GSM725379	dose 0 cGy (control), 24 hrs, replicate 2
GSM725380	dose 10cGy, 0 hrs (control), replicate 1
GSM725381	dose 10cGy, 0 hrs (control), replicate 2
GSM725384	dose 10cGy, 8 hrs, replicate 1
GSM725385	dose 10cGy, 8 hrs, replicate 2
GSM725386	dose 10cGy, 24 hrs , replicate 1
GSM725387	dose 10cGy, 24 hrs, replicate 2
GSM725388	dose 100cGy, 0 hrs (control), replicate 1
GSM725389	dose 100cGy, 0 hrs (control), replicate 2
GSM725392	dose 100cGy, 8 hrs, replicate 1
GSM725393	dose 100cGy, 8 hrs, replicate 2
GSM725394	dose 100cGy, 24 hrs , replicate 1
GSM725395	dose 100cGy, 24 hrs, replicate 2

Πίνακας 4:Επιλεγμένα δείγματα από την σειρά δεδομένων GSE29344.

Ακολουθούμε παρόμοια διαδικασία με την επεξεργασία της σειράς δεδομένων GSE52918 και προμηθευόμαστε τα πρωτογενή δεδομένα από τη σελίδα του GEOτα οποία θα τα επεξεργαστούμε με το limma στο Rstudio. Εφόσον έχουμε κατεβάσει τα απαραίτητα πακέτα προς επεξεργασία των βιολογικών μας δεδομένων του Bioconductor αρκεί αυτή τη φορά να φορτώσουμε μόνο τις βιβλιοθήκες για τη τρέχουσα συνεδρία.

```
>source("http://www.bioconductor.org/biocLite.R")
```

```
>library("statmod")
```

```
>library ("limma")
```

Θα πρέπει να ορίσουμε το νέο περιβάλλον εργασίας για τη τρέχουσα συνεδρία από το οποίο θα αντλούνται τα δεδομένα και αυτό γίνεται με την εντολή

```
>setwd("C:/Users/Illumina/GSE29344")
```

Επόμενο βήμα είναι η δημιουργία ενός αντικειμένου "EListRaw" ή αλλιώς RawExpressionList με τη μέθοδο read.ilmn(). Φυσικά πρέπει πρώτα να επεξεργαστούμε το αρχείο των raw data ανάλογα προκειμένου να διαβαστεί σαν data matrix από το Rstudio. Αναθέτουμε σε μια μεταβλητή *x* το αντικείμενο **EListRaw** της μεθόδου **read.ilmn()** ρυθμίζοντας τις παραμέτρους της μεθόδου **files,probeid, annotation, expr, other.columns,sep,quote, verbose**, ώστε να προσαρμόζονται στα δικά μας πρωτογενή δεδομένα

```
>x<-read.ilmn(files="GSE29344_non-normalized_data.txt",  
ctrlfiles=NULL,probeid="ID_REF",annotation=c("TargetID",  
"SYMBOL"),expr="GSM",other.columns="Detection",sep="\t", quote="\'",  
verbose=TRUE)
```

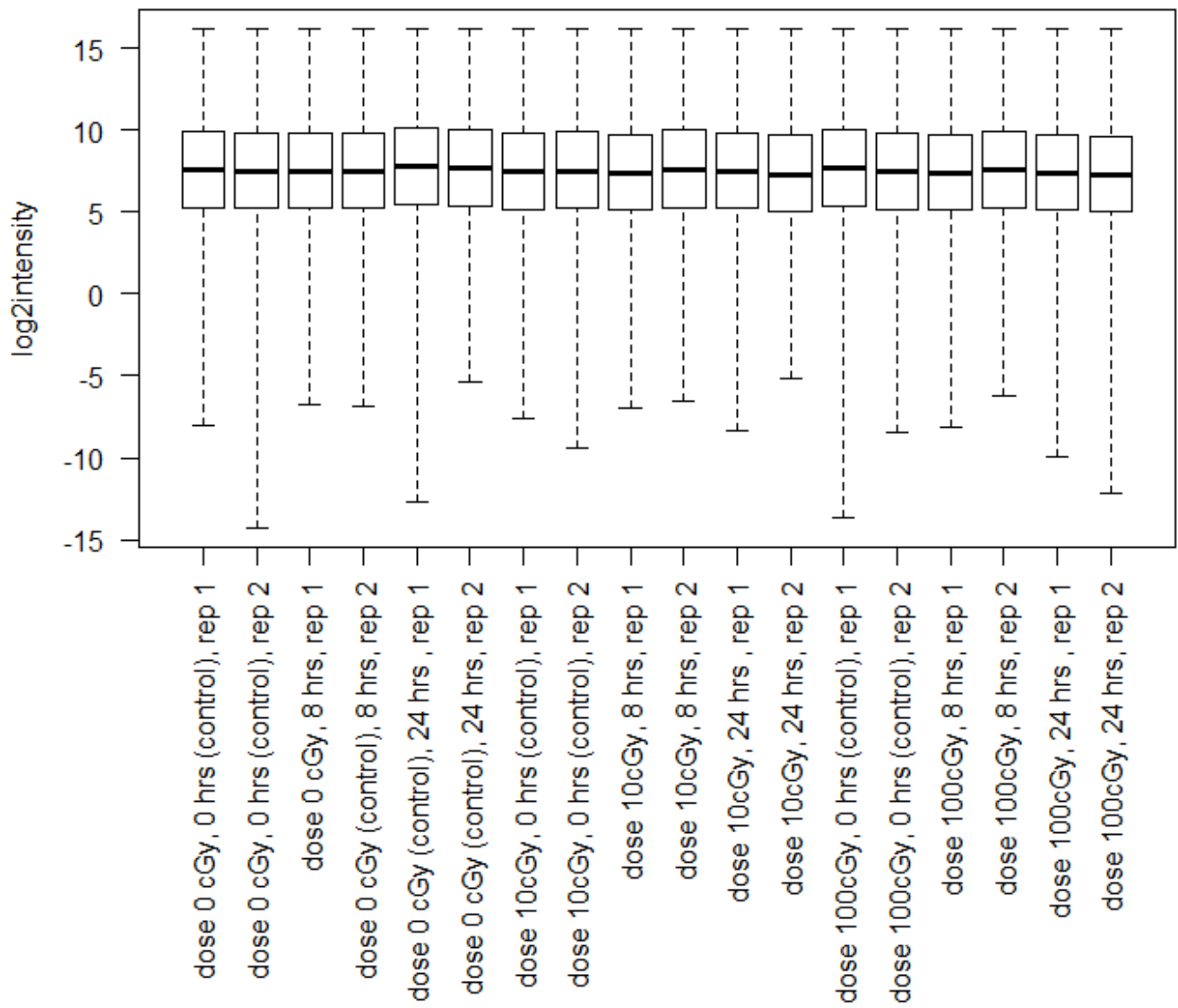
Μπορούμε πλέον να κάνουμε κανονικοποίηση των δεδομένων και αυτό επιλέξαμε να το κάνουμε με τη μέθοδο **neqc()** που υλοποιεί την **normexp()**. Ειδικότερα η **neqc()** δημιουργεί ένα αντικείμενο **EList (ExpressionList)** το οποίο περιέχει τις λογαριθμημένες εντάσεις αφού έχουν υποστεί διόρθωση υποβάθρου και κανονικοποίηση. Αναθέτουμε στη μεταβλητή **normalized.data.gse29344** τα κανονικοποιημένα πλέον δεδομένα

```
>normalized.data.gse29344<-  
neqc(x,status=NULL,negctrl="negative",regular="regular",offset=16,robust=FALSE  
,detection.p="Detection")
```

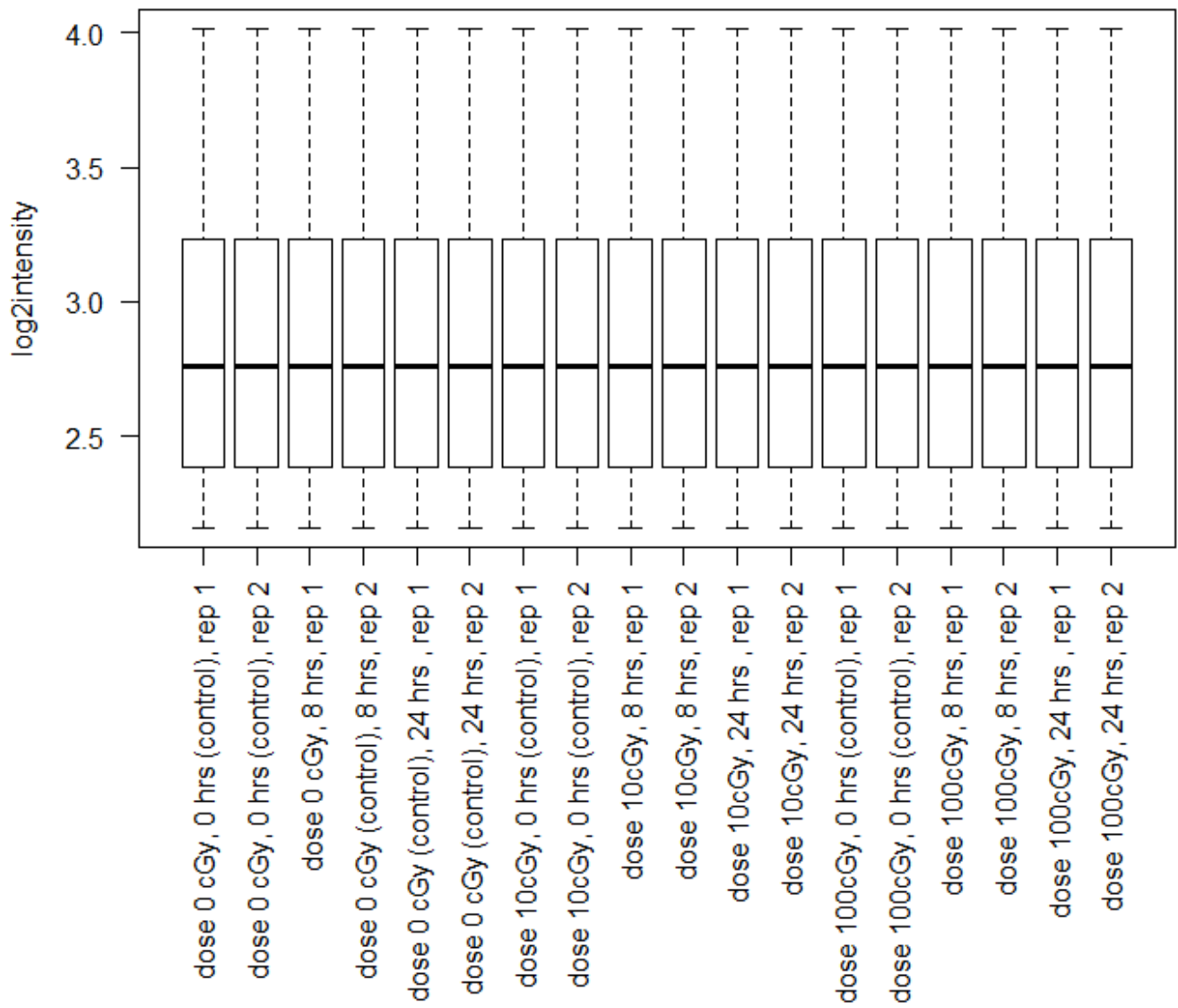
Η συνιστώσα *E* περιέχει την τιμή έκφρασης του κάθε probe. Στην εικόνα βλέπουμε ένα boxplot των δεδομένων πριν και μετά την κανονικοποίηση που δημιουργήθηκαν με τις εντολές:

```
> boxplot(log2(y$E),range=0,ylab="log2intensity",las=2, names=c("dose 0 cGy, 0 hrs (control), rep 1","dose 0 cGy, 0 hrs (control), rep 2","dose 0 cGy, 8 hrs, rep 1","dose 0 cGy (control), 8 hrs, rep 2","dose 0 cGy (control), 24 hrs , rep 1","dose 0 cGy (control), 24 hrs, rep 2","dose 10cGy, 0 hrs (control), rep 1","dose 10cGy, 0 hrs (control), rep 2","dose 10cGy, 8 hrs, rep 1","dose 10cGy, 8 hrs, rep 2","dose 10cGy, 24 hrs , rep 1","dose 10cGy, 24 hrs, rep 2","dose 100cGy, 0 hrs (control), rep 1","dose 100cGy, 0 hrs (control), rep 2","dose 100cGy, 8 hrs, rep 1","dose 100cGy, 8 hrs, rep 2","dose 100cGy, 24 hrs , rep 1","dose 100cGy, 24 hrs, rep 2"))
```

```
> boxplot(log2(x$E),range=0,ylab="log2intensity",las=2, names=c("dose 0 cGy, 0 hrs (control), rep 1","dose 0 cGy, 0 hrs (control), rep 2","dose 0 cGy, 8 hrs, rep 1","dose 0 cGy (control), 8 hrs, rep 2","dose 0 cGy (control), 24 hrs , rep 1","dose 0 cGy (control), 24 hrs, rep 2","dose 10cGy, 0 hrs (control), rep 1","dose 10cGy, 0 hrs (control), rep 2","dose 10cGy, 8 hrs, rep 1","dose 10cGy, 8 hrs, rep 2","dose 10cGy, 24 hrs , rep 1","dose 10cGy, 24 hrs, rep 2","dose 100cGy, 0 hrs (control), rep 1","dose 100cGy, 0 hrs (control), rep 2","dose 100cGy, 8 hrs, rep 1","dose 100cGy, 8 hrs, rep 2","dose 100cGy, 24 hrs , rep 1","dose 100cGy, 24 hrs, rep 2"))
```



Εικόνα 35: Βοχplot δεδομένων σειράς GSE29344 πριν τη κανονικοποίηση.



Εικόνα 36: Βοχplot δεδομένων σειράς GSE29344 μετά τη κανονικοποίηση με τη μέθοδο `negc()`

Η σειρά δεδομένων GSE59861 προέκυψε από τη μελέτη : «mRNA Expression data from AG 1522 human fibroblasts post X-ray radiation» των Hou J, Wang F, Kong P, Yu PK et al.πραγματοποιήθηκε στο Center of Medical Physics and Technology, Hefei Institutes of Physical Science, Chinese Academy of Sciences και δημοσιεύθηκε 30 Ιουλίου του 2014 στο περιοδικό PLoS One.

Αντικείμενο της μελέτης τους ήταν η ακτινοπροσαρμοστική απόκριση (Radioadaptive response (RAR)) των κυττάρων θηλαστικών που αφορά το φαινόμενο όπου μικρές δόσεις ακτινοβολίας διαφοροποιούν τη γονιδιακή έκφραση των κυττάρων προστατεύοντας τα από επόμενη μεγάλη δόση ακτινοβολίας. Μελέτησαν την RAR στην κυτταρική σειρά, AG 1522 ινοβλαστών ανθρώπινων κυττάρων με δόσεις 5 cGy και 2 Gy. Η εν λόγω μελέτη περιλαμβάνει 40 συνολικά δείγματα στα οποία πραγματοποιήθηκε άμεση ακτινοβόληση με χαμηλή δόση ακτινοβολίας στα 5cGy και 2Gy. Κρατήσαμε για τη μελέτη μας τα δεδομένα που αφορούν τις χαμηλές και υψηλές δόσεις ακτινοβολίας που αντιστοιχούν στα παρακάτω 22 δείγματα:

<u>GSM1448246</u>	radiation sham, biological rep1 (PrimeView)
<u>GSM1448247</u>	radiation sham, biological rep2 (PrimeView)
<u>GSM1448248</u>	radiation sham, biological rep3 (PrimeView)
<u>GSM1448249</u>	radiation sham, biological rep4 (PrimeView)
<u>GSM1448253</u>	6 h post 2 Gy radiation, biological rep1 (PrimeView)
<u>GSM1448254</u>	6 h post 2 Gy radiation, biological rep2 (PrimeView)
<u>GSM1448255</u>	6 h post 2 Gy radiation, biological rep3 (PrimeView)
<u>GSM1448256</u>	12 h post 2 Gy radiation, biological rep1 (PrimeView)
<u>GSM1448257</u>	12 h post 2 Gy radiation, biological rep2 (PrimeView)
<u>GSM1448258</u>	12 h post 2 Gy radiation, biological rep3 (PrimeView)
<u>GSM1448259</u>	24 h post 2 Gy radiation, biological rep1 (PrimeView)
<u>GSM1448260</u>	24 h post 2 Gy radiation, biological rep2 (PrimeView)
<u>GSM1448261</u>	24 h post 2 Gy radiation, biological rep3 (PrimeView)
<u>GSM1448265</u>	6 h post 5 cGy radiation, biological rep1 (PrimeView)
<u>GSM1448266</u>	6 h post 5 cGy radiation, biological rep2 (PrimeView)
<u>GSM1448267</u>	6 h post 5 cGy radiation, biological rep3 (PrimeView)
<u>GSM1448268</u>	12 h post 5 cGy radiation, biological rep1 (PrimeView)
<u>GSM1448269</u>	12 h post 5 cGy radiation, biological rep2 (PrimeView)
<u>GSM1448270</u>	12 h post 5 cGy radiation, biological rep3 (PrimeView)

<u>GSM1448271</u>	24 h post 5 cGy radiation, biological rep1 (PrimeView)
<u>GSM1448272</u>	24 h post 5 cGy radiation, biological rep2 (PrimeView)
<u>GSM1448273</u>	24 h post 5 cGy radiation, biological rep3 (PrimeView)

Πίνακας 5: Πίνακας δειγμάτων Affymetrix.

Χρησιμοποιήθηκαν δεδομένα που αντλήθηκαν από την πλατφόρμα GPL16043 που αντιστοιχεί στο chip GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs) της Affymetrix και φαίνεται στη παρακάτω Εικόνα 37.



Εικόνα 37: GeneChip® PrimeView™.

Φορτώνουμε τη σουίτα του Bioconductor και τις βιβλιοθήκες του BiocLite για την Affymetrix και ορίζουμε το περιβάλλον εργασίας:

```
>source("https://bioconductor.org/biocLite.R")
```

```
>biocLite()
```

```
>biocLite("affy")
```

```
>source("https://bioconductor.org/biocLite.R")
```

```
>library("affy")
```

Χρήση του τελευταίου Brainarray CDF που βασίζεται πχ στους ορισμούς των γονιδίων κατά TAIR με την παρακάτω εντολή:

```
>library("primeviewhsensgcdf")
```

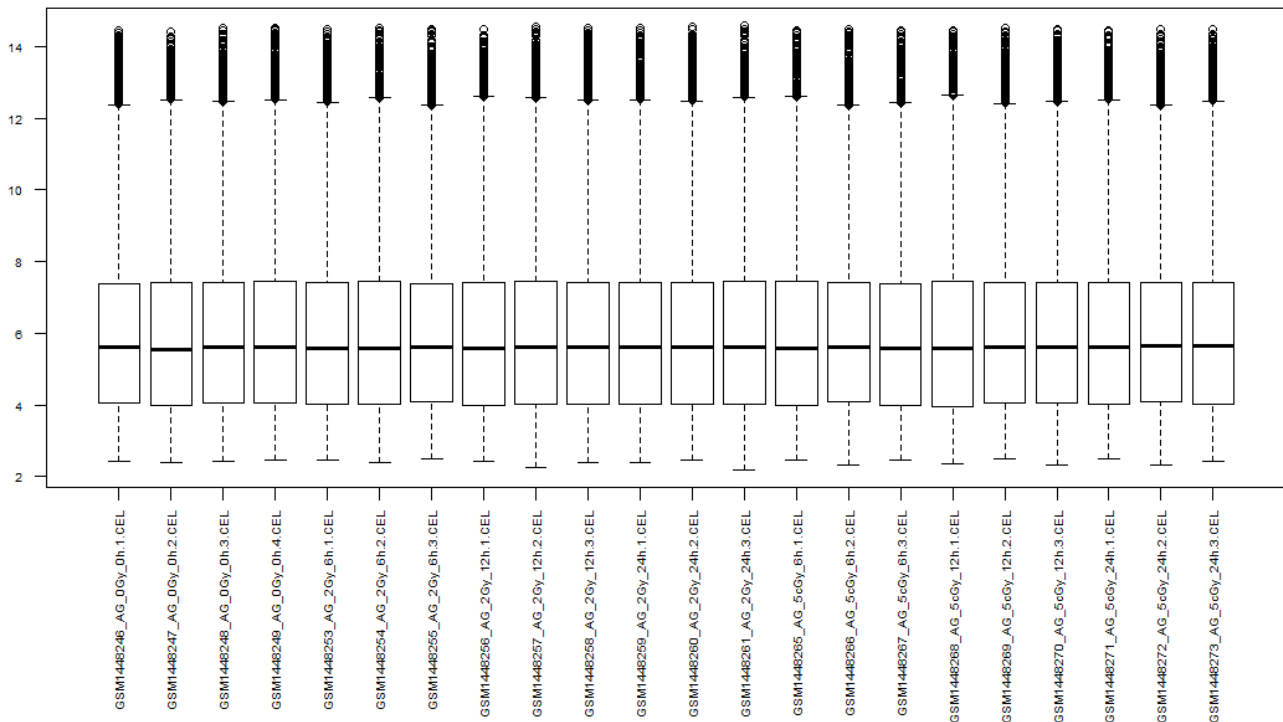
```
>setwd("C:\\Users\\Affymetrix\\GSE59861")
```

```
>eset.rma=justRMA(cdfname ="primeviewhsensgcdf")
```

```
>write.table(eset.rma, file="normAffy.txt", quote=F, sep="\t", row.names=TRUE)
```

Η ανάλυση όπως παρατηρούμε έγινε λίγο διαφορετικά από ότι στην Illumina με τον αλγόριθμο Robust MultiArray Analysis (RMA). Ο αλγόριθμος RMA είναι η μέθοδος που ακολουθούμε εδώ για να πάμε από επίπεδο ανιχνευτών σε τιμές γονιδιακής έκφρασης. Αυτή η μέθοδος είναι διαφορετική από τις μεθόδους της Affymetrix επειδή αγνοεί εντελώς τις τιμές των 22 ανιχνευτών MM (Bolstad et al., 2003; Irizarry et al., 2003). Οι εφευρέτες του αλγόριθμου υποστηρίζουν ότι οι ανιχνευτές MM εισάγουν περισσότερο θόρυβο. Παρότι αναγνωρίζουν ότι οι ανιχνευτές MM παρέχουν χρήσιμες πληροφορίες, μέχρι τη στιγμή της δημοσίευσης της μεθόδου, δεν βρήκαν κάποιο παραγωγικό τρόπο για να τους χρησιμοποιήσουν. Ο αλγόριθμος λειτουργεί προσαρμόζοντας τα σήματα για το θόρυβο υποβάθρου με μια πρωτογενή κλίμακα έντασης, η οποία δεν οδηγεί σε αρνητικές τιμές διόρθωσης υποβάθρου. Στη συνέχεια λαμβάνεται η \log_2 μετασχηματισμένη τιμή κάθε PM ανιχνευτή, διορθωμένη από το θόρυβο υποβάθρου, και αυτές οι τιμές κανονικοποιούνται χρησιμοποιώντας κανονικοποίηση ποσοστημορίων (Bolstad et al., 2003; Irizarry et al., 2003). Η ανάλυση των μικροσυστοιχιών προς εξέταση, διεξάγεται στη συνέχεια με κανονικοποίηση επί των ποσοστημορίων.

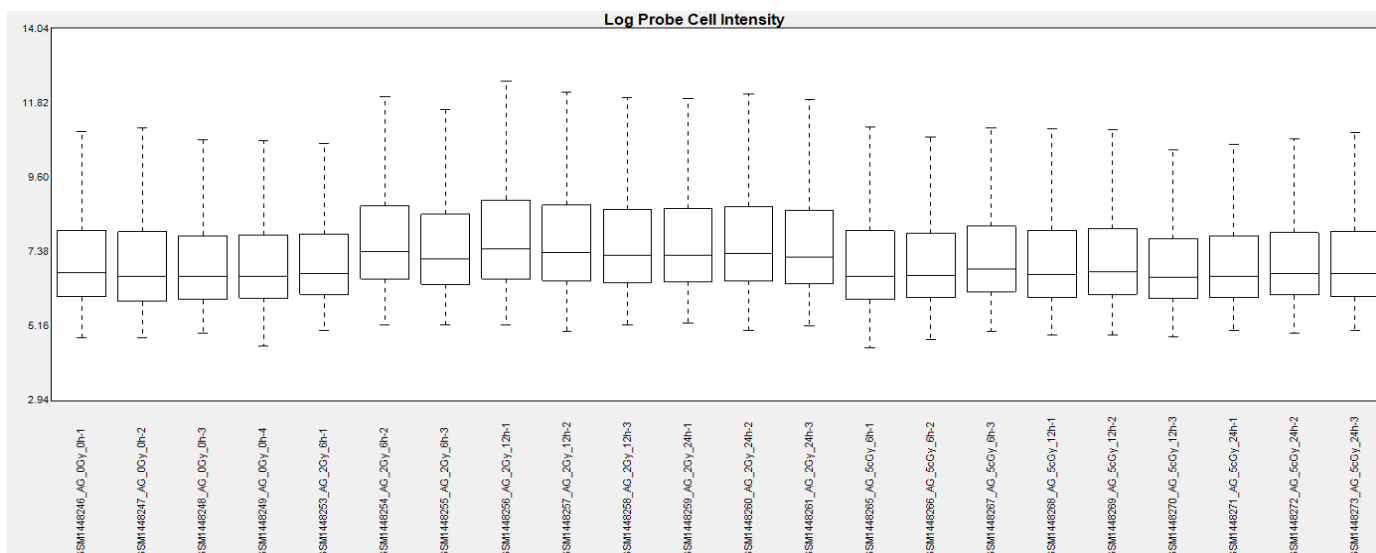
Επίσης για πληρότητα θα παραθέσουμε ένα boxplot των κανονικοποιημένων δεδομένων από το Rstudio, ωστόσο δε θα είναι η βασική μέθοδος οπτικοποίησης των δεδομένων μας.



Εικόνα 38: Boxplot κανονικοποιημένων δεδομένων (17.593 παρατηρήσεις σε 22 μεταβλητές) της Affymetrix.

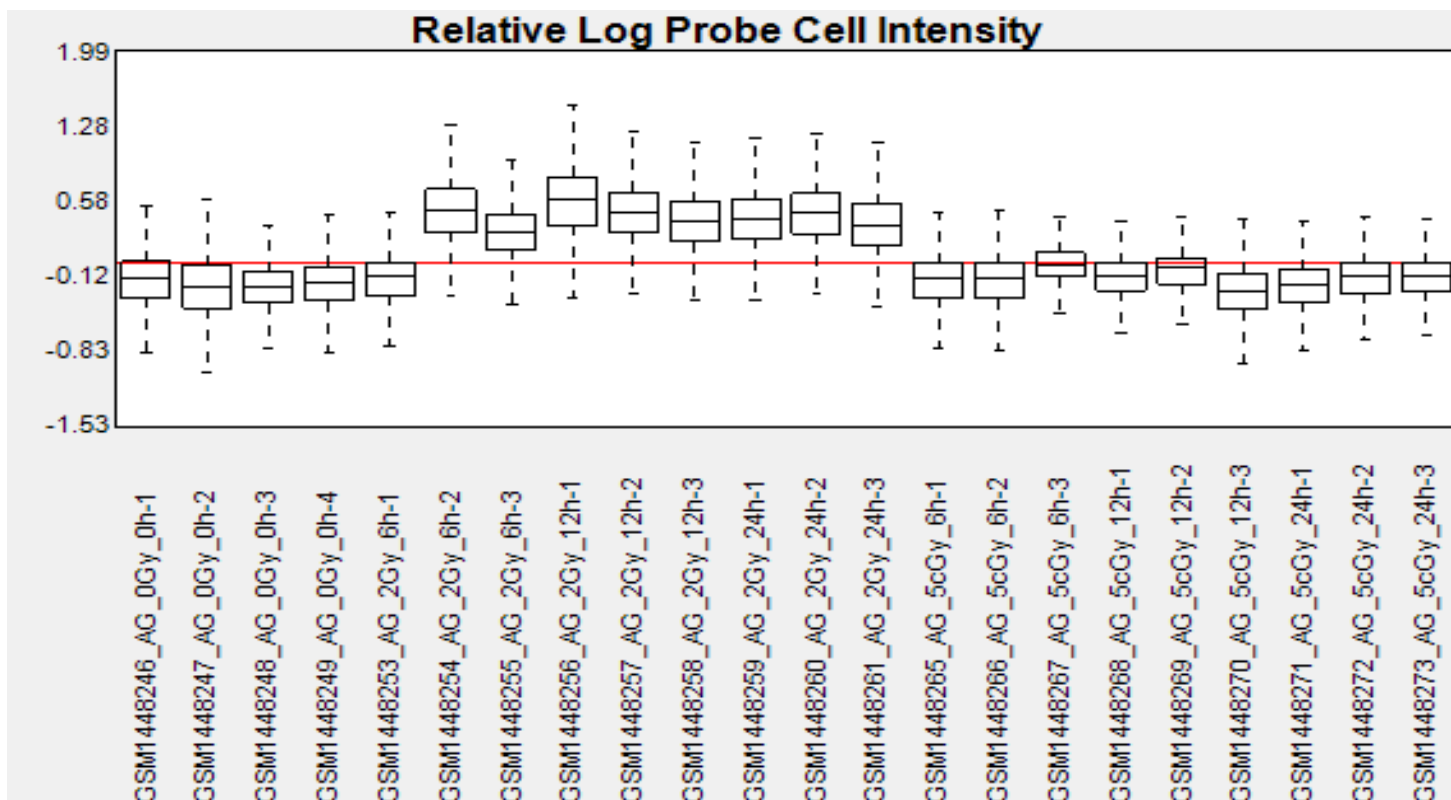
Όπως αναφέραμε η οπτικοποίηση θα γίνει με διαφορετικό τρόπο. Αυτή θα γίνει με το λογισμικό Expression Console της Affymetrix (<https://www.affymetrix.com/support/>). Εισάγουμε τα αρχεία .CEL και εφαρμόζουμε σε αυτά τον αλγόριθμο RMA και έχουμε τα ζητούμενα αρχεία .CHP. Παρακάτω δίνεται μια συνοπτική περιγραφή του κάθε γραφήματος των αρχείων των κανονικοποιημένων εκφράσεων .CHP και η οπτική του αναπαράσταση.

Box Plot Probe Cell Intensity: Το boxplot των λογαριθμημένων εντάσεων των probes δημιουργείται από τα αρχεία CEL που περιέχουν τις τιμές των εντάσεων του κάθε probe. Έχουμε boxplot για κάθε συστοιχία πριν την επεξεργασία οπότε αναμένουμε διαφορές στις κατανομές κάτι το οποίο εύκολα διαπιστώνουμε στην Εικόνα 39.



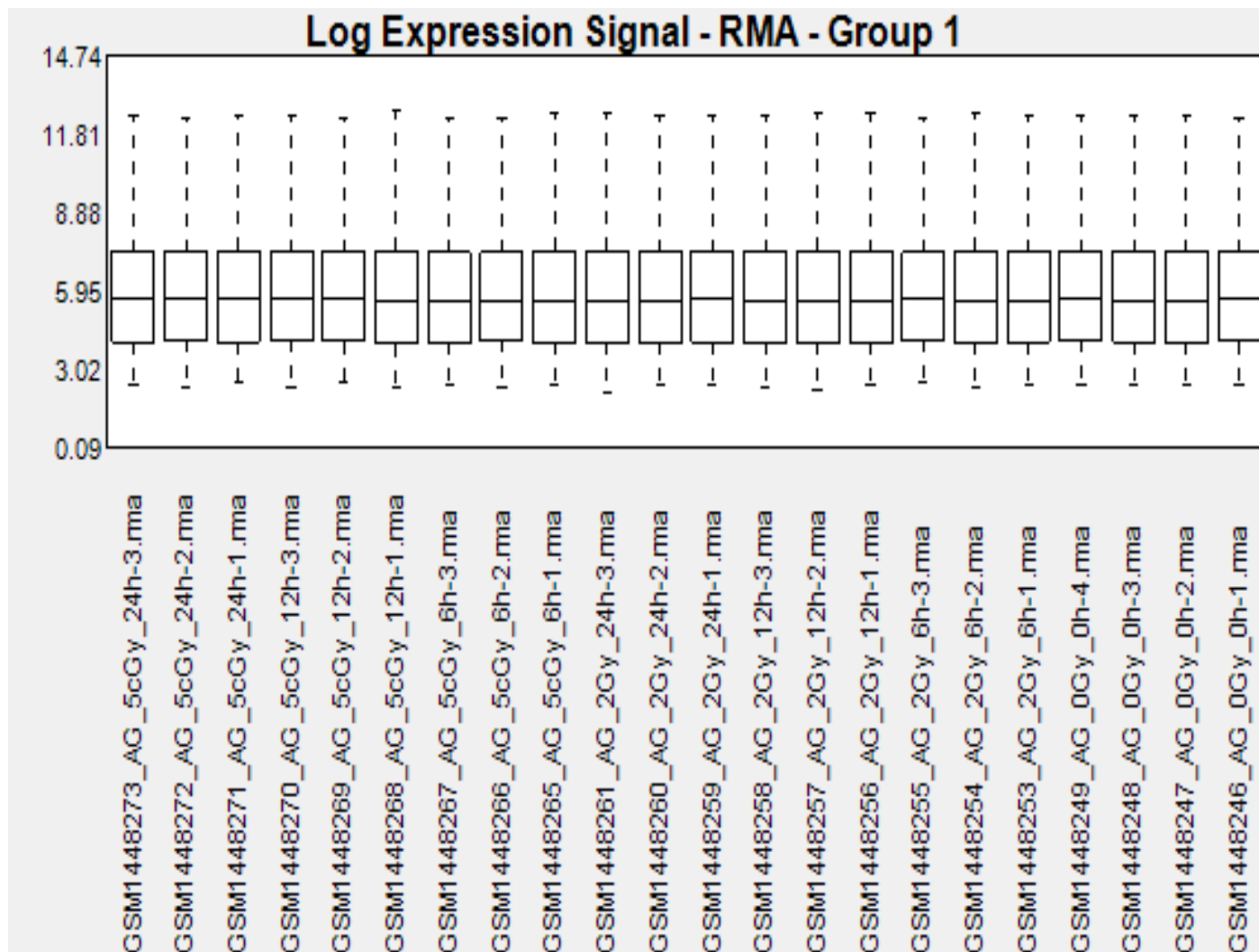
Εικόνα 39: Box Plot Probe Cell Intensity πριν την επεξεργασία.

Relative Probe Cell Intensity: Αφορά την κατανομή του λόγου των λογαριθμημένων εντάσεων του κάθε probe προς τη διάμεσο των εντάσεων των probes όλου του γκρουπ των δειγμάτων. Έτσι είναι εύκολο να διαπιστωθεί αν υπάρχει κάποιο ιδιαίτερα αποκλίνων δείγμα ώστε να αφαιρεθεί από τη μελέτη κάτι το οποίο δε χρειάζεται όπως διαπιστώνουμε στην Εικόνα 40.



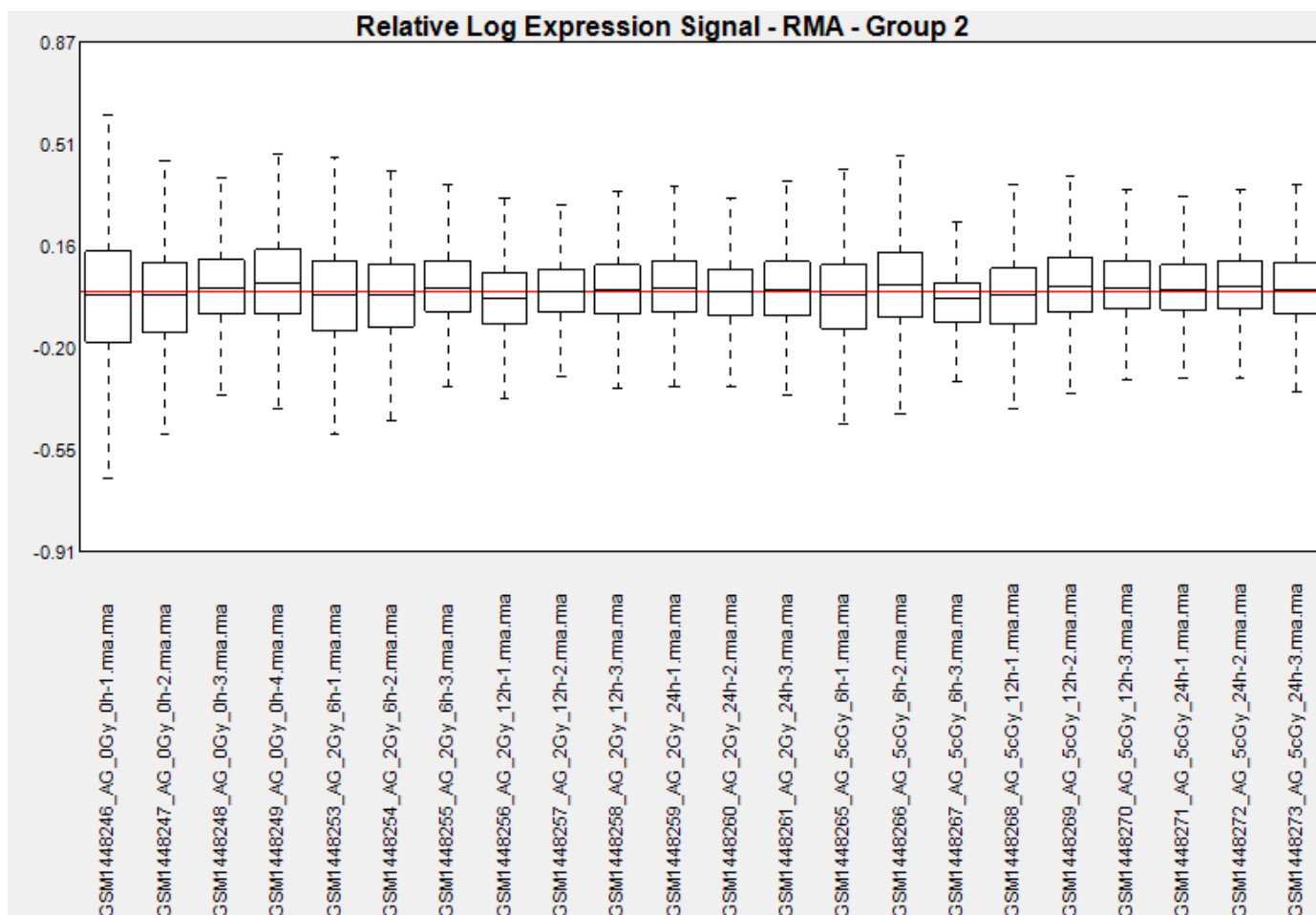
Εικόνα 40: Relative Probe Cell Intensity

Ακολουθεί το boxplot των κανονικοποιημένων πλέον τιμών με RMA που προέρχεται πλέον από τα CHP αρχεία. (Εικόνα 41).



Εικόνα 41:Κανονικοποιημένα δεδομένα με RMA της GSE59861.

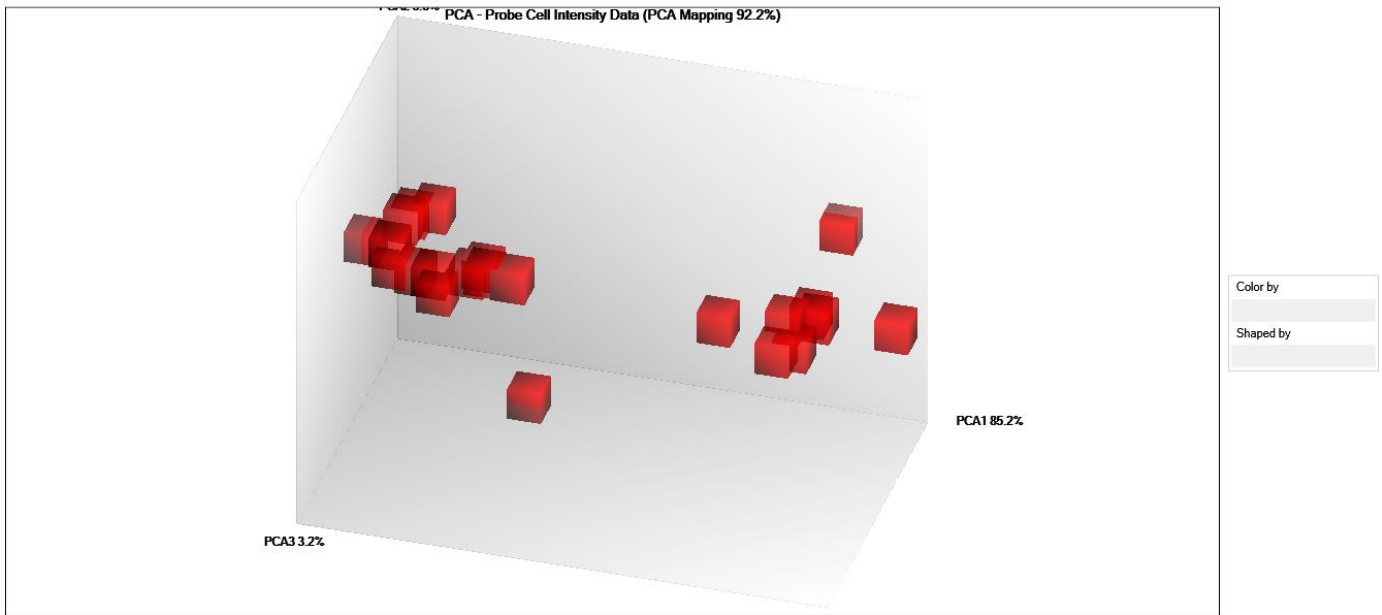
Box Plot Relative Signal : Το Relative Signal Box Plot συνοψίζει την κατανομή του λογαριθμημένου λόγου του σήματος του κάθε «probe set» προς τη διάμεσο του σήματος των «probe set» όλων των δειγμάτων .Έτσι μπορούμε να ξεχωρίσουμε οπτικά κάποια αποκλίνον δείγμα.



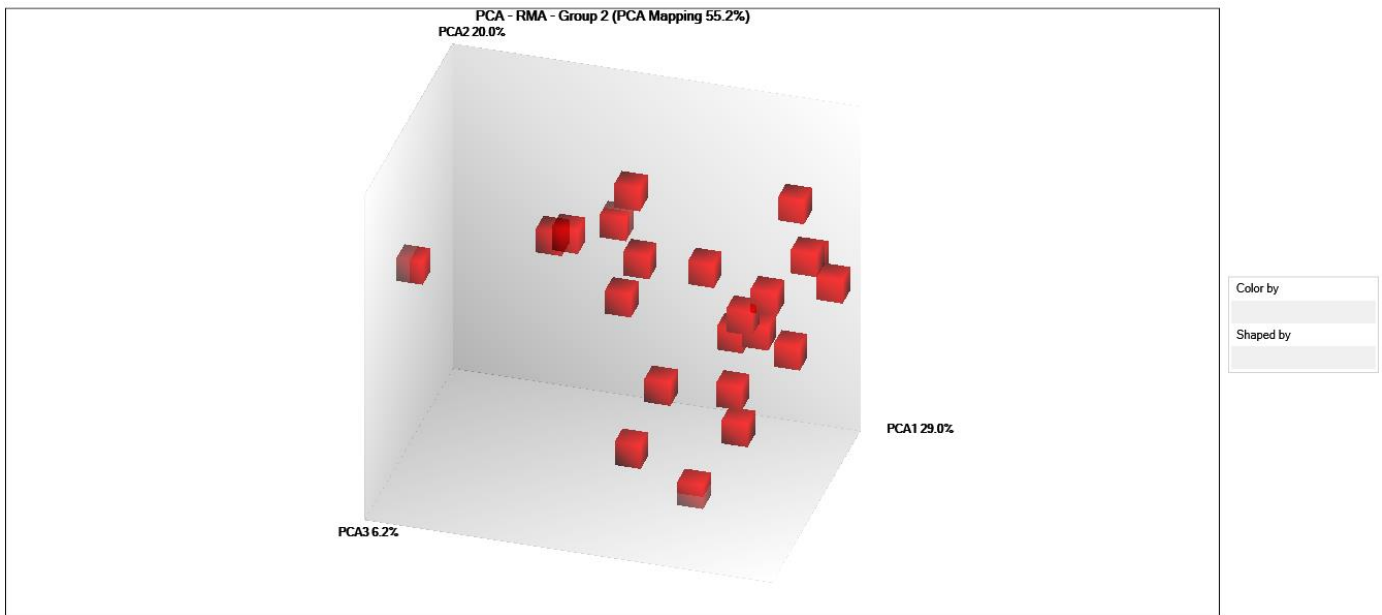
Εικόνα 42: Box Plot Relative Signal της GSE59861. Δε διαπιστώνεται κάποιο ιδιαίτερα αποκλίνον δείγμα.

Η ανάλυση PCA (Principal Component Analysis) δίνει μια οπτική εικόνα των συσχετίσεων της έκφρασης μεταξύ των συστοιχιών: τα δεδομένα προβάλλονται σε διάφορους άξονες που ταξινομούνται ως προς τη μείωση της σημαντικότητας. Το πρώτο κύριο στοιχείο (PC1) εξηγεί τις περισσότερες από τις παραλλαγές της έκφρασης. (Στην Εικόνα 43, ο PC1 εξηγεί σχεδόν το 85,2% της διακύμανσης). Συστάδες γονιδίων σε μια γραφική παράσταση PCA παρουσιάζουν ισχυρή συσχέτιση σημάτων έκφρασης. Η ανάλυση αυτή προτείνεται πριν και μετά την εξομάλυνση.

Το Expression Console μας δίνει τη δυνατότητα να εκτελέσουμε ανάλυση κύριων συνιστωσών (PCA) σε δεδομένα έντασης κυψελών ανίχνευσης (CEL) ή σήματος (CHP). Η PCA προσδιορίζει ένα νέο σύνολο μεταβλητών (PCA1, PCA2, PCA3) που αντιπροσωπεύουν το μεγαλύτερο μέρος της διακύμανσης στο αρχικό σύνολο δεδομένων. Το πρώτο κύριο στοιχείο (PCA1) καταγράφει όσο το δυνατόν μεγαλύτερη μεταβλητότητα στα δεδομένα. Το PCA2 συλλαμβάνει όσο το δυνατόν περισσότερο την υπόλοιπη μεταβλητότητα (που δεν υπολογίζεται από το PCA1). Το PCA3 συλλαμβάνει όσο το δυνατόν περισσότερο την υπόλοιπη μεταβλητότητα (που δεν υπολογίζεται από το PCA2). Τα δεδομένα σχετικά με την ένταση Probe cell προ-επεξεργάζονται για ανάλυση PCA λόγω περιορισμών μνήμης. Μια μη τυχαία μέθοδος δειγματοληψίας επιλέγει 50.000 σύνολα ανιχνευτών για ανάλυση.

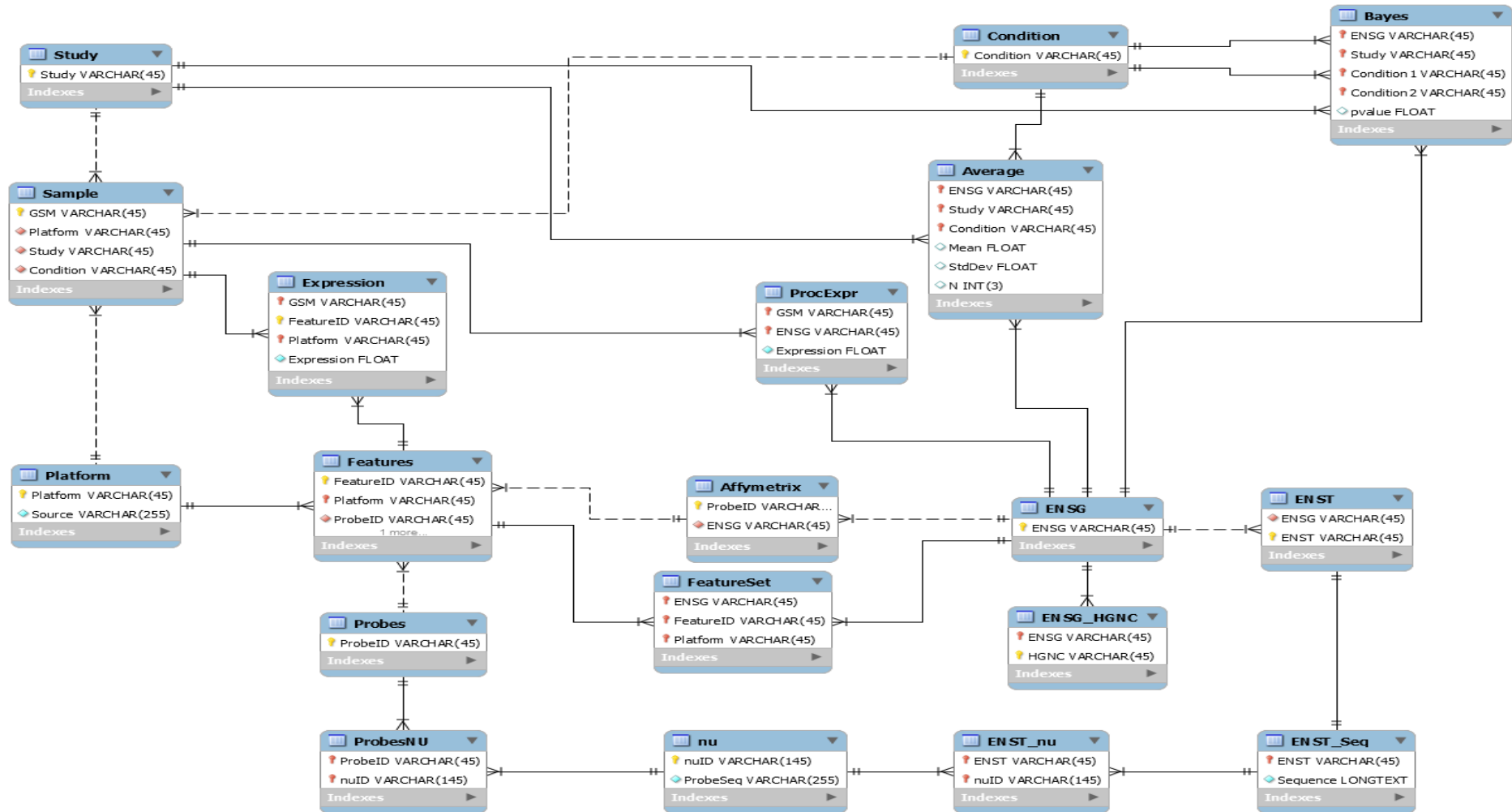


Εικόνα 43: Ανάλυση κύριων συνιστωσών (PCA) σε δεδομένα έντασης κυψελών ανίχνευσης (CEL). Συστάδες γονιδίων παρουσιάζουν ισχυρή συσχέτιση σημάτων έκφρασης.



Εικόνα 44: Ανάλυση κύριων συνιστωσών (PCA) σε δεδομένα .CHP.

Δημιουργήσαμε τη βάση δεδομένων “Microarrays” (Εικόνα 45):



Εικόνα 45: Διάγραμμα Συσχέτισης Οντοτήτων βάσης δεδομένων Microarrays.

Η δημιουργία της Βάσης έγινε με το MySQL Workbench ειδικότερα με το παρακάτω script :

```
mysql -u Christine -h carneades.bioacademy.gr -***** microarrays --local-  
infile
```

```
CREATE DATABASE microarrays;
```

```
Grant all ON microarrays.* TO Christine@localhost IDENTIFIED BY '159753aaa!'  
WITH GRANT OPTION;
```

```
UPDATE user SET File_priv='Y' WHERE User='Christine';
```

```
FLUSH privileges;
```

```
CREATE TABLE ENSG (
```

```
ENSG VARCHAR(45) NOT NULL,
```

```
PRIMARY KEY (ENSG)
```

```
);
```

```
load          data          local          infile  
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\GSE6968\\ENSG.txt"  
into table ENSG;
```

```
CREATE TABLE ENSG_HGNC (
```

```
ENSG VARCHAR(45) NOT NULL,
```

```
HGNC VARCHAR(45) NOT NULL,
```

```
PRIMARY KEY (ENSG, HGNC)
```

```
);
```

```
load          data          infile  
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\Manifest  
files\\ENSG_HGNC.txt" into table ENSG_HGNC;
```

```
CREATE TABLE ENST (
```

```
ENSG VARCHAR(45) NOT NULL,
```

```

ENST VARCHAR(45) NOT NULL,
PRIMARY KEY (ENST)
);

load          data          local          infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\GSE6968\\ENST.txt"
into table ENSG;

CREATE TABLE ENST_Seq (
    ENST VARCHAR(45) NOT NULL,
    Sequence LONGTEXT NOT NULL,
    PRIMARY KEY (ENST)
);

load          data          local          infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\GSE6968\\ENST_Seq.txt"
into table ENST_Seq;

CREATE TABLE ENST_nu (
    ENST VARCHAR(45) NOT NULL,
    nuID VARCHAR(145) NOT NULL,
    PRIMARY KEY (ENST,nuID)
);

select distinct ENST, nuID from nu,ENST_Seq where Sequence like
CONCAT("%",ProbeSeq,"%") limit 5 INTO OUTFILE
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\Manifest
files\\ENST_nu.txt" FIELDS TERMINATED BY '\\t' ENCLOSED BY '' LINES TERMINATED
BY '\\n';

load          data          local          infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\Manifest
files\\ENST_nu.txt" into table ENST_nu;

```

```

CREATE TABLE nu (
    nuID VARCHAR(145) NOT NULL,
    ProbeSeq VARCHAR(255) NOT NULL,
    PRIMARY KEY (nuID)
);

load          data          local          infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\Manifest
files\\nu.sort.txt" into table nu;

CREATE TABLE ProbesNU (
    ProbeID VARCHAR(45) NOT NULL,
    nuID VARCHAR(145) NOT NULL,
    PRIMARY KEY (ProbeID, nuID)
);

load          data          local          infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\Manifest
files\\ProbesNUUniq.nocr.txt.txt" into table ProbesNU;

CREATE TABLE Features (
    FeatureID VARCHAR(45) NOT NULL,
    Platform VARCHAR(45) NOT NULL,
    ProbeID VARCHAR(45) NOT NULL,
    PRIMARY KEY (FeatureID, Platform)
);

load          data          local          infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\Manifest
files\\Platform Tables\\Agilent\\Features_Table\\Features_Table_Agilent.txt"
into table Features;

```

```
load          data          local          infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\Manifest
files\\Platform Tables\\Illumina\\Features_Table\\Features_Table_Illumina.txt"
into table Features;
```

```
load          data          local          infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\Manifest
files\\Platform Tables\\Affymetrix\\Features_Table_Affymetrix.txt" into table
Features;
```

```
CREATE TABLE Expression (
    GSM VARCHAR(45) NOT NULL,
    FeatureID VARCHAR(45) NOT NULL,
    Platform VARCHAR(45) NOT NULL,
    Expression FLOAT NOT NULL,
    PRIMARY KEY (GSM, FeatureID, Platform)
);
```

```
load          data          local          infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Expression_Table\\Expression_Ta
ble_All.txt" into table Expression;
```

```
load          data          local          infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Expression_Table\\Affymetrix\\E
xpression.txt" into table Expression;
```

```
load          data          local          infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Expression_Table\\Illumina\\Ill
umina_expression_table.txt" into table Expression;
```

```
CREATE TABLE ProbeSet (
    ENSG VARCHAR(45) NOT NULL,
    FeatureID VARCHAR(45) NOT NULL,
    Platform VARCHAR(45) NOT NULL,
```

```

PRIMARY KEY (ENSG, FeatureID, Platform)
);

select  ENSG,FeatureID,Platform  from  ENST,  ENST_nu,ProbesNU,Features  WHERE
ENST.ENST      =ENST_nu.ENST      AND      ENST_nu.nuID=ProbesNU.nuID      AND
ProbesNU.ProbeID=Features.ProbeID  INTO  OUTFILE  "/tmp/ProbeSet.txt"  FIELDS
TERMINATED BY '\t' ENCLOSED BY '' LINES TERMINATED BY '\n';

load      data      local      infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\Manifest
files\\Other_Tables\\ProbeSet.txt" into table ProbeSet;

CREATE TABLE ProcExpr (
    GSM VARCHAR(45) NOT NULL,
    ENSG VARCHAR(45) NOT NULL,
    Expression FLOAT NOT NULL,
    PRIMARY KEY (GSM, ENSG)
);

load      data      local      infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\Manifest
files\\Other_Tables\\ProcExpr.txt" into table ProcExpr;

CREATE TABLE Platform (
    Platform VARCHAR(45) NOT NULL,
    Source VARCHAR(255) NOT NULL,
    PRIMARY KEY (Platform)
);

load      data      local      infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\Manifest
files\\Other_Tables\\Platform_Table.txt" into table Platform;

CREATE TABLE Sample (

```



```

GSM VARCHAR(45) NOT NULL,
Platform VARCHAR(45) NOT NULL,
Study VARCHAR(45) NOT NULL,
Condition VARCHAR(45) NOT NULL,
PRIMARY KEY (GSM)
);

load          data          local          infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\Manifest
files\\Other_Tables\\Sample_Table.txt" into table Sample;

CREATE TABLE Condition (
    Condition VARCHAR(45) NOT NULL,
    PRIMARY KEY (Condition)
);

load          data          local          infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\Manifest
files\\Other_Tables\\Condition.txt" into table Condition;

CREATE TABLE Study (
    Study VARCHAR(45) NOT NULL,
    PRIMARY KEY (Study)
);

load          data          local          infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\Manifest
files\\Other_Tables\\Study.txt" into table Study;

CREATE TABLE Affymetrix (
    ProbeID VARCHAR(45) NOT NULL,
    ENSG VARCHAR(45) NOT NULL,

```

PRIMARY KEY (ProbeID)

);

```
load data local infile
"C:\\Users\\Christine\\Documents\\Microarrays\\Analyses\\Manifest
files\\AffyTable\\Affymetrix.txt" into table Affymetrix;
```

Με απλές εντολές στο τερματικό cmd, χειριζόμαστε τη βάση δεδομένων που δημιουργήσαμε, όπως για παράδειγμα παραθέτουμε μερικές :

```
>show databases;
```

Μπορούμε να δούμε ότι στη βάση μπορούν να μπουν όσοι προέρχονται από το “carneades” ή βρίσκονται στο τοπικό δίκτυο και γενικά να έχουμε εικόνα ποιοί έχουν πρόσβαση σε αυτήν.

```
>use Microarrays;
```

Έτσι επιλέγουμε τη βάση δεδομένων που μας ενδιαφέρει.

```
>show tables;
```

Έτσι βλέπουμε πόσα tables περιέχει η βάση δεδομένων που επιλέξαμε.

```
>desc Bayes;
```

Μπορούμε επίσης να διαπιστώσουμε την αρχιτεκτονική του κάθε table, κοινώς τα «metadata». Η παραπάνω εντολή δίνει τα metadata του Bayes κοκ

Περιεχόμενα βάσης δεδομένων

Προκειμένου να βάλουμε τα δεδομένα μας στη βάση δεδομένων δημιουργήθηκε η ανάγκη για την εύρεση μιας μοναδικής ονομασίας των αλληλουχιών του εκάστοτε τσιπ που εξετάζεται προκειμένου να μπορεί να επιτευχθεί σύγκριση της έκφρασης ανάμεσα σε διαφορετικές πλατφόρμες. Αυτό, πέρα από την ευκολία ανάγνωσής και επεξεργασίας που προσφέρει, λύνει και το πρόβλημα που δημιουργείται στις μικροσυστοιχίες όπου ένα probe μπορεί να αντιστοιχίζεται σε ένα ή περισσότερα probeset και εμείς χρειαζόμαστε για κάθε γονίδιο του κάθε δείγματος μία τιμή. Επιλέχθηκε μια αρκετά διαδεδομένη ονοματολογία γονιδίων, η ENSGxxxxxxx. Όλα τα ENSG ID's πάρθηκαν από το Ensembl και ειδικότερα από το Biomart. Το Ensembl είναι ένα πρόγραμμα περιήγησης για τα γονιδιώματα των σπονδυλωτών που υποστηρίζει την έρευνα σε συγκριτικές μελέτες γονιδιωματικής. Το Ensembl περιλαμβάνει εργαλεία όπως BLAST, BLAT, Biomart και VEP για όλα τα υποστηριζόμενα είδη.

Στο Ensembl αρχικά επιλέγουμε μια βάση δεδομένων στη δική μας περίπτωση την Ensembl Genes 86 από όπου θα αντλήσουμε τα ENSG ID's, και φυσικά επιλέγουμε σαν είδος προς μελέτη τον άνθρωπο (Homo Sapiens Genes(GRC.h38.p3)) και κρατάμε τα μοναδικά αποτελέσματα μόνο. Έτσι εξάγουμε ένα αρχείο με τις ζητούμενες ονομασίες αλληλουχιών το ENSG.txt .

Δημιουργούμε ένα script για συντακτική ανάλυση (parser) για να "καθαρίσουμε" το αρχείο και να το διαβάσουμε

```
<?php
$line=file("ENSG.txt");
$linenum=count($line);
for($i=0;$i<$linenum;$i++){
    $line[$i]=rtrim($line[$i],"\n");
    $arr=explode("\t",$line[$i]);
    //echo "$arr[1]";
    if($arr[1]=="") {
        $arr[1]=$arr[0];
    }
    echo "$arr[0]\t$arr[1]\n";
}
?>
```

Λάβαμε λοιπόν τα δεδομένα μας έπειτα από την κατάλληλη αντιστοίχιση που αναφέρθηκε παραπάνω και τα εισάγαμε στη βάση δεδομένων Microarrays (Εικόνα 45) και εφαρμόσαμε τη στατιστική μέθοδο Empirical Bayes. Η μέθοδος αυτή είναι μια διαδικασία που δίνει στατιστικά αποτελέσματα στα οποία η εκτίμηση της κατανομής δίνεται από τα δεδομένα που εισάγονται σε αντίθεση με τη κλασσική Bayesian μέθοδο όπου εισάγεται το είδος κατανομής προτού εισαχθεί οποιοσδήποτε όγκος δεδομένων

Λάβαμε εν συνεχεία τα unadjusted p-values για κάθε ένα από τα γονίδια και έτσι δημιουργήθηκε ο πίνακας Bayes. Είσοδος στην βάση δεδομένων γίνεται μέσω του XAAMP, που αναφέρθηκε πιο πάνω, που το

χρησιμοποιούμε για να έχουμε τον εξυπηρετητή της MySQL (MySQL client). Μετά την εγκατάσταση του XAAMP μεταβαίνουμε στις προχωρημένες ρυθμίσεις των μεταβλητών περιβάλλοντος (environmental variables) προκειμένου να εισάγουμε στο PATH τα php.exe και mysql.exe δεδομένου ότι προγραμματιστικά έτσι δημιουργήθηκε η βάση. Ανατρέχουμε πάντα και στα ανάλογα directory μέσω του τερματικού cmd.

Όσον αφορά τη πλατφόρμα GeneChip της Affymetrix, δε χρειάστηκε ιδιαίτερη επεξεργασία δεδομένου ότι το ίδιο το τσιπ περιλαμβάνει από μόνο του τις ζητούμενες ονομασίες, συνεπώς δημιουργήθηκε ένα script για συντακτική ανάλυση (parser) για την πλατφόρμα της Affymetrix με το παρακάτω script απλά για την εισαγωγή και μορφοποίηση των δεδομένων της:

```
<?php
$line=file("normAffyGSE59861.txt");
$linenum=count($line);
for($i=1;$i<$linenum;$i++){
    $arr1=explode("\t",$line[$i]);
    $ProbeID=str_replace("\"","",$arr1[0]);
    $arr2=explode("_",$ProbeID);
    echo
"$ProbeID\tGeneChip®_PrimeView_Human_Gene_Expression_Arr\t$ProbeID\n";
}
?>
```

Κατά αντιστοιχία, έναν για την Illumina:

```
<?php
$line=file("normalized_52918(HumanHT-12_V4_0_R2_15002873_B).txt");
$linenum=count($line);

$line[0]=rtrim($line[0]);
```

```

$line[0]=str_replace("\'", "'", $line[0]);
//echo "$line[0]\n";
$arr1=explode("\t", $line[0]);
//print_r($arr1);
$arr1num=count($arr1);
for($i=0;$i<$arr1num;$i++){
    $arr2=explode("_", $arr1[$i]);
    // echo "$arr1[$i]\t$arr2[0]\n";
    $GSM[$i]=$arr2[0];
}
//print_r($GSM);
$GSMnum=count($GSM);

for($i=1;$i<$linenum;$i++){
    $line[$i]=rtrim($line[$i]);
    $arr1=explode("\t", $line[$i]);
    $ProbeID=str_replace("\'", "'", $arr1[0]);
    for($j=0;$j<$GSMnum;$j++){
        $k=$j+1;
        echo "$GSM[$j]\t$ProbeID\tHumanHT-
12_V4_0_R2_15002873_B\t$arr1[$k]\n";
    }
}
?>

```

Και τέλος ένας για Agilent :

```
<?php
$line=file("Agilent-014850_alt.txt");
$linenum=count($line);
for($i=23;$i<$linenum;$i++){
    $line[$i]=rtrim($line[$i],"\n");
    $arr=explode("\t",$line[$i]);
    echo "$arr[0]\tAgilent-014850\t$arr[4]\n";
}
?>
```

Θα προχωρήσουμε σε μια επεξήγηση των tables:

Ο table **ENSG_HGNC** περιλαμβάνει όλα τα ENSG IDs του ανθρώπινου οργανισμού με τα αντίστοιχα HGNCs (gene names) από τη βάση HGNC. Στήλες του και primary keys είναι τα ENSG ID,HGNC. Τα δεδομένα του αντλήθηκαν από το <http://www.genenames.org/cgi-bin/download>. Μέσω του παρακάτω script (ENSG.edit.php) διατηρηθήκαν μόνο τα χρήσιμα στοιχεία, δηλαδή το ENSG και το gene name από το αρχείο του HGNC.

```
<?php
$line=file("ENSG.txt");
$linenum=count($line);
for($i=0;$i<$linenum;$i++){
    $line[$i]=rtrim($line[$i],"\n");
    $arr=explode("\t",$line[$i]);
    //echo "$arr[1]";
    if($arr[1]=="") {
        $arr[1]=$arr[0];
    }
}
```

```

    }

    echo "$arr[0]\t$arr[1]\n";
}
?>

```

Ο table **ENST** περιλαμβάνει όλα τα transcripts (ENST IDs) και τα γονίδια (ENSG IDs) στα οποία αντιστοιχούν. Τα δεδομένα του αντλήθηκαν από το Biomart. '<http://www.ensembl.org/biomart/martview/>'

Ο Table **ENST_Seq** περιλαμβάνει όλα τα transcripts (ENST IDs) και οι αντίστοιχες αλληλουχίες τους από Biomart. Στήλες του είναι οι ENST ID, ENST Sequence. Ο συγκεκριμένος πίνακας δημιουργήθηκε με το παρακάτω script (ENST_Seq.php) το οποίο αφαιρεί από το format FASTA το new line character '\n' και έτσι έχουμε ενιαία αλληλουχία ανά transcript στα οποία μπορούμε να εφαρμόσουμε και συγκρίσεις Strings, πχ να δούμε αν ένα probe είναι substring μιας ακολουθίας.

```

<?php

//php -d memory_limit=2G ENST.php

$line=file("ENST.fas");

$linenum=count($line);

$flag=0;

for($i=0;$i<$linenum;$i++){

    $line[$i]=rtrim($line[$i]);

    //echo "$line[$i]\n";

    if(!strncmp($line[$i], ">", 1)){

        if($flag) {

            echo "$ENST\t$seq\n";

        }

        //echo "HEADER!\n";

        $ENST=str_replace(">", "", $line[$i]);

        //echo "$ENST\n";

```

```

        $seq="";
        $flag=1;
    }
    else {
        $seq.=$line[$i];
    }
}
echo "$ENST\t$seq\n";
?>

```

Ο Table **nu** περιέχει όλες τις αλληλουχίες των ολιγονουκλεοτιδίων όλων των probes από όλα τα chips Illumina και της Agilent με τα αντίστοιχα transcripts που κωδικοποιήσαμε σε μορφή nuID (όπου nu_ID η 64-bit αναπαράσταση της αλληλουχίας του ολιγονουκλεοτιδίου, όπου μια διαδοχική τριπλέτα αναπαρίσταται με ένα χαρακτήρα) που αναγνωρίζουν. Στήλες του είναι οι nuID και Probe Sequence. Δημιουργήθηκε με το script (seq2id.php) που ακολουθεί και κωδικοποιεί μία αλληλουχία και τη μετατρέπει σε ένα μοναδικό κωδικό, ο οποίος έχει μικρότερο μέγεθος οπότε τρέχουν πιο γρήγορα οι εντολές της SQL.

```

<?php
// $id=seq2id($seq);

function seq2id($seq) {
    for($i=65;$i<91;$i++){
        $j=$i-65;
        $code[$j]=chr($i);
    }
    for($i=97;$i<123;$i++){
        $j=$i-71;
        $code[$j]=chr($i);
    }
}

```



```

}
for($i=48;$i<58;$i++){
    $j=$i+4;
    $code[$j]=chr($i);
}
$code[62]="_";
$code[63]=".";

$seq = strtoupper($seq);

$nc2nm["A"]=0;
$nc2nm["C"]=1;
$nc2nm["G"]=2;
$nc2nm["T"]=3;
$nc2nm["U"]=3;

$len=strlen($seq);

for($i=0;$i<$len;$i++){
    $num[$i]=$nc2nm[$seq[$i]];
}

if($len%3) {
    $appLen=3-$len%3;

```

```

    for($i=$len;$i<$len+$appLen;$i++){
        $num[$i]=0;
    }
}
else {
    $appLen=0;
}
$len+=$appLen;

$ncol=$len/3;
$checkSum=0;
for($i=0;$i<$ncol;$i++){
    $code64[$i]=0;
    for($j=0;$j<3;$j++){
        $exp=2-$j;
        $k=$i*3+$j;
        $code64[$i]+=$num[$k]*pow(4,$exp);
    }
    $checkSum+=$code64[$i];
}

$res = $checkSum%21;

$checkCode = $code[$res * 3 + $appLen];

```

```

$id=$checkCode;

for($i=0;$i<$ncol;$i++){
    $id.=$code[$code64[$i]];
}

return $id;
}
?>

```

Ο Table ENST_nu περιέχει όλα τα transcripts (ENST IDs) και τις αντίστοιχες μοναδικές υπογραφές τους (nu IDs). Στήλες του είναι τα ENST ID και nu ID. Το SQL query για τη δημιουργία αυτού πίνακα αντιστοιχίζει τις ακολουθίες των probes από τον πίνακα nu με τις ακολουθίες των Transcripts από τον πίνακα ENST_Seq, και όπου υπάρχει απόλυτη αντιστοίχιση, μας επιστρέφει το Transcript ID (ENST) του πίνακα ENST_Seq και το αντίστοιχο nuID του πίνακα nu.

Ο Table ProbesNU περιλαμβάνει όλα τα μοναδικά και sorted nuIDs από τα probes της Illumina και της Agilent και τα αντίστοιχα τους ProbeIDs και επί της ουσίας υλοποιεί μέσω του παρακάτω script (Probes.php) την επεξεργασία όλων των Manifest files των πλατφορμών και την εξαγωγή των απαραίτητων στοιχείων, δηλαδή του ProbeID και της ακολουθίας του, η οποία μετατρέπεται παράλληλα σε μορφή nuID με το script seq2id.php.

```

<?php

//php -d memory_limit=512M Probes.php

include("seq2id.php");

$fp1 = fopen("Probes.txt", "w");

```

```

$fp2 = fopen("nu.txt", "w");

/***** Parsing Agilent *****/

$manifest=array("Agilent-014850", "Agilent-026652","Agilent-012391","Agilent-
014850_alt");

$startline=array(18,18,21,23);

$probecol=array(1,1,4,4);

$seqcol=array(16,16,19,19);

$manifestnum=count($manifest);

for($j=0;$j<$manifestnum;$j++){

    $line=file("Platform Tables\Agilent\\$manifest[$j].txt");

    $linenum=count($line);

    for($i=$startline[$j];$i<$linenum;$i++){

        $line[$i]=rtrim($line[$i],"\n");

        $line[$i]=rtrim($line[$i],"\r");

        $arr=explode("\t",$line[$i]);

        if($arr[$seqcol[$j]]!="") {

            $nuid=seq2id($arr[$seqcol[$j]]);

            fwrite($fp1,$arr[$probecol[$j]]."\t$nuid\n");

            fwrite($fp2,"$nuid\t".$arr[$seqcol[$j]]."\n");

        }

    }

}

}

```

```

$line=file("Platform Tables\Illumina\HumanWG-6\Human_WG-6.csv");
$linenum=count($line);
for($i=1;$i<$linenum;$i++){
    $arr=explode(",",$line[$i]);
    $nuid=seq2id($arr[9]);
    fwrite($fp1,"$arr[1]\t$nuid\n");
    fwrite($fp2,"$nuid\t$arr[9]\n");
}

```

```

$manifest=array("Platform Tables\Illumina\HumanRef-
8_V3_0_R3_11282963_A.txt", "Platform Tables\Illumina\HumanWG-6\HumanWG-
6_V2_0_R4_11223189_A.txt", "Platform Tables\Illumina\HumanWG-6\HumanWG-
6_V3_0_R3_11282955_A.txt", "Platform Tables\Illumina\HumanHT-12\HumanHT-
12_V3_0_R3_11283641_A.txt", "Platform Tables\Illumina\HumanHT-12\HumanHT-
12_V4_0_R2_15002873_B.txt");
$manifestnum=count($manifest);
for($i=0;$i<$manifestnum;$i++){
    $marr=explode("\\",$manifest[$i]);
    $marrnum=count($marr);
    if($marrnum==1){
        $source=$marr[0];
    }
    else {
        $source=$marr[1];
    }
    $source=str_replace(".txt","",$source);
}

```

```

//echo "$source\n";
//echo "$manifest[$i]\n";
$line=file($manifest[$i]);
//echo "$line[0]";
$linenum=count($line);
//echo "$linenum\n";
$probes=0;
$controls=0;
for ($j=0;$j<$linenum;$j++){
    //echo $line[$j];
    $line[$j]=rtrim($line[$j]);
    if(!strcmp("[Probes]", $line[$j])) {
        $probes=1;
        $controls=0;
        $j+=2;
        $line[$j]=rtrim($line[$j]);
    }
    else if(!strcmp("[Controls]", $line[$j])) {
        $probes=0;
        $controls=1;
        $j+=2;
        $line[$j]=rtrim($line[$j]);
    }
    else if(!strcmp("[Columns]", $line[$j])) {

```

```

        break;
    }

    $arr=explode("\t",$line[$j]);
    //echo "###$line[$j]\n";
    if($probes && !$controls) {
        $nuid=seq2id($arr[17]);
        //echo "$arr[13]\t$source\t$arr[17]\n";
        fwrite($fp1,"$arr[13]\t$nuid\n");
        fwrite($fp2,"$nuid\t$arr[17]\n");
    }
    else if(!$probes && $controls) {
        $nuid=seq2id($arr[5]);
        //echo "$arr[0]\t$source\t$arr[5]\n";
        fwrite($fp1,"$arr[0]\t$nuid\n");
        fwrite($fp2,"$nuid\t$arr[5]\n");
    }
}

fclose($fp1);
fclose($fp2);
?>

```

Ο Table Expression περιέχει την κανονικοποιημένη τιμή έκφρασης που αντιστοιχεί σε κάθε FeatureID για κάθε GSM και για κάθε Plattform. Στήλες του τα GSM, FeatureID, Plattform, Expression. Με τα κάτωθι script σε

R έγινε η δημιουργία και εξαγωγή του FeatureID (τα rows & columns σχηματίζουν ένα μοναδικό αριθμό για κάθε probe στην Agilent, το FeatureID (ή SpotID)) και επεξεργασία για κατάλληλη μορφή.

```
setwd("C:/Users/*link to the working directory")

# create the list of files in the working directory by the pattern in the
expression data file titles

list012391<-list.files(path = , pattern = "012391")

#function "funct" to manipulate the data

#Usage: funct(x), where x is the list of normalized expression data files

funct_012391<-function(x) {

#Read in the data and create SpotID empty matrix with rows equal to the number
of probes in the normalized.txt

exprs<-read.table(x, header=T)

spotID<-matrix(nrow=nrow(exprs),ncol = 1)

#Convert Col/Row content to numeric characters and store them in exprs matrix

exprs[, 1] <- as.numeric(as.character( exprs[, 1] ))

exprs[, 2] <- as.numeric(as.character( exprs[, 2] ))

#Loop for each probe and calculate the SpotID value, then store in in the
SpotID matrix and bind it with the data

for (i in 1:nrow(spotID))

{spotID[i,1]<-215*(exprs[i,1]-1)+exprs[i,2]}

expression<-cbind(spotID,exprs)

#Create expression2 matrix, containing only the GSMs and the expressions

expression2<-expression[,-(1:3)]

#Set number of samples

number_of_samples<-ncol(expression2)
```



```

#Make final format matrix with as many rows as number of samples * number of
probes in the platform

expression_formatted<-matrix(ncol=4,nrow=number_of_samples*nrow(expression2))

#Fill first column with the SpotIDs

expression_formatted[,2]<-expression[,1]

#Blocks are the pointers that separate the GSMs

blocks<-seq(0, nrow(expression_formatted), nrow(expression))

#Loop to fill the 2nd column of the matrix with the GSM blocks and the 3rd
with the GSM ID

for (i in 1:ncol(expression2))

expression_formatted[blocks[i]+1:blocks[2],4]<-expression2[,i]

expression_formatted[blocks[i]+1:blocks[2],1]<-
strsplit(colnames(expression2[i]), "_")[[1]][1]

expression_formatted[,3]<-"Agilent-012391"

}

#Set appropriate column names and write the output tables to a file

#colnames(expression_formatted)<-c("SpotID","Expression","GSM")

write.table(expression_formatted, file=paste0("Expression_Table",x), quote=F,
row.names=F, col.names=F, sep = "\t")

}

#Use lapply function apply the above function 'funct' to all the normalized
files files listed

lapply(list012391, funct)

```

```

setwd("C:/Users/*link to the working directory")

# create the list of files in the working directory by the pattern in the
expression data file titles

list014850<-list.files(path = , pattern = "014850")

#function "funct" to manipulate the data

#Usage: funct(x), where x is the list of normalized expression data files

funct_014850<-function(x) {

#Read in the data and create SpotID empty matrix with rows equal to the number
of probes in the normalized.txt

exprs<-read.table(x, header=T)

spotID<-matrix(nrow=nrow(exprs),ncol = 1)

#Convert Col/Row content to numeric characters and store them in exprs matrix

exprs[, 1] <- as.numeric(as.character( exprs[, 1] ))

exprs[, 2] <- as.numeric(as.character( exprs[, 2] ))

#Loop for each probe and calculate the SpotID value, then store in in the
SpotID matrix and bind it with the data

for (i in 1:nrow(spotID))

{spotID[i,1]<-85*(exprs[i,1]-1)+exprs[i,2]}

expression<-cbind(spotID,exprs)

#Create expression2 matrix, containing only the GSMs and the expressions

expression2<-expression[,-(1:3)]

#Set number of samples

number_of_samples<-ncol(expression2)

```

```

#Make final format matrix with as many rows as number of samples * number of
probes in the platform

expression_formatted<-matrix(ncol=4,nrow=number_of_samples*nrow(expression2))

#Fill first column with the SpotIDs

expression_formatted[,2]<-expression[,1]

#Blocks are the pointers that separate the GSMs

blocks<-seq(0, nrow(expression_formatted), nrow(expression))

#Loop to fill the 2nd column of the matrix with the GSM blocks and the 3rd
with the GSM ID

for (i in 1:ncol(expression2))

{

expression_formatted[blocks[i]+1:blocks[2],4]<-expression2[,i]

expression_formatted[blocks[i]+1:blocks[2],1]<-
strsplit(colnames(expression2[i]), "_")[[1]][1]

expression_formatted[,3]<-"Agilent-014850"

}

#Set appropriate column names and write the output tables to a file

#colnames(expression_formatted)<-c("SpotID", "Expression", "GSM")

write.table(expression_formatted, file=paste0("Expression_Table",x), quote=F,
row.names=F, col.names=F, sep = "\t")

}

#Use lapply function apply the above function 'funct' to all the normalized
files files listed

lapply(list014850, funct)

```

```

setwd("C:/Users/*link to the working directory")

# create the list of files in the working directory by the pattern β€□vsβ€™
in the expression data file titles

list026652<-list.files(path = , pattern = "026652")

#function "funct" to manipulate the data

#Usage: funct(x), where x is the list of normalized expression data files

funct_026652<-function(x) {

#Read in the data and create SpotID empty matrix with rows equal to the number
of probes in the normalized.txt

exprs<-read.table(x, header=T)

spotID<-matrix(nrow=nrow(exprs),ncol = 1)

#Convert Col/Row content to numeric characters and store them in exprs matrix

exprs[, 1] <- as.numeric(as.character( exprs[, 1] ))

exprs[, 2] <- as.numeric(as.character( exprs[, 2] ))

#Loop for each probe and calculate the SpotID value, then store in in the
SpotID matrix and bind it with the data

for (i in 1:nrow(spotID))

{spotID[i,1]<-85*(exprs[i,1]-1)+exprs[i,2]}

expression<-cbind(spotID,exprs)

#Create expression2 matrix, containing only the GSMs and the expressions

expression2<-expression[,-(1:3)]

```

```

#Set number of samples
number_of_samples<-ncol(expression2)

#Make final format matrix with as many rows as number of samples * number of
probes in the platform

expression_formatted<-matrix(ncol=4,nrow=number_of_samples*nrow(expression2))

#Fill first column with the SpotIDs

expression_formatted[,2]<-expression[,1]

#Blocks are the pointers that separate the GSMs

blocks<-seq(0, nrow(expression_formatted), nrow(expression))

#Loop to fill the 2nd column of the matrix with the GSM blocks and the 3rd
with the GSM ID

for (i in 1:ncol(expression2))
{
expression_formatted[blocks[i]+1:blocks[2],4]<-expression2[,i]

expression_formatted[blocks[i]+1:blocks[2],1]<-
strsplit(colnames(expression2[i]), "_")[[1]][1]

expression_formatted[,3]<-"Agilent-026652"

}

#Set appropriate column names and write the output tables to a file

#colnames(expression_formatted)<-c("SpotID","Expression","GSM")

write.table(expression_formatted, file=paste0("Expression_Table",x), quote=F,
row.names=F, col.names=F, sep = "\t")

}

#Use lapply function apply the above function 'funct' to all the normalized
files files listed

lapply(list026652, funct_026652)

```

Ο Table **Features** περιέχει όλα τα FeatureIDs για κάθε platform και τα αντίστοιχα ProbeIDs τους. Εδώ έγινε η συντακτική ανάλυση (παρσάρισμα) των Manifest files από τα platforms και διατηρήθηκε το Feature ID για τα Agilent data ή το Probe ID για τα Illumina data στην 1η στήλη, συμπληρώθηκε το platform name στη 2η στήλη και διατηρήθηκε το Probe ID και εισήχθη στην 3η στήλη. Για τα Illumina data και τα Affymetrix data, η 1η και η 3η στήλη είναι ίδιες αφού τα Illumina data δεν έχουν feature ID. Στήλες του τα FeatureID, Platform και ProbeID.

Ο Table **FeatureSet** περιέχει την αντιστοίχιση όλων των Feature IDs από κάθε platform με τα ENSG IDs τους. Αυτό έγινε με την εξής σχετικά πολύπλοκη διαδικασία, αντιστοιχίζουμε αρχικά τα ENST του πίνακα ENST με τα ENST, του πίνακα ENST_nu, έτσι καταφέρνουμε να αντιστοιχήσουμε κάθε ENSG από τον πίνακα ENST με ένα nuID από το πίνακα ENST_nu. Στη συνέχεια αντιστοιχίζουμε τα nuID του πίνακα ENST_nu με τα nuID από τον πίνακα Probes_nu και αντιστοιχίζονται τα ENSG της προηγούμενης αντιστοίχισης με τα ProbeIDs του πίνακα ProbesNU. Τέλος αντιστοιχίζουμε τα ProbeIDs του πίνακα ProbesNU με τα ProbeIDs του πίνακα Features και έτσι αντιστοιχίζουμε τα ENSG με το αντίστοιχο FeatureID και το Platform από τον πίνακα Features. Με αυτή τη διαδικασία έχουμε δημιουργήσει τα στοιχεία του πίνακα Features με στήλες τα ENSG ID, FeatureID και Platform.

Ο Table **ProcExpr** περιέχει την αντιστοίχιση κάθε ενός ENSG ID από κάθε GSM με ένα υπολογισμένο ένα σταθμισμένο μέσο (με βάση την διάμεσο) στην έκφραση μέσω του αλγόριθμου Tukey. Στήλες του τα GSM, ENSG και Expression.

Ο Αλγόριθμος One-Step Tukey υπολογίζει μια νέα σταθμισμένη μέση τιμή στις τιμές έκφρασης με καλύτερη προσέγγιση, αφού πολλαπλασιάζει τις τιμές έκφρασης με ένα συντελεστή βαρύτητας ανάλογα τη θέση τους σε σχέση με τη διάμεσο, αποφεύγοντας έτσι τις σχετικά απομακρυσμένες τιμές που αλλοιώνουν την πραγματική τιμή της έκφρασης και υπολογίζεται ως εξής:

Υπολογίζουμε τη διάμεσο M των τιμών x_i , $i=1,\dots,n$, έπειτα υπολογίζουμε την απόλυτη απόκλιση της διαμέσου S , ως τη διάμεσο των απολύτων τιμών των αποστάσεων από την M . Η ομοιόμορφη μέτρηση της απόστασης κάθε τιμής i από το κέντρο, είναι:

$$u_i = \frac{x_i - M}{cS + \varepsilon}, i = 1, \dots, n$$

όπου: c (προεπιλογή $c = 5$) και ε μια πολύ μικρή τιμή για την αποφυγή της διαίρεσης δια του μηδέν (προεπιλογή $\varepsilon = 0.0001$). Στη συνέχεια υπολογίζουμε τα βάρη με τον τύπο που ακολουθεί:

$$w(u_i) = \begin{cases} (1 - u_i^2)^2, & |u_i| \leq 1 \\ 0, & |u_i| > 1 \end{cases}$$

Και τέλος υπολογίζουμε τη βεβαρημένη μέση τιμή:

$$T_{bi} = \frac{\sum_{i=1}^n w(u_i) x_i}{\sum_{i=1}^n w(u_i)}$$

Ο Table **Sample** περιέχει όλα τα GSM από κάθε platform και τα GSE με τα αντίστοιχα conditions. Δημιουργήθηκε χειροκίνητα στο excel και στήλες του είναι τα GSM, Platform, Study (GSE) και Condition (High - Low - Ctrl) (Πίνακας 6).

GSM	Platform Name	Study	Condition
GSM725372	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Control
GSM725373	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Control
GSM725374	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Control
GSM725375	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Control
GSM725376	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Control

GSM725377	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Control
GSM725378	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Control
GSM725379	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Control
GSM725380	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Control
GSM725381	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Control
GSM725384	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Low(10cGy)
GSM725385	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Low(10cGy)
GSM725386	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Low(10cGy)
GSM725387	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Low(10cGy)
GSM725388	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Control
GSM725389	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Control

GSM725392	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Low(100cGy)
GSM725393	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Low(100cGy)
GSM725394	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Low(100cGy)
GSM725395	Illumina HumanRef-8 v3.0 expression beadchip	GSE29344	Low(100cGy)
GSM1278163	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Control
GSM1278164	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Control
GSM1278165	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Control
GSM1278166	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Control
GSM1278167	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Control
GSM1278168	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Control
GSM1278169	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Control

GSM1278170	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Control
GSM1278171	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Low(10cGy)
GSM1278172	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Low(10cGy)
GSM1278173	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Low(10cGy)
GSM1278174	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Low(10cGy)
GSM1278175	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Low(10cGy)
GSM1278176	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Low(10cGy)
GSM1278177	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Low(10cGy)
GSM1278178	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Low(10cGy)
GSM1278179	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Low(200cGy)
GSM1278180	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Low(200cGy)

GSM1278181	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Low(200cGy)
GSM1278182	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Low(200cGy)
GSM1278183	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Low(200cGy)
GSM1278184	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Low(200cGy)
GSM1278185	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Low(200cGy)
GSM1278186	Illumina HumanHT-12 V4.0 expression beadchip	GSE52918	Low(200cGy)
GSM1448246	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	Control
GSM1448247	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	Control
GSM1448248	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	Control
GSM1448249	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	Control
GSM1448250	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	High
GSM1448251	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	High
GSM1448252	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	High
GSM1448253	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	High
GSM1448254	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	High
GSM1448255	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	High

GSM1448256	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	High
GSM1448257	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	High
GSM1448258	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	High
GSM1448259	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	High
GSM1448260	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	High
GSM1448261	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	High
GSM1448262	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	Low
GSM1448263	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	Low
GSM1448264	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	Low
GSM1448265	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	Low
GSM1448266	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	Low
GSM1448267	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	Low
GSM1448268	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	Low
GSM1448269	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	Low
GSM1448270	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	Low
GSM1448271	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	Low
GSM1448272	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	Low
GSM1448273	GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GSE59861	Low

Πίνακας 6: Πίνακας Sample

Ο Table Plattform περιέχει κάθε πλατφόρμα και το αντίστοιχο GPL του. Επίσης δημιουργήθηκε χειροκίνητα στο Excel και φαίνεται παρακάτω (Πίνακας 7). Στήλες του είναι τα Plattform και Source.

Platform Name	GPL
Illumina HumanHT-12 V4.0 expression beadchip	GPL10558
Illumina HumanRef-8 v3.0 expression beadchip	GPL6883
GeneChip® PrimeView™ Human Gene Expression Array (with External spike-in RNAs)	GPL16043

Πίνακας 7: Πίνακας Plattform.

Ο Table **Average** περιέχει για κάθε γονίδιο το GSE του και το μέγεθος δείγματος του GSE, τη μέση τιμή έκφρασης του, και τη τυπική του απόκλιση σε κάθε κατάσταση. Εδώ το SQL query δημιουργεί όλα τα περιγραφικά στατιστικά που χρειάζονται για την δημιουργία του πίνακα Average. Στήλες του είναι τα ENSG, Study (GSE), Condition, Mean, SD, Sample size (N).

Ο Table **Bayes** περιέχει όλα τα ENSG από κάθε GSE σε όλες τις καταστάσεις τους (ανά δύο) και τα αντίστοιχα (unadjusted) p-values. Στήλες του είναι τα ENSG, Study (GSE), Condition1, Condition2, p-value.

ΜΕΤΑ-ΑΝΑΛΥΣΗ

Έπειτα από την επεξεργασία των δεδομένων μας στη βάση δεδομένων Microarrays (Εικόνα 45), χωρίσαμε τα δεδομένα μας σε δύο group ωρών (0-16hrs) και (16-24hrs) και υλοποιήσαμε την μετα-ανάλυση με το STATA (Boston and Sumner, 2003). Υλοποιήθηκαν οι μέθοδοι διόρθωσης πολλαπλών ερωτημάτων: Bonferroni, Bonferroni Step-Down, Benjamini-Hochberg, Holland και Sidak και δε βρέθηκαν στατιστικά σημαντικά διαφορικά εκφρασμένα γονίδια. Επιλέχθηκαν στην πορεία όπως αναλύεται σε κάθε μελέτη τα γονίδια με **p-value<0.01** για τις υψηλές δόσεις (Πίνακας 8) και για τις χαμηλές δόσεις (Πίνακας 9).

ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΑΝΑΖΗΤΗΣΗ

Για τη σύγκριση των πειραματικών αποτελεσμάτων με την υπάρχουσα βιβλιογραφία πάνω στο θέμα που μελετάμε, χρησιμοποιήθηκαν και τα γονίδια που βρήκαμε στο Quertle (Giglia, 2011; Hymel, 2011) για μελέτες ή πειράματα, σχετικά με την έκθεση σε ιοντίζουσες ακτινοβολίες, στα οποία δε χρησιμοποιήθηκαν δεδομένα μικροσυστοιχιών ή αυτά δε συμπεριλαμβάνονταν στο GEO. Τα διαφορεικά εκφραζόμενα γονίδια στις μελέτες αυτές (Πίνακας 17, Πίνακας 18) θα συγκριθούν στη συνέχεια με τα δικά μας αποτελέσματα. Εκτελέσαμε την αναζήτηση: \$protein low dose radiation human. Για να είναι δυνατή η σύγκριση των ευρημάτων από τις μελέτες γονιδιακής απόκρισης που βρέθηκαν αντιστοιχίσαμε την ονοματολογία που χρησιμοποίησαν με το τελευταίο όνομα στο HGNC (Bruford et al., 2008; Povey et al., 2001) για κάθε γονίδιο, αναζητώντας το πρωταρχικό γονίδιο ή το μετάγραφο του στο GeneCards (Safran et al., 2010) και το GenBank (Kiehl et al., 2009) ή χρησιμοποιώντας την εφαρμογή BLAT (Kent, 2002) στο Ensembl (Cunningham et al., 2015) Επιπλέον χρησιμοποιώντας τη λίστα cancer gene sensus την οποία κατεβάσαμε από τον κατάλογο COSMIC (<http://cancer.sanger.ac.uk/cosmic>) (Forbes et al., 2008; Forbes et al., 2011), σχεδιάστηκε το ανάλογο διάγραμμα Venn (Εικόνα 52) με τα γονίδια που προέκυψαν από αυτή την ανάλυση και σε συνδυασμό με τη βιβλιογραφία.

ΑΠΟΤΕΛΕΣΜΑΤΑ

Από τη μελέτη της μετα-ανάλυσης προέκυψαν οι δύο λίστες των διαφορεικά εκφρασμένων γονιδίων η πρώτη για τις χαμηλές δόσεις ακτινοβολίας (Πίνακας 8) και η δεύτερη για τις υψηλές (Πίνακας 9). Έπειτα χρησιμοποιήθηκε το WebGestalt (WEB-based GENE SeT AnaLysis Toolkit) (Zhang et al., 2005) διαδικτυακό εργαλείο για την ερμηνεία των λιστών των διαφορεικά αυτών εκφρασμένων γονιδίων. Το WebGestalt είναι σχεδιασμένο για τη πρακτική ανάλυση γενωμικών, πρωτεωμικών και μεγάλης κλίμακας γενετικών μελετών (Zhang et al., 2005) από τις οποίες παράγεται μεγάλος αριθμός λιστών γονιδίων (πχ διαφορεικά εκφρασμένων γονιδίων, συνεκφρασμένων γονιδίων κλπ) μέσω εκθετικών κατανομών και FDR ή άλλων μεθόδων εντοπίζει τους υπερεκπροσωπημένους όρους από γονιδιακές λίστες όπως: GO (Gene Ontology, 2015), KEGG (Kanehisa, 2002), WikiPathways (Kutmon et al., 2016), TransFac (Matys et al., 2003) και Pathway Commons (Cerami et al., 2011). Το WebGestalt επιλέχθηκε ανάμεσα σε άλλα αντίστοιχα διαδικτυακά εργαλεία ανάλυσης έπειτα από μελέτη και διαπίστωση ότι διαθέτει ενημερωμένες λίστες γονιδίων, βιβλιοθήκες για όλα τα δημοφιλή διαθέσιμα chip των microarrays, πολύ καλή οπτικοποίηση των αποτελεσμάτων με αρκετούς διαφορετικούς τρόπους και πλήθος λειτουργιών που δε καταφέραμε να βρούμε σε άλλες αναζητήσεις. Ο κάθε πίνακας παρέχει:

C: τον αριθμό γονιδίων από τη λίστα αναφοράς στην κατηγορία GO,

O: τον αριθμό γονιδίων στο σύνολο διαφορεικά εκφρασμένων γονιδίων,

E: τον αναμενόμενο αριθμό παρατηρήσεων στην κατηγορία διαφορεικά εκφρασμένων και στη λίστα αναφοράς,

R: το λόγο εμπλουτισμού ($R=O/E$),

RawP: την τιμή p από την υπεργεωμετρική δοκιμή,

AdjP: τιμή p προσαρμοσμένη με διόρθωση πολλαπλών ερωτημάτων.

**ΑΠΟΤΕΛΕΣΜΑΤΑ ΜΕΤΑ-ΑΝΑΛΥΣΗΣ ΓΙΑ ΥΨΗΛΕΣ ΚΑΙ ΧΑΜΗΛΕΣ ΔΟΣΕΙΣ ΙΟΝΤΙΖΟΥΣΑΣ
ΑΚΤΙΝΟΒΟΛΙΑΣ**

Γονίδια που ξεχώρισαν έπειτα από έκθεση σε υψηλή δόση ακτινοβολίας στο group 16-24 ωρών

ENSG	HGNC	Description	Diff	P-value
ENSG00000168906	MAT2A	methionine adenosyltransferase 2A [Source:HGNC Symbol;Acc:HGNC:6904]	1.60496	2.60E-06
ENSG00000138764	CCNG2	cyclin G2 [Source:HGNC Symbol;Acc:HGNC:1593]	-1.60152	0.000012
ENSG00000111432	FZD10	frizzled class receptor 10 [Source:HGNC Symbol;Acc:HGNC:4039]	1.42858	0.000015
ENSG00000148572	NRBF2	nuclear receptor binding factor 2 [Source:HGNC Symbol;Acc:HGNC:19692]	-1.54437	0.000023
ENSG00000249884	RNF103-CHMP3	RNF103-CHMP3 readthrough [Source:HGNC Symbol;Acc:HGNC:38847]	-1.53889	0.000023
ENSG00000115561	CHMP3	charged multivesicular body protein 3 [Source:HGNC Symbol;Acc:HGNC:29865]	-1.37543	0.000028
ENSG00000075785	RAB7A	RAB7A, member RAS oncogene family [Source:HGNC Symbol;Acc:HGNC:9788]	-1.34139	0.000037
ENSG00000260735			1.68997	0.000039
ENSG00000282091			1.68997	0.000039
ENSG00000168077	SCARA3	scavenger receptor class A member 3 [Source:HGNC Symbol;Acc:HGNC:19000]	1.33404	0.000043
ENSG00000172216	CEBPB	CCAAT/enhancer binding protein beta [Source:HGNC Symbol;Acc:HGNC:1834]	-1.33193	0.000044
ENSG00000168314	MOBP	myelin-associated oligodendrocyte basic protein [Source:HGNC Symbol;Acc:HGNC:7189]	1.40343	0.000054
ENSG00000047056	WDR37	WD repeat domain 37 [Source:HGNC Symbol;Acc:HGNC:31406]	1.30456	0.000072
ENSG00000141232	TOB1	transducer of ERBB2, 1 [Source:HGNC Symbol;Acc:HGNC:11979]	-1.73714	0.000078
ENSG00000112697	TMEM30A	transmembrane protein 30A [Source:HGNC Symbol;Acc:HGNC:16667]	-1.26883	0.00009
ENSG00000229097	CALM2P2	calmodulin 2 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:1447]	-1.58273	0.000093
ENSG00000165689	SDCCAG3	serologically defined colon cancer antigen 3 [Source:HGNC Symbol;Acc:HGNC:10667]	1.26327	0.000106
ENSG00000131876	SNRPA1	small nuclear ribonucleoprotein polypeptide A' [Source:HGNC Symbol;Acc:HGNC:11152]	2.71065	0.000116
ENSG00000128191	DGCR8	DGCR8, microprocessor complex subunit [Source:HGNC Symbol;Acc:HGNC:2847]	1.49382	0.000121
ENSG00000196757	ZNF700	zinc finger protein 700 [Source:HGNC Symbol;Acc:HGNC:25292]	1.36036	0.000121
ENSG00000188554	NBR1	NBR1, autophagy cargo receptor [Source:HGNC Symbol;Acc:HGNC:6746]	-1.24802	0.000125
ENSG00000273174			-1.53396	0.000131
ENSG00000145945	FAM50B	family with sequence similarity 50 member B [Source:HGNC Symbol;Acc:HGNC:18789]	1.33924	0.000144
ENSG00000136280	CCM2	CCM2 scaffolding protein [Source:HGNC Symbol;Acc:HGNC:21708]	1.21508	0.000151
ENSG00000205352	PRR13	proline rich 13 [Source:HGNC Symbol;Acc:HGNC:24528]	-1.33332	0.000153
ENSG00000184304	PRKD1	protein kinase D1 [Source:HGNC Symbol;Acc:HGNC:9407]	1.23349	0.000156
ENSG00000189067	LITAF	lipopolysaccharide induced TNF factor [Source:HGNC Symbol;Acc:HGNC:16841]	-1.38259	0.000159
ENSG00000211891	IGHE	immunoglobulin heavy constant epsilon [Source:HGNC Symbol;Acc:HGNC:5522]	-1.50153	0.000165
ENSG00000227468			-1.50153	0.000165
ENSG00000276192	IGHE	immunoglobulin heavy constant epsilon [Source:HGNC Symbol;Acc:HGNC:5522]	-1.50153	0.000165
ENSG00000118990	GLRXP3	glutaredoxin pseudogene 3 [Source:HGNC Symbol;Acc:HGNC:34049]	-1.48571	0.000182
ENSG00000147119	CHST7	carbohydrate sulfotransferase 7 [Source:HGNC Symbol;Acc:HGNC:13817]	-1.19961	0.000185
ENSG00000167088	SNRPD1	small nuclear ribonucleoprotein D1 polypeptide [Source:HGNC Symbol;Acc:HGNC:11158]	1.20534	0.00019
ENSG00000213050	TPM3P1	tropomyosin 3 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:16350]	-1.47779	0.000193
ENSG00000168038	ULK4	unc-51 like kinase 4 [Source:HGNC Symbol;Acc:HGNC:15784]	-1.3056	0.000215
ENSG00000136114	THSD1	thrombospondin type 1 domain containing 1 [Source:HGNC Symbol;Acc:HGNC:17754]	1.28732	0.000226
ENSG00000229539			1.43959	0.000266
ENSG00000171469	ZNF561	zinc finger protein 561 [Source:HGNC Symbol;Acc:HGNC:28684]	-1.15924	0.000274
ENSG00000138741	TRPC3	transient receptor potential cation channel subfamily C member 3 [Source:HGNC Symbol;Acc:HGNC:12335]	1.45993	0.000276
ENSG00000168685	IL7R	interleukin 7 receptor [Source:HGNC Symbol;Acc:HGNC:6024]	1.34491	0.000307

ENSG0000012817	KDM5D	lysine demethylase 5D [Source:HGNC Symbol;Acc:HGNC:11115]	-2.41678	0.000328
ENSG00000187242	KRT12	keratin 12 [Source:HGNC Symbol;Acc:HGNC:6414]	2.44839	0.000348
ENSG00000176681	LRRC37A	leucine rich repeat containing 37A [Source:HGNC Symbol;Acc:HGNC:29069]	-4.02556	0.000364
ENSG00000261192	RNF126P1	ring finger protein 126 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:30340]	-3.98456	0.000389
ENSG00000167904	TMEM68	transmembrane protein 68 [Source:HGNC Symbol;Acc:HGNC:26510]	-1.46421	0.00042
ENSG00000183527	PSMG1	proteasome assembly chaperone 1 [Source:HGNC Symbol;Acc:HGNC:3043]	1.11024	0.000441
ENSG00000161860	SYCE2	synaptonemal complex central element protein 2 [Source:HGNC Symbol;Acc:HGNC:27411]	1.30632	0.000443
ENSG00000215571	GRK6P1	G protein-coupled receptor kinase 6 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:4547]	7.72632	0.000449
ENSG00000213366	GSTM2	glutathione S-transferase mu 2 [Source:HGNC Symbol;Acc:HGNC:4634]	1.23072	0.000463
ENSG00000143442	POGZ	pogo transposable element with ZNF domain [Source:HGNC Symbol;Acc:HGNC:18801]	1.12234	0.000466
ENSG00000240771	ARHGEF25	Rho guanine nucleotide exchange factor 25 [Source:HGNC Symbol;Acc:HGNC:30275]	-1.10671	0.000486
ENSG00000136040	PLXNC1	plexin C1 [Source:HGNC Symbol;Acc:HGNC:9106]	-1.11064	0.000487
ENSG00000213614	HEXA	hexosaminidase subunit alpha [Source:HGNC Symbol;Acc:HGNC:4878]	-3.84626	0.000488
ENSG00000164520	RAET1E	retinoic acid early transcript 1E [Source:HGNC Symbol;Acc:HGNC:16793]	1.11664	0.000495
ENSG00000100612	DHRS7	dehydrogenase/reductase 7 [Source:HGNC Symbol;Acc:HGNC:21524]	-1.10804	0.000496
ENSG00000131355	ADGRE3	adhesion G protein-coupled receptor E3 [Source:HGNC Symbol;Acc:HGNC:23647]	-1.12347	0.000497
ENSG00000187808	SOWAHD	soosonowah ankyrin repeat domain family member D [Source:HGNC Symbol;Acc:HGNC:32960]	-1.21936	0.000499
ENSG00000040633	PHF23	PHD finger protein 23 [Source:HGNC Symbol;Acc:HGNC:28428]	-1.09704	0.0005
ENSG00000164687	FABP5	fatty acid binding protein 5 [Source:HGNC Symbol;Acc:HGNC:3560]	1.20753	0.0005
ENSG00000234964	FABP5P7	fatty acid binding protein 5 pseudogene 7 [Source:HGNC Symbol;Acc:HGNC:31070]	1.20753	0.0005
ENSG00000072182	ASIC4	acid sensing ion channel subunit family member 4 [Source:HGNC Symbol;Acc:HGNC:21263]	-1.11488	0.000501
ENSG00000174353	STAG3L3	stromal antigen 3-like 3 (pseudogene) [Source:HGNC Symbol;Acc:HGNC:33845]	1.21111	0.000501
ENSG00000129116	PALLD	palladin, cytoskeletal associated protein [Source:HGNC Symbol;Acc:HGNC:17068]	1.1016	0.000502
ENSG00000254889			1.35375	0.000515
ENSG00000255025			1.35375	0.000515
ENSG00000164649	CDCA7L	cell division cycle associated 7 like [Source:HGNC Symbol;Acc:HGNC:30777]	2.35712	0.000518
ENSG00000281519			1.35244	0.000521
ENSG00000281782			1.35244	0.000521
ENSG00000084112	SSH1	slingshot protein phosphatase 1 [Source:HGNC Symbol;Acc:HGNC:30579]	1.09967	0.000546
ENSG00000159131	GART	phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase [Source:HGNC Symbol;Acc:HGNC:4163]	1.10775	0.000547
ENSG00000169564	PCBP1	poly(rC) binding protein 1 [Source:HGNC Symbol;Acc:HGNC:8647]	-1.11075	0.00055
ENSG00000048392	RRM2B	ribonucleotide reductase regulatory TP53 inducible subunit M2B [Source:HGNC Symbol;Acc:HGNC:17296]	-1.09618	0.000579
ENSG00000111729	CLEC4A	C-type lectin domain family 4 member A [Source:HGNC Symbol;Acc:HGNC:13257]	-1.18058	0.000625
ENSG00000107745	MICU1	mitochondrial calcium uptake 1 [Source:HGNC Symbol;Acc:HGNC:1530]	-1.88276	0.000638
ENSG00000186973	FAM183A	family with sequence similarity 183 member A [Source:HGNC Symbol;Acc:HGNC:34347]	1.18407	0.000638
ENSG00000089053	ANAPC5	anaphase promoting complex subunit 5 [Source:HGNC Symbol;Acc:HGNC:15713]	1.07347	0.000639
ENSG00000245060	LINC00847	long intergenic non-protein coding RNA 847 [Source:HGNC Symbol;Acc:HGNC:45050]	-6.32375	0.00065
ENSG00000135956	TMEM127	transmembrane protein 127 [Source:HGNC Symbol;Acc:HGNC:26038]	-1.07952	0.000655
ENSG00000244041	LINC01011	long intergenic non-protein coding RNA 1011 [Source:HGNC Symbol;Acc:HGNC:33812]	1.17478	0.000658
ENSG00000136068	FLNB	filamin B [Source:HGNC Symbol;Acc:HGNC:3755]	1.0851	0.000668
ENSG00000196597	ZNF782	zinc finger protein 782 [Source:HGNC Symbol;Acc:HGNC:33110]	2.22006	0.000691
ENSG00000166860	ZBTB39	zinc finger and BTB domain containing 39 [Source:HGNC Symbol;Acc:HGNC:29014]	1.27345	0.000692
ENSG00000227115	LINC01630	long intergenic non-protein coding RNA 1630 [Source:HGNC Symbol;Acc:HGNC:52295]	-6.096	0.000704
ENSG00000136003	ISCU	iron-sulfur cluster assembly enzyme [Source:HGNC Symbol;Acc:HGNC:29882]	-1.78229	0.000714

ENSG00000105058	FAM32A	family with sequence similarity 32 member A [Source:HGNC Symbol;Acc:HGNC:24563]	-1.60415	0.000733
ENSG00000128928	IVD	isovaleryl-CoA dehydrogenase [Source:HGNC Symbol;Acc:HGNC:6186]	1.07748	0.000755
ENSG00000138185	ENTPD1	ectonucleoside triphosphate diphosphohydrolase 1 [Source:HGNC Symbol;Acc:HGNC:3363]	-1.16146	0.000755
ENSG00000114770	ABCC5	ATP binding cassette subfamily C member 5 [Source:HGNC Symbol;Acc:HGNC:56]	-1.07064	0.000759
ENSG00000163507	KIAA1524	KIAA1524 [Source:HGNC Symbol;Acc:HGNC:29302]	1.08153	0.00077
ENSG00000280873	DGKD	diacylglycerol kinase delta [Source:HGNC Symbol;Acc:HGNC:2851]	1.22325	0.000772
ENSG00000084693	AGBL5	ATP/GTP binding protein like 5 [Source:HGNC Symbol;Acc:HGNC:26147]	-1.0629	0.0008
ENSG00000259834			1.29683	0.00082
ENSG00000226979	LTA	lymphotoxin alpha [Source:HGNC Symbol;Acc:HGNC:6709]	1.17321	0.000821
ENSG00000244694	PTCHD4	patched domain containing 4 [Source:HGNC Symbol;Acc:HGNC:21345]	-2.23616	0.000862
ENSG00000227765	MYL12BP2	myosin light chain 12B pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:51650]	-1.28728	0.000867
ENSG00000225050			5.54758	0.00088
ENSG00000215859			1.28965	0.000883
ENSG00000070061	IKBKAP	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase complex-associated protein [Source:HGNC Symbol;Acc:HGNC:5959]	1.49722	0.000887
ENSG00000268095	ZNF649-AS1	ZNF649 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:51285]	1.87112	0.000906
ENSG00000274221	CCL3	C-C motif chemokine ligand 3 [Source:HGNC Symbol;Acc:HGNC:10627]	-3.48437	0.000922
ENSG00000278567	CCL3	C-C motif chemokine ligand 3 [Source:HGNC Symbol;Acc:HGNC:10627]	-3.48437	0.000922
ENSG00000110958	PTGES3	prostaglandin E synthase 3 [Source:HGNC Symbol;Acc:HGNC:16049]	2.14229	0.000932
ENSG00000005175	RPAP3	RNA polymerase II associated protein 3 [Source:HGNC Symbol;Acc:HGNC:26151]	-1.04232	0.000944
ENSG00000160072	ATAD3B	ATPase family, AAA domain containing 3B [Source:HGNC Symbol;Acc:HGNC:24007]	1.0367	0.00099
ENSG00000269586	CT45A10	cancer/testis antigen family 45 member A10 [Source:HGNC Symbol;Acc:HGNC:51263]	1.0358	0.000991
ENSG00000162992	NEUROD1	neuronal differentiation 1 [Source:HGNC Symbol;Acc:HGNC:7762]	1.15645	0.001011
ENSG00000065615	CY5R4	cytochrome b5 reductase 4 [Source:HGNC Symbol;Acc:HGNC:20147]	-1.14165	0.001025
ENSG00000155097	ATP6V1C1	ATPase H+ transporting V1 subunit C1 [Source:HGNC Symbol;Acc:HGNC:856]	-1.34064	0.001066
ENSG00000116151	MORN1	MORN repeat containing 1 [Source:HGNC Symbol;Acc:HGNC:25852]	1.03083	0.001072
ENSG00000174827	PDZK1	PDZ domain containing 1 [Source:HGNC Symbol;Acc:HGNC:8821]	1.13623	0.001073
ENSG00000130775	THEMIS2	thymocyte selection associated family member 2 [Source:HGNC Symbol;Acc:HGNC:16839]	-1.74828	0.001078
ENSG00000171161	ZNF672	zinc finger protein 672 [Source:HGNC Symbol;Acc:HGNC:26179]	-1.02715	0.001083
ENSG00000145246	ATP10D	ATPase phospholipid transporting 10D (putative) [Source:HGNC Symbol;Acc:HGNC:13549]	1.04117	0.001093
ENSG00000140691	ARMC5	armadillo repeat containing 5 [Source:HGNC Symbol;Acc:HGNC:25781]	1.02325	0.0011
ENSG00000123219	CENPK	centromere protein K [Source:HGNC Symbol;Acc:HGNC:29479]	1.02793	0.001141
ENSG00000255302	EID1	EP300 interacting inhibitor of differentiation 1 [Source:HGNC Symbol;Acc:HGNC:1191]	1.02836	0.001142
ENSG00000196243	LINC00615	long intergenic non-protein coding RNA 615 [Source:HGNC Symbol;Acc:HGNC:26343]	1.25374	0.001153
ENSG00000280088			-1.24943	0.001187
ENSG00000164576	SAP30L	SAP30 like [Source:HGNC Symbol;Acc:HGNC:25663]	-1.02756	0.001188
ENSG00000124209	RAB22A	RAB22A, member RAS oncogene family [Source:HGNC Symbol;Acc:HGNC:9764]	-1.02519	0.001196
ENSG00000057657	PRDM1	PR/SET domain 1 [Source:HGNC Symbol;Acc:HGNC:9346]	-1.20162	0.001201
ENSG00000182253	SYNM	synemin [Source:HGNC Symbol;Acc:HGNC:24466]	1.03222	0.001203
ENSG00000224706	RPS17P13	ribosomal protein S17 pseudogene 13 [Source:HGNC Symbol;Acc:HGNC:36664]	-1.24619	0.001204
ENSG00000117155	SSX2IP	SSX family member 2 interacting protein [Source:HGNC Symbol;Acc:HGNC:16509]	1.01635	0.001208
ENSG00000171126	KCNGB3	potassium voltage-gated channel modifier subfamily G member 3 [Source:HGNC Symbol;Acc:HGNC:18306]	1.01864	0.001238
ENSG00000255725	TDGP1	thymine-DNA glycosylase pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:49891]	1.24928	0.001238
ENSG00000095627	TDRD1	tudor domain containing 1 [Source:HGNC Symbol;Acc:HGNC:11712]	3.13841	0.001239
ENSG00000171484			-3.32545	0.001243
ENSG00000280094	OR1B1	olfactory receptor family 1 subfamily B member 1 (gene/pseudogene) [Source:HGNC	-3.32545	0.001243

		Symbol;Acc:HGNC:8181]		
ENSG00000102054	RBBP7	RB binding protein 7, chromatin remodeling factor [Source:HGNC Symbol;Acc:HGNC:9890]	1.01058	0.001284
ENSG00000267706			-1.24177	0.001286
ENSG00000127334	DYRK2	dual specificity tyrosine phosphorylation regulated kinase 2 [Source:HGNC Symbol;Acc:HGNC:3093]	1.15937	0.0013
ENSG00000164645	C7orf62	chromosome 7 open reading frame 62 [Source:HGNC Symbol;Acc:HGNC:22402]	-1.01325	0.001304
ENSG00000184319	RPL23AP82	ribosomal protein L23a pseudogene 82 [Source:HGNC Symbol;Acc:HGNC:33730]	-1.10076	0.001316
ENSG00000115806	GORASP2	golgi reassembly stacking protein 2 [Source:HGNC Symbol;Acc:HGNC:17500]	1.01632	0.001331
ENSG00000257335	MGAM	maltase-glucoamylase [Source:HGNC Symbol;Acc:HGNC:7043]	-1.01374	0.001334
ENSG00000173214	MFSD4B	major facilitator superfamily domain containing 4B [Source:HGNC Symbol;Acc:HGNC:21053]	-1.01113	0.001347
ENSG00000255073	ZFP91-CNTF	ZFP91-CNTF readthrough (NMD candidate) [Source:HGNC Symbol;Acc:HGNC:33441]	1.10404	0.001352
ENSG00000255413			3.272	0.001378
ENSG00000118518	RNF146	ring finger protein 146 [Source:HGNC Symbol;Acc:HGNC:21336]	-1.00219	0.001404
ENSG00000255439			1.10546	0.001417
ENSG00000108883	EFTUD2	elongation factor Tu GTP binding domain containing 2 [Source:HGNC Symbol;Acc:HGNC:30858]	1.01475	0.001426
ENSG00000137312	FLOT1	flotillin 1 [Source:HGNC Symbol;Acc:HGNC:3757]	-0.99889	0.001442
ENSG00000204482	LST1	leukocyte specific transcript 1 [Source:HGNC Symbol;Acc:HGNC:14189]	-1.18439	0.001448
ENSG00000204103	MAFB	MAF bZIP transcription factor B [Source:HGNC Symbol;Acc:HGNC:6408]	0.997086	0.00146
ENSG00000215007	DNAJA1P3	DnaJ heat shock protein family (Hsp40) member A1 pseudogene 3 [Source:HGNC Symbol;Acc:HGNC:39339]	-1.21985	0.001465
ENSG00000136161	RCBTB2	RCC1 and BTB domain containing protein 2 [Source:HGNC Symbol;Acc:HGNC:1914]	1.28696	0.001487
ENSG00000197921	HES5	hes family bHLH transcription factor 5 [Source:HGNC Symbol;Acc:HGNC:19764]	1.08651	0.001541
ENSG00000186148			1.20969	0.001577
ENSG00000196792	STRN3	striatin 3 [Source:HGNC Symbol;Acc:HGNC:15720]	-0.98587	0.001598
ENSG00000131724	IL13RA1	interleukin 13 receptor subunit alpha 1 [Source:HGNC Symbol;Acc:HGNC:5974]	-1.00042	0.001601
ENSG00000141968	VAV1	vav guanine nucleotide exchange factor 1 [Source:HGNC Symbol;Acc:HGNC:12657]	-1.10699	0.001605
ENSG00000232307	DAOA-AS1	DAOA antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:30243]	1.08262	0.001607
ENSG00000160445	ZER1	zyg-11 related cell cycle regulator [Source:HGNC Symbol;Acc:HGNC:30960]	-1.03227	0.00161
ENSG00000185989	RASA3	RAS p21 protein activator 3 [Source:HGNC Symbol;Acc:HGNC:20331]	-2.00496	0.001613
ENSG00000127831	VIL1	villin 1 [Source:HGNC Symbol;Acc:HGNC:12690]	-1.6126	0.001635
ENSG00000214021	TTL3	tubulin tyrosine ligase like 3 [Source:HGNC Symbol;Acc:HGNC:24483]	-0.9891	0.001643
ENSG00000250903	GMDS-AS1	GMDS antisense RNA 1 (head to head) [Source:HGNC Symbol;Acc:HGNC:48993]	1.20447	0.001646
ENSG00000163961	RNF168	ring finger protein 168 [Source:HGNC Symbol;Acc:HGNC:26661]	-1.37318	0.001659
ENSG00000270757	HSPE1-MOB4	HSPE1-MOB4 readthrough [Source:HGNC Symbol;Acc:HGNC:49184]	-1.49483	0.001668
ENSG00000172530	BANP	BTG3 associated nuclear protein [Source:HGNC Symbol;Acc:HGNC:13450]	-0.98225	0.001677
ENSG00000117616	RSRP1	arginine and serine rich protein 1 [Source:HGNC Symbol;Acc:HGNC:25234]	-1.07653	0.001684
ENSG00000177943	MAMDC4	MAM domain containing 4 [Source:HGNC Symbol;Acc:HGNC:24083]	-1.0755	0.001693
ENSG00000140319	SRP14	signal recognition particle 14 [Source:HGNC Symbol;Acc:HGNC:11299]	-0.99517	0.001715
ENSG00000239332	LINC01119	long intergenic non-protein coding RNA 1119 [Source:HGNC Symbol;Acc:HGNC:49262]	-4.37277	0.001736
ENSG00000254858	MPV17L2	MPV17 mitochondrial inner membrane protein like 2 [Source:HGNC Symbol;Acc:HGNC:28177]	1.02559	0.001749
ENSG00000156875	MFSD14A	major facilitator superfamily domain containing 14A [Source:HGNC Symbol;Acc:HGNC:23363]	-0.99159	0.001758
ENSG00000261556	SMG1P7	SMG1P7, nonsense mediated mRNA decay associated PI3K related kinase pseudogene 7 [Source:HGNC Symbol;Acc:HGNC:49864]	1.19981	0.001799
ENSG00000138376	BARD1	BRCA1 associated RING domain 1 [Source:HGNC Symbol;Acc:HGNC:952]	1.37109	0.001802
ENSG00000105723	GSK3A	glycogen synthase kinase 3 alpha [Source:HGNC Symbol;Acc:HGNC:4616]	-1.0694	0.001818
ENSG00000236496	GPS2P1	G protein pathway suppressor 2 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:49622]	-1.19538	0.001819
ENSG00000236044	FABP5P2	fatty acid binding protein 5 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:31060]	1.07939	0.001823

ENSG0000004777	ARHGAP33	Rho GTPase activating protein 33 [Source:HGNC Symbol;Acc:HGNC:23085]	0.981514	0.001828
ENSG00000132824	SERINC3	serine incorporator 3 [Source:HGNC Symbol;Acc:HGNC:11699]	-0.9731	0.001849
ENSG00000179447			-1.2734	0.001862
ENSG00000086065	CHMP5	charged multivesicular body protein 5 [Source:HGNC Symbol;Acc:HGNC:26942]	-0.97837	0.001865
ENSG00000206379	FLOT1	flotillin 1 [Source:HGNC Symbol;Acc:HGNC:3757]	-1.06411	0.001867
ENSG00000206480	FLOT1	flotillin 1 [Source:HGNC Symbol;Acc:HGNC:3757]	-1.06411	0.001867
ENSG00000224740	FLOT1	flotillin 1 [Source:HGNC Symbol;Acc:HGNC:3757]	-1.06411	0.001867
ENSG00000230143	FLOT1	flotillin 1 [Source:HGNC Symbol;Acc:HGNC:3757]	-1.06411	0.001867
ENSG00000232280	FLOT1	flotillin 1 [Source:HGNC Symbol;Acc:HGNC:3757]	-1.06411	0.001867
ENSG00000236271	FLOT1	flotillin 1 [Source:HGNC Symbol;Acc:HGNC:3757]	-1.06411	0.001867
ENSG00000144711	IQSEC1	IQ motif and Sec7 domain 1 [Source:HGNC Symbol;Acc:HGNC:29112]	-0.97687	0.00188
ENSG00000236814			-1.1871	0.001893
ENSG00000119950	MXI1	MAX interactor 1, dimerization protein [Source:HGNC Symbol;Acc:HGNC:7534]	-2.00827	0.001896
ENSG00000137393	RNF144B	ring finger protein 144B [Source:HGNC Symbol;Acc:HGNC:21578]	-0.96995	0.001905
ENSG00000139629	GALNT6	polypeptide N-acetylgalactosaminyltransferase 6 [Source:HGNC Symbol;Acc:HGNC:4128]	0.97341	0.001916
ENSG00000259705			1.19318	0.001917
ENSG00000184741	OR5AQ1P	olfactory receptor family 5 subfamily AQ member 1 pseudogene [Source:HGNC Symbol;Acc:HGNC:15259]	1.25125	0.001919
ENSG00000146457	WTAP	Wilms tumor 1 associated protein [Source:HGNC Symbol;Acc:HGNC:16846]	-1.1287	0.001943
ENSG00000197324	LRP10	LDL receptor related protein 10 [Source:HGNC Symbol;Acc:HGNC:14553]	-1.03255	0.001958
ENSG00000171385	KCND3	potassium voltage-gated channel subfamily D member 3 [Source:HGNC Symbol;Acc:HGNC:6239]	-0.97069	0.001971
ENSG00000184208	C22orf46	chromosome 22 open reading frame 46 [Source:HGNC Symbol;Acc:HGNC:26294]	-0.98017	0.002008
ENSG00000128989	ARPP19	cAMP regulated phosphoprotein 19 [Source:HGNC Symbol;Acc:HGNC:16967]	-0.96416	0.002049
ENSG00000101336	HCK	HCK proto-oncogene, Src family tyrosine kinase [Source:HGNC Symbol;Acc:HGNC:4840]	-0.96594	0.002059
ENSG00000248699			-1.17542	0.002063
ENSG00000114302	PRKAR2A	protein kinase cAMP-dependent type II regulatory subunit alpha [Source:HGNC Symbol;Acc:HGNC:9391]	1.09604	0.002071
ENSG00000134884	ARGLU1	arginine and glutamate rich 1 [Source:HGNC Symbol;Acc:HGNC:25482]	0.973074	0.002098
ENSG00000159556	ISL2	ISL LIM homeobox 2 [Source:HGNC Symbol;Acc:HGNC:18524]	-0.96651	0.002103
ENSG00000170365	SMAD1	SMAD family member 1 [Source:HGNC Symbol;Acc:HGNC:6767]	1.87403	0.002123
ENSG00000063854	HAGH	hydroxyacylglutathione hydrolase [Source:HGNC Symbol;Acc:HGNC:4805]	-0.95984	0.002127
ENSG00000251138			-1.17637	0.002128
ENSG00000119865	CNRIP1	cannabinoid receptor interacting protein 1 [Source:HGNC Symbol;Acc:HGNC:24546]	0.964157	0.00214
ENSG00000112210	RAB23	RAB23, member RAS oncogene family [Source:HGNC Symbol;Acc:HGNC:14263]	0.957784	0.002141
ENSG00000013441	CLK1	CDC like kinase 1 [Source:HGNC Symbol;Acc:HGNC:2068]	-0.96382	0.002181
ENSG00000168807	SNTB2	syntrophin beta 2 [Source:HGNC Symbol;Acc:HGNC:11169]	0.954586	0.002201
ENSG00000117318	ID3	inhibitor of DNA binding 3, HLH protein [Source:HGNC Symbol;Acc:HGNC:5362]	0.966287	0.002202
ENSG00000173457	PPP1R14B	protein phosphatase 1 regulatory inhibitor subunit 14B [Source:HGNC Symbol;Acc:HGNC:9057]	1.04727	0.002272
ENSG00000274042	GOLGA6L10	golgin A6 family-like 10 [Source:HGNC Symbol;Acc:HGNC:37228]	1.16178	0.002283
ENSG00000127220	ABHD8	abhydrolase domain containing 8 [Source:HGNC Symbol;Acc:HGNC:23759]	-1.0901	0.002285
ENSG00000123416	TUBA1B	tubulin alpha 1b [Source:HGNC Symbol;Acc:HGNC:18809]	1.932	0.002289
ENSG00000264743	DPRXP4	divergent-paired related homeobox pseudogene 4 [Source:HGNC Symbol;Acc:HGNC:32170]	-3.01711	0.002297
ENSG00000091073	DTX2	deltex E3 ubiquitin ligase 2 [Source:HGNC Symbol;Acc:HGNC:15973]	-0.95497	0.002319
ENSG00000132326	PER2	period circadian clock 2 [Source:HGNC Symbol;Acc:HGNC:8846]	0.964844	0.002319
ENSG00000280381			1.16316	0.00232
ENSG00000100564	PIGH	phosphatidylinositol glycan anchor biosynthesis class H [Source:HGNC Symbol;Acc:HGNC:8964]	0.95167	0.002332

ENSG00000229415	SFTA3	surfactant associated 3 [Source:HGNC Symbol;Acc:HGNC:18387]	-1.15989	0.002334
ENSG00000257520		surfactant associated 3 [Source:EntrezGene;Acc:253970]	-1.15989	0.002334
ENSG00000114742	WDR48	WD repeat domain 48 [Source:HGNC Symbol;Acc:HGNC:30914]	-0.95394	0.00234
ENSG00000188647	PTAR1	protein prenyltransferase alpha subunit repeat containing 1 [Source:HGNC Symbol;Acc:HGNC:30449]	1.04391	0.002378
ENSG00000115649	CNPPD1	cyclin Pas1/PHO80 domain containing 1 [Source:HGNC Symbol;Acc:HGNC:25220]	-0.94631	0.00238
ENSG00000162639	HENMT1	HEN1 methyltransferase homolog 1 [Source:HGNC Symbol;Acc:HGNC:26400]	0.949173	0.00238
ENSG00000134058	CDK7	cyclin dependent kinase 7 [Source:HGNC Symbol;Acc:HGNC:1778]	-0.95203	0.002381
ENSG00000115446	UNC50	unc-50 inner nuclear membrane RNA binding protein [Source:HGNC Symbol;Acc:HGNC:16046]	-1.45862	0.002382
ENSG00000163823	CCR1	C-C motif chemokine receptor 1 [Source:HGNC Symbol;Acc:HGNC:1602]	-0.95162	0.002386
ENSG00000150760	DOCK1	dedicator of cytokinesis 1 [Source:HGNC Symbol;Acc:HGNC:2987]	-1.96075	0.002387
ENSG00000236801	RPL24P8	ribosomal protein L24 pseudogene 8 [Source:HGNC Symbol;Acc:HGNC:36409]	1.15671	0.002399
ENSG00000234814	SVILP1	supervillin pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:44959]	-1.15484	0.002431
ENSG00000249345			1.15751	0.002431
ENSG00000280759	AP2A2	adaptor related protein complex 2 alpha 2 subunit [Source:HGNC Symbol;Acc:HGNC:562]	1.03654	0.002464
ENSG00000281385	AP2A2	adaptor related protein complex 2 alpha 2 subunit [Source:HGNC Symbol;Acc:HGNC:562]	1.03654	0.002464
ENSG00000172548	NIPAL4	NIPA like domain containing 4 [Source:HGNC Symbol;Acc:HGNC:28018]	-3.93533	0.002465
ENSG00000147862	NFIB	nuclear factor I B [Source:HGNC Symbol;Acc:HGNC:7785]	0.943468	0.00247
ENSG00000000938	FGR	FGR proto-oncogene, Src family tyrosine kinase [Source:HGNC Symbol;Acc:HGNC:3697]	-0.94537	0.002486
ENSG00000138433	CIR1	corepressor interacting with RBPJ, 1 [Source:HGNC Symbol;Acc:HGNC:24217]	-1.30258	0.002523
ENSG00000231643			3.90894	0.002523
ENSG00000112081	SRSF3	serine and arginine rich splicing factor 3 [Source:HGNC Symbol;Acc:HGNC:10785]	0.948303	0.002525
ENSG00000226856			1.03305	0.002534
ENSG00000022567	SLC45A4	solute carrier family 45 member 4 [Source:HGNC Symbol;Acc:HGNC:29196]	-0.94429	0.002541
ENSG00000124635	HIST1H2BJ	histone cluster 1 H2B family member j [Source:HGNC Symbol;Acc:HGNC:4761]	-1.66752	0.002611
ENSG00000125844	RRBP1	ribosome binding protein 1 [Source:HGNC Symbol;Acc:HGNC:10448]	4.32233	0.002618
ENSG00000249532			1.15064	0.00262
ENSG00000243479	MNX1-AS1	MNX1 antisense RNA 1 (head to head) [Source:HGNC Symbol;Acc:HGNC:48954]	-1.14718	0.002634
ENSG00000108666	C17orf75	chromosome 17 open reading frame 75 [Source:HGNC Symbol;Acc:HGNC:30173]	1.06632	0.002635
ENSG00000176654	NANOGP1	Nanog homeobox pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:23099]	1.43556	0.002682
ENSG00000174946	GPR171	G protein-coupled receptor 171 [Source:HGNC Symbol;Acc:HGNC:30057]	0.932497	0.002698
ENSG00000120709	FAM53C	family with sequence similarity 53 member C [Source:HGNC Symbol;Acc:HGNC:1336]	-0.94591	0.002705
ENSG00000177192	PUS1	pseudouridylyl synthase 1 [Source:HGNC Symbol;Acc:HGNC:15508]	0.93784	0.002706
ENSG00000197943	PLCG2	phospholipase C gamma 2 [Source:HGNC Symbol;Acc:HGNC:9066]	-0.93104	0.002724
ENSG00000213903	LTB4R	leukotriene B4 receptor [Source:HGNC Symbol;Acc:HGNC:6713]	0.934965	0.002724
ENSG00000185950	IRS2	insulin receptor substrate 2 [Source:HGNC Symbol;Acc:HGNC:6126]	-0.93465	0.002728
ENSG00000089737	DDX24	DEAD-box helicase 24 [Source:HGNC Symbol;Acc:HGNC:13266]	-1.05236	0.002732
ENSG00000125730	C3	complement C3 [Source:HGNC Symbol;Acc:HGNC:1318]	-0.93079	0.002744
ENSG00000011295	TTC19	tetratricopeptide repeat domain 19 [Source:HGNC Symbol;Acc:HGNC:26006]	1.67402	0.002748
ENSG00000177576	C18orf32	chromosome 18 open reading frame 32 [Source:HGNC Symbol;Acc:HGNC:31690]	-1.98605	0.002761
ENSG00000205846	CLEC6A	C-type lectin domain family 6 member A [Source:HGNC Symbol;Acc:HGNC:14556]	1.87371	0.002766
ENSG00000136527	TRA2B	transformer 2 beta homolog (Drosophila) [Source:HGNC Symbol;Acc:HGNC:10781]	1.52372	0.002773
ENSG00000138111	MFS13A	major facilitator superfamily domain containing 13A [Source:HGNC Symbol;Acc:HGNC:26196]	-1.86545	0.002792
ENSG00000197582	GPX1P1	glutathione peroxidase pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:4560]	-2.9219	0.002801
ENSG00000169155	ZBTB43	zinc finger and BTB domain containing 43 [Source:HGNC Symbol;Acc:HGNC:17908]	0.93188	0.002802
ENSG00000184210	DGAT2L6	diacylglycerol O-acyltransferase 2 like 6 [Source:HGNC Symbol;Acc:HGNC:23250]	3.76954	0.002867

ENSG00000168152	THAP9	THAP domain containing 9 [Source:HGNC Symbol;Acc:HGNC:23192]	0.932648	0.002879
ENSG00000154654	NCAM2	neural cell adhesion molecule 2 [Source:HGNC Symbol;Acc:HGNC:7657]	1.03729	0.002892
ENSG00000280252			1.13955	0.002899
ENSG00000054148	PHPT1	phosphohistidine phosphatase 1 [Source:HGNC Symbol;Acc:HGNC:30033]	-1.01234	0.002906
ENSG00000173208	ABCD2	ATP binding cassette subfamily D member 2 [Source:HGNC Symbol;Acc:HGNC:66]	0.939028	0.002916
ENSG00000132823	OSER1	oxidative stress responsive serine rich 1 [Source:HGNC Symbol;Acc:HGNC:16105]	-0.92528	0.002934
ENSG00000164325	TMEM174	transmembrane protein 174 [Source:HGNC Symbol;Acc:HGNC:28187]	1.05397	0.002942
ENSG00000115464	USP34	ubiquitin specific peptidase 34 [Source:HGNC Symbol;Acc:HGNC:20066]	-0.92552	0.002971
ENSG00000102032	RENBP	renin binding protein [Source:HGNC Symbol;Acc:HGNC:9959]	-1.44215	0.002978
ENSG00000101843	PSMD10	proteasome 26S subunit, non-ATPase 10 [Source:HGNC Symbol;Acc:HGNC:9555]	-0.93316	0.003004
ENSG00000166128	RAB8B	RAB8B, member RAS oncogene family [Source:HGNC Symbol;Acc:HGNC:30273]	-0.93593	0.003006
ENSG00000229009	TMPRSS11GP	transmembrane protease, serine 11G, pseudogene [Source:HGNC Symbol;Acc:HGNC:42983]	-1.12468	0.003012
ENSG00000054116	TRAPP3	trafficking protein particle complex 3 [Source:HGNC Symbol;Acc:HGNC:19942]	-0.98213	0.003043
ENSG00000119878	CRIP1	CXXC repeat containing interactor of PDZ3 domain [Source:HGNC Symbol;Acc:HGNC:14312]	-0.92019	0.003123
ENSG00000124091	GCNT7	glucosaminyl (N-acetyl) transferase family member 7 [Source:HGNC Symbol;Acc:HGNC:16099]	3.67122	0.003151
ENSG00000243716	NPIP5	nuclear pore complex interacting protein family member B5 [Source:HGNC Symbol;Acc:HGNC:37233]	1.02793	0.003151
ENSG00000215158			1.0012	0.003153
ENSG00000244270	RPL32P29	ribosomal protein L32 pseudogene 29 [Source:HGNC Symbol;Acc:HGNC:35771]	-2.85738	0.00321
ENSG00000187806	TMEM202	transmembrane protein 202 [Source:HGNC Symbol;Acc:HGNC:33733]	-0.92041	0.003248
ENSG00000245694	CRNDE	colorectal neoplasia differentially expressed [Source:HGNC Symbol;Acc:HGNC:37078]	1.03117	0.00326
ENSG00000131873	CHSY1	chondroitin sulfate synthase 1 [Source:HGNC Symbol;Acc:HGNC:17198]	1.5399	0.003281
ENSG00000231305			3.62772	0.00329
ENSG00000124615	MOCS1	molybdenum cofactor synthesis 1 [Source:HGNC Symbol;Acc:HGNC:7190]	-0.90996	0.003298
ENSG00000120937	NPPB	natriuretic peptide B [Source:HGNC Symbol;Acc:HGNC:7940]	0.928624	0.003306
ENSG00000074696	HACD3	3-hydroxyacyl-CoA dehydratase 3 [Source:HGNC Symbol;Acc:HGNC:24175]	0.912815	0.003314
ENSG00000099901	RANBP1	RAN binding protein 1 [Source:HGNC Symbol;Acc:HGNC:9847]	1.32037	0.003314
ENSG00000179918	SEPHS2	selenophosphate synthetase 2 [Source:HGNC Symbol;Acc:HGNC:19686]	3.61974	0.003317
ENSG00000138463	DIRC2	disrupted in renal carcinoma 2 [Source:HGNC Symbol;Acc:HGNC:16628]	-1.23914	0.003349
ENSG00000154822	PLCL2	phospholipase C like 2 [Source:HGNC Symbol;Acc:HGNC:9064]	-0.91465	0.003383
ENSG00000187905	LRRC74B	leucine rich repeat containing 74B [Source:HGNC Symbol;Acc:HGNC:34301]	1.27802	0.003383
ENSG00000115828	QPCT	glutaminyl-peptide cyclotransferase [Source:HGNC Symbol;Acc:HGNC:9753]	-0.91259	0.003424
ENSG00000261610			1.10941	0.00343
ENSG00000170855	TRIAP1	TP53 regulated inhibitor of apoptosis 1 [Source:HGNC Symbol;Acc:HGNC:26937]	-0.9165	0.003437
ENSG00000277224	HIST1H2BF	histone cluster 1 H2B family member f [Source:HGNC Symbol;Acc:HGNC:4752]	-1.87194	0.00347
ENSG00000172273	HINFP	histone H4 transcription factor [Source:HGNC Symbol;Acc:HGNC:17850]	-0.90933	0.003484
ENSG00000127586	CHTF18	chromosome transmission fidelity factor 18 [Source:HGNC Symbol;Acc:HGNC:18435]	0.912868	0.003508
ENSG00000273529	HES5	hes family bHLH transcription factor 5 [Source:HGNC Symbol;Acc:HGNC:19764]	1.10592	0.00351
ENSG00000143222	UFC1	ubiquitin-fold modifier conjugating enzyme 1 [Source:HGNC Symbol;Acc:HGNC:26941]	-1.13715	0.003536
ENSG00000123066	MED13L	mediator complex subunit 13 like [Source:HGNC Symbol;Acc:HGNC:22962]	-0.90349	0.003544
ENSG00000168079	SCARA5	scavenger receptor class A member 5 [Source:HGNC Symbol;Acc:HGNC:28701]	2.03566	0.003559
ENSG00000139233	LLPH	LLP homolog, long-term synaptic facilitation [Source:HGNC Symbol;Acc:HGNC:28229]	1.3737	0.003583
ENSG00000213406	ANXA2P1	annexin A2 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:538]	-3.53813	0.003605
ENSG00000261678	SCRT1	scratch family transcriptional repressor 1 [Source:HGNC Symbol;Acc:HGNC:15950]	0.90138	0.00363
ENSG00000225693	LAGE3P1	L antigen family member 3 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:49909]	1.09846	0.003645
ENSG00000282607	MGAM	maltase-glucoamylase [Source:HGNC Symbol;Acc:HGNC:7043]	-0.9941	0.003682

ENSG00000278259	MYO19	myosin XIX [Source:HGNC Symbol;Acc:HGNC:26234]	0.899591	0.003708
ENSG00000097046	CDC7	cell division cycle 7 [Source:HGNC Symbol;Acc:HGNC:1745]	0.901385	0.003724
ENSG00000178951	ZBTB7A	zinc finger and BTB domain containing 7A [Source:HGNC Symbol;Acc:HGNC:18078]	0.902919	0.003727
ENSG00000107036	RIC1	RIC1 homolog, RAB6A GEF complex partner 1 [Source:HGNC Symbol;Acc:HGNC:17686]	-1.03438	0.003732
ENSG00000206433	LST1	leukocyte specific transcript 1 [Source:HGNC Symbol;Acc:HGNC:14189]	-1.41827	0.003782
ENSG00000223465	LST1	leukocyte specific transcript 1 [Source:HGNC Symbol;Acc:HGNC:14189]	-1.41827	0.003782
ENSG00000226182	LST1	leukocyte specific transcript 1 [Source:HGNC Symbol;Acc:HGNC:14189]	-1.41827	0.003782
ENSG00000230791	LST1	leukocyte specific transcript 1 [Source:HGNC Symbol;Acc:HGNC:14189]	-1.41827	0.003782
ENSG00000231048	LST1	leukocyte specific transcript 1 [Source:HGNC Symbol;Acc:HGNC:14189]	-1.41827	0.003782
ENSG00000234514	LST1	leukocyte specific transcript 1 [Source:HGNC Symbol;Acc:HGNC:14189]	-1.41827	0.003782
ENSG00000235915	LST1	leukocyte specific transcript 1 [Source:HGNC Symbol;Acc:HGNC:14189]	-1.41827	0.003782
ENSG00000169203	NPIPB12	nuclear pore complex interacting protein family, member B12 [Source:HGNC Symbol;Acc:HGNC:37491]	0.994279	0.003854
ENSG00000120656	TAF12	TATA-box binding protein associated factor 12 [Source:HGNC Symbol;Acc:HGNC:11545]	-0.98006	0.003856
ENSG00000130309	COLGALT1	collagen beta(1-O)galactosyltransferase 1 [Source:HGNC Symbol;Acc:HGNC:26182]	0.899181	0.003869
ENSG00000232268	OR5211	olfactory receptor family 52 subfamily I member 1 [Source:HGNC Symbol;Acc:HGNC:15220]	-1.82818	0.003909
ENSG00000263070			-2.76229	0.003937
ENSG00000188186	LAMTOR4	late endosomal/lysosomal adaptor, MAPK and MTOR activator 4 [Source:HGNC Symbol;Acc:HGNC:33772]	-1.78955	0.003954
ENSG00000113303	BTNL8	butyrophilin like 8 [Source:HGNC Symbol;Acc:HGNC:26131]	-0.89432	0.003969
ENSG00000151025	GPR158	G protein-coupled receptor 158 [Source:HGNC Symbol;Acc:HGNC:23689]	0.903979	0.003986
ENSG00000267179			1.55919	0.004049
ENSG00000278500			0.982548	0.004059
ENSG00000152291	TGOLN2	trans-golgi network protein 2 [Source:HGNC Symbol;Acc:HGNC:15450]	0.894623	0.004078
ENSG00000242028	HYPK	huntingtin interacting protein K [Source:HGNC Symbol;Acc:HGNC:18418]	0.974675	0.004084
ENSG00000174501	ANKRD36C	ankyrin repeat domain 36C [Source:HGNC Symbol;Acc:HGNC:32946]	0.970576	0.004085
ENSG00000000460	C1orf112	chromosome 1 open reading frame 112 [Source:HGNC Symbol;Acc:HGNC:25565]	0.914159	0.004087
ENSG00000097096	SYDE2	synapse defective Rho GTPase homolog 2 [Source:HGNC Symbol;Acc:HGNC:25841]	0.973273	0.004094
ENSG00000237575	PYY2	peptide YY 2 (pseudogene) [Source:HGNC Symbol;Acc:HGNC:9749]	1.0195	0.004151
ENSG00000182117	NOP10	NOP10 ribonucleoprotein [Source:HGNC Symbol;Acc:HGNC:14378]	-0.89037	0.004164
ENSG00000206505	HLA-A	major histocompatibility complex, class I, A [Source:HGNC Symbol;Acc:HGNC:4931]	2.73522	0.004176
ENSG00000216649	GAGE12E	G antigen 12E [Source:HGNC Symbol;Acc:HGNC:31905]	-2.73518	0.004176
ENSG00000223980	HLA-A	major histocompatibility complex, class I, A [Source:HGNC Symbol;Acc:HGNC:4931]	2.73522	0.004176
ENSG00000231834	HLA-A	major histocompatibility complex, class I, A [Source:HGNC Symbol;Acc:HGNC:4931]	2.73522	0.004176
ENSG00000080603	SRCAP	Snf2 related CREBBP activator protein [Source:HGNC Symbol;Acc:HGNC:16974]	-0.88886	0.00418
ENSG00000277273	CDK7	cyclin dependent kinase 7 [Source:HGNC Symbol;Acc:HGNC:1778]	-0.97679	0.00418
ENSG00000187239	FNBP1	formin binding protein 1 [Source:HGNC Symbol;Acc:HGNC:17069]	-0.98239	0.004222
ENSG00000198658	ABHD17AP1	abhydrolase domain containing 17A pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:28394]	1.01995	0.004227
ENSG00000204625	HCG9	HLA complex group 9 (non-protein coding) [Source:HGNC Symbol;Acc:HGNC:21243]	-1.75905	0.004261
ENSG00000188133	TMEM215	transmembrane protein 215 [Source:HGNC Symbol;Acc:HGNC:33816]	-2.7249	0.004271
ENSG00000242600	MBL1P	mannose binding lectin 1, pseudogene [Source:HGNC Symbol;Acc:HGNC:6921]	-0.96876	0.004279
ENSG00000111241	FGF6	fibroblast growth factor 6 [Source:HGNC Symbol;Acc:HGNC:3684]	-1.76052	0.004298
ENSG00000198185	ZNF334	zinc finger protein 334 [Source:HGNC Symbol;Acc:HGNC:15806]	1.81514	0.004318
ENSG00000281335	MIF-AS1	MIF antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:27669]	1.07545	0.004324
ENSG00000188646			-1.08023	0.004327
ENSG00000127870	RNF6	ring finger protein 6 [Source:HGNC Symbol;Acc:HGNC:10069]	-1.14405	0.004328
ENSG00000101251	SEL1L2	SEL1L2 ERAD E3 ligase adaptor subunit [Source:HGNC Symbol;Acc:HGNC:15897]	3.36429	0.004348

ENSG00000228106			1.07436	0.004354
ENSG00000112941	PAPD7	poly(A) RNA polymerase D7, non-canonical [Source:HGNC Symbol;Acc:HGNC:16705]	1.155	0.004398
ENSG00000172752	COL6A5	collagen type VI alpha 5 chain [Source:HGNC Symbol;Acc:HGNC:26674]	0.885044	0.0044
ENSG00000168214	RBPJ	recombination signal binding protein for immunoglobulin kappa J region [Source:HGNC Symbol;Acc:HGNC:5724]	-0.88656	0.004421
ENSG00000123407	HOXC12	homeobox C12 [Source:HGNC Symbol;Acc:HGNC:5124]	1.30328	0.004432
ENSG00000015133	CCDC88C	coiled-coil domain containing 88C [Source:HGNC Symbol;Acc:HGNC:19967]	0.965344	0.004449
ENSG00000074800	ENO1	enolase 1 [Source:HGNC Symbol;Acc:HGNC:3350]	0.883873	0.004468
ENSG00000255965			1.07164	0.004487
ENSG00000175497	DPP10	dipeptidyl peptidase like 10 [Source:HGNC Symbol;Acc:HGNC:20823]	1.06942	0.00452
ENSG00000119421	NDUFA8	NADH:ubiquinone oxidoreductase subunit A8 [Source:HGNC Symbol;Acc:HGNC:7692]	-1.84991	0.004521
ENSG00000175538	KCNE3	potassium voltage-gated channel subfamily E regulatory subunit 3 [Source:HGNC Symbol;Acc:HGNC:6243]	-0.8885	0.004525
ENSG00000150556	LYPD6B	LY6/PLAUR domain containing 6B [Source:HGNC Symbol;Acc:HGNC:27018]	0.88817	0.004529
ENSG00000227268	KLLN	killin, p53-regulated DNA replication inhibitor [Source:HGNC Symbol;Acc:HGNC:37212]	-3.3038	0.004656
ENSG00000143622	RIT1	Ras like without CAAX 1 [Source:HGNC Symbol;Acc:HGNC:10023]	-0.88031	0.004662
ENSG00000133640	LRR1Q1	leucine rich repeats and IQ motif containing 1 [Source:HGNC Symbol;Acc:HGNC:25708]	0.916971	0.004664
ENSG00000143799	PARP1	poly(ADP-ribose) polymerase 1 [Source:HGNC Symbol;Acc:HGNC:270]	3.29793	0.004688
ENSG00000196741	LINC01560	long intergenic non-protein coding RNA 1560 [Source:HGNC Symbol;Acc:HGNC:27333]	1.07069	0.00469
ENSG00000156500	FAM122C	family with sequence similarity 122C [Source:HGNC Symbol;Acc:HGNC:25202]	-0.87697	0.004701
ENSG00000280477	RASA3	RAS p21 protein activator 3 [Source:HGNC Symbol;Acc:HGNC:20331]	-2.67812	0.004731
ENSG00000234844	CDC42P2	cell division cycle 42 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:33952]	-1.06305	0.004739
ENSG00000157224	CLDN12	claudin 12 [Source:HGNC Symbol;Acc:HGNC:2034]	-0.95167	0.004749
ENSG00000162736	NCSTN	nicastrin [Source:HGNC Symbol;Acc:HGNC:17091]	-0.87555	0.004764
ENSG00000134897	BIVM	basic, immunoglobulin-like variable motif containing [Source:HGNC Symbol;Acc:HGNC:16034]	0.959378	0.004794
ENSG00000243981			1.05992	0.004863
ENSG00000130592	LSP1	lymphocyte-specific protein 1 [Source:HGNC Symbol;Acc:HGNC:6707]	1.7746	0.004886
ENSG00000271975			-3.25929	0.004902
ENSG00000077235	GTF3C1	general transcription factor IIIC subunit 1 [Source:HGNC Symbol;Acc:HGNC:4664]	-0.88038	0.004953
ENSG00000254154			-2.65673	0.004959
ENSG00000275723	NAPEPLD	N-acyl phosphatidylethanolamine phospholipase D [Source:HGNC Symbol;Acc:HGNC:21683]	1.05716	0.004988
ENSG00000213240			1.06259	0.004995
ENSG00000131080	EDA2R	ectodysplasin A2 receptor [Source:HGNC Symbol;Acc:HGNC:17756]	-2.83843	0.005047
ENSG00000116584	ARHGEF2	Rho/Rac guanine nucleotide exchange factor 2 [Source:HGNC Symbol;Acc:HGNC:682]	-1.96901	0.005075
ENSG00000156113	KCNMA1	potassium calcium-activated channel subfamily M alpha 1 [Source:HGNC Symbol;Acc:HGNC:6284]	-0.87396	0.005091
ENSG00000154473	BUB3	BUB3, mitotic checkpoint protein [Source:HGNC Symbol;Acc:HGNC:1151]	1.03476	0.005094
ENSG00000165512	ZNF22	zinc finger protein 22 [Source:HGNC Symbol;Acc:HGNC:13012]	1.16921	0.005101
ENSG00000213190	MLLT11	myeloid/lymphoid or mixed-lineage leukemia; translocated to, 11 [Source:HGNC Symbol;Acc:HGNC:16997]	1.37496	0.005136
ENSG00000165861	ZFYVE1	zinc finger FYVE-type containing 1 [Source:HGNC Symbol;Acc:HGNC:13180]	-2.2368	0.005194
ENSG00000177125	ZBTB34	zinc finger and BTB domain containing 34 [Source:HGNC Symbol;Acc:HGNC:31446]	-0.95052	0.005198
ENSG00000248223			1.05266	0.005217
ENSG00000270091			1.05128	0.005258
ENSG00000245059			1.05664	0.005264
ENSG00000142798	HSPG2	heparan sulfate proteoglycan 2 [Source:HGNC Symbol;Acc:HGNC:5273]	-1.21407	0.005266
ENSG00000188312	CENPP	centromere protein P [Source:HGNC Symbol;Acc:HGNC:32933]	0.87114	0.00527
ENSG00000230837	RPL31P2	ribosomal protein L31 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:16575]	1.0558	0.005284

ENSG00000176428	VPS37D	VPS37D, ESCRT-I subunit [Source:HGNC Symbol;Acc:HGNC:18287]	-2.33903	0.005299
ENSG00000124762	CDKN1A	cyclin dependent kinase inhibitor 1A [Source:HGNC Symbol;Acc:HGNC:1784]	-1.58931	0.005302
ENSG00000213453	FTH1P3	ferritin heavy chain 1 pseudogene 3 [Source:HGNC Symbol;Acc:HGNC:3990]	-2.62601	0.005308
ENSG00000175920	DOK7	docking protein 7 [Source:HGNC Symbol;Acc:HGNC:26594]	0.86547	0.005343
ENSG00000136122	BORA	bora, aurora kinase A activator [Source:HGNC Symbol;Acc:HGNC:24724]	1.72017	0.005348
ENSG00000139372	TDG	thymine DNA glycosylase [Source:HGNC Symbol;Acc:HGNC:11700]	1.0762	0.005383
ENSG00000109163	GNRHR	gonadotropin releasing hormone receptor [Source:HGNC Symbol;Acc:HGNC:4421]	-1.36199	0.005384
ENSG00000139880	CDH24	cadherin 24 [Source:HGNC Symbol;Acc:HGNC:14265]	0.859182	0.005423
ENSG00000147416	ATP6V1B2	ATPase H+ transporting V1 subunit B2 [Source:HGNC Symbol;Acc:HGNC:854]	-1.07942	0.005448
ENSG00000166123	GPT2	glutamic--pyruvic transaminase 2 [Source:HGNC Symbol;Acc:HGNC:18062]	0.861425	0.005471
ENSG00000226549	SCDP1	stearoyl-CoA desaturase pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:10572]	1.04503	0.005479
ENSG00000100865	CINP	cyclin dependent kinase 2 interacting protein [Source:HGNC Symbol;Acc:HGNC:23789]	-0.86112	0.005503
ENSG00000095713	CRTAC1	cartilage acidic protein 1 [Source:HGNC Symbol;Acc:HGNC:14882]	-0.85725	0.005538
ENSG00000070785	EIF2B3	eukaryotic translation initiation factor 2B subunit gamma [Source:HGNC Symbol;Acc:HGNC:3259]	3.15647	0.005541
ENSG00000102804	TSC22D1	TSC22 domain family member 1 [Source:HGNC Symbol;Acc:HGNC:16826]	-0.94541	0.005543
ENSG00000231169	EEF1B2P1	eukaryotic translation elongation factor 1 beta 2 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:3207]	-1.04184	0.005558
ENSG00000178974	FBXO34	F-box protein 34 [Source:HGNC Symbol;Acc:HGNC:20201]	-0.85786	0.005561
ENSG00000260526			-1.04243	0.005573
ENSG00000113722	CDX1	caudal type homeobox 1 [Source:HGNC Symbol;Acc:HGNC:1805]	-1.69097	0.005575
ENSG00000175868	CALCB	calcitonin related polypeptide beta [Source:HGNC Symbol;Acc:HGNC:1438]	1.73565	0.005577
ENSG00000179840	PIK3CD-AS1	PIK3CD antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:32346]	-2.60281	0.005588
ENSG00000176742	OR51V1	olfactory receptor family 51 subfamily V member 1 [Source:HGNC Symbol;Acc:HGNC:19597]	1.68971	0.005596
ENSG00000185710	SMG1P4	SMG1P4, nonsense mediated mRNA decay associated PI3K related kinase pseudogene 4 [Source:HGNC Symbol;Acc:HGNC:49861]	1.0397	0.005617
ENSG00000227040			-0.93451	0.005621
ENSG00000135926	TMBIM1	transmembrane BAX inhibitor motif containing 1 [Source:HGNC Symbol;Acc:HGNC:23410]	-0.85904	0.005637
ENSG00000116062	MSH6	mutS homolog 6 [Source:HGNC Symbol;Acc:HGNC:7329]	0.861449	0.005644
ENSG00000244879	GABPB1-AS1	GABPB1 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:44157]	1.03929	0.005659
ENSG00000259969			-1.0382	0.005683
ENSG00000167985	SDHAF2	succinate dehydrogenase complex assembly factor 2 [Source:HGNC Symbol;Acc:HGNC:26034]	-1.03883	0.005749
ENSG00000112218	GPR63	G protein-coupled receptor 63 [Source:HGNC Symbol;Acc:HGNC:13302]	-0.85896	0.005766
ENSG00000170889	RPS9	ribosomal protein S9 [Source:HGNC Symbol;Acc:HGNC:10442]	-2.58805	0.005775
ENSG00000237214			-2.58805	0.005775
ENSG00000274005	RPS9	ribosomal protein S9 [Source:HGNC Symbol;Acc:HGNC:10442]	-2.58805	0.005775
ENSG00000274626	RPS9	ribosomal protein S9 [Source:HGNC Symbol;Acc:HGNC:10442]	-2.58805	0.005775
ENSG00000274646	RPS9	ribosomal protein S9 [Source:HGNC Symbol;Acc:HGNC:10442]	-2.58805	0.005775
ENSG00000274950	RPS9	ribosomal protein S9 [Source:HGNC Symbol;Acc:HGNC:10442]	-2.58805	0.005775
ENSG00000275323	RPS9	ribosomal protein S9 [Source:HGNC Symbol;Acc:HGNC:10442]	-2.58805	0.005775
ENSG00000277079	RPS9	ribosomal protein S9 [Source:HGNC Symbol;Acc:HGNC:10442]	-2.58805	0.005775
ENSG00000277359	RPS9	ribosomal protein S9 [Source:HGNC Symbol;Acc:HGNC:10442]	-2.58805	0.005775
ENSG00000278081	RPS9	ribosomal protein S9 [Source:HGNC Symbol;Acc:HGNC:10442]	-2.58805	0.005775
ENSG00000278270	RPS9	ribosomal protein S9 [Source:HGNC Symbol;Acc:HGNC:10442]	-2.58805	0.005775
ENSG00000172172	MRPL13	mitochondrial ribosomal protein L13 [Source:HGNC Symbol;Acc:HGNC:14278]	0.861401	0.00578
ENSG00000159055	MIS18A	MIS18 kinetochore protein A [Source:HGNC Symbol;Acc:HGNC:1286]	2.21234	0.005799
ENSG00000183747	ACSM2A	acyl-CoA synthetase medium-chain family member 2A [Source:HGNC Symbol;Acc:HGNC:32017]	-1.70518	0.005813

ENSG00000255245	FXVD6-FXVD2	FXVD6-FXVD2 readthrough [Source:HGNC Symbol;Acc:HGNC:39978]	-1.03687	0.005828
ENSG00000176165	FOXG1	forkhead box G1 [Source:HGNC Symbol;Acc:HGNC:3811]	0.97752	0.005895
ENSG00000115665	SLC5A7	solute carrier family 5 member 7 [Source:HGNC Symbol;Acc:HGNC:14025]	-1.31436	0.005932
ENSG00000234506	LINC01506	long intergenic non-protein coding RNA 1506 [Source:HGNC Symbol;Acc:HGNC:51187]	-1.27613	0.005932
ENSG00000170846			0.937871	0.005935
ENSG00000105856	HBP1	HMG-box transcription factor 1 [Source:HGNC Symbol;Acc:HGNC:23200]	-0.85448	0.005952
ENSG00000139330	KERA	keratocan [Source:HGNC Symbol;Acc:HGNC:6309]	-0.85562	0.005961
ENSG00000145832	SLC25A48	solute carrier family 25 member 48 [Source:HGNC Symbol;Acc:HGNC:30451]	-0.84996	0.005965
ENSG00000127957	PMS2P3	PMS1 homolog 2, mismatch repair system component pseudogene 3 [Source:HGNC Symbol;Acc:HGNC:9128]	0.924525	0.006021
ENSG00000215717	TMEM167B	transmembrane protein 167B [Source:HGNC Symbol;Acc:HGNC:30187]	-0.85699	0.006047
ENSG00000166135	HIF1AN	hypoxia inducible factor 1 alpha subunit inhibitor [Source:HGNC Symbol;Acc:HGNC:17113]	-0.85123	0.006048
ENSG00000263366	ABHD17AP5	abhydrolase domain containing 17A pseudogene 5 [Source:HGNC Symbol;Acc:HGNC:34043]	0.924627	0.006114
ENSG00000225423	TNPO1P1	transportin 1 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:45120]	1.0273	0.00615
ENSG00000184258	CDR1	cerebellar degeneration related protein 1 [Source:HGNC Symbol;Acc:HGNC:1798]	-0.84582	0.006156
ENSG00000168066	SF1	splicing factor 1 [Source:HGNC Symbol;Acc:HGNC:12950]	0.85565	0.006205
ENSG00000269235	ZNF350-AS1	ZNF350 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:48598]	-1.02609	0.006212
ENSG00000276469	GPR179	G protein-coupled receptor 179 [Source:HGNC Symbol;Acc:HGNC:31371]	-2.55058	0.006279
ENSG00000035862	TIMP2	TIMP metalloproteinase inhibitor 2 [Source:HGNC Symbol;Acc:HGNC:11821]	-0.84735	0.006286
ENSG00000144559	TAMM41	TAM41 mitochondrial translocator assembly and maintenance homolog [Source:HGNC Symbol;Acc:HGNC:25187]	1.59313	0.006296
ENSG00000103994	ZNF106	zinc finger protein 106 [Source:HGNC Symbol;Acc:HGNC:12886]	-0.86171	0.006306
ENSG00000146859	TMEM140	transmembrane protein 140 [Source:HGNC Symbol;Acc:HGNC:21870]	-2.67636	0.006307
ENSG00000145833	DDX46	DEAD-box helicase 46 [Source:HGNC Symbol;Acc:HGNC:18681]	0.841831	0.006312
ENSG00000089351	GRAMD1A	GRAM domain containing 1A [Source:HGNC Symbol;Acc:HGNC:29305]	-0.84758	0.006317
ENSG00000197565	COL4A6	collagen type IV alpha 6 chain [Source:HGNC Symbol;Acc:HGNC:2208]	1.26755	0.006331
ENSG00000278372	MYO19	myosin XIX [Source:HGNC Symbol;Acc:HGNC:26234]	0.922508	0.006365
ENSG00000282319			-2.17793	0.006415
ENSG00000143499	SMYD2	SET and MYND domain containing 2 [Source:HGNC Symbol;Acc:HGNC:20982]	0.95722	0.006419
ENSG00000130703	OSBPL2	oxysterol binding protein like 2 [Source:HGNC Symbol;Acc:HGNC:15761]	-1.10295	0.006501
ENSG00000122254	HS3ST2	heparan sulfate-glucosamine 3-sulfotransferase 2 [Source:HGNC Symbol;Acc:HGNC:5195]	-0.84727	0.006505
ENSG00000071889	FAM3A	family with sequence similarity 3 member A [Source:HGNC Symbol;Acc:HGNC:13749]	-0.8429	0.006529
ENSG00000257950	P2RX5-TAX1BP3	P2RX5-TAX1BP3 readthrough (NMD candidate) [Source:HGNC Symbol;Acc:HGNC:49191]	-0.91706	0.006583
ENSG00000277044	OPRL1	opioid related nociceptin receptor 1 [Source:HGNC Symbol;Acc:HGNC:8155]	-0.91046	0.006645
ENSG00000204622	HLA-J	major histocompatibility complex, class I, J (pseudogene) [Source:HGNC Symbol;Acc:HGNC:4967]	-1.01539	0.006706
ENSG00000223439	HLA-J	major histocompatibility complex, class I, J (pseudogene) [Source:HGNC Symbol;Acc:HGNC:4967]	-1.01539	0.006706
ENSG00000226586	HLA-J	major histocompatibility complex, class I, J (pseudogene) [Source:HGNC Symbol;Acc:HGNC:4967]	-1.01539	0.006706
ENSG00000228220	HLA-J	major histocompatibility complex, class I, J (pseudogene) [Source:HGNC Symbol;Acc:HGNC:4967]	-1.01539	0.006706
ENSG00000228524	HLA-J	major histocompatibility complex, class I, J (pseudogene) [Source:HGNC Symbol;Acc:HGNC:4967]	-1.01539	0.006706
ENSG00000234074	HLA-J	major histocompatibility complex, class I, J (pseudogene) [Source:HGNC Symbol;Acc:HGNC:4967]	-1.01539	0.006706
ENSG00000235309	HLA-J	major histocompatibility complex, class I, J (pseudogene) [Source:HGNC Symbol;Acc:HGNC:4967]	-1.01539	0.006706
ENSG00000164237	CMBL	carboxymethylenebutenolidase homolog [Source:HGNC Symbol;Acc:HGNC:25090]	-1.56363	0.006709
ENSG00000251062			1.0153	0.006716
ENSG00000226933	NRBF2P2	nuclear receptor binding factor 2 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:44605]	-1.01922	0.006725

ENSG00000143546	S100A8	S100 calcium binding protein A8 [Source:HGNC Symbol;Acc:HGNC:10498]	-0.8451	0.00675
ENSG00000227258	SMIM2-AS1	SMIM2 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:42674]	-1.01492	0.006769
ENSG00000164087	POC1A	POC1 centriolar protein A [Source:HGNC Symbol;Acc:HGNC:24488]	0.992195	0.006843
ENSG00000094755	GABRP	gamma-aminobutyric acid type A receptor pi subunit [Source:HGNC Symbol;Acc:HGNC:4089]	2.98529	0.006874
ENSG00000228801			2.98374	0.006888
ENSG00000118271	TTR	transthyretin [Source:HGNC Symbol;Acc:HGNC:12405]	0.831821	0.006917
ENSG00000140379	BCL2A1	BCL2 related protein A1 [Source:HGNC Symbol;Acc:HGNC:991]	-0.83165	0.006929
ENSG00000281660	DUPD1	dual specificity phosphatase and pro isomerase domain containing 1 [Source:HGNC Symbol;Acc:HGNC:23481]	-2.50513	0.006954
ENSG00000187912	CLEC17A	C-type lectin domain family 17 member A [Source:HGNC Symbol;Acc:HGNC:34520]	1.63498	0.00696
ENSG00000171103	TRMT61B	tRNA methyltransferase 61B [Source:HGNC Symbol;Acc:HGNC:26070]	0.894783	0.006961
ENSG00000153283	CD96	CD96 molecule [Source:HGNC Symbol;Acc:HGNC:16892]	0.831483	0.006973
ENSG00000136867	SLC31A2	solute carrier family 31 member 2 [Source:HGNC Symbol;Acc:HGNC:11017]	-1.39931	0.006982
ENSG00000163464	CXCR1	C-X-C motif chemokine receptor 1 [Source:HGNC Symbol;Acc:HGNC:6026]	-0.83553	0.007035
ENSG00000154342	WNT3A	Wnt family member 3A [Source:HGNC Symbol;Acc:HGNC:15983]	0.83518	0.007048
ENSG00000158483	FAM86C1	family with sequence similarity 86 member C1 [Source:HGNC Symbol;Acc:HGNC:25561]	1.08154	0.007119
ENSG00000136518	ACTL6A	actin like 6A [Source:HGNC Symbol;Acc:HGNC:24124]	0.830495	0.007142
ENSG00000005073	HOXA11	homeobox A11 [Source:HGNC Symbol;Acc:HGNC:5101]	-0.83609	0.00715
ENSG00000217261	POM121L4P	POM121 transmembrane nucleoporin like 4, pseudogene [Source:HGNC Symbol;Acc:HGNC:19326]	1.00736	0.007183
ENSG00000250536	ABHD17AP3	abhydrolase domain containing 17A pseudogene 3 [Source:HGNC Symbol;Acc:HGNC:34041]	0.9031	0.00721
ENSG00000133488	SEC14L4	SEC14 like lipid binding 4 [Source:HGNC Symbol;Acc:HGNC:20627]	0.827713	0.007231
ENSG00000160221	C21orf33	chromosome 21 open reading frame 33 [Source:HGNC Symbol;Acc:HGNC:1273]	0.904504	0.007233
ENSG00000276612			0.904504	0.007233
ENSG00000280071			0.904504	0.007233
ENSG00000185269	NOTUM	NOTUM, palmitoleoyl-protein carboxylesterase [Source:HGNC Symbol;Acc:HGNC:27106]	-0.82666	0.007269
ENSG00000109046	WSB1	WD repeat and SOCS box containing 1 [Source:HGNC Symbol;Acc:HGNC:19221]	1.08023	0.007285
ENSG00000100156	SLC16A8	solute carrier family 16 member 8 [Source:HGNC Symbol;Acc:HGNC:16270]	0.837338	0.007291
ENSG00000070610	GBA2	glucosylceramidase beta 2 [Source:HGNC Symbol;Acc:HGNC:18986]	-2.00557	0.007359
ENSG00000244537	KRTAP4-2	keratin associated protein 4-2 [Source:HGNC Symbol;Acc:HGNC:18900]	1.26038	0.007392
ENSG00000135111	TBX3	T-box 3 [Source:HGNC Symbol;Acc:HGNC:11602]	-0.83361	0.007396
ENSG00000258472			1.00195	0.007396
ENSG00000143951	WDPCP	WD repeat containing planar cell polarity effector [Source:HGNC Symbol;Acc:HGNC:28027]	-0.8361	0.007444
ENSG00000164938	TP53INP1	tumor protein p53 inducible nuclear protein 1 [Source:HGNC Symbol;Acc:HGNC:18022]	-1.4561	0.007478
ENSG00000214290	COLCA2	colorectal cancer associated 2 [Source:HGNC Symbol;Acc:HGNC:26978]	0.902422	0.007485
ENSG00000262145	ATP23	ATP23 metallopeptidase and ATP synthase assembly factor homolog [Source:HGNC Symbol;Acc:HGNC:29452]	1.00109	0.007566
ENSG00000125910	S1PR4	sphingosine-1-phosphate receptor 4 [Source:HGNC Symbol;Acc:HGNC:3170]	-1.22386	0.00759
ENSG00000184774	MGAT4EP	MGAT4 family member E, pseudogene [Source:HGNC Symbol;Acc:HGNC:49418]	2.9102	0.007592
ENSG00000042753	AP2S1	adaptor related protein complex 2 sigma 1 subunit [Source:HGNC Symbol;Acc:HGNC:565]	0.821259	0.007617
ENSG00000214602	CTBP2P5	C-terminal binding protein 2 pseudogene 5 [Source:HGNC Symbol;Acc:HGNC:45197]	-0.99903	0.007617
ENSG00000171793	CTPS1	CTP synthase 1 [Source:HGNC Symbol;Acc:HGNC:2519]	2.90727	0.007622
ENSG00000172081	MOB3A	MOB kinase activator 3A [Source:HGNC Symbol;Acc:HGNC:29802]	-0.82805	0.007667
ENSG00000145808	ADAMTS19	ADAM metallopeptidase with thrombospondin type 1 motif 19 [Source:HGNC Symbol;Acc:HGNC:17111]	1.65318	0.00768
ENSG00000070404	FSTL3	folistatin like 3 [Source:HGNC Symbol;Acc:HGNC:3973]	0.821432	0.007721
ENSG00000166908	PIP4K2C	phosphatidylinositol-5-phosphate 4-kinase type 2 gamma [Source:HGNC Symbol;Acc:HGNC:23786]	-1.63154	0.007729

ENSG00000166979	EVA1C	eva-1 homolog C [Source:HGNC Symbol;Acc:HGNC:13239]	-0.82734	0.007746
ENSG00000091009	RBM27	RNA binding motif protein 27 [Source:HGNC Symbol;Acc:HGNC:29243]	0.999328	0.007754
ENSG00000255419	MSH5-SAPCD1	MSH5-SAPCD1 readthrough (NMD candidate) [Source:HGNC Symbol;Acc:HGNC:41994]	0.895456	0.007782
ENSG00000264619	MSH5-SAPCD1	MSH5-SAPCD1 readthrough (NMD candidate) [Source:HGNC Symbol;Acc:HGNC:41994]	0.895456	0.007782
ENSG00000265750			-2.45341	0.007818
ENSG00000206428	AIF1	allograft inflammatory factor 1 [Source:HGNC Symbol;Acc:HGNC:352]	-0.90104	0.007844
ENSG00000234836	AIF1	allograft inflammatory factor 1 [Source:HGNC Symbol;Acc:HGNC:352]	-0.90104	0.007844
ENSG00000235588	AIF1	allograft inflammatory factor 1 [Source:HGNC Symbol;Acc:HGNC:352]	-0.90104	0.007844
ENSG00000235985	AIF1	allograft inflammatory factor 1 [Source:HGNC Symbol;Acc:HGNC:352]	-0.90104	0.007844
ENSG00000237727	AIF1	allograft inflammatory factor 1 [Source:HGNC Symbol;Acc:HGNC:352]	-0.90104	0.007844
ENSG00000144455	SUMF1	sulfatase modifying factor 1 [Source:HGNC Symbol;Acc:HGNC:20376]	-1.6614	0.007864
ENSG00000100439	ABHD4	abhydrolase domain containing 4 [Source:HGNC Symbol;Acc:HGNC:20154]	-0.82337	0.007879
ENSG00000130175	PRKCSH	protein kinase C substrate 80K-H [Source:HGNC Symbol;Acc:HGNC:9411]	-0.82537	0.007882
ENSG00000134531	EMP1	epithelial membrane protein 1 [Source:HGNC Symbol;Acc:HGNC:3333]	0.831584	0.007896
ENSG00000229585	RPL21P44	ribosomal protein L21 pseudogene 44 [Source:HGNC Symbol;Acc:HGNC:33820]	0.992919	0.007896
ENSG00000196166	C8orf86	chromosome 8 open reading frame 86 [Source:HGNC Symbol;Acc:HGNC:33774]	2.87832	0.007926
ENSG00000218537	MIF-AS1	MIF antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:27669]	0.897685	0.007926
ENSG00000170906	NDUFA3	NADH:ubiquinone oxidoreductase subunit A3 [Source:HGNC Symbol;Acc:HGNC:7686]	-0.82489	0.007931
ENSG00000167941	SOST	sclerostin [Source:HGNC Symbol;Acc:HGNC:13771]	1.06796	0.007941
ENSG00000162650	ATXN7L2	ataxin 7 like 2 [Source:HGNC Symbol;Acc:HGNC:28713]	0.821042	0.007945
ENSG00000121766	ZCCHC17	zinc finger CCHC-type containing 17 [Source:HGNC Symbol;Acc:HGNC:30246]	1.84971	0.00798
ENSG00000167769	ACER1	alkaline ceramidase 1 [Source:HGNC Symbol;Acc:HGNC:18356]	0.858215	0.008016
ENSG00000165556	CDX2	caudal type homeobox 2 [Source:HGNC Symbol;Acc:HGNC:1806]	-0.81407	0.00805
ENSG00000133048	CHI3L1	chitinase 3 like 1 [Source:HGNC Symbol;Acc:HGNC:1932]	-0.8953	0.008057
ENSG00000125850	OVOL2	ovo like zinc finger 2 [Source:HGNC Symbol;Acc:HGNC:15804]	1.91879	0.008061
ENSG00000223611	SUPT20HL2	SPT20 homolog, SAGA complex component-like 2 [Source:HGNC Symbol;Acc:HGNC:31797]	-2.86158	0.008108
ENSG00000142512	SIGLEC10	sialic acid binding Ig like lectin 10 [Source:HGNC Symbol;Acc:HGNC:15620]	-0.89434	0.008112
ENSG00000148229	POLE3	DNA polymerase epsilon 3, accessory subunit [Source:HGNC Symbol;Acc:HGNC:13546]	4.29888	0.00815
ENSG00000168569	TMEM223	transmembrane protein 223 [Source:HGNC Symbol;Acc:HGNC:28464]	0.813018	0.008219
ENSG00000162433	AK4	adenylate kinase 4 [Source:HGNC Symbol;Acc:HGNC:363]	0.824183	0.008239
ENSG00000261803	LINC02140	long intergenic non-protein coding RNA 2140 [Source:HGNC Symbol;Acc:HGNC:53000]	0.886494	0.008241
ENSG00000183559	C10orf120	chromosome 10 open reading frame 120 [Source:HGNC Symbol;Acc:HGNC:25707]	1.60709	0.008256
ENSG00000172073	TEX37	testis expressed 37 [Source:HGNC Symbol;Acc:HGNC:26341]	0.815507	0.008271
ENSG00000146223	RPL7L1	ribosomal protein L7 like 1 [Source:HGNC Symbol;Acc:HGNC:21370]	0.892941	0.008292
ENSG00000227132			0.994113	0.008315
ENSG00000229126			0.994113	0.008315
ENSG00000229199			0.994113	0.008315
ENSG00000231474			0.994113	0.008315
ENSG00000234063			0.994113	0.008315
ENSG00000261272	MUC22	mucin 22 [Source:HGNC Symbol;Acc:HGNC:39755]	0.994113	0.008315
ENSG00000281434	NDUFA10	NADH:ubiquinone oxidoreductase subunit A10 [Source:HGNC Symbol;Acc:HGNC:7684]	1.03067	0.008365
ENSG00000172349	IL16	interleukin 16 [Source:HGNC Symbol;Acc:HGNC:5980]	0.895278	0.008435
ENSG00000261448			-0.98293	0.008442
ENSG00000261402			-1.50128	0.008492
ENSG00000231705			1.33591	0.008537
ENSG00000125734	GPR108	G protein-coupled receptor 108 [Source:HGNC Symbol;Acc:HGNC:17829]	-0.81914	0.008586

ENSG00000121152	NCAPH	non-SMC condensin I complex subunit H [Source:HGNC Symbol;Acc:HGNC:1112]	1.23763	0.008596
ENSG00000165006	UBAP1	ubiquitin associated protein 1 [Source:HGNC Symbol;Acc:HGNC:12461]	-0.81658	0.008611
ENSG00000233827	FAM58BP	family with sequence similarity 58 member B, pseudogene [Source:HGNC Symbol;Acc:HGNC:27373]	-2.81625	0.008631
ENSG00000113312	TTC1	tetratricopeptide repeat domain 1 [Source:HGNC Symbol;Acc:HGNC:12391]	-0.80872	0.008661
ENSG00000113595	TRIM23	tripartite motif containing 23 [Source:HGNC Symbol;Acc:HGNC:660]	-1.09553	0.008753
ENSG00000184924	PTRHD1	peptidyl-tRNA hydrolase domain containing 1 [Source:HGNC Symbol;Acc:HGNC:33782]	-0.80639	0.008828
ENSG00000153391	INO80C	INO80 complex subunit C [Source:HGNC Symbol;Acc:HGNC:26994]	-0.87576	0.008868
ENSG00000167550	RHEBL1	Ras homolog enriched in brain like 1 [Source:HGNC Symbol;Acc:HGNC:21166]	1.09311	0.008885
ENSG00000135220	UGT2A3	UDP glucuronosyltransferase family 2 member A3 [Source:HGNC Symbol;Acc:HGNC:28528]	0.8092	0.008891
ENSG00000224628			0.981219	0.008901
ENSG00000279921			0.982602	0.00898
ENSG00000280130			0.982602	0.00898
ENSG00000145743	FBXL17	F-box and leucine rich repeat protein 17 [Source:HGNC Symbol;Acc:HGNC:13615]	-0.81024	0.009057
ENSG00000183307	CECR6	cat eye syndrome chromosome region, candidate 6 [Source:HGNC Symbol;Acc:HGNC:1844]	-1.18184	0.009068
ENSG00000217791	ASS1P9	argininosuccinate synthetase 1 pseudogene 9 [Source:HGNC Symbol;Acc:HGNC:772]	0.976155	0.00907
ENSG00000265479	DTX2P1-UPK3BP1-PMS2P11	DTX2P1-UPK3BP1-PMS2P11 readthrough, transcribed pseudogene [Source:HGNC Symbol;Acc:HGNC:42360]	0.883859	0.009091
ENSG00000224215			-1.04126	0.009117
ENSG00000062822	POLD1	DNA polymerase delta 1, catalytic subunit [Source:HGNC Symbol;Acc:HGNC:9175]	0.882675	0.009146
ENSG00000125968	ID1	inhibitor of DNA binding 1, HLH protein [Source:HGNC Symbol;Acc:HGNC:5360]	1.53713	0.009184
ENSG00000257524			1.44893	0.009187
ENSG00000152520	PAN3	PAN3 poly(A) specific ribonuclease subunit [Source:HGNC Symbol;Acc:HGNC:29991]	-0.8076	0.009197
ENSG00000166508	MCM7	minichromosome maintenance complex component 7 [Source:HGNC Symbol;Acc:HGNC:6950]	1.83742	0.009201
ENSG00000091482	SMPX	small muscle protein, X-linked [Source:HGNC Symbol;Acc:HGNC:11122]	-0.8061	0.009242
ENSG00000186340	THBS2	thrombospondin 2 [Source:HGNC Symbol;Acc:HGNC:11786]	-1.56772	0.009263
ENSG00000165521	EML5	echinoderm microtubule associated protein like 5 [Source:HGNC Symbol;Acc:HGNC:18197]	-1.61342	0.009352
ENSG00000183918	SH2D1A	SH2 domain containing 1A [Source:HGNC Symbol;Acc:HGNC:10820]	-0.87142	0.009361
ENSG00000159899	NPR2	natriuretic peptide receptor 2 [Source:HGNC Symbol;Acc:HGNC:7944]	-1.32468	0.009383
ENSG00000227590	ATP5G1P5	ATP synthase, H+ transporting, mitochondrial Fo complex subunit C1 (subunit 9) pseudogene 5 [Source:HGNC Symbol;Acc:HGNC:39508]	0.970524	0.009391
ENSG00000101337	TM9SF4	transmembrane 9 superfamily member 4 [Source:HGNC Symbol;Acc:HGNC:30797]	-0.80574	0.009404
ENSG00000130413	STK33	serine/threonine kinase 33 [Source:HGNC Symbol;Acc:HGNC:14568]	-0.80496	0.009436
ENSG00000175793	SFN	stratifin [Source:HGNC Symbol;Acc:HGNC:10773]	0.801796	0.009463
ENSG00000149346	SLX4IP	SLX4 interacting protein [Source:HGNC Symbol;Acc:HGNC:16225]	1.41725	0.009494
ENSG00000154415	PPP1R3A	protein phosphatase 1 regulatory subunit 3A [Source:HGNC Symbol;Acc:HGNC:9291]	2.74575	0.009532
ENSG00000137845	ADAM10	ADAM metallopeptidase domain 10 [Source:HGNC Symbol;Acc:HGNC:188]	-1.1141	0.009563
ENSG00000153786	ZDHHC7	zinc finger DHHC-type containing 7 [Source:HGNC Symbol;Acc:HGNC:18459]	-0.87463	0.00957
ENSG00000108175	ZMIZ1	zinc finger MIZ-type containing 1 [Source:HGNC Symbol;Acc:HGNC:16493]	-1.11326	0.009605
ENSG00000136014	USP44	ubiquitin specific peptidase 44 [Source:HGNC Symbol;Acc:HGNC:20064]	-0.80811	0.009702
ENSG00000176971	FIBIN	fin bud initiation factor homolog (zebrafish) [Source:HGNC Symbol;Acc:HGNC:33747]	-2.72998	0.00975
ENSG00000181260	MTHFD2P7	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2, methenyltetrahydrofolate cyclohydrolase pseudogene 7 [Source:HGNC Symbol;Acc:HGNC:48865]	-0.96756	0.009814
ENSG00000135617	PRADC1	protease associated domain containing 1 [Source:HGNC Symbol;Acc:HGNC:16047]	0.813756	0.009815
ENSG00000255537			0.864628	0.009842
ENSG00000196323	ZBTB44	zinc finger and BTB domain containing 44 [Source:HGNC Symbol;Acc:HGNC:25001]	-0.79595	0.009889
ENSG00000113643	RARS	arginyl-tRNA synthetase [Source:HGNC Symbol;Acc:HGNC:9870]	0.798795	0.009912

ENSG00000100554	ATP6V1D	ATPase H+ transporting V1 subunit D [Source:HGNC Symbol;Acc:HGNC:13527]	-0.80259	0.009954
ENSG00000173875	ZNF791	zinc finger protein 791 [Source:HGNC Symbol;Acc:HGNC:26895]	1.10754	0.009975

Πίνακας 8: Γονίδια που ξεχώρισαν ως διαφορετικά εκφρασμένα μέσω μετα-ανάλυσης, με βάση την τιμή p -value < 0.01, έπειτα από έκθεση σε υψηλή δόση ακτινοβολίας (>0.5Gy) στο διάστημα 16-24 ωρών με το αντίστοιχο HGNC σύμβολο με την αντίστοιχη περιγραφή και οι ακριβείς τιμές του υπολογισμένου p -value και της διαφοράς στην έκφραση

Γονίδια που ξεχώρισαν έπειτα από έκθεση σε χαμηλή δόση ακτινοβολίας στο group 16-24

ωρών

ENSG	HGNC	Description	Diff	P-value
ENSG00000151948	GLT1D1	glycosyltransferase 1 domain containing 1 [Source:HGNC Symbol;Acc:HGNC:26483]	-1.17952	0.000065
ENSG00000120952	PRAMEF2	PRAME family member 2 [Source:HGNC Symbol;Acc:HGNC:28841]	-1.26611	0.000067
ENSG00000221052	MIR1266	microRNA 1266 [Source:HGNC Symbol;Acc:HGNC:35334]	1.98709	0.00013
ENSG00000273174			-2.06243	0.000168
ENSG00000282673	PRB1	proline rich protein BstNI subfamily 1 [Source:HGNC Symbol;Acc:HGNC:9337]	-1.17678	0.000169
ENSG00000213050	TPM3P1	tropomyosin 3 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:16350]	-2.05204	0.000174
ENSG00000240125	RPL23AP75	ribosomal protein L23a pseudogene 75 [Source:HGNC Symbol;Acc:HGNC:35975]	1.8851	0.000228
ENSG00000173875	ZNF791	zinc finger protein 791 [Source:HGNC Symbol;Acc:HGNC:26895]	1.64781	0.000242
ENSG00000263069			-1.90991	0.000358
ENSG00000268153			-1.90871	0.000365
ENSG00000167850	CD300C	CD300c molecule [Source:HGNC Symbol;Acc:HGNC:19320]	-1.17237	0.000403
ENSG00000217791	ASS1P9	argininosuccinate synthetase 1 pseudogene 9 [Source:HGNC Symbol;Acc:HGNC:772]	1.86877	0.000442
ENSG00000256235	SMIM3	small integral membrane protein 3 [Source:HGNC Symbol;Acc:HGNC:30248]	-1.25525	0.000468
ENSG00000231341	VDAC1P6	voltage dependent anion channel 1 pseudogene 6 [Source:HGNC Symbol;Acc:HGNC:37481]	-1.86187	0.00047
ENSG00000254593	OR7E126P	olfactory receptor family 7 subfamily E member 126 pseudogene [Source:HGNC Symbol;Acc:HGNC:15304]	1.85171	0.000483
ENSG0000021488	SLC7A9	solute carrier family 7 member 9 [Source:HGNC Symbol;Acc:HGNC:11067]	0.999887	0.000487
ENSG00000241721	SUMO1P1	SUMO1 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:33148]	-1.81546	0.000591
ENSG00000282817	PRAMEF2	PRAME family member 2 [Source:HGNC Symbol;Acc:HGNC:28841]	-1.15308	0.000612
ENSG00000235082	SUMO1P3	SUMO1 pseudogene 3 [Source:HGNC Symbol;Acc:HGNC:33150]	-1.80852	0.000614
ENSG00000152932	RAB3C	RAB3C, member RAS oncogene family [Source:HGNC Symbol;Acc:HGNC:30269]	-0.9766	0.000711
ENSG00000207956	MIR579	microRNA 579 [Source:HGNC Symbol;Acc:HGNC:32835]	1.67894	0.000716
ENSG00000169432	SCN9A	sodium voltage-gated channel alpha subunit 9 [Source:HGNC Symbol;Acc:HGNC:10597]	-0.96696	0.000717
ENSG00000162992	NEUROD1	neuronal differentiation 1 [Source:HGNC Symbol;Acc:HGNC:7762]	1.032	0.000783
ENSG00000275517	PRB1	proline rich protein BstNI subfamily 1 [Source:HGNC Symbol;Acc:HGNC:9337]	-1.02406	0.000877
ENSG00000231898			-1.72024	0.000996
ENSG00000196757	ZNF700	zinc finger protein 700 [Source:HGNC Symbol;Acc:HGNC:25292]	1.00771	0.001023
ENSG0000024526	DEPDC1	DEP domain containing 1 [Source:HGNC Symbol;Acc:HGNC:22949]	0.937391	0.001096
ENSG00000159055	MIS18A	MIS18 kinetochore protein A [Source:HGNC Symbol;Acc:HGNC:1286]	1.47427	0.001128
ENSG00000144426	NBEAL1	neurobeachin like 1 [Source:HGNC Symbol;Acc:HGNC:20681]	0.933208	0.001137
ENSG00000163507	KIAA1524	KIAA1524 [Source:HGNC Symbol;Acc:HGNC:29302]	0.918755	0.001249
ENSG00000184566			-1.67522	0.001264
ENSG00000231824	AKAIN1	A-kinase anchor inhibitor 1 [Source:HGNC Symbol;Acc:HGNC:28285]	-1.57565	0.001275
ENSG00000204363	SPANXN5	SPANX family member N5 [Source:HGNC Symbol;Acc:HGNC:33178]	4.82299	0.001286
ENSG00000056972	TRAF3IP2	TRAF3 interacting protein 2 [Source:HGNC Symbol;Acc:HGNC:1343]	0.919693	0.001313
ENSG00000267980			1.56974	0.001318
ENSG00000158427	TMSB15B	thymosin beta 15B [Source:HGNC Symbol;Acc:HGNC:28612]	0.983051	0.001398
ENSG00000269226			0.983051	0.001398
ENSG00000187773	FAM69C	family with sequence similarity 69 member C [Source:HGNC Symbol;Acc:HGNC:31729]	-0.90911	0.001487
ENSG00000107745	MICU1	mitochondrial calcium uptake 1 [Source:HGNC Symbol;Acc:HGNC:1530]	-1.38804	0.001491
ENSG00000241549	GUSBP2	glucuronidase, beta pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:18792]	0.978147	0.001505
ENSG00000234515	PPP1R2P1	protein phosphatase 1 regulatory inhibitor subunit 2 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:9289]	-1.36065	0.001521

ENSG00000235513			-1.63584	0.001579
ENSG00000100629	CEP128	centrosomal protein 128 [Source:HGNC Symbol;Acc:HGNC:20359]	0.90968	0.001595
ENSG00000258352			1.62661	0.001604
ENSG00000229196			4.47613	0.001612
ENSG00000173404	INSM1	INSM transcriptional repressor 1 [Source:HGNC Symbol;Acc:HGNC:6090]	-0.90029	0.001631
ENSG00000170426	SDR9C7	short chain dehydrogenase/reductase family 9C member 7 [Source:HGNC Symbol;Acc:HGNC:29958]	1.01762	0.00167
ENSG00000099797	TECR	trans-2,3-enoyl-CoA reductase [Source:HGNC Symbol;Acc:HGNC:4551]	0.962805	0.001741
ENSG00000214362	RPS3AP36	ribosomal protein S3a pseudogene 36 [Source:HGNC Symbol;Acc:HGNC:35596]	-1.61406	0.00179
ENSG00000215183	MSMP	microseminoprotein, prostate associated [Source:HGNC Symbol;Acc:HGNC:29663]	-1.13142	0.001822
ENSG00000153292	ADGRF1	adhesion G protein-coupled receptor F1 [Source:HGNC Symbol;Acc:HGNC:18990]	-0.88181	0.001882
ENSG00000253614			1.50346	0.001909
ENSG00000282027			1.50346	0.001909
ENSG00000213601	KRT18P19	keratin 18 pseudogene 19 [Source:HGNC Symbol;Acc:HGNC:33387]	1.20315	0.001948
ENSG00000182095	TNRC18	trinucleotide repeat containing 18 [Source:HGNC Symbol;Acc:HGNC:11962]	-1.02245	0.002041
ENSG00000163374	YY1AP1	YY1 associated protein 1 [Source:HGNC Symbol;Acc:HGNC:30935]	-0.87613	0.002047
ENSG00000207697	MIR573	microRNA 573 [Source:HGNC Symbol;Acc:HGNC:32829]	1.48936	0.002065
ENSG00000158023	WDR66	WD repeat domain 66 [Source:HGNC Symbol;Acc:HGNC:28506]	-0.88178	0.00215
ENSG00000125798	FOXA2	forkhead box A2 [Source:HGNC Symbol;Acc:HGNC:5022]	-0.8671	0.002232
ENSG00000167825	OR5I1	olfactory receptor family 5 subfamily I member 1 [Source:HGNC Symbol;Acc:HGNC:8347]	-0.87283	0.002236
ENSG00000245060	LINC00847	long intergenic non-protein coding RNA 847 [Source:HGNC Symbol;Acc:HGNC:45050]	-4.03799	0.002257
ENSG00000259025			1.47228	0.002271
ENSG00000259769			1.47228	0.002271
ENSG00000148308	GTF3C5	general transcription factor IIIC subunit 5 [Source:HGNC Symbol;Acc:HGNC:4668]	1.0247	0.002275
ENSG00000215991	MIR208B	microRNA 208b [Source:HGNC Symbol;Acc:HGNC:33669]	1.46922	0.00231
ENSG00000225670	CADM3-AS1	CADM3 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:40812]	1.46804	0.002325
ENSG00000128872	TMOD2	tropomodulin 2 [Source:HGNC Symbol;Acc:HGNC:11872]	-1.18582	0.002353
ENSG00000260223			-1.55779	0.002355
ENSG00000198056	PRIM1	primase (DNA) subunit 1 [Source:HGNC Symbol;Acc:HGNC:9369]	1.09048	0.002389
ENSG00000172273	HINFP	histone H4 transcription factor [Source:HGNC Symbol;Acc:HGNC:17850]	-0.86235	0.002404
ENSG00000135956	TMEM127	transmembrane protein 127 [Source:HGNC Symbol;Acc:HGNC:26038]	-0.85882	0.002434
ENSG00000273814	OCLN	occludin [Source:HGNC Symbol;Acc:HGNC:8104]	0.92574	0.002494
ENSG00000055044	NOP58	NOP58 ribonucleoprotein [Source:HGNC Symbol;Acc:HGNC:29926]	0.860037	0.00252
ENSG00000260268	LINC00919	long intergenic non-protein coding RNA 919 [Source:HGNC Symbol;Acc:HGNC:48610]	1.0037	0.002521
ENSG00000236740			-1.45155	0.002548
ENSG00000277013			1.53735	0.002549
ENSG00000199031	MIR371A	microRNA 371a [Source:HGNC Symbol;Acc:HGNC:31785]	-1.44645	0.002622
ENSG00000253117	OC90	otoconin 90 [Source:HGNC Symbol;Acc:HGNC:8100]	1.44532	0.002638
ENSG00000080200	CRYBG3	crystallin beta-gamma domain containing 3 [Source:HGNC Symbol;Acc:HGNC:34427]	0.853406	0.002689
ENSG00000236336			-1.4419	0.002689
ENSG00000213214	ARHGEF35	Rho guanine nucleotide exchange factor 35 [Source:HGNC Symbol;Acc:HGNC:33846]	-1.52335	0.002713
ENSG00000107187	LHX3	LIM homeobox 3 [Source:HGNC Symbol;Acc:HGNC:6595]	-0.8482	0.002718
ENSG00000131351	HAUS8	HAUS augmin like complex subunit 8 [Source:HGNC Symbol;Acc:HGNC:30532]	0.855519	0.002723
ENSG00000145945	FAM50B	family with sequence similarity 50 member B [Source:HGNC Symbol;Acc:HGNC:18789]	0.90684	0.002772
ENSG00000267179			0.906331	0.00288
ENSG00000243040	RBMV2FP	RNA binding motif protein, Y-linked, family 2, member F pseudogene [Source:HGNC Symbol;Acc:HGNC:23891]	0.912311	0.002883
ENSG00000005059	MCUB	mitochondrial calcium uniporter dominant negative beta subunit [Source:HGNC	1.26486	0.002891

		Symbol;Acc:HGNC:26076]		
ENSG00000132970	WASF3	WAS protein family member 3 [Source:HGNC Symbol;Acc:HGNC:12734]	-0.85363	0.002912
ENSG00000171723	GPHN	gephyrin [Source:HGNC Symbol;Acc:HGNC:15465]	0.841033	0.00294
ENSG00000259969			-1.50847	0.002969
ENSG00000165512	ZNF22	zinc finger protein 22 [Source:HGNC Symbol;Acc:HGNC:13012]	0.842003	0.002989
ENSG00000267706			-0.9804	0.002997
ENSG00000197822	OCLN	occludin [Source:HGNC Symbol;Acc:HGNC:8104]	0.907147	0.003021
ENSG00000204637			-1.50263	0.003039
ENSG00000229009	TMPRSS11GP	transmembrane protease, serine 11G, pseudogene [Source:HGNC Symbol;Acc:HGNC:42983]	-1.50199	0.003077
ENSG00000273567	REXO1L3P	REX1, RNA exonuclease 1 homolog-like 3, pseudogene [Source:HGNC Symbol;Acc:HGNC:32240]	1.41758	0.003077
ENSG00000197956	S100A6	S100 calcium binding protein A6 [Source:HGNC Symbol;Acc:HGNC:10496]	-0.83846	0.003107
ENSG00000171649	ZIK1	zinc finger protein interacting with K protein 1 [Source:HGNC Symbol;Acc:HGNC:33104]	0.838902	0.003114
ENSG00000276595	PTCHD3	patched domain containing 3 [Source:HGNC Symbol;Acc:HGNC:24776]	-1.19047	0.003158
ENSG00000013293	SLC7A14	solute carrier family 7 member 14 [Source:HGNC Symbol;Acc:HGNC:29326]	-0.84261	0.003267
ENSG00000216663			-1.49172	0.003307
ENSG00000279418	LINC00244	long intergenic non-protein coding RNA 244 [Source:HGNC Symbol;Acc:HGNC:13245]	-1.40294	0.003337
ENSG00000125686	MED1	mediator complex subunit 1 [Source:HGNC Symbol;Acc:HGNC:9234]	-0.83386	0.003342
ENSG00000067842	ATP2B3	ATPase plasma membrane Ca2+ transporting 3 [Source:HGNC Symbol;Acc:HGNC:816]	-1.17	0.003417
ENSG00000251733	SCARNA8	small Cajal body-specific RNA 8 [Source:HGNC Symbol;Acc:HGNC:32564]	1.39755	0.003438
ENSG00000197182	MIRLET7BHG	MIRLET7B host gene [Source:HGNC Symbol;Acc:HGNC:37189]	0.969715	0.003441
ENSG00000235065	RPL24P2	ribosomal protein L24 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:16597]	1.39369	0.003512
ENSG00000260391			-2.29551	0.003657
ENSG00000204252	HLA-DOA	major histocompatibility complex, class II, DO alpha [Source:HGNC Symbol;Acc:HGNC:4936]	-0.89242	0.003713
ENSG00000149636	DSN1	DSN1 homolog, MIS12 kinetochore complex component [Source:HGNC Symbol;Acc:HGNC:16165]	0.82573	0.003761
ENSG00000147647	DPYS	dihydropyrimidinase [Source:HGNC Symbol;Acc:HGNC:3013]	-0.81835	0.003827
ENSG00000197566	ZNF624	zinc finger protein 624 [Source:HGNC Symbol;Acc:HGNC:29254]	-0.83523	0.003935
ENSG00000145863	GABRA6	gamma-aminobutyric acid type A receptor alpha6 subunit [Source:HGNC Symbol;Acc:HGNC:4080]	-3.44719	0.00397
ENSG00000148200	NR6A1	nuclear receptor subfamily 6 group A member 1 [Source:HGNC Symbol;Acc:HGNC:7985]	-0.81871	0.004034
ENSG00000274286	ADRA2B	adrenoceptor alpha 2B [Source:HGNC Symbol;Acc:HGNC:282]	-0.82114	0.004062
ENSG00000251580			1.44259	0.004101
ENSG00000106462	EZH2	enhancer of zeste 2 polycomb repressive complex 2 subunit [Source:HGNC Symbol;Acc:HGNC:3527]	0.816244	0.004147
ENSG00000186973	FAM183A	family with sequence similarity 183 member A [Source:HGNC Symbol;Acc:HGNC:34347]	1.23871	0.004183
ENSG00000187944	C2orf66	chromosome 2 open reading frame 66 [Source:HGNC Symbol;Acc:HGNC:33809]	0.812169	0.004214
ENSG00000241131	LINC02032	long intergenic non-protein coding RNA 2032 [Source:HGNC Symbol;Acc:HGNC:52866]	3.3894	0.004229
ENSG00000171792	RHNO1	RAD9-HUS1-RAD1 interacting nuclear orphan 1 [Source:HGNC Symbol;Acc:HGNC:28206]	0.874195	0.004234
ENSG00000160072	ATAD3B	ATPase family, AAA domain containing 3B [Source:HGNC Symbol;Acc:HGNC:24007]	1.10645	0.004347
ENSG00000232780	GAPDHP74	glyceraldehyde-3-phosphate dehydrogenase pseudogene 74 [Source:HGNC Symbol;Acc:HGNC:50647]	-1.3548	0.004352
ENSG00000251510			-1.75711	0.004478
ENSG00000145868	FBXO38	F-box protein 38 [Source:HGNC Symbol;Acc:HGNC:28844]	-0.80244	0.004577
ENSG00000224578	HNRNPA1P48	heterogeneous nuclear ribonucleoprotein A1 pseudogene 48 [Source:HGNC Symbol;Acc:HGNC:48778]	-1.41892	0.004618
ENSG00000233304	LINC01346	long intergenic non-protein coding RNA 1346 [Source:HGNC Symbol;Acc:HGNC:50563]	1.25861	0.004629
ENSG00000196597	ZNF782	zinc finger protein 782 [Source:HGNC Symbol;Acc:HGNC:33110]	1.00819	0.00464
ENSG00000248223			1.41197	0.004767
ENSG00000145451	GLRA3	glycine receptor alpha 3 [Source:HGNC Symbol;Acc:HGNC:4328]	0.801336	0.004777

ENSG00000204444	APOM	apolipoprotein M [Source:HGNC Symbol;Acc:HGNC:13916]	0.801998	0.004807
ENSG00000227581			1.33648	0.004812
ENSG00000215571	GRK6P1	G protein-coupled receptor kinase 6 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:4547]	3.26263	0.004883
ENSG00000269556	TMEM185A	transmembrane protein 185A [Source:HGNC Symbol;Acc:HGNC:17125]	0.793714	0.004912
ENSG00000169684	CHRNA5	cholinergic receptor nicotinic alpha 5 subunit [Source:HGNC Symbol;Acc:HGNC:1959]	0.809749	0.004929
ENSG00000278690			2.10254	0.004955
ENSG00000282071			2.10254	0.004955
ENSG00000146722			0.939921	0.005044
ENSG00000275576			1.32503	0.005124
ENSG00000156671	SAMD8	sterile alpha motif domain containing 8 [Source:HGNC Symbol;Acc:HGNC:26320]	-0.79241	0.005154
ENSG00000205755	CRLF2	cytokine receptor-like factor 2 [Source:HGNC Symbol;Acc:HGNC:14281]	-0.79027	0.005187
ENSG00000072182	ASIC4	acid sensing ion channel subunit family member 4 [Source:HGNC Symbol;Acc:HGNC:21263]	-0.78829	0.005235
ENSG00000270623	SPCS2P2	signal peptidase complex subunit 2 homolog (S. cerevisiae) pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:45235]	1.31809	0.005322
ENSG00000162695	SLC30A7	solute carrier family 30 member 7 [Source:HGNC Symbol;Acc:HGNC:19306]	0.812007	0.005328
ENSG00000228523			-1.39328	0.005329
ENSG00000221525	MIR1185-1	microRNA 1185-1 [Source:HGNC Symbol;Acc:HGNC:35257]	-1.3127	0.005481
ENSG00000221614	MIR1185-2	microRNA 1185-2 [Source:HGNC Symbol;Acc:HGNC:35254]	-1.3127	0.005481
ENSG00000261572			1.38407	0.005506
ENSG00000007516	BAIAP3	BAI1 associated protein 3 [Source:HGNC Symbol;Acc:HGNC:948]	-0.78764	0.00551
ENSG00000028277	POU2F2	POU class 2 homeobox 2 [Source:HGNC Symbol;Acc:HGNC:9213]	-1.11917	0.005556
ENSG00000223583			1.30813	0.00562
ENSG00000087365	SF3B2	splicing factor 3b subunit 2 [Source:HGNC Symbol;Acc:HGNC:10769]	1.00634	0.005646
ENSG00000100124	ANKRD54	ankyrin repeat domain 54 [Source:HGNC Symbol;Acc:HGNC:25185]	0.786063	0.005653
ENSG00000149294	NCAM1	neural cell adhesion molecule 1 [Source:HGNC Symbol;Acc:HGNC:7656]	-0.77818	0.005703
ENSG00000207865			1.30509	0.005714
ENSG00000270416	REXO1L4P	REX1, RNA exonuclease 1 homolog-like 4, pseudogene [Source:HGNC Symbol;Acc:HGNC:32241]	1.30321	0.005773
ENSG00000271381	REXO1L9P	REX1, RNA exonuclease 1 homolog-like 9, pseudogene [Source:HGNC Symbol;Acc:HGNC:44432]	1.30321	0.005773
ENSG00000274343			1.30321	0.005773
ENSG00000275928	REXO1L11P	REX1, RNA exonuclease 1 homolog-like 11, pseudogene [Source:HGNC Symbol;Acc:HGNC:44434]	1.30321	0.005773
ENSG00000276006	REXO1L12P	REX1, RNA exonuclease 1 homolog-like 12, pseudogene [Source:HGNC Symbol;Acc:HGNC:44435]	1.30321	0.005773
ENSG00000276540	REXO1L10P	REX1, RNA exonuclease 1 homolog-like 10, pseudogene [Source:HGNC Symbol;Acc:HGNC:44433]	1.30321	0.005773
ENSG00000276998	REXO1L2P	REX1, RNA exonuclease 1 homolog-like 2, pseudogene [Source:HGNC Symbol;Acc:HGNC:32239]	1.30321	0.005773
ENSG00000277422	REXO1L6P	REX1, RNA exonuclease 1 homolog-like 6, pseudogene [Source:HGNC Symbol;Acc:HGNC:32243]	1.30321	0.005773
ENSG00000277436	REXO1L5P	REX1, RNA exonuclease 1 homolog-like 5, pseudogene [Source:HGNC Symbol;Acc:HGNC:32242]	1.30321	0.005773
ENSG00000276291	FRG1HP	FSDH region gene 1 family member H, pseudogene [Source:HGNC Symbol;Acc:HGNC:51767]	1.30065	0.005854
ENSG00000196792	STRN3	striatin 3 [Source:HGNC Symbol;Acc:HGNC:15720]	-0.77512	0.005933
ENSG00000259439	LINC01833	long intergenic non-protein coding RNA 1833 [Source:HGNC Symbol;Acc:HGNC:52644]	0.904428	0.005993
ENSG00000132326	PER2	period circadian clock 2 [Source:HGNC Symbol;Acc:HGNC:8846]	0.990249	0.006045
ENSG00000157322	CLEC18A	C-type lectin domain family 18 member A [Source:HGNC Symbol;Acc:HGNC:30388]	0.957753	0.006047
ENSG00000214286	PDCL3P3	phosducin-like 3 pseudogene 3 [Source:HGNC Symbol;Acc:HGNC:44504]	-1.36638	0.006051
ENSG00000171049	FPR2	formyl peptide receptor 2 [Source:HGNC Symbol;Acc:HGNC:3827]	-0.77728	0.00607
ENSG00000145919	BOD1	biorientation of chromosomes in cell division 1 [Source:HGNC Symbol;Acc:HGNC:25114]	-1.64868	0.006071
ENSG00000276231	PIK3R6	phosphoinositide-3-kinase regulatory subunit 6 [Source:HGNC Symbol;Acc:HGNC:27101]	-0.77879	0.006075

ENSG00000182600	C2orf82	chromosome 2 open reading frame 82 [Source:HGNC Symbol;Acc:HGNC:33763]	-0.77432	0.006111
ENSG00000180318	ALX1	ALX homeobox 1 [Source:HGNC Symbol;Acc:HGNC:1494]	-0.77782	0.006161
ENSG00000227620	ALG1L8P	asparagine-linked glycosylation 1-like 8, pseudogene [Source:HGNC Symbol;Acc:HGNC:44377]	-1.29061	0.006183
ENSG00000135974	C2orf49	chromosome 2 open reading frame 49 [Source:HGNC Symbol;Acc:HGNC:28772]	-1.68541	0.006225
ENSG00000150275	PCDH15	protocadherin related 15 [Source:HGNC Symbol;Acc:HGNC:14674]	-0.7772	0.00627
ENSG00000148229	POLE3	DNA polymerase epsilon 3, accessory subunit [Source:HGNC Symbol;Acc:HGNC:13546]	1.70457	0.006297
ENSG00000197092	GOLGA6L16P	golgin A6 family-like 16, pseudogene [Source:HGNC Symbol;Acc:HGNC:38117]	1.28599	0.00634
ENSG00000223698	GOLGA6L11P	golgin A6 family-like 11, pseudogene [Source:HGNC Symbol;Acc:HGNC:23958]	1.28599	0.00634
ENSG00000213201	FABP5P10	fatty acid binding protein 5 pseudogene 10 [Source:HGNC Symbol;Acc:HGNC:31066]	2.61133	0.00638
ENSG00000261829			1.35486	0.006428
ENSG00000276971	TRPV6	transient receptor potential cation channel subfamily V member 6 [Source:HGNC Symbol;Acc:HGNC:14006]	0.826878	0.006479
ENSG00000005102	MEOX1	mesenchyme homeobox 1 [Source:HGNC Symbol;Acc:HGNC:7013]	-0.77375	0.006541
ENSG00000180747	SMG1P3	SMG1P3, nonsense mediated mRNA decay associated PI3K related kinase pseudogene 3 [Source:HGNC Symbol;Acc:HGNC:49860]	0.895405	0.006545
ENSG00000141579	ZNF750	zinc finger protein 750 [Source:HGNC Symbol;Acc:HGNC:25843]	0.762413	0.006561
ENSG00000139734	DIAPH3	diaphanous related formin 3 [Source:HGNC Symbol;Acc:HGNC:15480]	0.768867	0.006581
ENSG00000175985	PLEKHD1	pleckstrin homology and coiled-coil domain containing D1 [Source:HGNC Symbol;Acc:HGNC:20148]	1.09753	0.006589
ENSG00000187242	KRT12	keratin 12 [Source:HGNC Symbol;Acc:HGNC:6414]	0.96799	0.006731
ENSG00000168393	DTYMK	deoxythymidylate kinase [Source:HGNC Symbol;Acc:HGNC:3061]	0.817161	0.006737
ENSG00000235919	ASH1L-AS1	ASH1L antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:44146]	-0.88424	0.006788
ENSG00000196782	MAML3	mastermind like transcriptional coactivator 3 [Source:HGNC Symbol;Acc:HGNC:16272]	-1.13442	0.006809
ENSG00000122254	HS3ST2	heparan sulfate-glucosamine 3-sulfotransferase 2 [Source:HGNC Symbol;Acc:HGNC:5195]	-0.76281	0.006836
ENSG00000240590	RPSAP48	ribosomal protein SA pseudogene 48 [Source:HGNC Symbol;Acc:HGNC:36646]	1.27183	0.006847
ENSG00000167613	LAIR1	leukocyte associated immunoglobulin like receptor 1 [Source:HGNC Symbol;Acc:HGNC:6477]	-1.98979	0.006882
ENSG00000206203	TSSK2	testis specific serine kinase 2 [Source:HGNC Symbol;Acc:HGNC:11401]	-1.34353	0.006912
ENSG00000255404			-1.33812	0.007026
ENSG00000123473	STIL	SCL/TAL1 interrupting locus [Source:HGNC Symbol;Acc:HGNC:10879]	0.765637	0.007038
ENSG00000184557	SOCS3	suppressor of cytokine signaling 3 [Source:HGNC Symbol;Acc:HGNC:19391]	0.755377	0.007054
ENSG00000213918	DNASE1	deoxyribonuclease 1 [Source:HGNC Symbol;Acc:HGNC:2956]	-0.76151	0.007056
ENSG00000274185	LILRA4	leukocyte immunoglobulin like receptor A4 [Source:HGNC Symbol;Acc:HGNC:15503]	-0.88407	0.007065
ENSG00000275210			-0.88407	0.007065
ENSG00000275893			-0.88407	0.007065
ENSG00000276798	LILRA4	leukocyte immunoglobulin like receptor A4 [Source:HGNC Symbol;Acc:HGNC:15503]	-0.88407	0.007065
ENSG00000104903	LYL1	LYL1, basic helix-loop-helix family member [Source:HGNC Symbol;Acc:HGNC:6734]	-1.38103	0.007085
ENSG00000254726	MEX3A	mex-3 RNA binding family member A [Source:HGNC Symbol;Acc:HGNC:33482]	-1.239	0.007086
ENSG00000118160	SLC8A2	solute carrier family 8 member A2 [Source:HGNC Symbol;Acc:HGNC:11069]	-0.87655	0.007129
ENSG00000109103	UNC119	unc-119 lipid binding chaperone [Source:HGNC Symbol;Acc:HGNC:12565]	-0.81081	0.007145
ENSG00000224435	NF1P6	neurofibromin 1 pseudogene 6 [Source:HGNC Symbol;Acc:HGNC:7771]	-1.26306	0.00718
ENSG00000277529	NF1P10	neurofibromin 1 pseudogene 10 [Source:HGNC Symbol;Acc:HGNC:51737]	-1.26306	0.00718
ENSG00000198976	MIR429	microRNA 429 [Source:HGNC Symbol;Acc:HGNC:13784]	1.26076	0.00727
ENSG00000272779			0.876643	0.007296
ENSG00000235400			-1.33498	0.00745
ENSG00000260539			-2.91179	0.007576
ENSG00000176007	FAM220BP	family with sequence similarity 220 member B, pseudogene [Source:HGNC Symbol;Acc:HGNC:23678]	1.25273	0.007593
ENSG00000111196	MAGOHB	mago homolog B, exon junction complex core component [Source:HGNC Symbol;Acc:HGNC:25504]	0.837808	0.007729
ENSG00000263293		thyroid cancer-associated transcript 158 [Source:EntrezGene;Acc:102724508]	-1.31854	0.007807

ENSG00000254013	MAP2K1P1	mitogen-activated protein kinase 1 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:6841]	2.88932	0.007808
ENSG00000234661	CHL1-AS1	CHL1 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:40148]	-1.32376	0.007819
ENSG00000159556	ISL2	ISL LIM homeobox 2 [Source:HGNC Symbol;Acc:HGNC:18524]	-0.75366	0.007899
ENSG00000187997	C17orf99	chromosome 17 open reading frame 99 [Source:HGNC Symbol;Acc:HGNC:34490]	-0.80048	0.007902
ENSG00000120328	PCDHB12	protocadherin beta 12 [Source:HGNC Symbol;Acc:HGNC:8683]	0.928158	0.007965
ENSG00000223939	PPP1R2P1	protein phosphatase 1 regulatory inhibitor subunit 2 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:9289]	-1.31149	0.007991
ENSG00000226914	PPP1R2P1	protein phosphatase 1 regulatory inhibitor subunit 2 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:9289]	-1.31149	0.007991
ENSG00000227435	PPP1R2P1	protein phosphatase 1 regulatory inhibitor subunit 2 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:9289]	-1.31149	0.007991
ENSG00000233498	PPP1R2P1	protein phosphatase 1 regulatory inhibitor subunit 2 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:9289]	-1.31149	0.007991
ENSG00000234642	PPP1R2P1	protein phosphatase 1 regulatory inhibitor subunit 2 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:9289]	-1.31149	0.007991
ENSG00000236093	PPP1R2P1	protein phosphatase 1 regulatory inhibitor subunit 2 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:9289]	-1.31149	0.007991
ENSG00000236488	PPP1R2P1	protein phosphatase 1 regulatory inhibitor subunit 2 pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:9289]	-1.31149	0.007991
ENSG00000276999			-1.24136	0.008074
ENSG00000254211	LINC01485	long intergenic non-protein coding RNA 1485 [Source:HGNC Symbol;Acc:HGNC:51138]	-1.30973	0.008078
ENSG00000237541	HLA-DQA2	major histocompatibility complex, class II, DQ alpha 2 [Source:HGNC Symbol;Acc:HGNC:4943]	-0.73974	0.008198
ENSG00000173699	SPATA3	spermatogenesis associated 3 [Source:HGNC Symbol;Acc:HGNC:17884]	-0.73949	0.008225
ENSG00000260778			1.23674	0.008277
ENSG00000274753			1.23674	0.008277
ENSG00000175643	RMI2	RecQ mediated genome instability 2 [Source:HGNC Symbol;Acc:HGNC:28349]	0.940336	0.008299
ENSG00000005073	HOXA11	homeobox A11 [Source:HGNC Symbol;Acc:HGNC:5101]	-0.74204	0.008389
ENSG00000156502	SUPV3L1	Suv3 like RNA helicase [Source:HGNC Symbol;Acc:HGNC:11471]	-0.7452	0.008449
ENSG00000240542	KRTAP9-1	keratin associated protein 9-1 [Source:HGNC Symbol;Acc:HGNC:18912]	1.23252	0.008467
ENSG00000255559	ZNF252P-AS1	ZNF252P antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:27821]	-0.86395	0.00856
ENSG00000205677	LYPLA2P1	lysophospholipase II pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:21069]	-1.22906	0.008626
ENSG00000217281	LYPLA2P1	lysophospholipase II pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:21069]	-1.22906	0.008626
ENSG00000228285	LYPLA2P1	lysophospholipase II pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:21069]	-1.22906	0.008626
ENSG00000232376	LYPLA2P1	lysophospholipase II pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:21069]	-1.22906	0.008626
ENSG00000232639	LYPLA2P1	lysophospholipase II pseudogene 1 [Source:HGNC Symbol;Acc:HGNC:21069]	-1.22906	0.008626
ENSG00000231177	LINC00852	long intergenic non-protein coding RNA 852 [Source:HGNC Symbol;Acc:HGNC:29904]	-0.87826	0.008678
ENSG00000196243	LINC00615	long intergenic non-protein coding RNA 615 [Source:HGNC Symbol;Acc:HGNC:26343]	0.856828	0.008692
ENSG00000198923			-1.29921	0.008771
ENSG00000131037	EPS8L1	EPS8 like 1 [Source:HGNC Symbol;Acc:HGNC:21295]	0.738153	0.00882
ENSG00000225308	ASS1P11	argininosuccinate synthetase 1 pseudogene 11 [Source:HGNC Symbol;Acc:HGNC:761]	-1.22451	0.00884
ENSG00000262943	ALOX12P2	arachidonate 12-lipoxygenase pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:432]	0.87133	0.008841
ENSG00000197106	SLC6A17	solute carrier family 6 member 17 [Source:HGNC Symbol;Acc:HGNC:31399]	0.734208	0.008879
ENSG00000089682	RBM41	RNA binding motif protein 41 [Source:HGNC Symbol;Acc:HGNC:25617]	-0.7358	0.008915
ENSG00000206630	SNORD60	small nucleolar RNA, C/D box 60 [Source:HGNC Symbol;Acc:HGNC:10217]	1.22258	0.008932
ENSG00000281010		Small nucleolar RNA snoR1 [Source:RFAM;Acc:RF00345]	1.22258	0.008932
ENSG00000140030	GPR65	G protein-coupled receptor 65 [Source:HGNC Symbol;Acc:HGNC:4517]	-0.73681	0.008941
ENSG00000162736	NCSTN	nicastatin [Source:HGNC Symbol;Acc:HGNC:17091]	-0.74206	0.008955
ENSG00000254967			1.22104	0.009006
ENSG00000118507	AKAP7	A-kinase anchoring protein 7 [Source:HGNC Symbol;Acc:HGNC:377]	0.763326	0.009036
ENSG00000276794			0.787778	0.009064
ENSG00000048462	TNFRSF17	TNF receptor superfamily member 17 [Source:HGNC Symbol;Acc:HGNC:11913]	-1.14773	0.009127

ENSG00000277606			0.854325	0.009186
ENSG00000278019			0.854325	0.009186
ENSG00000174989	FBXW8	F-box and WD repeat domain containing 8 [Source:HGNC Symbol;Acc:HGNC:13597]	-0.85216	0.009192
ENSG00000183307	CECR6	cat eye syndrome chromosome region, candidate 6 [Source:HGNC Symbol;Acc:HGNC:1844]	-1.12847	0.009214
ENSG00000257052			-1.31115	0.009237
ENSG00000137393	RNF144B	ring finger protein 144B [Source:HGNC Symbol;Acc:HGNC:21578]	-0.72838	0.009245
ENSG00000282989			-1.21542	0.009282
ENSG00000231464			-1.03226	0.009344
ENSG00000178764	ZHX2	zinc fingers and homeoboxes 2 [Source:HGNC Symbol;Acc:HGNC:18513]	-1.32883	0.009396
ENSG00000186678			1.21182	0.009463
ENSG00000204361	NXPE2	neurexophilin and PC-esterase domain family member 2 [Source:HGNC Symbol;Acc:HGNC:26331]	1.28872	0.009473
ENSG00000213068			-1.27637	0.009523
ENSG00000227563	RNF11P2	ring finger protein 11 pseudogene 2 [Source:HGNC Symbol;Acc:HGNC:16535]	-1.27337	0.009678
ENSG00000251095			1.28146	0.009698
ENSG00000213123	TCTEX1D2	Tctex1 domain containing 2 [Source:HGNC Symbol;Acc:HGNC:28482]	0.7787	0.009709
ENSG00000233441	CYP2AB1P	cytochrome P450 family 2 subfamily AB member 1, pseudogene [Source:HGNC Symbol;Acc:HGNC:39975]	-1.20672	0.009725
ENSG00000225953	SATB2-AS1	SATB2 antisense RNA 1 [Source:HGNC Symbol;Acc:HGNC:26490]	-1.12927	0.009748
ENSG00000245317			1.2057	0.009778
ENSG00000167104	BPIFB6	BPI fold containing family B member 6 [Source:HGNC Symbol;Acc:HGNC:16504]	-0.83957	0.009782
ENSG00000231840			-2.19712	0.009801
ENSG00000256120			-1.36746	0.009822
ENSG00000237697			2.72425	0.00983
ENSG00000257943			1.62302	0.009858
ENSG00000258684	BMS1P16	BMS1, ribosome biogenesis factor pseudogene 16 [Source:HGNC Symbol;Acc:HGNC:49161]	0.841633	0.009861
ENSG00000258780	BMS1P15	BMS1, ribosome biogenesis factor pseudogene 15 [Source:HGNC Symbol;Acc:HGNC:49160]	0.841633	0.009861
ENSG00000104499	GML	glycosylphosphatidylinositol anchored molecule like [Source:HGNC Symbol;Acc:HGNC:4375]	-0.72484	0.009882
ENSG00000181499	OR6T1	olfactory receptor family 6 subfamily T member 1 [Source:HGNC Symbol;Acc:HGNC:14848]	-0.91179	0.00989
ENSG00000188486	H2AFX	H2A histone family member X [Source:HGNC Symbol;Acc:HGNC:4739]	0.721296	0.009893
ENSG00000227907			-2.71822	0.009916
ENSG00000171793	CTPS1	CTP synthase 1 [Source:HGNC Symbol;Acc:HGNC:2519]	2.71585	0.00995

Πίνακας 9: Γονίδια που ξεχώρισαν ως διαφορετικά εκφρασμένα μέσω μετα-ανάλυσης, με βάση την τιμή p-value<0.01, έπειτα από έκθεση σε χαμηλή δόση ακτινοβολίας(<0.5Gy) στο διάστημα 16-24 ωρών με το αντίστοιχο HGNC σύμβολο με την αντίστοιχη περιγραφή και οι ακριβείς τιμές του υπολογισμένου p-value και της διαφοράς στην έκφραση.

Η Γονιδιακή Οντολογία αποτελεί ένα ελεγχόμενο λεξιλόγιο γονιδιακών όρων που είναι δομημένο. Η GO ορίζει τις έννοιες που χρησιμοποιούνται για να περιγράψουν τη λειτουργία των γονιδίων και τις σχέσεις μεταξύ αυτών των εννοιών (Ashburner et al., 2000; Gene Ontology, 2015). Η Γονιδιακή Οντολογία διαιρείται σε τρεις οντολογίες οι οποίες παρέχουν πληροφορίες κοινές για όλα τα είδη οργανισμών. Αυτές οι τρεις οντολογίες είναι οι: βιολογική διαδικασία (biological process), μοριακή λειτουργία (molecular function) και κυτταρική σύσταση (cellular component). Το πιο σημαντικό στοιχείο αυτών των οντολογιών-απόψεων είναι η ορθογωνιότητα, η διασφάλιση δηλαδή ύπαρξης κάθε όρου μόνο σε μία από τις τρεις απόψεις. Η ιδιότητα αυτή εγγυάται τη μοναδικότητα ενός γνωρίσματος που αποδίδεται σε ένα γονίδιο. Με απλά λόγια, όταν ένας όρος Γονιδιακής Οντολογίας (GO όρος) αποδίδεται σε ένα γονίδιο, αυτόματα αποκλείεται η ύπαρξη ενός αντίστοιχου όρου από άλλη οντολογία-άποψη που να φέρει πανομοιότυπες ιδιότητες. Η οντολογία μπορεί να είναι ένας Κατευθυνόμενος ακυκλικός γράφος (DAG, Directed Acyclic Graph) με τους κόμβους να αναπαριστούν τα αντικείμενα (και τα εσωτερικά χαρακτηριστικά τους) και τις ακμές τις σχέσεις μεταξύ τους. (Gene Ontology, 2015).

Υψηλές δόσεις/(16-24)hrs)

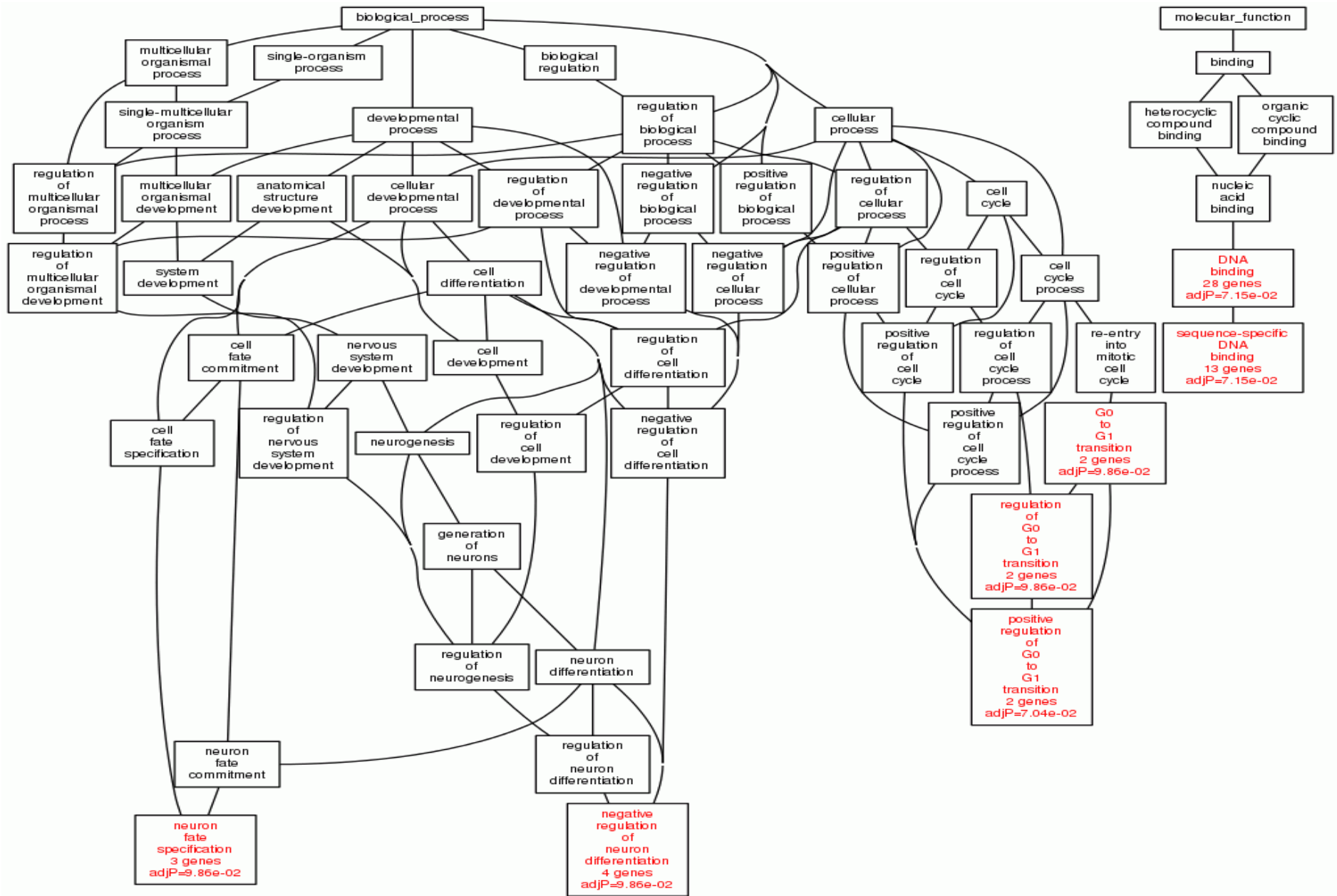
Θα ξεκινήσουμε την ανάλυση με τη λίστα γονιδίων που αφορά τα διαφορετικά εκφρασμένα γονίδια στις υψηλές δόσεις ακτινοβολίας με απόκριση στο διάστημα μεταξύ 16-24 ωρών. Σε αυτό το διάστημα ωρών έχουμε 25539 γονίδια στις συστοιχίες μας που συμμετέχουν σε τουλάχιστον μια μελέτη ($df > 0$) και επίσης η μετα-ανάλυση έδωσε μια λίστα 553 διαφορεικά εκφρασμένων γονιδίων με $p\text{-value} < 0.01$ (unadjusted). Σε αυτή τη λίστα στατιστικώς σημαντικών διαφορετικά εκφρασμένων γονιδίων ξεκινήσαμε την ανάλυση στο WebGestalt (Wang et al., 2017) ρυθμίζοντας τις κατάλληλες παραμέτρους (πχ είδος οργανισμού κλπ) με Gene Ontology Analysis (Wang et al., 2017). Αναγνώρισε 386 μοναδικά Entrez Gene IDs τα οποία και χρησιμοποίησε για την ανάλυση. Χρησιμοποιήθηκε η υπεργεωμετρική στατιστική μέθοδος. Καθώς ελέγχουμε πολλαπλά σελ γονιδίων ταυτόχρονα οφείλουμε να αναπροσαρμόσουμε το $p\text{-value}$ και αυτό στο συγκεκριμένο εργαλείο γίνεται με την συνάρτηση της R, $p.adjust$ που παρέχει 5 διαφορετικές μεθόδους υπολογισμού του $p\text{-value}$ σε τεστ πολλαπλών ερωτημάτων. Η μέθοδος που εδώ επιλέξαμε είναι αυτή των Benjamini και Hochberg. Δε βρέθηκε υπερεκπροσώπηση στατιστικώς σημαντικών διαφορετικά εκφρασμένων γονιδίων με τα γονίδια των συστοιχιών μας.

Χαμηλές δόσεις/(16-24)hrs)

Θα συνεχίσουμε την ανάλυση με τη λίστα γονιδίων που αφορά τα διαφορετικά εκφρασμένα γονίδια στις χαμηλές δόσεις ακτινοβολίας με απόκριση στο διάστημα μεταξύ 16-24 ωρών. Σε αυτό το διάστημα ωρών έχουμε 26727 γονίδια στις συστοιχίες μας που συμμετέχουν σε τουλάχιστον μια μελέτη ($df > 0$) και επίσης η μετα-ανάλυση έδωσε μια λίστα 218 διαφορετικά εκφρασμένων γονιδίων με $p\text{-value} < 0.01$ (unadjusted). Σε

αυτήν τη λίστα στατιστικώς σημαντικών διαφορικά εκφρασμένων γονιδίων ξεκινήσαμε την ανάλυση στο WebGestalt (Wang et al., 2017) ρυθμίζοντας τις κατάλληλες παραμέτρους (πχ είδος οργανισμού κλπ) με Gene Ontology Analysis (Wang et al., 2017). Αναγνώρισε 127 μοναδικά Entrez Gene IDs τα οποία και χρησιμοποίησε για την ανάλυση. Χρησιμοποιήθηκε η υπεργεωμετρική στατιστική μέθοδος. Καθώς ελέγχουμε πολλαπλά σετ γονιδίων ταυτόχρονα οφείλουμε να αναπροσαρμόσουμε το p-value και αυτό στο συγκεκριμένο εργαλείο γίνεται με τη συνάρτηση της R, p.adjust που παρέχει 5 διαφορετικές μεθόδους υπολογισμού του p-value σε τεστ πολλαπλών ερωτημάτων. Η μέθοδος που εδώ επιλέξαμε είναι αυτή των Benjamini και Hochberg (Ferreira, 2007).

Τα αποτελέσματα εμπλουτισμού GO για τις γονιδιακές οντολογίες δώσανε γράφους στην Μοριακή Λειτουργία (MF) που περιγράφει τη βιοχημική δραστηριότητα ενός γονιδιακού προϊόντος χωρίς να καθορίζει το χρόνο ή το χώρο όπου αυτή (η δραστηριότητα) συνέβη και στην οντολογία της Βιολογικής διαδικασίας (BP) που αποτελείται από ένα ή περισσότερα στάδια λειτουργιών. Υπάρχει μία σαφής σχέση μεταξύ των οντολογιών- απόψεων της Βιολογικής διαδικασίας και της Μοριακής Λειτουργίας. Με τιμή $adjP < 0.1$ προέκυψε ο παρακάτω ακυκλικός γράφος (Εικόνα 46):



Εικόνα 46: : Control-Low (16-24hrs),GO, adjP<0.1. Ρίζα της οντολογίας δεξιά είναι ο όρος: GO:0003674, molecular_function και αριστερά ο όρος: GO:0008150, biological_process

Η βάση δεδομένων KEGG PATHWAY είναι μια βάση δεδομένων των διαγραμμάτων σύνδεσης, και είναι ο-πυρήνας της πηγής του KEGG. Είναι μια συλλογή χαρτών οδών που ενσωματώνει πολλές οντότητες που περιλαμβάνουν γονίδια, πρωτεΐνες, RNAs, χημικές ενώσεις, γλυκάνες και χημικές αντιδράσεις, καθώς και γονίδια ασθενειών και στοχευμένα φάρμακα, τα οποία αποθηκεύονται ως μεμονωμένες καταχωρίσεις στις άλλες βάσεις δεδομένων του KEGG (Kanehisa, 2002). Οι χάρτες οδών ταξινομούνται στις παρακάτω ενότητες:

- Μεταβολισμός
- Επεξεργασία γενετικών πληροφοριών (μεταγραφή, μετάφραση, αντιγραφή και επιδιόρθωση DNA κλπ.)
- Επεξεργασία περιβαλλοντικών πληροφοριών (μεμβρανική μεταφορά (membrane transport)), μεταγωγή σήματος (signal transduction), κλπ.)
- Κυτταρικές διεργασίες (κυτταρική αύξηση, κυτταρικός θάνατος, λειτουργίες κυτταρικής μεμβράνης, κλπ.)
- Συστήματα οργάνων (ανοσοποιητικό σύστημα, ενδοκρινικό σύστημα, νευρικό σύστημα, κλπ.)
- Ανθρώπινες ασθένειες
- Ανάπτυξη φαρμάκων

Υψηλές δόσεις (16-24 hrs)

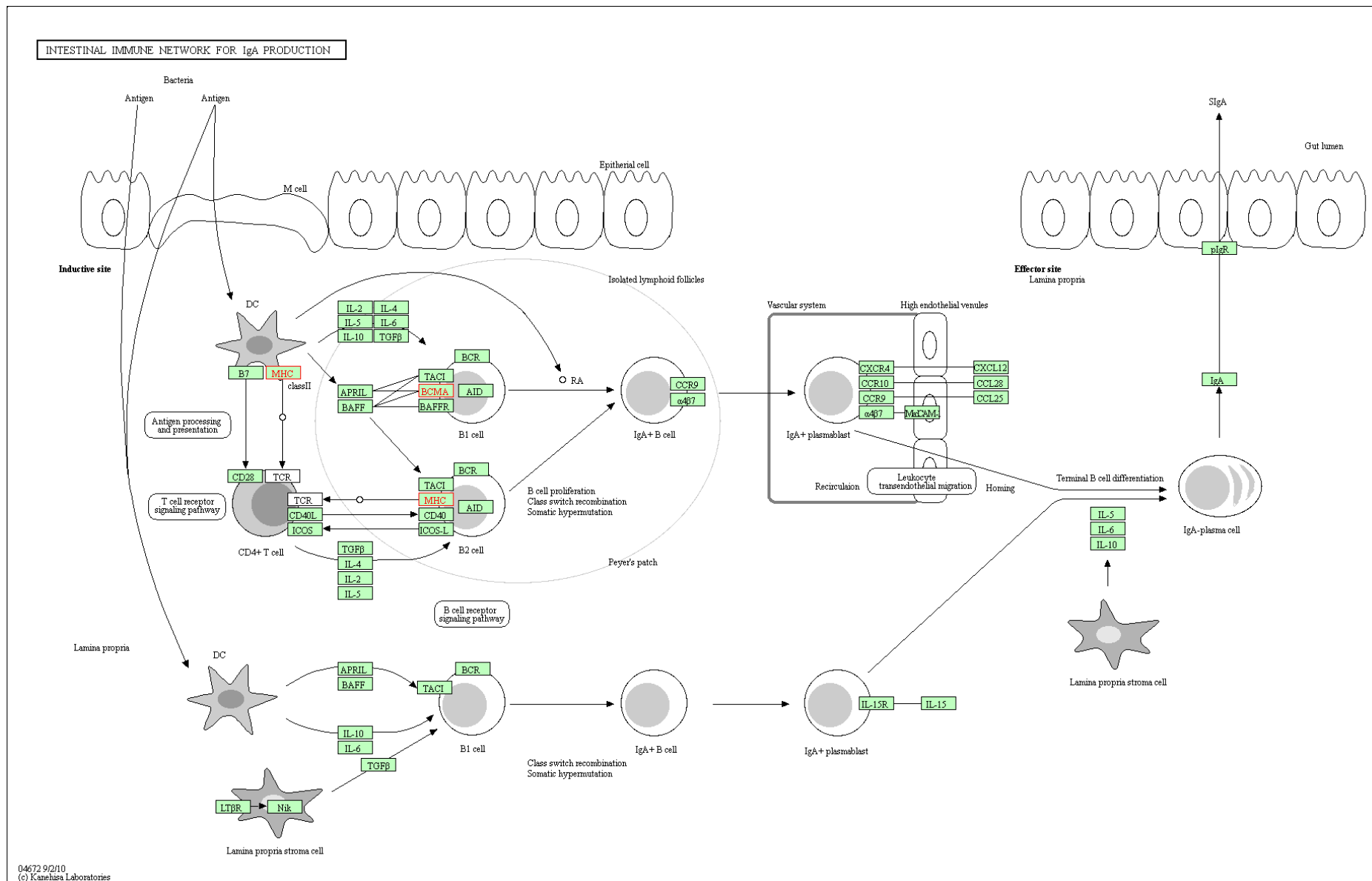
Θα ξεκινήσουμε την ανάλυση με τη λίστα γονιδίων που αφορά τα διαφορετικά εκφρασμένα γονίδια στις υψηλές δόσεις ακτινοβολίας με απόκριση στο διάστημα μεταξύ 16-24 ωρών. Σε αυτό το διάστημα ωρών έχουμε 25539 γονίδια στις συστοιχίες μας που συμμετέχουν σε τουλάχιστον μια μελέτη ($df > 0$) και επίσης η μετα-ανάλυση έδωσε μια λίστα 553 διαφορεικά εκφρασμένων γονιδίων με $p\text{-value} < 0.01$ (unadjusted). Σε αυτήν τη λίστα στατιστικώς σημαντικών διαφορετικά εκφρασμένων γονιδίων ξεκινήσαμε την ανάλυση στο WebGestalt (Wang et al., 2017) ρυθμίζοντας τις κατάλληλες παραμέτρους (πχ είδος οργανισμού κλπ) με Gene Ontology Analysis (Wang et al., 2017). Αναγνώρισε 386 μοναδικά Entrez Gene IDs τα οποία και χρησιμοποίησε για την ανάλυση. Χρησιμοποιήθηκε η υπεργεωμετρική στατιστική μέθοδος. Καθώς ελέγχουμε πολλαπλά σετ γονιδίων ταυτόχρονα οφείλουμε να αναπροσαρμόσουμε το $p\text{-value}$ και αυτό στο συγκεκριμένο εργαλείο γίνεται με την συνάρτηση της R, $p.adjust$ που παρέχει 5 διαφορετικές μεθόδους υπολογισμού του $p\text{-value}$ σε τεστ πολλαπλών ερωτημάτων. Η μέθοδος που εδώ επιλέξαμε είναι αυτή των Benjamini και Hochberg (Ferreira, 2007). Δε βρέθηκε υπερεκπροσώπηση στατιστικώς σημαντικών διαφορετικά εκφρασμένων γονιδίων με τα γονίδια των συστοιχιών μας.

Χαμηλές δόσεις/(16-24)hrs)

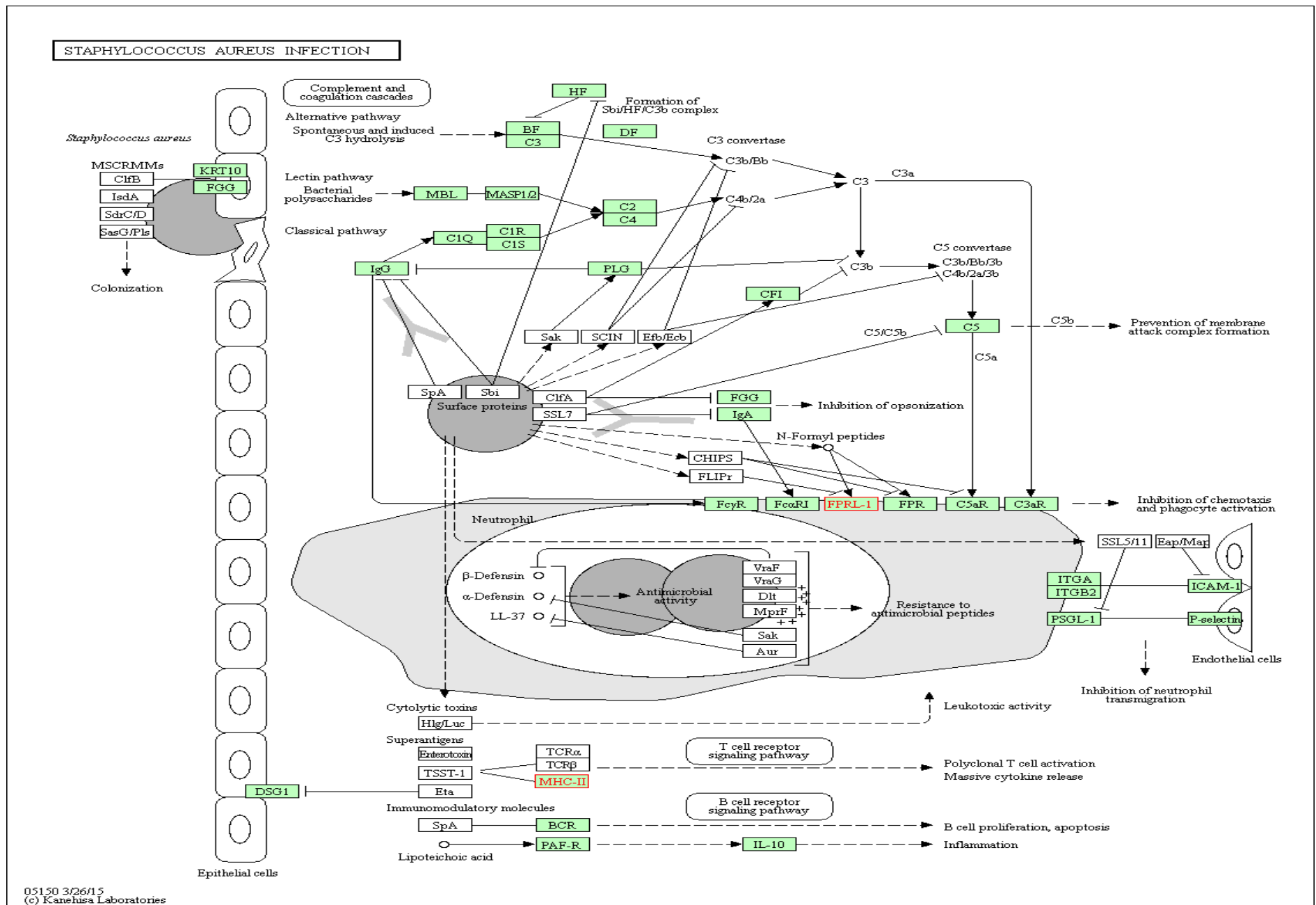
Θα συνεχίσουμε την ανάλυση με τη λίστα γονιδίων αναφερόμενοι στα διαφορετικά εκφρασμένα γονίδια στις χαμηλές δόσεις ακτινοβολίας με απόκριση το διάστημα μεταξύ 16-24 ωρών. Σε αυτό το διάστημα ωρών έχουμε 26727 γονίδια στις συστοιχίες μας που συμμετέχουν σε τουλάχιστον μια μελέτη ($df > 0$) και επίσης η μετα-ανάλυση έδωσε μια λίστα 218 διαφορεικά εκφρασμένων γονιδίων με $p\text{-value} < 0.01$ (unadjusted). Σε αυτήν τη λίστα στατιστικώς σημαντικών διαφορεικά εκφρασμένων γονιδίων ξεκινήσαμε την ανάλυση στο WebGestalt (Wang et al., 2017) ρυθμίζοντας τις κατάλληλες παραμέτρους (πχ είδος οργανισμού κλπ) με KEGG. Αναγνώρισε 127 μοναδικά Entrez Gene IDs τα οποία και χρησιμοποίησε για την ανάλυση. Χρησιμοποιήθηκε η υπεργεωμετρική στατιστική μέθοδος. Καθώς ελέγχουμε πολλαπλά σετ γονιδίων ταυτόχρονα οφείλουμε να αναπροσαρμόσουμε το $p\text{-value}$ και αυτό στο συγκεκριμένο εργαλείο γίνεται με την συνάρτηση της R, $p.adjust$ που παρέχει 5 διαφορετικές μεθόδους υπολογισμού του $p\text{-value}$ σε τεστ πολλαπλών ερωτημάτων. Η μέθοδος που εδώ επιλέξαμε είναι αυτή των Benjamini και Hochberg (Ferreira, 2007). Με $adjP < 0.1$ προέκυψαν αρκετοί υπερεκπροσωπημένοι όροι όπως φαίνεται στο παρακάτω πίνακα και κρατήσαμε τα τρία με το χαμηλότερο $p\text{-value}$.

PathwayName	#Gene	EntrezGene	Statistics
Intestinal immune network for IgA production	3	3118 608 3111	C=46;O=3;E=0.32;R=9.39;rawP=0.0040;adjP=0.0560
Staphylococcus aureus infection	3	3118 3111 2358	C=52;O=3;E=0.36;R=8.30;rawP=0.0056;adjP=0.0560
Pyrimidine metabolism	4	1841 54107 1807 5557	C=95;O=4;E=0.66;R=6.06;rawP=0.0043;adjP=0.0560
Cell adhesion molecules (CAMs)	4	3118 3111 4684 100506658	C=129;O=4;E=0.90;R=4.46;rawP=0.0126;adjP=0.0756
Maturity onset diabetes of the young	2	3170 4760	C=24;O=2;E=0.17;R=12.00;rawP=0.0120;adjP=0.0756
Systemic lupus erythematosus	3	3118 3111 3014	C=92;O=3;E=0.64;R=4.69;rawP=0.0263;adjP=0.0789
DNA replication	2	54107 5557	C=35;O=2;E=0.24;R=8.23;rawP=0.0245;adjP=0.0789
Allograft rejection	2	3118 3111	C=34;O=2;E=0.24;R=8.47;rawP=0.0232;adjP=0.0789
Graft-versus-host disease	2	3118 3111	C=36;O=2;E=0.25;R=8.00;rawP=0.0259;adjP=0.0789
Asthma	2	3118 3111	C=28;O=2;E=0.19;R=10.28;rawP=0.0161;adjP=0.0789
Type I diabetes mellitus	2	3118 3111	C=40;O=2;E=0.28;R=7.20;rawP=0.0315;adjP=0.0859

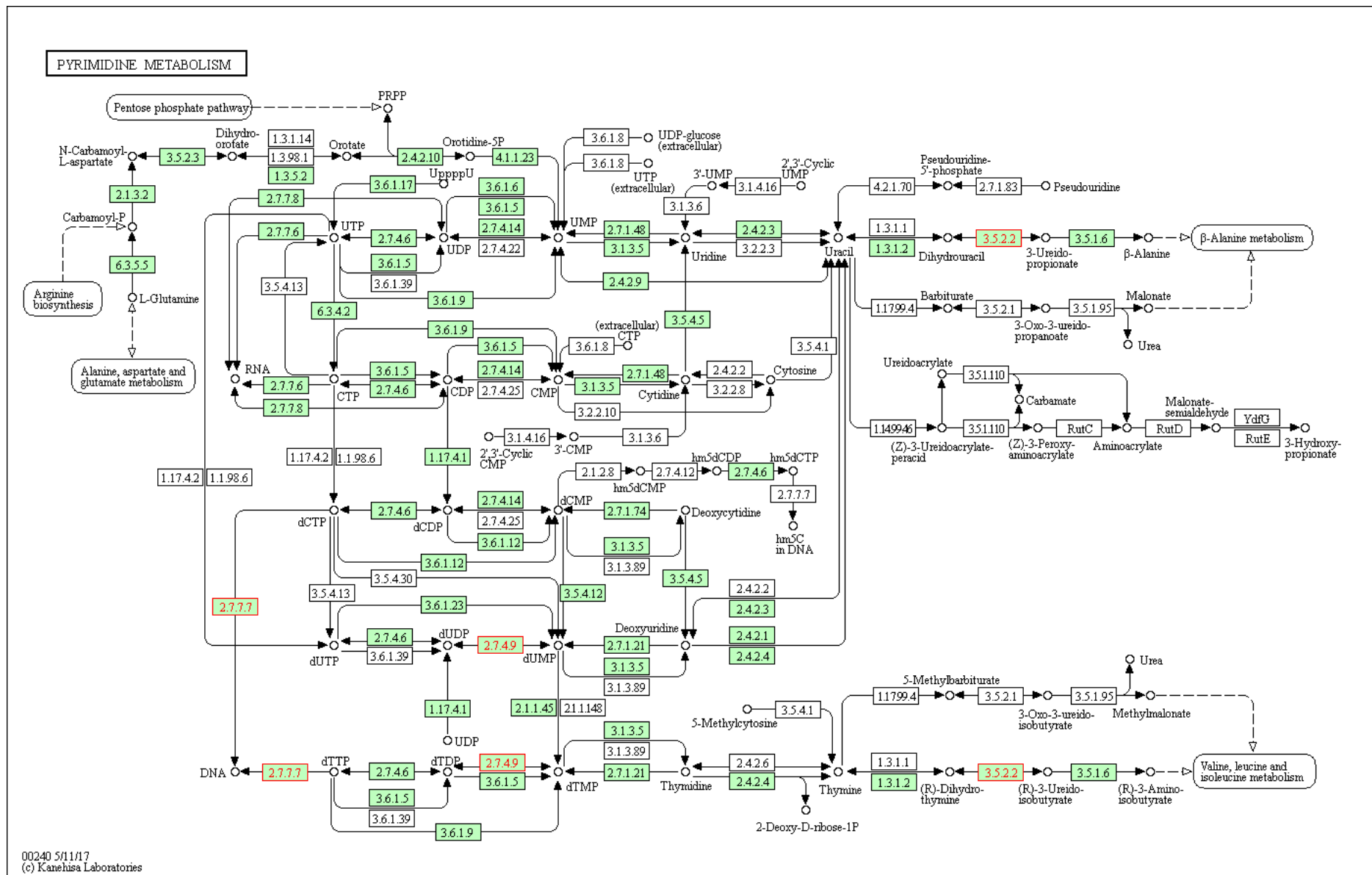
Πίνακας 10: Control-Low,(16-24)hrs, adjP<0.1,KEGG



Εικόνα 47: Εντερικό ανοσοποιητικό δίκτυο για παραγωγή IgA, Control-Low,(16-24)hrs, adjP<0.1,KEGG



Εικόνα 48: Λοίμωξη από *Staphylococcus aureus*, Control-Low,(16-24)hrs, adjP<0.1,KEGG



Εικόνα 49: Μεταβολισμός Πυριμιδίων, Control-Low, (16-24)hrs, $\text{adjP} < 0.1$, KEGG

ΑΝΑΛΥΣΗ TRANSCRIPTION FACTOR TARGET

Επόμενη ανάλυση ήταν η υπερεκπροσώπηση με Transcription Factor Target Analysis με διαχωρισμό σε ποια ανέβηκε και σε ποια έπεσε η έκφραση τους. (up/down regulated). Οι μεταγραφικοί παράγοντες είναι πρωτεΐνες με συγκεκριμένη τρισδιάστατη δομή, οι οποίες μπορούν να αλληλεπιδρούν με το DNA αναγνωρίζοντας ειδικές αλληλουχίες πάνω σε 3' αμετάφραστο άκρο. Οι μεταγραφικοί παράγοντες παίζουν κεντρικό ρόλο σε έναν οργανισμό, που αφορά στη δράση τους κατά τη διάρκεια της κυτταρικής αύξησης και διαφοροποίησης. Ανάλογα την έκφραση των μεταγραφικών παραγόντων και των γονιδίων-στόχων ,μπορούμε να εξάγουμε διάφορα συμπεράσματα. Αυτό πραγματοποιείται με την ανάλυση αυτή.

Υψηλές δόσεις (16-24 hrs)

Θα ξεκινήσουμε την ανάλυση με τη λίστα γονιδίων που αφορά τα διαφορετικά εκφρασμένα γονίδια που ανεβαίνει η έκφραση τους στις υψηλές δόσεις ακτινοβολίας με απόκριση στο διάστημα μεταξύ 16-24 ωρών. Σε αυτό το διάστημα ωρών έχουμε 12581 γονίδια στις συστοιχίες μας που συμμετέχουν σε τουλάχιστον μια μελέτη (df>0) και επίσης η μετα-ανάλυση έδωσε μια λίστα 273 διαφορετικά εκφρασμένων γονιδίων με p-value<0.01 (unadjusted). Σε αυτήν τη λίστα στατιστικώς σημαντικών διαφορικά υπέρ-εκφρασμένων γονιδίων ξεκινήσαμε την ανάλυση στο WebGestalt (Wang et al., 2017) ρυθμίζοντας τις κατάλληλες παραμέτρους (πχ είδος οργανισμού κλπ) με Transcription Factor Target Analysis. Αναγνώρισε 179 μοναδικά Entrez Gene IDs τα οποία και χρησιμοποίησε για την ανάλυση Χρησιμοποιήθηκε η υπεργεωμετρική στατιστική μέθοδος. Καθώς ελέγχουμε πολλαπλά σετ γονιδίων ταυτόχρονα οφείλουμε να αναπροσαρμόσουμε το p-value και πάλι με Benjamini και Hochberg (Ferreira, 2007). Βρέθηκε υπερεκπροσώπηση στατιστικώς σημαντικών διαφορικά υπέρ-εκφρασμένων γονιδίων με τα γονίδια των συστοιχιών μας, για adjP<0.1

Database:Transcription Target Name:hsa_V\$CDPCR1_01 ID:DB_ID:1912		
C=56; O=7; E=1.12; R=6.25; rawP=0.0001; adjP=0.0370		
UserID	Gene Symbol	Gene Name
ENSG00000175497	DPP10	dipeptidyl-peptidase 10 (non-functional)
ENSG00000159131	GART	phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase
ENSG00000176165	FOXP1	forkhead box G1
ENSG00000147862	NFIB	nuclear factor I/B
ENSG00000162650	ATXN7L2	ataxin 7-like 2
ENSG00000127334	DYRK2	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 2
ENSG00000134531	EMP1	epithelial membrane protein 1

Database:Transcription Target Name:hsa_RRCCGTTA_UNKNOWN ID:DB_ID:2352		
C=46; O=6; E=0.92; R=6.52; rawP=0.0003; adjP=0.0555		
UserID	Gene Symbol	Gene Name
ENSG00000159131	GART	phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase
ENSG00000125968	ID1	inhibitor of DNA binding 1, dominant negative helix-loop-helix protein
ENSG00000167088	SNRPD1	small nuclear ribonucleoprotein D1 polypeptide 16kDa
ENSG00000168066	SF1	splicing factor 1
ENSG00000162639	HENMT1	HEN1 methyltransferase homolog 1 (Arabidopsis)
ENSG00000123416	TUBA1B	tubulin, alpha 1b

Πίνακας 11: Control-High (16-24hrs), TF, adjP<0.1, up-regulation

Θα συνεχίσουμε με τα γονίδια στα οποία μειώθηκε η εκφραση τους. Η μετα-ανάλυση έδωσε 280 γονίδια διαφορεικά υπό-εκφρασμένα με $p\text{-value}<0.01$ (unadjusted). Από αυτά αναγνωρίστηκαν τα 207 σε μοναδικά Entrez Gene IDs . Βρέθηκε υπερεκπροσώπηση στατιστικώς σημαντικών διαφορεικά υπέρ-εκφρασμένων γονιδίων με τα γονίδια των συστοιχιών μας, για $\text{adjP}<0.1$

Database:Transcription Target Name:hsa_V\$AP1_C ID:DB_ID:1987		
C=146; O=13; E=3.37; R=3.86; rawP=3.10e-05; adjP=0.0138		
UserID	Gene Symbol	Gene Name
ENSG00000113312	TTC1	tetratricopeptide repeat domain 1
ENSG00000122254	HS3ST2	heparan sulfate (glucosamine) 3-O-sulfotransferase 2
ENSG00000054116	TRAPPC3	trafficking protein particle complex 3
ENSG00000137393	RNF144B	ring finger protein 144B
ENSG00000144711	IQSEC1	IQ motif and Sec7 domain 1
ENSG00000091073	DTX2	deltex homolog 2 (Drosophila)
ENSG00000091482	SMPX	small muscle protein, X-linked
ENSG00000005073	HOXA11	homeobox A11
ENSG00000143622	RIT1	Ras-like without CAAX 1
ENSG00000057657	PRDM1	PR domain containing 1, with ZNF domain
ENSG00000147416	ATP6V1B2	ATPase, H ⁺ transporting, lysosomal 56/58kDa, V1 subunit B2
ENSG00000100439	ABHD4	abhydrolase domain containing 4
ENSG00000141232	TOB1	transducer of ERBB2, 1
Database: Transcription Target Name:hsa_V\$AP1_Q4_01 ID:DB_ID:2303		

C=138; O=11; E=3.19; R=3.45; rawP=0.0003; adjP=0.0444		
UserID	Gene Symbol	Gene Name
ENSG00000113312	TTC1	tetratricopeptide repeat domain 1
ENSG00000122254	HS3ST2	heparan sulfate (glucosamine) 3-O-sulfotransferase 2
ENSG00000137393	RNF144B	ring finger protein 144B
ENSG00000144711	IQSEC1	IQ motif and Sec7 domain 1
ENSG00000168214	RBPJ	recombination signal binding protein for immunoglobulin kappa J region
ENSG00000124762	CDKN1A	cyclin-dependent kinase inhibitor 1A (p21, Cip1)
ENSG00000005073	HOXA11	homeobox A11
ENSG00000143622	RIT1	Ras-like without CAAX 1
ENSG00000057657	PRDM1	PR domain containing 1, with ZNF domain
ENSG00000100439	ABHD4	abhydrolase domain containing 4
ENSG00000141232	TOB1	transducer of ERBB2, 1
Database:Transcription Target Name:hsa_V\$GABP_B ID:DB_ID:2059		
C=114; O=10; E=2.63; R=3.80; rawP=0.0003; adjP=0.0444		
UserID	Gene Symbol	Gene Name
ENSG00000135956	TMEM127	transmembrane protein 127
ENSG00000040633	PHF23	PHD finger protein 23
ENSG00000157224	CLDN12	claudin 12
ENSG00000075785	RAB7A	RAB7A, member RAS oncogene family
ENSG00000065615	CYB5R4	cytochrome b5 reductase 4
ENSG00000167985	SDHAF2	succinate dehydrogenase complex assembly factor 2
ENSG00000091073	DTX2	deltex homolog 2 (Drosophila)
ENSG00000136003	ISCU	iron-sulfur cluster scaffold homolog (E. coli)
ENSG00000169564	PCBP1	poly(rC) binding protein 1
ENSG00000119878	CRIPT	cysteine-rich PDZ-binding protein

Πίνακας 12: Πίνακας 9: Control-High (16-24hrs), TF, adjP<0.05, down-regulation

Χαμηλές δόσεις (16-24 hrs)

Θα συνεχίσουμε την ανάλυση με τη λίστα γονιδίων που αφορά τα διαφορικά εκφρασμένα γονίδια που ανεβαίνει η έκφραση τους στις χαμηλές δόσεις ακτινοβολίας με απόκριση στο διάστημα μεταξύ 16-24 ωρών. Σε αυτό το διάστημα ωρών έχουμε 13098 γονίδια στις συστοιχίες μας που συμμετέχουν σε τουλάχιστον μια μελέτη ($df > 0$) και επίσης η μετα-ανάλυση έδωσε μια λίστα 91 διαφορικά εκφρασμένων γονιδίων με $p\text{-value} < 0.01$ (unadjusted). Σε αυτή τη λίστα στατιστικώς σημαντικών διαφορικά υπέρ-εκφρασμένων γονιδίων ξεκινήσαμε την ανάλυση στο WebGestalt (Wang et al., 2017) ρυθμίζοντας τις κατάλληλες παραμέτρους (πχ είδος οργανισμού κλπ) με Transcription Factor Analysis. Αναγνώρισε 56 μοναδικά Entrez Gene IDs τα οποία και χρησιμοποίησε για την ανάλυση. Χρησιμοποιήθηκε η υπεργεωμετρική στατιστική μέθοδος. Καθώς ελέγχουμε πολλαπλά σεντ γονιδίων ταυτόχρονα οφείλουμε να αναπροσαρμόσουμε το $p\text{-value}$ με Benjamini και Hochberg (Ferreira, 2007). Δε βρέθηκε υπερεκπροσώπηση στατιστικώς σημαντικών διαφορικά υπό-εκφρασμένων γονιδίων με τα γονίδια των συστοιχιών μας.

Συνεχίζοντας την ανάλυση στα αποτελέσματα ακτινοβολίας με χαμηλές δόσεις στα γονίδια που μειώνεται η έκφραση τους έχουμε στις συστοιχίες μας 13629 γονίδια από τα οποία η μετα-ανάλυση έδωσε μια λίστα 127 διαφορικά εκφρασμένων γονιδίων με $p\text{-value} < 0.01$ (unadjusted), με τα 71 να αντιστοιχούν σε μοναδικά IDs. Δε βρέθηκε υπερεκπροσώπηση στατιστικώς σημαντικών διαφορικά υπό-εκφρασμένων γονιδίων με τα γονίδια των συστοιχιών μας

ΑΝΑΛΥΣΗ MICRORNA TARGET

Ακολούθησε η MicroRNA Target Analysis που είναι ανάλυση εμπλουτισμού για τους στόχους των MicroRNAs (Chaudhry et al., 2012). Η συμπληρωματικότητα μεταξύ της φύτρας του miRNA και των περιοχών 3' UTR (περιλαμβάνονται στη βάση Ensembl) του mRNA στόχου είναι το θεμελιώδες πλεονέκτημα για την υπολογιστική ανάλυση. Χωρίσαμε την ανάλυση ανάλογα αν το επίπεδο έκφρασης των γονιδίων ανεβαίνει ή μειώνεται.

Υψηλές δόσεις/(16-24)hrs

Θα ξεκινήσουμε την ανάλυση με τη λίστα γονιδίων που αφορά τα διαφορετικά εκφρασμένα γονίδια που ανεβαίνει η έκφραση τους στις υψηλές δόσεις ακτινοβολίας με απόκριση στο διάστημα μεταξύ 16-24 ωρών. Σε αυτό το διάστημα ωρών έχουμε 12581 γονίδια στις συστοιχίες μας που συμμετέχουν σε τουλάχιστον μια μελέτη ($df > 0$) και επίσης η μετα-ανάλυση έδωσε μια λίστα 273 διαφορεικά εκφρασμένων γονιδίων με $p\text{-value} < 0.01$ (unadjusted). Σε αυτή τη λίστα στατιστικώς σημαντικών διαφορικά υπέρ-εκφρασμένων γονιδίων ξεκινήσαμε την ανάλυση στο WebGestalt (Wang et al., 2017) ρυθμίζοντας τις κατάλληλες παραμέτρους (πχ είδος οργανισμού κλπ) με MicroRNA Target Analysis. Αναγνώρισε 179 μοναδικά Entrez Gene IDs τα οποία και χρησιμοποίησε για την ανάλυση. Χρησιμοποιήθηκε η υπεργεωμετρική στατιστική μέθοδος. Καθώς ελέγχουμε πολλαπλά σεντ γονιδίων ταυτόχρονα οφείλουμε να αναπροσαρμόσουμε το $p\text{-value}$ και αυτό στο συγκεκριμένο εργαλείο γίνεται με την συνάρτηση της R, $p.adjust$ που παρέχει 5 διαφορετικές μεθόδους υπολογισμού του $p\text{-value}$ σε τεστ πολλαπλών ερωτημάτων. Η μέθοδος που εδώ επιλέξαμε είναι αυτή των Benjamini και Hochberg (Ferreira, 2007). Βρέθηκε υπερεκπροσώπηση στατιστικώς σημαντικών διαφορικά υπέρ-εκφρασμένων γονιδίων με τα γονίδια των συστοιχιών μας με $adjP < 0.1$.

Database:microRNA		
Target Name:hsa_GTGCAAT,MIR-25,MIR-32,MIR-92,MIR-363,MIR-367 ID:DB_ID:668		
C=139; O=9; E=2.78; R=3.23; rawP=0.0019; adjP=0.0902		
UserID	Gene Symbol	Gene Name
ENSG00000175497	DPP10	dipeptidyl-peptidase 10 (non-functional)
ENSG00000112941	PAPD7	PAP associated domain containing 7
ENSG00000164649	CDCA7L	cell division cycle associated 7-like
ENSG00000111432	FZD10	frizzled family receptor 10
ENSG00000112210	RAB23	RAB23, member RAS oncogene family
ENSG00000176165	FOXP1	forkhead box G1
ENSG00000147862	NFIB	nuclear factor I/B
ENSG00000132326	PER2	period homolog 2 (Drosophila)
ENSG00000127334	DYRK2	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 2

Database:microRNA Target Name:hsa_TCTCTCC,MIR-185 ID:DB_ID:796		
C=60; O=6; E=1.20; R=5.00; rawP=0.0012; adjP=0.0902		
Index	UserID	Value
ENSG00000168066	SF1	splicing factor 1
ENSG00000145808	ADAMTS19	ADAM metallopeptidase with thrombospondin type 1 motif, 19
ENSG00000182253	SYNM	synemin, intermediate filament protein
ENSG00000213190	MLLT11	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 11
ENSG00000114302	PRKAR2A	protein kinase, cAMP-dependent, regulatory, type II, alpha
ENSG00000168906	MAT2A	methionine adenosyltransferase II, alpha

Πίνακας 13: : Control-High (16-24hrs), miRNA, adjP<0.1, up-regulation

Θα συνεχίσουμε με τα γονίδια στα οποία μειώθηκε η εκφραση τους. Η μετα-ανάλυση έδωσε 280 γονίδια διαφορετικά υπό-εκφρασμένα με p-value<0.01 (unadjusted). Από αυτά αναγνωρίστηκαν τα 207 σε μοναδικά Entrez Gene IDs.

A) Αποτελέσματα για adjP<0.05.

Database:microRNA Target Name:hsa_ATATGCA,MIR-448 ID:DB_ID:774		
C=97; O=10; E=2.24; R=4.46; rawP=7.58e-05; adjP=0.0090		
UserID	Gene Symbol	Gene Name
ENSG00000185950	IRS2	insulin receptor substrate 2
ENSG00000145743	FBXL17	F-box and leucine-rich repeat protein 17
ENSG00000013441	CLK1	CDC-like kinase 1
ENSG00000137845	ADAM10	ADAM metallopeptidase domain 10
ENSG00000075785	RAB7A	RAB7A, member RAS oncogene family
ENSG00000135111	TBX3	T-box 3
ENSG00000146457	WTAP	Wilms tumor 1 associated protein
ENSG00000123066	MED13L	mediator complex subunit 13-like
ENSG00000117616	C1orf63	chromosome 1 open reading frame 63
ENSG00000005073	HOXA11	homeobox A11
Database:microRNA Target Name:hsa_CTTTGTA,MIR-524 ID:DB_ID:802		
C=225; O=14; E=5.20; R=2.69; rawP=0.0007; adjP=0.0416		
UserID	Gene Symbol	Gene Name
ENSG00000156875	HIAT1	hippocampus abundant transcript 1
ENSG00000185950	IRS2	insulin receptor substrate 2

ENSG00000113595	TRIM23	tripartite motif containing 23
ENSG00000119950	MXI1	MAX interactor 1
ENSG00000150760	DOCK1	dedicator of cytokinesis 1
ENSG00000137845	ADAM10	ADAM metallopeptidase domain 10
ENSG00000164938	TP53INP1	tumor protein p53 inducible nuclear protein 1
ENSG00000103994	ZFP106	zinc finger protein 106 homolog (mouse)
ENSG00000120656	TAF12	TAF12 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 20kDa
ENSG00000057657	PRDM1	PR domain containing 1, with ZNF domain
ENSG00000108175	ZMIZ1	zinc finger, MIZ-type containing 1
ENSG00000115649	CNPPD1	cyclin Pas1/PH080 domain containing 1
ENSG00000128989	ARPP19	cAMP-regulated phosphoprotein, 19kDa
ENSG00000100439	ABHD4	abhydrolase domain containing 4

Πίνακας 14: Control-High (16-24hrs), miRNA, adjP<0.05, down-regulation

B) Αποτελέσματα για adjP<0.1

Database:microRNA Target Name:hsa_ATATGCA,MIR-448 ID:DB_ID:774		
C=97; O=10; E=2.24; R=4.46; rawP=7.58e-05; adjP=0.0090		
UserID	Gene Symbol	Gene Name
ENSG00000185950	IRS2	insulin receptor substrate 2
ENSG00000145743	FBXL17	F-box and leucine-rich repeat protein 17
ENSG00000013441	CLK1	CDC-like kinase 1
ENSG00000137845	ADAM10	ADAM metalloproteinase domain 10
ENSG00000075785	RAB7A	RAB7A, member RAS oncogene family
ENSG00000135111	TBX3	T-box 3
ENSG00000146457	WTAP	Wilms tumor 1 associated protein
ENSG00000123066	MED13L	mediator complex subunit 13-like
ENSG00000117616	C1orf63	chromosome 1 open reading frame 63
ENSG00000005073	HOXA11	homeobox A11
Database:microRNA Target Name:hsa_CTTTGTA,MIR-524 ID:DB_ID:802		
C=225; O=14; E=5.20; R=2.69; rawP=0.0007; adjP=0.0416		
UserID	Gene Symbol	Gene Name
ENSG00000156875	HIAT1	hippocampus abundant transcript 1
ENSG00000185950	IRS2	insulin receptor substrate 2
ENSG00000113595	TRIM23	tripartite motif containing 23
ENSG00000119950	MXI1	MAX interactor 1
ENSG00000150760	DOCK1	dedicator of cytokinesis 1
ENSG00000137845	ADAM10	ADAM metalloproteinase domain 10
ENSG00000164938	TP53INP1	tumor protein p53 inducible nuclear protein 1
ENSG00000103994	ZFP106	zinc finger protein 106 homolog (mouse)
ENSG00000120656	TAF12	TAF12 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 20kDa
ENSG00000057657	PRDM1	PR domain containing 1, with ZNF domain
ENSG00000108175	ZMIZ1	zinc finger, MIZ-type containing 1
ENSG00000115649	CNPPD1	cyclin Pas1/PH080 domain containing 1
ENSG00000128989	ARPP19	cAMP-regulated phosphoprotein, 19kDa
ENSG00000100439	ABHD4	abhydrolase domain containing 4
Database:microRNA Target Name:hsa_CTGAGCC,MIR-24 ID:DB_ID:760		
C=120; O=9; E=2.77; R=3.25; rawP=0.0018; adjP=0.0714		

UserID	Gene Symbol	Gene Name
ENSG00000141968	VAV1	vav 1 guanine nucleotide exchange factor
ENSG00000119950	MXI1	MAX interactor 1
ENSG00000040633	PHF23	PHD finger protein 23
ENSG00000070610	GBA2	glucosidase, beta (bile acid) 2
ENSG00000165556	CDX2	caudal type homeobox 2
ENSG00000154822	PLCL2	phospholipase C-like 2
ENSG00000124615	MOCS1	molybdenum cofactor synthesis 1
ENSG00000196323	ZBTB44	zinc finger and BTB domain containing 44
ENSG00000160445	ZER1	zer-1 homolog (C. elegans)
Database:microRNA Target Name:hsa_ATCATGA,MIR-433 ID:DB_ID:864		
C=61; O=6; E=1.41; R=4.26; rawP=0.0027; adjP=0.0803		
UserID	Gene Symbol	Gene Name
ENSG00000117616	C1orf63	chromosome 1 open reading frame 63
ENSG00000119950	MXI1	MAX interactor 1
ENSG00000108175	ZMIZ1	zinc finger, MIZ-type containing 1
ENSG00000086065	CHMP5	charged multivesicular body protein 5
ENSG00000156113	KCNMA1	potassium large conductance calcium-activated channel, subfamily M, alpha member 1
ENSG00000146457	WTAP	Wilms tumor 1 associated protein

Πίνακας 15: Control-High (16-24hrs), miRNA, adjP<0.1, down-regulation

Χαμηλές δόσεις/(16-24)hrs

Θα ξεκινήσουμε την ανάλυση με τη λίστα γονιδίων που αφορά τα διαφορετικά εκφρασμένα γονίδια που ανεβαίνει η έκφραση τους στις χαμηλές δόσεις ακτινοβολίας με απόκριση το διάστημα μεταξύ 16-24 ωρών. Σε αυτό το διάστημα ωρών έχουμε 13098 γονίδια στις συστοιχίες μας που συμμετέχουν σε τουλάχιστον μια μελέτη ($df > 0$) και επίσης η μετα-ανάλυση έδωσε μια λίστα 91 διαφορετικά εκφρασμένων γονιδίων με $p\text{-value} < 0.01$ (unadjusted). Σε αυτή τη λίστα στατιστικώς σημαντικών διαφορικά υπέρ-εκφρασμένων γονιδίων ξεκινήσαμε την ανάλυση στο WebGestalt (Wang et al., 2017) ρυθμίζοντας τις κατάλληλες παραμέτρους (πχ είδος οργανισμού κλπ) με MicroRNA Target Analysis. Αναγνώρισε 56 μοναδικά Entrez Gene IDs τα οποία και χρησιμοποίησε για την ανάλυση. Χρησιμοποιήθηκε η υπεργεωμετρική στατιστική μέθοδος. Καθώς ελέγχουμε πολλαπλά σεντ γονιδίων ταυτόχρονα οφείλουμε να αναπροσαρμόσουμε το $p\text{-value}$ με Benjamini και Hochberg (Ferreira, 2007). Δεν βρέθηκε υπερεκπροσώπηση στατιστικώς σημαντικών διαφορικά υπό-εκφρασμένων γονιδίων με τα γονίδια των συστοιχιών μας.

Συνεχίζοντας την ανάλυση στα αποτελέσματα ακτινοβολίας με χαμηλές δόσεις στα γονίδια που μειώνεται η έκφραση τους έχουμε στις συστοιχίες μας 13629 γονίδια από τα οποία η μετα-ανάλυση έδωσε μια λίστα 127 διαφορετικά εκφρασμένων γονιδίων με $p\text{-value} < 0.01$ (unadjusted) με τα 71 να αντιστοιχούν σε μοναδικά IDs

Με $\text{adj}P < 0.1$ προκύπτει ο παρακάτω πίνακας

Database:microRNA Target Name:hsa_TACGGGT,MIR-99A,MIR-100,MIR-99B ID:DB_ID:674		
C=13; O=2; E=0.10; R=20.25; rawP=0.0042; adjP=0.0840		
UserID	Gene Symbol	Gene Name
ENSG00000173404	INSM1	insulinoma-associated 1
ENSG00000122254	HS3ST2	heparan sulfate (glucosamine) 3-O-sulfotransferase 2

Πίνακας 16: Control-Low (16-24hrs), miRNA, $\text{adj}P < 0.1$, down-regulation

ΑΠΕΙΚΟΝΙΣΗ ΒΙΟΛΟΓΙΚΩΝ ΔΙΚΤΥΩΝ

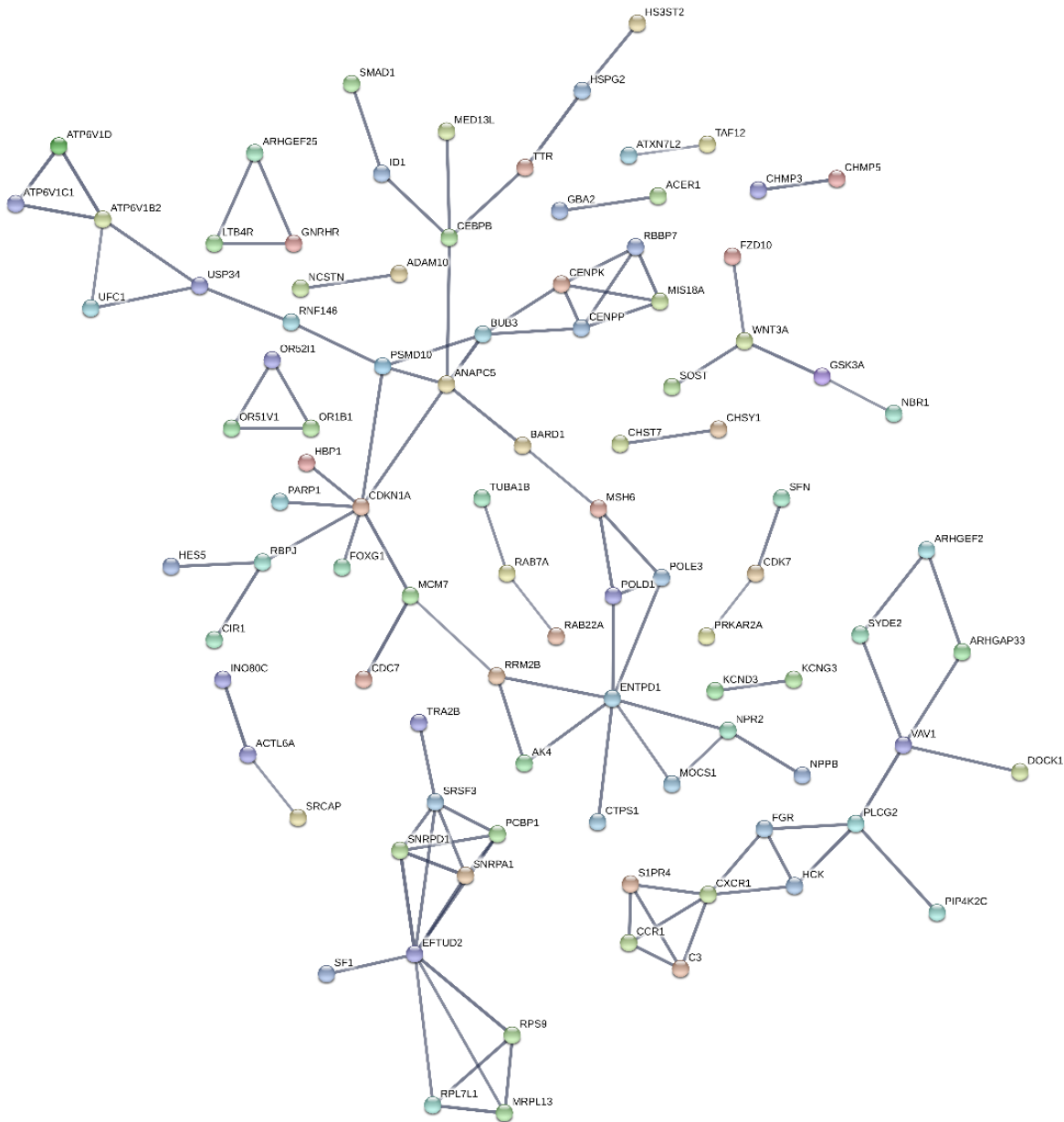
Η πρόβλεψη αλληλεπιδράσεων οι οποίες λαμβάνουν μέρος στα διαφορετικά επίπεδα του κυττάρου καθώς και η διερεύνηση του τρόπου με τον οποίο αυτές οι αλληλεπιδράσεις συνεργάζονται μεταξύ τους έτσι ώστε να επιτελέσουν τις κυτταρικές λειτουργίες διαφαίνονται με απεικονίσεις των βιολογικών δικτύων. Στο χαμηλότερο επίπεδο του κυττάρου υπάρχουν οι φυσικές αλληλεπιδράσεις οι οποίες ισοδυναμούν με σύνδεση των πρωτεϊνών (ή μιας πρωτεΐνης και ενός DNA μορίου) Την ανίχνευση των παραπάνω διαφορετικών ειδών αλληλεπιδράσεων καθώς και την εννοιολογική σύνδεσή τους θα πραγματοποιήσουμε χρησιμοποιώντας την βάση δεδομένων STRING v 10.0 (<https://string-db.org/>) (von Mering et al., 2005).

Κατά την διαδικασία που ακολουθήθηκε για κάθε μία από τις τέσσερις λίστες γονιδίων που προέκυψαν υπήρξαν γονίδια (ENSG IDs) που δεν αντιστοιχήθηκαν στη βάση δεδομένων.

Σχετικά με τις πηγές της πληροφορίας για τις αλληλεπιδράσεις δεν λήφθηκαν υπόψιν μέθοδοι εξόρυξης κειμένων από την βιβλιογραφία, ομάδες γονιδίων που συχνά παρατηρούνται στις γενομικές περιοχές η μία της άλλης αλλά και οικογένειες γονιδίων των οποίων τα μοτίβα εμφάνισης στο γονιδίωμα εμφανίζουν ομοιότητες. Για να αυξήσουμε την στατιστική σημαντικότητα των αλληλεπιδράσεων που θα απεικονιστούν θεωρήσαμε μεγάλο βαθμό εμπιστοσύνης (high confidence 0.700) ως ελάχιστο απαιτούμενο βαθμό αλληλεπίδρασης για κάθε ακμή. Τέλος στην απεικόνιση σημειώνεται ότι το πάχος κάθε ακμής υποδηλώνει την ισχύ του αποτελέσματος, ενώ αγνοήθηκαν οι κόμβοι που δεν συμμετείχαν σε κάποια διαδικασία. Τα στατιστικά στοιχεία κάθε βιολογικού δικτύου υποδηλώνουν την σημαντικότητα των ευρημάτων.

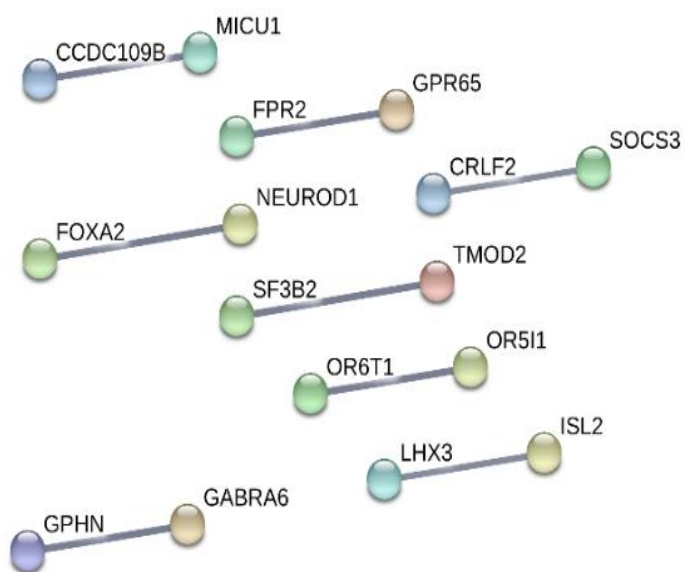
Για να είναι δυνατή η σύγκριση των ευρημάτων από τις μελέτες γονιδιακής απόκρισης που βρέθηκαν αντιστοιχίσαμε την ονοματολογία που χρησιμοποιήσαν με το τελευταίο όνομα στο HGNC για κάθε γονίδιο (von Mering et al., 2005).

Τα γονίδια αναπαρίστανται ως κόμβοι (nodes) και οι διασυνδέσεις μεταξύ τους ως ακμές στο δίκτυο (Εικόνα 35, Εικόνα 36).

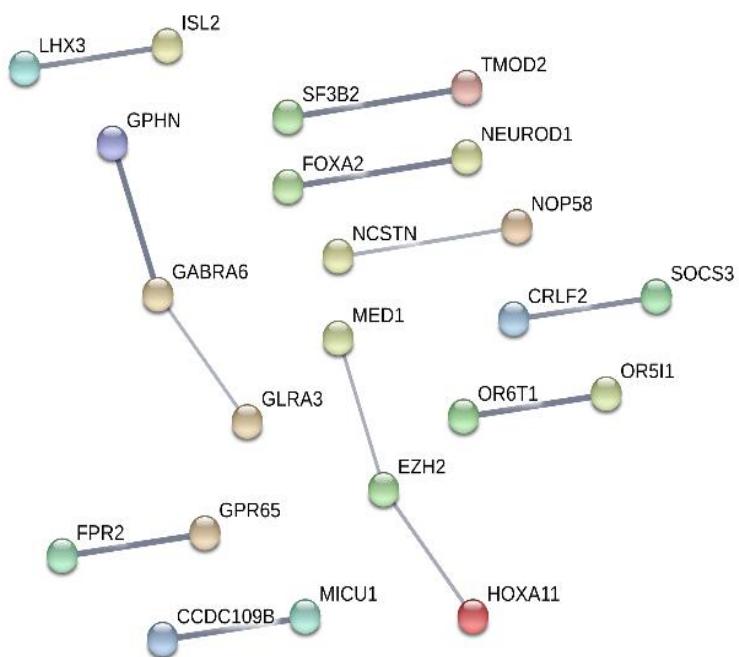


Εικόνα 50: Βιολογικό δίκτυο γονιδίων όπως προέκυψαν από το σύνολο των δειγμάτων που εκτέθηκαν σε υψηλή δόση ακτινοβολίας και συλλέχθηκαν στο διάστημα μεταξύ 16-24 ωρών.

A



B



Εικόνα 51: Βιολογικά δίκτυα διαφορεικά εκφρασμένων γονιδίων, όπως προέκυψαν από το σύνολο των δειγμάτων που εκτέθηκαν σε χαμηλή δόση ακτινοβολίας και συλλέχθηκαν στο διάστημα μεταξύ 16 και 24

ωρών, ορίζοντας την ουδό του score αλληλεπιδράσεων σε: A) 0.7 (υψηλή σημαντικότητα) και B) 0.4 (μέση σημαντικότητα).

ΓΟΝΙΔΙΑ ΜΕ ΑΠΟΚΡΙΣΗ ΣΕ ΙΟΝΤΙΖΟΥΣΑ ΑΚΤΙΝΟΒΟΛΙΑ ΒΑΣΕΙ ΒΙΒΛΙΟΓΡΑΦΙΑΣ

Παρακάτω παρατίθενται οι λίστες που περιέχουν τα γονίδια όπως προέκυψαν από ενδελεχή βιβλιογραφική αναζήτηση στο εργαλείο Quertle (Πίνακας 17, Πίνακας 18) Τα ENSG IDs έχουν αντιστοιχηθεί με το αντίστοιχο HGNC και την αντίστοιχη περιγραφή.

ENSG_ID	HGNC_Symbol	Description
ENSG00000163017	ACTG2	actin, gamma 2, smooth muscle, enteric [Source:HGNC Symbol;Acc:HGNC:145]
ENSG0000020181	ADGRA2	adhesion G protein-coupled receptor A2 [Source:HGNC Symbol;Acc:HGNC:17849]
ENSG00000205336	ADGRG1	adhesion G protein-coupled receptor G1 [Source:HGNC Symbol;Acc:HGNC:4512]
ENSG00000141385	AFG3L2	AFG3 like matrix AAA peptidase subunit 2 [Source:HGNC Symbol;Acc:HGNC:315]
ENSG00000164252	AGGF1	angiogenic factor with G-patch and FHA domains 1 [Source:HGNC Symbol;Acc:HGNC:24684]
ENSG00000164331	ANKRA2	ankyrin repeat family A member 2 [Source:HGNC Symbol;Acc:HGNC:13208]
ENSG00000104537	ANXA13	annexin A13 [Source:HGNC Symbol;Acc:HGNC:536]
ENSG00000175220	ARHGAP1	Rho GTPase activating protein 1 [Source:HGNC Symbol;Acc:HGNC:673]
ENSG00000107863	ARHGAP21	Rho GTPase activating protein 21 [Source:HGNC Symbol;Acc:HGNC:23725]
ENSG00000163399	ATP1A1	ATPase Na ⁺ /K ⁺ transporting subunit alpha 1 [Source:HGNC Symbol;Acc:HGNC:799]
ENSG00000054793	ATP9A	ATPase phospholipid transporting 9A (putative) [Source:HGNC Symbol;Acc:HGNC:13540]
ENSG00000138686	BBS7	Bardet-Biedl syndrome 7 [Source:HGNC Symbol;Acc:HGNC:18758]
ENSG00000185825	BCAP31	B-cell receptor-associated protein 31 [Source:HGNC Symbol;Acc:HGNC:16695]
ENSG00000134107	BHLHE40	basic helix-loop-helix family member e40 [Source:HGNC Symbol;Acc:HGNC:1046]
ENSG00000196072	BLOC1S2	biogenesis of lysosomal organelles complex 1 subunit 2 [Source:HGNC Symbol;Acc:HGNC:20984]
ENSG00000204217	BMPR2	bone morphogenetic protein receptor type 2 [Source:HGNC Symbol;Acc:HGNC:1078]
ENSG00000158636	EMSY	EMSY, BRCA2 Interacting Transcriptional Repressor [Source:HGNC Symbol;Acc:HGNC:18071]
ENSG00000164879	CA3	carbonic anhydrase 3 [Source:HGNC Symbol;Acc:HGNC:1374]
ENSG00000070808	CAMK2A	calcium/calmodulin dependent protein kinase II alpha [Source:HGNC Symbol;Acc:HGNC:1460]
ENSG00000110931	CAMKK2	calcium/calmodulin dependent protein kinase kinase 2 [Source:HGNC Symbol;Acc:HGNC:1470]
ENSG00000137500	CCDC90B	coiled-coil domain containing 90B [Source:HGNC Symbol;Acc:HGNC:28108]
ENSG00000113361	CDH6	cadherin 6 [Source:HGNC Symbol;Acc:HGNC:1765]
ENSG00000117724	CENPF	centromere protein F [Source:HGNC Symbol;Acc:HGNC:1857]
ENSG00000124177	CHD6	chromodomain helicase DNA binding protein 6 [Source:HGNC Symbol;Acc:HGNC:19057]
ENSG00000131873	CHSY1	chondroitin sulfate synthase 1 [Source:HGNC Symbol;Acc:HGNC:17198]
ENSG00000182973	CNOT10	CCR4-NOT transcription complex subunit 10 [Source:HGNC Symbol;Acc:HGNC:23817]
ENSG00000095794	CREM	cAMP responsive element modulator [Source:HGNC Symbol;Acc:HGNC:2352]
ENSG00000169826	CSGALNACT2	chondroitin sulfate N-acetylgalactosaminyltransferase 2 [Source:HGNC Symbol;Acc:HGNC:24292]
ENSG00000198730	CTR9	CTR9 homolog, Paf1/RNA polymerase II complex component [Source:HGNC Symbol;Acc:HGNC:16850]
ENSG00000110063	DCPS	decapping enzyme, scavenger [Source:HGNC Symbol;Acc:HGNC:29812]
ENSG00000281763	DHX36	DEAH-box helicase 36 [Source:HGNC Symbol;Acc:HGNC:14410]
ENSG00000174953		DEAH-box helicase 36 [Source:HGNC Symbol;Acc:HGNC:14410]
ENSG00000213918	DNASE1	deoxyribonuclease 1 [Source:HGNC Symbol;Acc:HGNC:2956]
ENSG00000167130	DOLPP1	dolichylidiphosphatase 1 [Source:HGNC Symbol;Acc:HGNC:29565]
ENSG00000175497	DPP10	dipeptidyl peptidase like 10 [Source:HGNC Symbol;Acc:HGNC:20823]
ENSG00000143196	DPT	dermatopontin [Source:HGNC Symbol;Acc:HGNC:3011]
ENSG00000149547	EI24	EI24, autophagy associated transmembrane protein [Source:HGNC Symbol;Acc:HGNC:13276]
ENSG00000108515	ENO3	enolase 3 [Source:HGNC Symbol;Acc:HGNC:3354]
ENSG00000145242	EPHA5	EPH receptor A5 [Source:HGNC Symbol;Acc:HGNC:3389]
ENSG00000065361	ERBB3	erb-b2 receptor tyrosine kinase 3 [Source:HGNC Symbol;Acc:HGNC:3431]

ENSG00000005238	FAM214B	family with sequence similarity 214 member B [Source:HGNC Symbol;Acc:HGNC:25666]
ENSG00000167196	FBX022	F-box protein 22 [Source:HGNC Symbol;Acc:HGNC:13593]
ENSG00000127951	FGL2	fibrinogen like 2 [Source:HGNC Symbol;Acc:HGNC:3696]
ENSG00000152661	GJA1	gap junction protein alpha 1 [Source:HGNC Symbol;Acc:HGNC:4274]
ENSG00000128266	GNAZ	G protein subunit alpha z [Source:HGNC Symbol;Acc:HGNC:4395]
ENSG00000125675	GRIA3	glutamate ionotropic receptor AMPA type subunit 3 [Source:HGNC Symbol;Acc:HGNC:4573]
ENSG00000170445	HARS	histidyl-tRNA synthetase [Source:HGNC Symbol;Acc:HGNC:4816]
ENSG00000172273	HINFP	histone H4 transcription factor [Source:HGNC Symbol;Acc:HGNC:17850]
ENSG00000275713	HIST1H2BH	histone cluster 1 H2B family member h [Source:HGNC Symbol;Acc:HGNC:4755]
ENSG00000101294	HM13	histocompatibility minor 13 [Source:HGNC Symbol;Acc:HGNC:16435]
ENSG00000277424	HMOX2	heme oxygenase 2 [Source:HGNC Symbol;Acc:HGNC:5014]
ENSG00000103415	HMOX2	heme oxygenase 2 [Source:HGNC Symbol;Acc:HGNC:5014]
ENSG00000100099	HPS4	HPS4, biogenesis of lysosomal organelles complex 3 subunit 2 [Source:HGNC Symbol;Acc:HGNC:15844]
ENSG00000235941	HSPA1A	heat shock protein family A (Hsp70) member 1A [Source:HGNC Symbol;Acc:HGNC:5232]
ENSG00000237724	HSPA1A	heat shock protein family A (Hsp70) member 1A [Source:HGNC Symbol;Acc:HGNC:5232]
ENSG00000234475	HSPA1A	heat shock protein family A (Hsp70) member 1A [Source:HGNC Symbol;Acc:HGNC:5232]
ENSG00000215328	HSPA1A	heat shock protein family A (Hsp70) member 1A [Source:HGNC Symbol;Acc:HGNC:5232]
ENSG00000204389	HSPA1A	heat shock protein family A (Hsp70) member 1A [Source:HGNC Symbol;Acc:HGNC:5232]
ENSG00000044574	HSPA5	heat shock protein family A (Hsp70) member 5 [Source:HGNC Symbol;Acc:HGNC:5238]
ENSG00000157423	HYDIN	HYDIN, axonemal central pair apparatus protein [Source:HGNC Symbol;Acc:HGNC:19368]
ENSG00000283022	HYDIN	HYDIN, axonemal central pair apparatus protein [Source:HGNC Symbol;Acc:HGNC:19368]
ENSG00000162783	IER5	immediate early response 5 [Source:HGNC Symbol;Acc:HGNC:5393]
ENSG00000129351	ILF3	interleukin enhancer binding factor 3 [Source:HGNC Symbol;Acc:HGNC:6038]
ENSG00000168264	IRF2BP2	interferon regulatory factor 2 binding protein 2 [Source:HGNC Symbol;Acc:HGNC:21729]
ENSG00000079616	KIF22	kinesin family member 22 [Source:HGNC Symbol;Acc:HGNC:6391]
ENSG00000142945	KIF2C	kinesin family member 2C [Source:HGNC Symbol;Acc:HGNC:6393]
ENSG00000205420	KRT6A	keratin 6A [Source:HGNC Symbol;Acc:HGNC:6443]
ENSG00000001497	LAS1L	LAS1 like, ribosome biogenesis factor [Source:HGNC Symbol;Acc:HGNC:25726]
ENSG00000086730	LAT2	linker for activation of T-cells family member 2 [Source:HGNC Symbol;Acc:HGNC:12749]
ENSG00000136167	LCP1	lymphocyte cytosolic protein 1 [Source:HGNC Symbol;Acc:HGNC:6528]
ENSG00000008735	MAPK8IP2	mitogen-activated protein kinase 8 interacting protein 2 [Source:HGNC Symbol;Acc:HGNC:6883]
ENSG00000073111	MCM2	minichromosome maintenance complex component 2 [Source:HGNC Symbol;Acc:HGNC:6944]
ENSG00000112118	MCM3	minichromosome maintenance complex component 3 [Source:HGNC Symbol;Acc:HGNC:6945]
ENSG00000166508	MCM7	minichromosome maintenance complex component 7 [Source:HGNC Symbol;Acc:HGNC:6950]
ENSG00000055732	MCOLN3	mucopolipin 3 [Source:HGNC Symbol;Acc:HGNC:13358]
ENSG00000112282	MED23	mediator complex subunit 23 [Source:HGNC Symbol;Acc:HGNC:2372]
ENSG00000147065	MSN	moesin [Source:HGNC Symbol;Acc:HGNC:7373]
ENSG00000030304	MUSK	muscle associated receptor tyrosine kinase [Source:HGNC Symbol;Acc:HGNC:7525]
ENSG00000121579	NAA50	N(alpha)-acetyltransferase 50, NatE catalytic subunit [Source:HGNC Symbol;Acc:HGNC:29533]
ENSG00000184117	NIPSNAP1	nipsnap homolog 1 [Source:HGNC Symbol;Acc:HGNC:7827]
ENSG00000165271	NOL6	nucleolar protein 6 [Source:HGNC Symbol;Acc:HGNC:19910]
ENSG00000180530	NRIP1	nuclear receptor interacting protein 1 [Source:HGNC Symbol;Acc:HGNC:8001]
ENSG00000185652	NTF3	neurotrophin 3 [Source:HGNC Symbol;Acc:HGNC:8023]
ENSG00000154814	OXNAD1	oxidoreductase NAD binding domain containing 1 [Source:HGNC Symbol;Acc:HGNC:25128]
ENSG00000146247	PHIP	pleckstrin homology domain interacting protein [Source:HGNC Symbol;Acc:HGNC:15673]
ENSG00000118762	PKD2	polycystin 2, transient receptor potential cation channel [Source:HGNC Symbol;Acc:HGNC:9009]
ENSG00000163344	PMVK	phosphomevalonate kinase [Source:HGNC Symbol;Acc:HGNC:9141]
ENSG00000105287	PRKD2	protein kinase D2 [Source:HGNC Symbol;Acc:HGNC:17293]
ENSG00000100462	PRMT5	protein arginine methyltransferase 5 [Source:HGNC Symbol;Acc:HGNC:10894]

ENSG00000185246	PRPF39	pre-mRNA processing factor 39 [Source:HGNC Symbol;Acc:HGNC:20314]
ENSG00000147224	PRPS1	phosphoribosyl pyrophosphate synthetase 1 [Source:HGNC Symbol;Acc:HGNC:9462]
ENSG00000100994	PYGB	glycogen phosphorylase B [Source:HGNC Symbol;Acc:HGNC:9723]
ENSG00000162924	REL	REL proto-oncogene, NF-kB subunit [Source:HGNC Symbol;Acc:HGNC:9954]
ENSG00000143333	RGS16	regulator of G-protein signaling 16 [Source:HGNC Symbol;Acc:HGNC:9997]
ENSG0000018189	RUFY3	RUN and FYVE domain containing 3 [Source:HGNC Symbol;Acc:HGNC:30285]
ENSG00000047634	SCML1	sex comb on midleg-like 1 (Drosophila) [Source:HGNC Symbol;Acc:HGNC:10580]
ENSG00000099937	SERPIND1	serpin family D member 1 [Source:HGNC Symbol;Acc:HGNC:4838]
ENSG00000096717	SIRT1	sirtuin 1 [Source:HGNC Symbol;Acc:HGNC:14929]
ENSG00000114544	SLC41A3	solute carrier family 41 member 3 [Source:HGNC Symbol;Acc:HGNC:31046]
ENSG00000145147	SLIT2	slit guidance ligand 2 [Source:HGNC Symbol;Acc:HGNC:11086]
ENSG00000273173	SNURF	SNRPN upstream reading frame [Source:HGNC Symbol;Acc:HGNC:11171]
ENSG00000167642	SPINT2	serine peptidase inhibitor, Kunitz type 2 [Source:HGNC Symbol;Acc:HGNC:11247]
ENSG00000168610	STAT3	signal transducer and activator of transcription 3 [Source:HGNC Symbol;Acc:HGNC:11364]
ENSG00000243244	STON1	stonin 1 [Source:HGNC Symbol;Acc:HGNC:17003]
ENSG00000055208	TAB2	TGF-beta activated kinase 1/MAP3K7 binding protein 2 [Source:HGNC Symbol;Acc:HGNC:17075]
ENSG00000064115	TM7SF3	transmembrane 7 superfamily member 3 [Source:HGNC Symbol;Acc:HGNC:23049]
ENSG00000100216	TOMM22	translocase of outer mitochondrial membrane 22 [Source:HGNC Symbol;Acc:HGNC:18002]
ENSG00000115129	TP53I3	tumor protein p53 inducible protein 3 [Source:HGNC Symbol;Acc:HGNC:19373]
ENSG00000198467	TPM2	tropomyosin 2 (beta) [Source:HGNC Symbol;Acc:HGNC:12011]
ENSG00000100271	TTL1	tubulin tyrosine ligase like 1 [Source:HGNC Symbol;Acc:HGNC:1312]
ENSG00000176014	TUBB6	tubulin beta 6 class V [Source:HGNC Symbol;Acc:HGNC:20776]
ENSG00000140553	UNC45A	unc-45 myosin chaperone A [Source:HGNC Symbol;Acc:HGNC:30594]
ENSG00000173915	USMG5	up-regulated during skeletal muscle growth 5 homolog (mouse) [Source:HGNC Symbol;Acc:HGNC:30889]
ENSG00000165280	VCP	valosin containing protein [Source:HGNC Symbol;Acc:HGNC:12666]
ENSG00000158195	WASF2	WAS protein family member 2 [Source:HGNC Symbol;Acc:HGNC:12733]
ENSG00000178252	WDR6	WD repeat domain 6 [Source:HGNC Symbol;Acc:HGNC:12758]
ENSG00000176871	WSB2	WD repeat and SOCS box containing 2 [Source:HGNC Symbol;Acc:HGNC:19222]
ENSG00000154767	XPC	XPC complex subunit, DNA damage recognition and repair factor [Source:HGNC Symbol;Acc:HGNC:12816]
ENSG00000152422	XRCC4	X-ray repair cross complementing 4 [Source:HGNC Symbol;Acc:HGNC:12831]
ENSG00000023041	ZDHHC6	zinc finger DHHC-type containing 6 [Source:HGNC Symbol;Acc:HGNC:19160]
ENSG00000163867	ZMYM6	zinc finger MYM-type containing 6 [Source:HGNC Symbol;Acc:HGNC:13050]
ENSG00000160961	ZNF333	zinc finger protein 333 [Source:HGNC Symbol;Acc:HGNC:15624]
ENSG00000198131	ZNF544	zinc finger protein 544 [Source:HGNC Symbol;Acc:HGNC:16759]
ENSG00000196381	ZNF781	zinc finger protein 781 [Source:HGNC Symbol;Acc:HGNC:26745]
ENSG00000130449	ZSWIM6	zinc finger SWIM-type containing 6 [Source:HGNC Symbol;Acc:HGNC:29316]

Πίνακας 17: Διαφορικά εκφραζόμενα γονίδια που προκύπτουν από την βιβλιογραφία για τις χαμηλές δόσεις (<0.5Gy).

ENSG_ID	HGNC_Symbol	Description
ENSG00000141385	AFG3L2	AFG3 like matrix AAA peptidase subunit 2 [Source:HGNC Symbol;Acc:HGNC:315]
ENSG00000164331	ANKRA2	ankyrin repeat family A member 2 [Source:HGNC Symbol;Acc:HGNC:13208]
ENSG00000196072	BLOC1S2	biogenesis of lysosomal organelles complex 1 subunit 2 [Source:HGNC Symbol;Acc:HGNC:20984]
ENSG00000070808	CAMK2A	calcium/calmodulin dependent protein kinase II alpha [Source:HGNC Symbol;Acc:HGNC:1460]
ENSG00000137500	CCDC90B	coiled-coil domain containing 90B [Source:HGNC Symbol;Acc:HGNC:28108]
ENSG00000117724	CENPF	centromere protein F [Source:HGNC Symbol;Acc:HGNC:1857]
ENSG00000182973	CNOT10	CCR4-NOT transcription complex subunit 10 [Source:HGNC Symbol;Acc:HGNC:23817]
ENSG00000110063	DCPS	decapping enzyme, scavenger [Source:HGNC Symbol;Acc:HGNC:29812]
ENSG00000149547	EI24	EI24, autophagy associated transmembrane protein [Source:HGNC Symbol;Acc:HGNC:13276]
ENSG00000108515	ENO3	enolase 3 [Source:HGNC Symbol;Acc:HGNC:3354]
ENSG00000005238	FAM214B	family with sequence similarity 214 member B [Source:HGNC Symbol;Acc:HGNC:25666]
ENSG00000167196	FBX022	F-box protein 22 [Source:HGNC Symbol;Acc:HGNC:13593]
ENSG00000170445	HARS	histidyl-tRNA synthetase [Source:HGNC Symbol;Acc:HGNC:4816]
ENSG00000101294	HM13	histocompatibility minor 13 [Source:HGNC Symbol;Acc:HGNC:16435]
ENSG00000277424	HMOX2	heme oxygenase 2 [Source:HGNC Symbol;Acc:HGNC:5014]
ENSG00000103415	HMOX2	heme oxygenase 2 [Source:HGNC Symbol;Acc:HGNC:5014]
ENSG00000100099	HPS4	HPS4, biogenesis of lysosomal organelles complex 3 subunit 2 [Source:HGNC Symbol;Acc:HGNC:15844]
ENSG00000044574	HSPA5	heat shock protein family A (Hsp70) member 5 [Source:HGNC Symbol;Acc:HGNC:5238]
ENSG00000162783	IER5	immediate early response 5 [Source:HGNC Symbol;Acc:HGNC:5393]
ENSG00000129351	ILF3	interleukin enhancer binding factor 3 [Source:HGNC Symbol;Acc:HGNC:6038]
ENSG00000168264	IRF2BP2	interferon regulatory factor 2 binding protein 2 [Source:HGNC Symbol;Acc:HGNC:21729]
ENSG00000086730	LAT2	linker for activation of T-cells family member 2 [Source:HGNC Symbol;Acc:HGNC:12749]
ENSG00000136167	LCP1	lymphocyte cytosolic protein 1 [Source:HGNC Symbol;Acc:HGNC:6528]
ENSG00000147065	MSN	moesin [Source:HGNC Symbol;Acc:HGNC:7373]
ENSG00000121579	NAA50	N(alpha)-acetyltransferase 50, NatE catalytic subunit [Source:HGNC Symbol;Acc:HGNC:29533]
ENSG00000185652	NTF3	neurotrophin 3 [Source:HGNC Symbol;Acc:HGNC:8023]
ENSG00000114544	SLC41A3	solute carrier family 41 member 3 [Source:HGNC Symbol;Acc:HGNC:31046]
ENSG00000167642	SPINT2	serine peptidase inhibitor, Kunitz type 2 [Source:HGNC Symbol;Acc:HGNC:11247]
ENSG00000064115	TM7SF3	transmembrane 7 superfamily member 3 [Source:HGNC Symbol;Acc:HGNC:23049]
ENSG00000115129	TP53I3	tumor protein p53 inducible protein 3 [Source:HGNC Symbol;Acc:HGNC:19373]
ENSG00000176014	TUBB6	tubulin beta 6 class V [Source:HGNC Symbol;Acc:HGNC:20776]
ENSG00000173915	USMG5	up-regulated during skeletal muscle growth 5 homolog (mouse) [Source:HGNC Symbol;Acc:HGNC:30889]
ENSG00000178252	WDR6	WD repeat domain 6 [Source:HGNC Symbol;Acc:HGNC:12758]
ENSG00000176871	WSB2	WD repeat and SOCS box containing 2 [Source:HGNC Symbol;Acc:HGNC:19222]
ENSG00000154767	XPC	XPC complex subunit, DNA damage recognition and repair factor [Source:HGNC Symbol;Acc:HGNC:12816]
ENSG00000130449	ZSWIM6	zinc finger SWIM-type containing 6 [Source:HGNC Symbol;Acc:HGNC:29316]

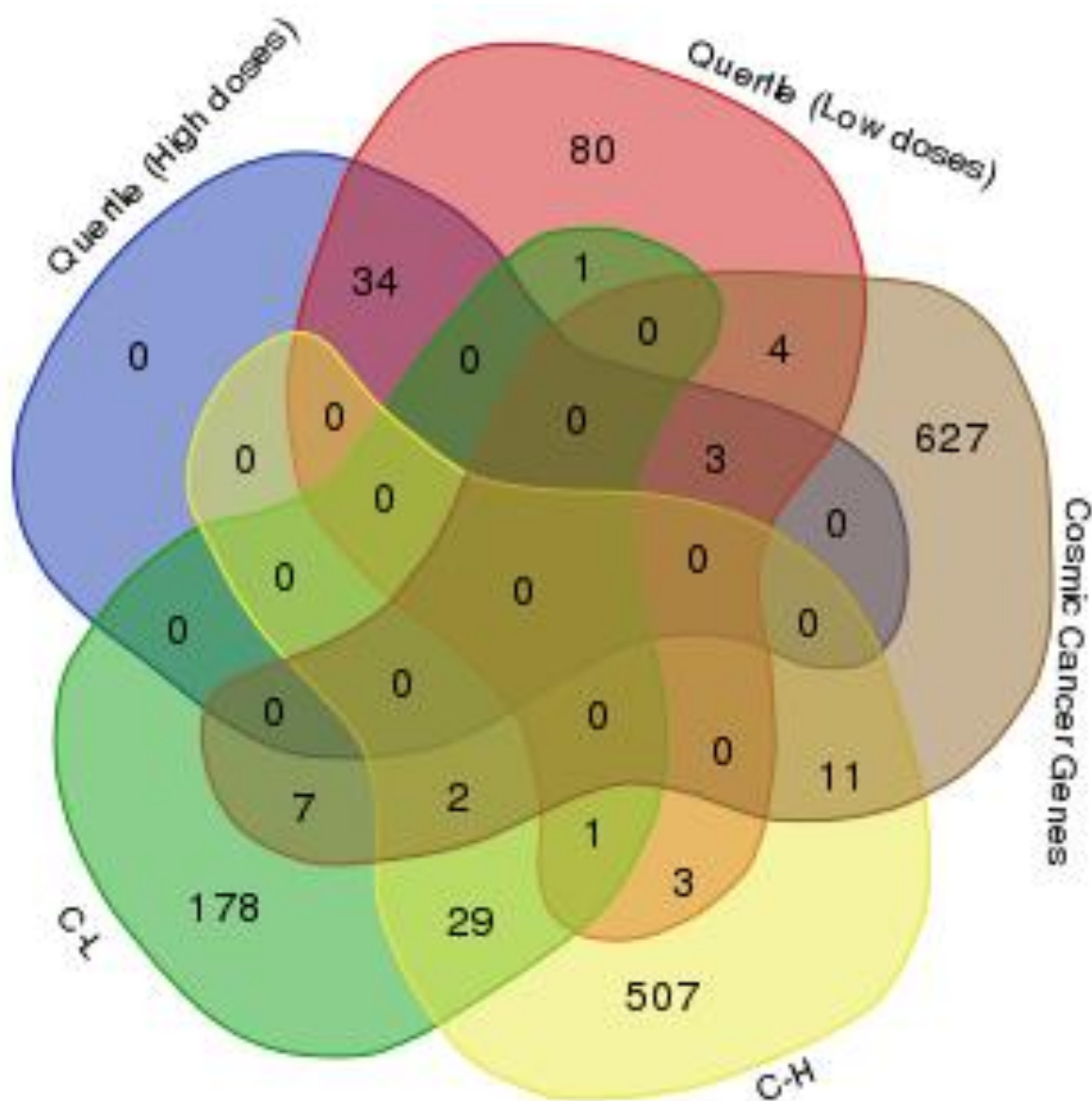
Πίνακας 18: Διαφορικά εκφραζόμενα γονίδια που προκύπτουν από την βιβλιογραφία για τις υψηλές δόσεις (>0.5Gy)

ΚΟΙΝΟΣ ΓΕΩΜΕΤΡΙΚΟΣ ΤΟΠΟΣ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

List names	number of elements	number of unique elements
C-H	553	553
C-L	218	218
Cosmic Cancer Genes	654	654
Quertle (High doses)	37	37
Quertle (Low doses)	126	126
Overall number of unique elements	1487	

Πίνακας 19: Δεδομένα εισαγωγής στο Venn-διάγραμμα

Η κατασκευή του διαγράμματος Venn (Εικόνα 52) έγινε με τα δεδομένα που παρουσιάζονται στον Πίνακας 19.



Εικόνα 52: Διάγραμμα Venn μεταξύ των διαφορετικά εκφρασμένων γονιδίων της παρούσας μελέτης, των γονιδίων από τη βιβλιογραφία και τέλος τη λίστα γονιδίων που σχετίζονται με καρκίνο.

Βρέθηκαν 11 γονίδια κοινά στις υψηλές δόσεις με τα γονίδια σχετιζόμενα με καρκίνο (

Πίνακας 20), 2 γονίδια κοινά εκφρασμένα στις υψηλές και χαμηλές δόσεις και σχετιζόμενα με καρκίνο (Πίνακας 21), 7 κοινά γονίδια διαφορετικά εκφρασμένα με χαμηλή δόση που σχετίζονται με τον καρκίνο. (Πίνακας 22).

Gene stable ID	HGNC symbol	Gene description
ENSG00000057657	PRDM1	PR/SET domain 1 [Source:HGNC Symbol;Acc:HGNC:9346]
ENSG00000070404	FSTL3	follistatin like 3 [Source:HGNC Symbol;Acc:HGNC:3973]
ENSG00000112081	SRSF3	serine and arginine rich splicing factor 3 [Source:HGNC Symbol;Acc:HGNC:10785]
ENSG00000116062	MSH6	mutS homolog 6 [Source:HGNC Symbol;Acc:HGNC:7329]
ENSG00000135111	TBX3	T-box 3 [Source:HGNC Symbol;Acc:HGNC:11602]
ENSG00000147862	NFIB	nuclear factor I B [Source:HGNC Symbol;Acc:HGNC:7785]
ENSG00000165556	CDX2	caudal type homeobox 2 [Source:HGNC Symbol;Acc:HGNC:1806]
ENSG00000167985	SDHAF2	succinate dehydrogenase complex assembly factor 2 [Source:HGNC Symbol;Acc:HGNC:26034]
ENSG00000168685	IL7R	interleukin 7 receptor [Source:HGNC Symbol;Acc:HGNC:6024]
ENSG00000204103	MAFB	MAF bZIP transcription factor B [Source:HGNC Symbol;Acc:HGNC:6408]
ENSG00000213190	MLLT11	myeloid/lymphoid or mixed-lineage leukemia; translocated to, 11 [Source:HGNC Symbol;Acc:HGNC:16997]

Πίνακας 20: Κοινός τύπος γονιδίων που είναι διαφορετικά εκφρασμένα στις υψηλές δόσεις στη μελέτη μας και γονίδια που σχετίζονται με τον καρκίνο.

Gene stable ID	HGNC symbol	Gene description
ENSG00000005073	HOXA11	homeobox A11 [Source:HGNC Symbol;Acc:HGNC:5101]
ENSG00000135956	TMEM127	transmembrane protein 127 [Source:HGNC Symbol;Acc:HGNC:26038]

Πίνακας 21: Κοινός τύπος γονιδίων που είναι διαφορετικά εκφρασμένα στις υψηλές και στις χαμηλές δόσεις στη μελέτη μας με γονίδια που σχετίζονται με τον καρκίνο.

Gene stable ID	HGNC symbol	Gene description
ENSG00000048462	TNFRSF17	TNF receptor superfamily member 17 [Source:HGNC Symbol;Acc:HGNC:11913]
ENSG00000067842	ATP2B3	ATPase plasma membrane Ca ²⁺ transporting 3 [Source:HGNC Symbol;Acc:HGNC:816]
ENSG00000104903	LYL1	LYL1, basic helix-loop-helix family member [Source:HGNC Symbol;Acc:HGNC:6734]

<u>ENSG00000106462</u>	<u>EZH2</u>	enhancer of zeste 2 polycomb repressive complex 2 subunit [Source:HGNC Symbol;Acc:HGNC:3527]
<u>ENSG00000123473</u>	<u>STIL</u>	SCL/TAL1 interrupting locus [Source:HGNC Symbol;Acc:HGNC:10879]
<u>ENSG00000171723</u>	<u>GPHN</u>	gephyrin [Source:HGNC Symbol;Acc:HGNC:15465]
<u>ENSG00000175643</u>	<u>RMI2</u>	RecQ mediated genome instability 2 [Source:HGNC Symbol;ι μεAcc:HGNC:28349]

Πίνακας 22: Κοινά διαφορετικά εκφρασμένα γονίδια στις χαμηλές δόσεις που σχετίζονται με καρκίνο.

ΣΥΖΗΤΗΣΗ-ΠΡΟΟΠΤΙΚΕΣ

Λόγω του όγκου των αποτελεσμάτων παραθέτουμε ένα συνοπτικό πίνακα για να διακρίνουμε σε ποιες αναλύσεις είχαμε αποτελέσματα:

Control/High (16-24)hrs		
Type of Enrichment/S.L.	0.05	0.1
GO	-	-
KEGG	-	-
TF	+UP +DOWN	+UP +DOWN
MicroRNA	-UP +DOWN	+UP +DOWN
Control/Low (16-24)hrs (<0.5Gy)		
Type of Enrichment/S.L.	0.05	0.1
GO	-	+
KEGG	-	+
TF	-UP -DOWN	-UP -DOWN
MicroRNA	-UP -DOWN	-UP +DOWN

Πίνακας 23: Συνοπτική παράθεση αποτελεσμάτων αναλύσεων WebGestalt.

Διαπιστώνουμε με μια πρώτη ματιά ότι έχουμε μεγαλύτερη απόκριση στο διάστημα 16-24 ωρών στις χαμηλές δόσεις (<0.5Gy) συγκριτικά με τις υψηλές δόσεις. Ενδεχομένως το συγκεκριμένο αποτέλεσμα να οφείλεται στην ραδιοπροσαρμοστική απόκριση (RAR) που συναντούμε στα κύτταρα και αναφέρεται στο φαινόμενο όπου μια προηγούμενη χαμηλή αρχική δόση μειώνει τη βιολογική επίπτωση επακόλουθων υψηλών δόσεων (Blaese et al., 2003; Kim et al., 2011). Η ακτινοβολία μεταβάλλει τα προφίλ γονιδιακής έκφρασης και προστατεύει τα κύτταρα από μετέπειτα έκθεση. Για κλινική ακτινοθεραπεία, το RAR μπορεί να αξιοποιηθεί για την αποφυγή ή τη μείωση των υπερβολικών βλαβών σε φυσιολογικά σωματικά κύτταρα που προκαλούνται από την ακτινοβολία. Μια πιθανή εξήγηση είναι επίσης ότι η ρύθμιση στα γονίδια μετά την ακτινοβόληση γίνεται νωρίτερα χρονικά και στο διάστημα αυτό βρισκόμαστε σε μια κατάσταση ηρεμίας στο κύτταρο και έχουν πάψει οι μηχανισμοί καταστολής ή απόπτωσης.

Στην ανάλυση της υπεργεωμετρικής κατανομής με $\text{adj}P < 0.1$, παρατηρήθηκε υπερεκπροσώπηση κάποιων μεταγραφικών παραγόντων που ελέγχουν γονίδια τα οποία εκφράζονται διαφορικά σε υψηλές δόσεις (Πίνακας 11).

Ειδικότερα είχαμε υπερεκπροσώπηση σε 7 υπέρ-εκφρασμένα γονίδια που οι ρυθμιστικές τους αλληλουχίες αναγνωρίζονται από το μεταγραφικό παράγοντα CUX1. Ο τελευταίος αποτελεί μέλος της οικογένειας των μεταγραφικών παραγόντων Homeodomain και περιέχει τρεις επαναλαμβανόμενες περιοχές προσδέσεως DNA, που ονομάζονται *Cut* (CR). Λειτουργεί ως παράγοντας προσδιορισμού της τύχης του κυττάρου κατά την ανάπτυξη. Τα γονίδια-στόχοι του που υπερεκφράζονται μετά από υψηλές δόσεις, είναι τα:

- **DPP1** που καταλύει την αποφωσφορυλίωση της πυροφωσφορικής διακυλγλυκερόλης για να σχηματίσει φωσφατιδικό άλας και του Pi, το πυροφωσφορικό διακυλγλυκερόλης, το φωσφατιδικό και η διακυλγλυκερόλη έχουν ρόλους ως μόρια σημάτων λιπιδίων σε ανώτερα ευκαρυωτικά κύτταρα συμμετέχοντας έτσι στη γονιδιακή ρύθμιση κυτταρικού επιπέδου.
- **GART** η πρωτεΐνη που κωδικοποιείται από αυτό το γονίδιο είναι ένα τριλειτουργικό πολυπεπτίδιο. Έχει ενεργότητα φωσφοριβοσυλγλυκιναμιδικής φορμυλοτρανσφεράσης, φωσφοριβοσυλγλυκιναμιδικής συνθετάσης και φωσφοριβοσυλαμινοϊμιδαζολικής συνθετάσης που απαιτούνται για την βιοσύνθεση πουρίνης *de novo*. Οι πουρίνες είναι κρίσιμες για τον ενεργειακό μεταβολισμό, την κυτταρική σηματοδότηση και την αναπαραγωγή των κυττάρων.
- **FOXG**, είναι ένα γονίδιο του οποίου η οικογένεια των μεταγραφικών παραγόντων συμμετέχει σε σημαντικές βιολογικές λειτουργίες σε πολλά είδη, από τον έλεγχο του κυτταρικού κύκλου μέχρι τη διαφοροποίηση των επιθηλίων και από την ανάπτυξη του πλακούντα μέχρι το σχηματισμό του έσω ωτός (Hannenhalli and Kaestner, 2009).
- **NFIB** (Πυρηνικός Παράγοντας I B) που είναι ένας μεταγραφικός παράγοντας.
- **DYRK2** που ανήκει σε μια οικογένεια πρωτεϊνικών κινασών των οποίων τα μέλη τεκμαίρεται ότι εμπλέκονται στην κυτταρική ανάπτυξη. Το DYRK2 (διπλής ειδικότητας ρυθμιζόμενη από τη φωσφορυλίωση τυροσίνης κινάση 2) είναι γονίδιο σχετικό με την έκφραση γονιδίων και τη ρύθμιση της ενεργότητας του TP53.
- **EMP1** που κωδικοποιεί την πρωτεΐνη επιθηλιακής μεμβράνης 1 και μειώνει σημαντικά τη βιωσιμότητα των κυττάρων, τον σχηματισμό αποικιών και τη μετανάστευση, αλλά ενισχύει την απόπτωση.

Επιπλέον, άλλα 6 γονίδια που υπέρ-εκφράζονται, φέρουν στην ρυθμιστική τους περιοχή, ένα μοτίβο που περιγράφεται από την συναινετική αλληλουχία RRCCGTTA της οποίας ο μεταγραφικός παράγοντας που την

αναγνωρίζει είναι άγνωστος. Ενδιαφέρον έχει το γονίδιο GART ,που περιέχει το μοτίβο αυτό στη ρυθμιστική του αλληλουχία, και το αναγνωρίζει και ο CUX1.

Όσον αφορά τα γονίδια που μειώθηκε η έκφραση τους, είχαμε υπερεκπροσώπηση σε 15 γονίδια που ρυθμίζονται από το μεταγραφικό παράγοντα AP1 και 10 γονίδια από τον GABP (Πίνακας 12). Ενδιαφέρον έχει το γονίδιο DTX2 που ρυθμίζεται και από τους δύο αυτούς μεταγραφικούς παράγοντες. Το γονίδιο αυτό λειτουργεί ως λιγάση ουμπικουιτίνης E3 (Takeyama et al., 2003) και αποτελεί ρυθμιστή του μονοπατιού μεταγωγής σήματος του Notch.

Η πρωτεΐνη-ενεργοποιητής 1 (AP-1) είναι ένας μεταγραφικός παράγοντας που ρυθμίζει τη γονιδιακή έκφραση σε απόκριση σε μια ποικιλία ερεθισμάτων, συμπεριλαμβανομένων κυτοκινών, αυξητικών παραγόντων, στρες και βακτηριακών και ιικών λοιμώξεων. Το AP-1 ελέγχει έναν αριθμό κυτταρικών διεργασιών περιλαμβανομένης της διαφοροποίησης, του πολλαπλασιασμού και της απόπτωσης. Η δομή του AP-1 είναι ένα ετεροδιμερές φερμουάρ λευκίνης που αποτελείται από πρωτεΐνες που ανήκουν στις οικογένειες c-Fos, c-Jun, ATF και JDP (Hess et al., 2004).

Από τα γονίδια που υποεκφράζονται και η έκφραση τους εξαρτάται από τον AP1, το σημαντικότερο είναι το **CDKN1A** γνωστό p21 που αποτελεί τον εξαρτώμενο από κυκλίνη αναστολέα κινάσης (CKI). Η πρωτεΐνη αυτή αποτελεί μείζονα στόχο του p53 και έτσι σχετίζεται με τη σύνδεση των βλαβών του DNA με την κυτταρική αναστολή (Bunz et al., 1998). Η πρωτεΐνη p21 δεσμεύεται στα σύμπλοκα κυκλίνης-CDK2, -CDK1 και -CDK4/6 και αναστέλλει τη δράση τους με συνέπεια να λειτουργεί ως ρυθμιστής της εξέλιξης του κυτταρικού κύκλου στη φάση G1 και S. Το p21 αλληλεπιδρά επίσης με το πυρηνικό αντιγόνο πολλαπλασιασμού των κυττάρων (PCNA), έναν βοηθητικό παράγοντα πολυμεράσης DNA και παίζει ρυθμιστικό ρόλο στην αντιγραφή DNA στη φάση S και στην επιδιόρθωση της βλάβης του DNA και συμμετέχει στην αναστολή απόπτωσης. Αυτή η πρωτεΐνη αναφέρθηκε ότι διασπάται ειδικά από κασπάσες τύπου CASP3, πράγμα που οδηγεί με τον τρόπο αυτό σε ενεργοποίηση του CDK2 και μπορεί να συμβάλει στην εκτέλεση της απόπτωσης μετά την ενεργοποίηση της κασπάσης. Ωστόσο, η p21 μπορεί να αναστείλει την απόπτωση και δεν προκαλεί από μόνη της τον κυτταρικό θάνατο. Μελέτες της διακοπής του κυτταρικού κύκλου που εξαρτάται από τον p53 σε απόκριση της βλάβης του DNA αναγνώρισαν την p21 ως τον κύριο μεσολαβητή της κατάρρευσης του κυτταρικού κύκλου. Τα κύτταρα με υψηλό p21 εισέρχονται σε κατάσταση G0 / ηρεμίας, ενώ εκείνα με χαμηλή p21 συνεχίζουν να πολλαπλασιάζονται (el-Deiry et al., 1993).

Ο αναστολέας p21 λοιπόν έχει διπλή ανασταλτική δράση. Από τη μία ενώνεται στο σύμπλεγμα της CDK2 με τις κυκλίνες A, D και E και οδηγεί στην απενεργοποίησή τους και από την άλλη έχει μια περιοχή σύνδεσης για τη συμπληρωματική πρωτεΐνη αναδιπλασιασμού PCNA, η οποία είναι σημαντική για τη σύνθεση του DNA και τη σύνδεση της πολυμεράσης με το DNA. Η πρόσδεση του αναστολέα αποτρέπει τον αναδιπλασιασμό του DNA από την πολυμεράση. Η συγκέντρωση του p21 ρυθμίζεται στο επίπεδο της μεταγραφής και κύριο ρόλο σ' αυτό παίζει η ογκοκατασταλτική πρωτεΐνη p53. Η p53 προωθεί τη μεταγραφή του p21C μέσω σύνδεσής

της στον υποκινητή του. Η ενεργοποίηση της p53 παρατηρείται κυρίως κατά την καταστροφή του DNA. Ο κύριος σκοπός της ρύθμισης είναι η επιδιόρθωση των βλαβών ή η προώθηση της απόπτωσης (Rodriguez and Meuth, 2006).

Στις υψηλές δόσεις στη ανάλυση στόχων miRNA βρέθηκε υπερεκπροσώπηση γονιδίων που είναι στόχοι της οικογένειας miRNA-92a και ειδικότερα τα miR-25,miR-32,miR-92,miR-363,miR-367 με κοινή φύτρα –GTGCAAT- (Πίνακας 13) στην συνέχεια είχαμε υπερεκπροσώπηση στα γονίδια που είναι στόχοι του miR-185 με κοινή φύτρα –TCTCTCC-. Αυτά τα miR ενδεχομένως υποεκφράζονται ως απόκριση στην ακτινοβολία άρα ανεβαίνουν οι στόχοι. Από ότι παρατηρούμε στους στόχους υπάρχουν γονίδια που σχετίζονται με την ογκογένεση όπως το TP53INP1 και το RAB23.

Το MicroRNA-92a ρυθμίζει αρνητικά τη φλεγμονώδη ανταπόκριση των υποδοχέων τύπου Toll-like (TLR) με τη στόχευση κινάσης MKK4 (Lai et al., 2013). Οι υποδοχείς τύπου Toll (TLRs) διαδραματίζουν κρίσιμο ρόλο στην εκκίνηση ανοσοαποκρίσεων κατά των παθογόνων μικροοργανισμών που εισβάλλουν. Η αύξηση της έκφρασης των γονιδίων-στόχων συνδυάζεται με τη μείωση τα επίπεδα του microRNA-92a (miRNA-92a). Έπειτα ακολουθεί η ενεργοποίηση της οδού JNK / c-Jun και η παραγωγή φλεγμονωδών κυτοκινών σε μακροφάγα όταν διεγείρονται με προσδέματα για TLR4. Περαιτέρω, η (MKK4) ενεργοποιεί την κινάση πρωτεΐνης JNK / κινάσης ενεργοποιούμενη από στρες, και στοχεύεται απευθείας από τα miR-92a. Μηχανισμοί φλεγμονώδους απόκρισης ενδεχομένως να ενεργοποιούνται στην περίπτωση μας με τα μειωμένα επίπεδα miRNA που διαφαινούνται από την αύξηση της έκφρασης των γονιδίων λόγω ακτινοβολίας με υψηλή δόση ακτινοβολίας. Όμοια αποτελέσματα ρύθμισης έχουμε στην συγκεκριμένη οικογένεια miR και στο καρκίνο του παχέος εντέρου που προτείνεται σα βιοδείκτης (Yang et al., 2014) Το miR-185 είναι γνωστό ότι καταστέλλει τον πολλαπλασιασμό όγκων με άμεση στόχευση των E2F6 και DNMT1 και έμμεσα αναπροσαρμόζοντας τον BRCA1. Συμμετέχει ρυθμίζοντας την έκφραση σε διάφορες διαδικασίες σχετικές με καταστολή όγκων, κυτταρικό θάνατο κλπ

Μεγάλο ενδιαφέρον παρουσιάζουν τα γονίδια **NFIB**, **FOXG1**, **DYRK2** ,**DPP10**. Στα γονίδια αυτά η υπερέκφραση τους ρυθμίζεται και από τον μεταγραφικό παράγοντα CUX1 και από την οικογένεια miR-92a. Ενδεχομένως απόκριση στην υψηλή δόση της ακτινοβολίας επιφέρει συνεργατική δράση του CUX1 που προσδένεται στις ρυθμιστικές τους αλληλουχίες και ανασταλτική δραστηριότητα στην οικογένεια miR-92a των οποίων είναι στόχοι. Παρόμοια αποτελέσματα συνδυασμένης δράσης διαπιστώνουμε και στο μεταγραφικό παράγοντα με μοτίβο συναινετικής αλληλουχίας RRCCGTTA με το miR-185 ;όσον αφορά τη ρύθμιση έκφρασης του γονιδίου **SF1**.

Στην μείωση της έκφρασης βρέθηκαν με $adjP < 0.05$ 10 γονίδια με υπερεκπροσώπηση που είναι στόχοι του miR-448 και 14 που είναι στόχοι του miR-524. Ενδεχομένως η πτώση στην έκφραση των γονιδίων να οφείλεται στην αποσιώπηση που επέφερε η αύξηση έκφρασης των παραπάνω miR (Πίνακας 14).

Στην μείωση της έκφρασης βρέθηκαν με $\text{adjP} < 0.1$ 10 γονίδια με υπερεκπροσώπηση που είναι στόχοι του miR-448 και 14 που είναι στόχοι του miR-524. Επιπλέον 6 γονίδια στόχοι του miR433 και 9 γονίδια στόχοι του miR-24. Ενδεχομένως η πτώση στην έκφραση των γονιδίων να οφείλεται στην αποσιώπηση που επέφερε η αύξηση έκφρασης των παραπάνω miR (Πίνακας 15).

Τα γονίδια **MXI1**, **ADAM10**, **IRS2**, **WTAP**, **C1orf63** ελέγχονται από συνδυασμό miR.

Αξιοσημείωτα είναι και τα **PRDM1**, **PHF23**, **ABHD4** που φαίνεται να ρυθμίζονται σε συνδυασμό με την ενεργοποίηση miR-524 που επιφέρει την αποσιώπηση τους και την καταστολή του μεταγραφικού παράγοντα AP1 καθώς και τα **RAB7A**, **HOXA11** που ρυθμίζονται ταυτόχρονα από το miR-448 και τον AP1.

Στην ανάλυση GO στις χαμηλές δόσεις είχαμε αποτελέσματα υπερεκπροσώπησης μόνο για $\text{adjP} < 0.01$. Αρχικά, σύμφωνα με το Gene Ontology βρέθηκε υπερεκπροσώπηση χαμηλές δόσεις των ακόλουθων όρων γονιδιακής οντολογίας: Θετική ρύθμιση της μετάβασης από G0 σε G1 (Εικόνα 36, 2γονίδια) , ρύθμιση της μετάβασης από G0 σε G1 (Εικόνα 36, 2 γονίδια) , προδιαγραφή τύχης νευρώνων (Εικόνα 36 , 3γονιδια) , μετάβαση από G0 σε G1 (Εικόνα 36 ,2 γονίδια) , αρνητική ρύθμιση της διαφοροποίησης των νευρώνων (Εικόνα 36 ,4 γονίδια) , αλληλουχία ειδικής πρόσδεσης DNA (Εικόνα 36 ,2 γονίδια) ,δεσμοί DNA (Εικόνα 36 ,2 γονίδια).

Η μετάβαση από την G0 στην G1 είναι μια διαδικασία διακοπής του κυτταρικού κύκλου που οδηγεί σε διακοπή κατά τη διάρκεια της φάσης G1, οπότε το κύτταρο εισέρχεται σε μια εξειδικευμένη κατάσταση ηρεμίας γνωστή ως G0 ή ηρεμία. Επιπλέον, το κύτταρο μπορεί να μεταβεί από την G1 σε μία αδρανή φάση την G0, από την οποία μπορεί να εξέλθει μέσω συγκεκριμένων σημάτων. Ένα από αυτά είναι οι αυξητικοί παράγοντες. Με άλλα λόγια η G1 είναι μια φάση περιορισμού που ελέγχεται ο κυτταρικός κύκλος.

Η προδιαγραφή της τύχης των νευρώνων αφορά τη διαδικασία στην οποία ένα κύτταρο καθίσταται ικανό να διαφοροποιείται αυτόνομα σε νευρώνα σε ένα περιβάλλον που είναι ουδέτερο σε σχέση με την αναπτυξιακή του πορεία Με τη διαδικασία “neuron specification”, η μοίρα των κυττάρων μπορεί και να αντιστραφεί πράγμα που υποδεικνύεται από τον άλλο όρο γονιδιακής οντολογίας.

Στην ανάλυση GO στις χαμηλές δόσεις διαπιστώσαμε ότι τα γονίδια από τα δεδομένα μας συμμετέχουν στην απόκριση στην μοριακή λειτουργία του κυττάρου και ειδικότερα 13 γονίδια συμμετέχουν στη διαδικασία της επιλεκτικής και μη ομοιοπολικής πρόσδεσης με DNA που συνθέτει συγκεκριμένα νουκλεοτίδια. (πχ εκκινητές). Η συγκεκριμένη διαδικασία έχει συνδεθεί και με τις βλάβες δίκλωνου DNA κατά την απόκριση των κυττάρων σε χαμηλές δόσεις ακτινοβολίας όπως οι ακτίνες X (Grokhovskii and Zubarev, 1990) καθώς και με τις υψηλές δόσεις (Hada and Georgakilas, 2008).

Στη βάση δεδομένων KEGG PATHWAY είχαμε υπερεκπροσώπηση όρων που αφορούν το μεταβολισμό των πυριμιδινών. Μεταξύ των γονιδίων είναι το DTYMK (δεοξυθυμιδυλική κινάση) είναι ένα γονίδιο κωδικοποίησης πρωτεΐνης. Μεταξύ των σχετικών οδών της είναι η βιοσύνθεση δεοξυριβονουκλεοτιδίων

πυριμιδίνης de novo. Επίσης το POLE3 που είναι πρωτεΐνη που συνδυάζεται εντός μεγαλύτερων ενζυμικών συμπλοκών για μεταγραφή, αντιγραφή και πακεταρίσμα του DNA. Έπειτα η διυδροπυριμιδινάση που είναι ένα ένζυμο που στον άνθρωπο κωδικοποιείται από το γονίδιο DPYS Η διυδροπυριμιδινάση καταλύει τη μετατροπή 5,6-διυδροουρακίλης σε 3-ουρείδοπροπιονικό σε μεταβολισμό πυριμιδίνης. Τέλος το PRIM1, η DNA πριμάση είναι η πολυμεράση που συνθέτει μικρούς RNA primers για τα θραύσματα Okazaki που παράγονται κατά τη διάρκεια ασυνεχούς αντιγραφής DNA. Ετεροδιμερές μιας μικρής υπομονάδας και μιας μεγάλης υπομονάδας. Ανήκει στην οικογένεια μικρής υπομονάδας πριμασών ευκαρυωτικού τύπου.

Τα νουκλεοτίδια πυριμιδίνης διαδραματίζουν κρίσιμο ρόλο στον κυτταρικό μεταβολισμό. Υπάρχουν δύο τρόποι για τη σύνθεση των πυριμιδινών. Τα νουκλεοτίδια μπορούν να ανακυκλωθούν με τις οδούς διάσωσης ή να συντίθενται de novo από μικρούς μεταβολίτες. η βιοσύνθεση de novo πυριμιδίνης είναι απαραίτητη για τον πολλαπλασιασμό των κυττάρων και για την κάλυψη της αυξημένης ζήτησης προδρόμων νουκλεϊκών οξέων και άλλων κυτταρικών συστατικών.

Τέλος στη ανάλυση για τις χαμηλές δόσεις στα γονίδια που αυξήθηκε η έκφραση τους βρέθηκε υπερεκπροσώπηση σε 2 γονίδια που είναι στόχοι της οικογένειας MIR-99A, MIR-100, MIR-99B (Πίνακας 16). Το γονίδιο (INSM1) που σχετίζεται με το ινσουλίνωμα και είναι ενδογενές και κωδικοποιεί μια πρωτεΐνη που περιέχει αμφότερα μια περιοχή δέσμευσης DNA δακτύλου ψευδαργύρου και μια περιοχή προορμόνης. Αυτό το γονίδιο είναι ένας ευαίσθητος δείκτης για τη νευροενδοκρινική διαφοροποίηση των ανθρώπινων πνευμονικών όγκων. Επίσης η θειική ηπαρίνη 3-O-σουλφοτρανσφεράση 2 είναι ένα ένζυμο που στον άνθρωπο κωδικοποιείται από το γονίδιο HS3ST2 Τα βιοσυνθετικά ένζυμα θειικής ηπαρίνης είναι βασικά συστατικά για τη δημιουργία διακεκριμένων λεπτών δομών θειικής ηπαράνης που διεξάγουν πολλαπλές βιολογικές δραστηριότητες. Το ένζυμο που κωδικοποιείται από αυτό το γονίδιο είναι μέλος της οικογένειας των βιοσυνθετικών ενζύμων της θειικής ηπαράνης. Πρόκειται για μια ενσωματωμένη μεμβρανική πρωτεΐνη τύπου II και έχει δραστηριότητα 3-O-σουλφοτρανσφεράσης θειικής ηπαρίνης. Αυτό το γονίδιο εκφράζεται κυρίως στον εγκέφαλο και μπορεί να παίζει ρόλο στο νευρικό σύστημα. Η υποέκφραση αυτών των γονιδίων ίσως να οφείλεται στην αποσιώπηση λόγω της αύξησης της οικογένειας αυτής των MiR.

Τα αποτελέσματα των υψηλών δόσεων θεωρούνται αναμενόμενα συνδεδεμένα με καρκίνο, ογκογένεση, ρύθμιση του p21 πλήθος γονιδίων που συμμετέχουν σε μεταμεταφραστική τροποποίηση. Τα αποτελέσματα των χαμηλών δόσεων παρουσιάζουν χαμηλό επίπεδο εμπιστοσύνης. και όχι ιδιαίτερα ξεκάθαρα αποτελέσματα. Αυτό μπορεί να οφείλεται σε πλήθος παραγόντων. Αρχικά η υπόθεση μας για απόκριση στο διάστημα 16-24 ώρες ίσως να μην είναι ενδεδειγμένη μέθοδος μελέτης. Στη συνέχεια το γεγονός ότι τα πειράματα έχουν διεξαχθεί από διαφορετικούς ερευνητές τοποθετεί διάφορα συστηματικά σφάλματα. Επιπλέον τυχαία και συστηματικά σφάλματα που συμβαίνουν στα πείραμα μικροσυστοιχιών και προκύπτουν από τη χρήση διαφορετικών φθοριζουσών ουσιών τη χρήση διαφορετικών πλατφορμών, τις διαφορετικές πειραματικές συνθήκες και την εισαγωγή θορύβου στα δεδομένα από το σαρωτή. Τέλος η

σύνθεση πολλών διαφορετικών μελετών στη μετα-ανάλυση που αντιστοιχούν σε διαφορετικές πλατφόρμες και time frames ίσως είναι ένας ακόμη λόγος.

Στο μέλλον θα πρέπει να αξιοποιηθούν τα καινοτόμα και προσεκτικά σχεδιασμένα in vivo πειράματα για την αντιμετώπιση των σημαντικών ζητημάτων που σχετίζονται με την επαγωγή, την επεξεργασία και τη βιολογική σημασία των χαμηλών δόσεων ακτινοβολίας καθότι εισέρχονται συστηματικά στην καθημερινότητα.

BIBΛΙΟΓΡΑΦΙΑ

- Abrams, H.L. (1983). *Abrams angiography : vascular and interventional radiology*, 3rd edn (Boston: Little, Brown).
- Agarwal, B.K. (1991). *X-ray spectroscopy : an introduction*, 2nd edn (Berlin ; New York: Springer-Verlag).
- Alessio, N., Del Gaudio, S., Capasso, S., Di Bernardo, G., Cappabianca, S., Cipollaro, M., Peluso, G., and Galderisi, U. (2015). Low dose radiation induced senescence of human mesenchymal stromal cells and impaired the autophagy process. *Oncotarget* 6, 8155-8166.
- Allewell, N.M., Narhi, L.O., and Rayment, I. (2013). *Molecular biophysics for the life sciences* (New York: Springer).
- Altman, N. (2005). Replication, variation and normalisation in microarray experiments. *Appl Bioinformatics* 4, 33-44.
- Arora, R. (2008). *Herbal radiomodulators : applications in medicine, homeland defence and space* (Wallingford, UK ; Cambridge, MA: CABI).
- Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., *et al.* (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25, 25-29.
- Aten, J.A., Stap, J., Krawczyk, P.M., van Oven, C.H., Hoebe, R.A., Essers, J., and Kanaar, R. (2004). Dynamics of DNA double-strand breaks revealed by clustering of damaged chromosome domains. *Science* 303, 92-95.
- Bacq, Z.M. (1965). *Chemical protection against ionizing radiation* (Springfield, Ill.: Thomas).
- Bazyka, D.A., Muzalevska, K.D., Maznichenko, O.L., and Belyaev, O.A. (2014). Expression of gamma-H2AX histone in lymphocytes of the Chornobyl "Shelter" object staff exposed to ionizing radiation in occupational limits. *Probl Radiac Med Radiobiol* 19, 186-191.
- Belov, O.V., Batmunkh, M., Incerti, S., and Lkhagva, O. (2016). Radiation damage to neuronal cells: Simulating the energy deposition and water radiolysis in a small neural network. *Phys Med* 32, 1510-1520.
- Benford, F. (1938). The Law of Anomalous Numbers. *Proceedings of the American Philosophical Society* 78, 551-572.
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological)* 57, 289-300.
- Beyer, R.T. (1949). *Foundations of nuclear physics; facsimiles of thirteen fundamental studies as they were originally reported in the scientific journals* (New York: Dover Publications).
- Blaese, M.A., Santo-Hoeltje, L., and Rodemann, H.P. (2003). Sensitivity of human tumor cells to retinoids or combined treatment with retinoids and ionizing radiation is not dependent on RAR-beta 2 induction. *Strahlenther Onkol* 179, 401-409.
- Bolstad, B.M., Irizarry, R.A., Astrand, M., and Speed, T.P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 19, 185-193.
- Bond, K.R., and Duval, J.S. (1998). Aerial gamma-ray maps of regional surface concentrations of potassium, uranium, and thorium in Utah. In *Geophysical investigations map GP-1013* (Reston, Va.: USGS).
- Bonferroni, C.E. (1936). Teoria statistica delle classi e calcolo delle probabilità. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze* 8, 3-62.
- Bonner, W.M. (2003). Low-dose radiation: thresholds, bystander effects, and adaptive responses. *Proc Natl Acad Sci U S A* 100, 4973-4975.
- Boston, R.C., and Sumner, A.E. (2003). STATA: a statistical analysis system for examining biomedical data. *Adv Exp Med Biol* 537, 353-369.
- Bradač, G.B., and Oberson, R. (1983). *Angiography and computed tomography in cerebroarterial occlusive diseases*, 2nd rev. and expanded edn (Berlin ; New York: Springer-Verlag).
- Brazma, A. (2009). Minimum Information About a Microarray Experiment (MIAME)--successes, failures, challenges. *ScientificWorldJournal* 9, 420-423.
- Brenner, D.J., Doll, R., Goodhead, D.T., Hall, E.J., Land, C.E., Little, J.B., Lubin, J.H., Preston, D.L., Preston, R.J., Puskin, J.S., *et al.* (2003). Cancer risks attributable to low doses of ionizing radiation: assessing what we really know. *Proc Natl Acad Sci U S A* 100, 13761-13766.

Brooks, A.L., Hoel, D.G., and Preston, R.J. (2016). The role of dose rate in radiation cancer risk: evaluating the effect of dose rate at the molecular, cellular and tissue levels using key events in critical pathways following exposure to low LET radiation. *Int J Radiat Biol* 92, 405-426.

Bruford, E.A., Lush, M.J., Wright, M.W., Sneddon, T.P., Povey, S., and Birney, E. (2008). The HGNC Database in 2008: a resource for the human genome. *Nucleic Acids Res* 36, D445-448.

Bunz, F., Dutriaux, A., Lengauer, C., Waldman, T., Zhou, S., Brown, J.P., Sedivy, J.M., Kinzler, K.W., and Vogelstein, B. (1998). Requirement for p53 and p21 to sustain G2 arrest after DNA damage. *Science* 282, 1497-1501.

Cao, W., Epstein, C., Liu, H., DeLoughery, C., Ge, N., Lin, J., Diao, R., Cao, H., Long, F., Zhang, X., *et al.* (2004). Comparing gene discovery from Affymetrix GeneChip microarrays and Clontech PCR-select cDNA subtraction: a case study. *BMC Genomics* 5, 26.

Celicans, G.J. (1973). A review of ionizing radiation (Hyattsville, Md.: National Educational Press).

Cerami, E.G., Gross, B.E., Demir, E., Rodchenkov, I., Babur, O., Anwar, N., Schultz, N., Bader, G.D., and Sander, C. (2011). Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Res* 39, D685-690.

Cercek, L., Ebert, M., and Greene, D. (1968). RBE, OER and dose-rate effects with 14 MeV neutrons relative to 300 kVp x-rays in barley roots. *Int J Radiat Biol Relat Stud Phys Chem Med* 14, 453-462.

Chaudhry, M.A., Omaruddin, R.A., Kreger, B., de Toledo, S.M., and Azzam, E.I. (2012). Micro RNA responses to chronic or acute exposures to low dose ionizing radiation. *Mol Biol Rep* 39, 7549-7558.

Cincu, E., Manea, I., and Woods, M. (2008). Nuclear proficiency testing : 1st International Workshop on Proficiency Testing in Applications of the Ionizing Radiation and Nuclear Analytical Techniques in Industry, Medicine, and Environment, Bucharest, Romania, 6-9 October 2007 (Melville, N.Y.: American Institute of Physics).

Cunningham, F., Amode, M.R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., *et al.* (2015). Ensembl 2015. *Nucleic Acids Res* 43, D662-669.

De la Matta Martin, M., Herrera Gonzalez, A., Lopez Conejos, J.A., and Lopez Romero, J.L. (2011). [Utility of preoperative chest radiography in smokers undergoing transurethral resection of urinary bladder tumors]. *Rev Esp Anestesiol Reanim* 58, 203-210.

el-Deiry, W.S., Tokino, T., Velculescu, V.E., Levy, D.B., Parsons, R., Trent, J.M., Lin, D., Mercer, W.E., Kinzler, K.W., and Vogelstein, B. (1993). WAF1, a potential mediator of p53 tumor suppression. *Cell* 75, 817-825.

Fan, J.B., Gunderson, K.L., Bibikova, M., Yeakley, J.M., Chen, J., Wickham Garcia, E., Lebruska, L.L., Laurent, M., Shen, R., and Barker, D. (2006). [3] Illumina Universal Bead Arrays. In *Methods in Enzymology* (Academic Press), pp. 57-73.

Ferguson, C.D., Kazi, T., Birirā, J.d.t., and Monterey Institute of International Studies. Center for Nonproliferation Studies. (2003). Commercial radioactive sources : surveying the security risks (Monterey, Calif.: Monterey Institute of International Studies, Center for Nonproliferation Studies).

Ferreira, J.A. (2007). The Benjamini-Hochberg method in the case of discrete test statistics. *Int J Biostat* 3, Article 11.

Forbes, S.A., Bhamra, G., Bamford, S., Dawson, E., Kok, C., Clements, J., Menzies, A., Teague, J.W., Futreal, P.A., and Stratton, M.R. (2008). The Catalogue of Somatic Mutations in Cancer (COSMIC). *Curr Protoc Hum Genet Chapter 10*, Unit 10 11.

Forbes, S.A., Bindal, N., Bamford, S., Cole, C., Kok, C.Y., Beare, D., Jia, M., Shepherd, R., Leung, K., Menzies, A., *et al.* (2011). COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res* 39, D945-950.

Gaughran, E.R.L., Goudie, A.J., and Johnson & Johnson. (1974). Technical developments and prospects of sterilization by ionizing radiation : international conference, Vienna, Austria, April 1-4, 1974 (Montreal: Multiscience Publication).

Gene Ontology, C. (2015). Gene Ontology Consortium: going forward. *Nucleic Acids Res* 43, D1049-1056.

Georgakilas, A.G. (2008). Processing of DNA damage clusters in human cells: current status of knowledge. *Mol Biosyst* 4, 30-35.

Georgakilas, A.G. (2011). From chemistry of DNA damage to repair and biological significance. *Comprehending the future. Mutat Res* 711, 1-2.

Georgakilas, A.G. (2015). Bystander and non-targeted effects: a unifying model from ionizing radiation to cancer. *Cancer Lett* 356, 3-4.

Georgakilas, A.G., O'Neill, P., and Stewart, R.D. (2013). Induction and repair of clustered DNA lesions: what do we know so far? *Radiat Res* 180, 100-109.

- Georgakilas, A.G., Sideris, E.G., Sakelliou, L., and Kalfas, C.A. (1999). Low doses of alpha- and gamma-radiation enhance DNA thermal stability. *Biophys Chem* 80, 103-118.
- Georgakilas, A.G., Tsantoulis, P., Kotsinas, A., Michalopoulos, I., Townsend, P., and Gorgoulis, V.G. (2014). Are common fragile sites merely structural domains or highly organized "functional" units susceptible to oncogenic stress? *Cell Mol Life Sci* 71, 4519-4544.
- Giglia, E. (2011). Quertle and KNALIJ: searching PubMed has never been so easy and effective. *Eur J Phys Rehabil Med* 47, 687-690.
- Grokhovskii, S.L., and Zubarev, V.E. (1990). [The specific cleavage of double-stranded DNA induced by the ionization from the x-ray irradiation of the platinum atom]. *Dokl Akad Nauk SSSR* 313, 1500-1504.
- Hada, M., and Georgakilas, A.G. (2008). Formation of clustered DNA damage after high-LET irradiation: a review. *J Radiat Res* 49, 203-210.
- Hamada, N., and Fujimichi, Y. (2014). Classification of radiation effects for dose limitation purposes: history, current situation and future prospects. *J Radiat Res* 55, 629-640.
- Hannenhalli, S., and Kaestner, K.H. (2009). The evolution of Fox genes and their role in development and disease. *Nat Rev Genet* 10, 233-240.
- Hendee, W.R., Ibbott, G.S., and Hendee, E.G. (2005). *Radiation therapy physics*, 3rd edn (Hoboken, N.J.: J. Wiley).
- Hess, J., Angel, P., and Schorpp-Kistner, M. (2004). AP-1 subunits: quarrel and harmony among siblings. *J Cell Sci* 117, 5965-5973.
- Holland, C. (1987). Clancy Holland: Building a bill's momentum. Interview by Ann Gray. *Va Med* 114, 292-294.
- Holm, S. (1979). A simple sequentially rejective bonferroni test procedure. *Scand J Statist* 6, 65-70
- Hüttermann, J., Köhnlein, W., Téoule, R., Bertinchamps, A.J., Ahnström, G., and Commission of the European Communities. (1978). *Effects of ionizing radiation on DNA : physical, chemical and biological aspects* (Berlin ; New York: Springer-Verlag).
- Hymel, G.M. (2011). PubMed Central Inclusion, Quertle Indexing, Outbound Reference Linking, and Editorial Board Successions: Encouraging Developments in the IJTMB's Evolution. *Int J Ther Massage Bodywork* 4, 1-2.
- Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B., and Speed, T.P. (2003). Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 31, e15.
- Johnson, M.M. (2017). Full-Field Digital Mammography and Digital Breast Tomosynthesis. *Radiol Technol* 88, 299M-319M.
- Kal, H.B., and Barendsen, G.W. (1976). The OER at low dose-rates. *Br J Radiol* 49, 1049-1051.
- Kanehisa, M. (2002). The KEGG database. *Novartis Found Symp* 247, 91-101; discussion 101-103, 119-128, 244-152.
- Kent, W.J. (2002). BLAT--the BLAST-like alignment tool. *Genome Res* 12, 656-664.
- Kiehl, E., Walldorf, V., Klimpel, S., Al-Quraishy, S., and Mehlhorn, H. (2009). Genbank accession numbers of sequences of *Culicoides* species vectors of bluetongue virus in Germany. *Parasitol Res* 105, 293-295.
- Kim, W.Y., Lee, J.W., Park, Y.A., Choi, J.J., Sung, C.O., Song, S.Y., Choi, C.H., Kim, T.J., Huh, S.J., Kim, B.G., *et al.* (2011). RAR-beta expression is associated with early volumetric changes to radiation therapy in cervical cancer. *Gynecol Obstet Invest* 71, 11-18.
- Krisiuk, E.M. (1989). [Limitation of the population's exposure to natural sources of ionizing radiation]. *Gig Sanit*, 69-72.
- Kulinskaya, E., Dollinger, M.B., and Bjorkestol, K. (2011). Testing for homogeneity in meta-analysis I. The one-parameter case: standardized mean difference. *Biometrics* 67, 203-212.
- Kutmon, M., Riutta, A., Nunes, N., Hanspers, K., Willighagen, E.L., Bohler, A., Melius, J., Waagmeester, A., Sinha, S.R., Miller, R., *et al.* (2016). WikiPathways: capturing the full diversity of pathway knowledge. *Nucleic Acids Res* 44, D488-494.
- Lai, L., Song, Y., Liu, Y., Chen, Q., Han, Q., Chen, W., Pan, T., Zhang, Y., Cao, X., and Wang, Q. (2013). MicroRNA-92a negatively regulates Toll-like receptor (TLR)-triggered inflammatory response in macrophages by targeting MKK4 kinase. *J Biol Chem* 288, 7956-7967.
- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860-921.
- Lausted, C.G., Warren, C.B., Hood, L.E., and Lasky, S.R. (2006). Printing your own inkjet microarrays. *Methods Enzymol* 410, 168-189.
- Lehnert, S. (2008). *Biomolecular action of ionizing radiation* (New York: Taylor & Francis).

Lipshutz, R.J., Fodor, S.P., Gingeras, T.R., and Lockhart, D.J. (1999). High density synthetic oligonucleotide arrays. *Nat Genet* 21, 20-24.

Little, M.P., and Muirhead, C.R. (2000). Derivation of low-dose extrapolation factors from analysis of curvature in the cancer incidence dose response in Japanese atomic bomb survivors. *Int J Radiat Biol* 76, 939-953.

Lou, X.J., Schena, M., Horrigan, F.T., Lawn, R.M., and Davis, R.W. (2001). Expression monitoring using cDNA microarrays. A general protocol. *Methods Mol Biol* 175, 323-340.

Matys, V., Fricke, E., Geffers, R., Gossling, E., Haubrock, M., Hehl, R., Hornischer, K., Karas, D., Kel, A.E., Kel-Margoulis, O.V., *et al.* (2003). TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res* 31, 374-378.

McDonald, J.T., Briggs, C., Szelag, H., Peluso, M., Schneider, D., Perepletchikov, A., Klement, G.L., Tuerk, I., and Hlatky, L. (2014). Chronic low dose-rate radiation down-regulates transcription related to mitosis and chromosomal movement similar to acute high dose in prostate cells. *Int J Radiat Biol* 90, 231-240.

Mothersill, C., and Seymour, C.B. (1998). Cell-cell contact during gamma irradiation is not required to induce a bystander effect in normal human keratinocytes: evidence for release during irradiation of a signal controlling survival into the medium. *Radiat Res* 149, 256-262.

Muroya, Y., Plante, I., Azzam, E.I., Meesungnoen, J., Katsumura, Y., and Jay-Gerin, J.P. (2006). High-LET ion radiolysis of water: visualization of the formation and evolution of ion tracks and relevance to the radiation-induced bystander effect. *Radiat Res* 165, 485-491.

Narbutt, J., Cebula, B., Lesiak, A., Sysa-Jedrzejowska, A., Norval, M., Robak, T., and Smolewski, P. (2009). The effect of repeated exposures to low-dose UV radiation on the apoptosis of peripheral blood mononuclear cells. *Arch Dermatol* 145, 133-138.

Nikitaki, Z., Nikolov, V., Mavragani, I.V., Mladenov, E., Mangelis, A., Laskaratou, D.A., Fragkoulis, G.I., Hellweg, C.E., Martin, O.A., Emfietzoglou, D., *et al.* (2016). Measurement of complex DNA damage induction and repair in human cellular systems after exposure to ionizing radiations of varying linear energy transfer (LET). *Free Radic Res* 50, S64-S78.

Niu, S., Deboodt, P., Zeeb, H., International Atomic Energy Agency., International Labour Office., and World Health Organization. (2010). Approaches to attribution of detrimental health effects to occupational ionizing radiation exposure and their application in compensation programmes for cancer : a practical guide (Vienna: International Atomic Energy Agency ; Geneva : International Labour Office : International Health Organization).

Okoniewski, M.J., and Miller, C.J. (2008). Comprehensive analysis of affymetrix exon arrays using BioConductor. *PLoS Comput Biol* 4, e6.

Olivieri, G., Bodycote, J., and Wolff, S. (1984). Adaptive response of human lymphocytes to low concentrations of radioactive thymidine. *Science* 223, 594-597.

Pitman, S.C. (1997). Radiology imaging words and phrases : diagnostic imaging, interventional radiology, therapeutic radiology, nuclear medicine, neuroradiology, ultrasonography, computed tomography, magnetic resonance imaging (Modesto, Calif.: Health Professions Institute).

Pollack, J.R., Perou, C.M., Alizadeh, A.A., Eisen, M.B., Pergamenschikov, A., Williams, C.F., Jeffrey, S.S., Botstein, D., and Brown, P.O. (1999). Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* 23, 41-46.

Povey, S., Lovering, R., Bruford, E., Wright, M., Lush, M., and Wain, H. (2001). The HUGO Gene Nomenclature Committee (HGNC). *Hum Genet* 109, 678-680.

Rahmanian, S., Taleei, R., and Nikjoo, H. (2014). Radiation induced base excision repair (BER): a mechanistic mathematical approach. *DNA Repair (Amst)* 22, 89-103.

Ray, M., Yunis, R., Chen, X., and Rocke, D.M. (2012). Comparison of low and high dose ionising radiation using topological analysis of gene coexpression networks. *BMC Genomics* 13, 190.

Rodriguez, R., and Meuth, M. (2006). Chk1 and p21 cooperate to prevent apoptosis during DNA replication fork stress. *Mol Biol Cell* 17, 402-412.

Rothkamm, K., and Lobrich, M. (2003). Evidence for a lack of DNA double-strand break repair in human cells exposed to very low x-ray doses. *Proc Natl Acad Sci U S A* 100, 5057-5062.

Safran, M., Dalah, I., Alexander, J., Rosen, N., Iny Stein, T., Shmoish, M., Nativ, N., Bahir, I., Doniger, T., Krug, H., *et al.* (2010). GeneCards Version 3: the human gene integrator. Database (Oxford) 2010, baq020.

Schena, M., Shalon, D., Heller, R., Chai, A., Brown, P.O., and Davis, R.W. (1996). Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc Natl Acad Sci U S A* 93, 10614-10619.

- Schottenfeld, D., and Fraumeni, J.F. (2006). *Cancer epidemiology and prevention*, 3rd edn (Oxford ; New York: Oxford University Press).
- Seeram, E. (1997). *Computed tomography : a study guide and review* (Philadelphia: W.B. Saunders Co.).
- Shen, R., Fan, J.B., Campbell, D., Chang, W., Chen, J., Doucet, D., Yeakley, J., Bibikova, M., Wickham Garcia, E., McBride, C., *et al.* (2005). High-throughput SNP genotyping on universal bead arrays. *Mutat Res* 573, 70-82.
- Šidák, Z. (1967). Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. *Journal of the American Statistical Association* 62, 626-633.
- Simon, G. (1967). *X-ray diagnosis for clinical students and practitioners*, 2nd edn (London: Butterworths).
- Singleton, B.K., Griffin, C.S., and Thacker, J. (2002). Clustered DNA damage leads to complex genetic changes in irradiated human cells. *Cancer Res* 62, 6263-6269.
- Smith, G.M. (2010). What is a low dose? Response to "Reply to 'The RBE of low-LET radiations' " Reply to the Response to "Reply to 'The RBE of low-LET radiations' " Comment on 'Updated estimates of the proportion of childhood leukaemia incidence in Great Britain that may be caused by natural background ionising radiation' Reply to "Comment on 'Updated estimates of the proportion of childhood leukaemia incidence in Great Britain that may be caused by natural background ionising radiation' ". *J Radiol Prot* 30, 93-101; author reply 199-101.
- Smith, J., Mowla, S., and Prince, S. (2011). Basal transcription of the human TBX3 gene, a key developmental regulator which is overexpressed in several cancers, requires functional NF-Y and Sp1 sites. *Gene* 486, 41-46.
- Smyth, G.K., and Speed, T. (2003). Normalization of cDNA microarray data. *Methods* 31, 265-273.
- Stekel, D.J., Sarti, D., Trevino, V., Zhang, L., Salmon, M., Buckley, C.D., Stevens, M., Pallen, M.J., Penn, C., and Falciani, F. (2005). Analysis of host response to bacterial infection using error model based gene expression microarray experiments. *Nucleic Acids Res* 33, e53.
- Swenberg, C.E., Horneck, G., and Stassinopoulos, E.G. (1993). *Biological effects and physics of solar and galactic cosmic radiation / edited by Charles E. Swenberg, Gerda Horneck, and E.G. Stassinopoulos* (New York: Plenum Press).
- Takeyama, K., Aguiar, R.C., Gu, L., He, C., Freeman, G.J., Kutok, J.L., Aster, J.C., and Shipp, M.A. (2003). The BAL-binding protein BBAP and related Deltex family members exhibit ubiquitin-protein isopeptide ligase activity. *J Biol Chem* 278, 21930-21937.
- Taleei, R., Girard, P.M., and Nikjoo, H. (2015). DSB repair model for mammalian cells in early S and G1 phases of the cell cycle: application to damage induced by ionizing radiation of different quality. *Mutat Res Genet Toxicol Environ Mutagen* 779, 5-14.
- Trott, K.R., and Rosemann, M. (2000). Molecular mechanisms of radiation carcinogenesis and the linear, non-threshold dose response model of radiation risk estimation. *Radiat Environ Biophys* 39, 79-87.
- United Nations. Scientific Committee on the Effects of Atomic Radiation. (1988). *Sources, effects and risks of ionizing radiation : 1988 report to the General Assembly, with annexes* (New York: United Nations).
- Urbach, F., and American Society for Photobiology. (1992). *Biological responses to ultraviolet A radiation : a symposium* (Overland Park, Kan.: Valdenmar Pub. Co.).
- Vladimirov, V.G. (1989). [The biological effects in external exposure to low doses of ionizing radiation]. *Voen Med Zh*, 44-46.
- von Mering, C., Jensen, L.J., Snel, B., Hooper, S.D., Krupp, M., Foglierini, M., Jouffre, N., Huynen, M.A., and Bork, P. (2005). STRING: known and predicted protein-protein associations, integrated and transferred across organisms. *Nucleic Acids Res* 33, D433-437.
- Wang, J., Vasaikar, S., Shi, Z., Greer, M., and Zhang, B. (2017). WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res*.
- Wang, Y., Putnam, C.D., Kane, M.F., Zhang, W., Edelman, L., Russell, R., Carrion, D.V., Chin, L., Kucherlapati, R., Kolodner, R.D., *et al.* (2005). Mutation in Rpa1 results in defective DNA double-strand break repair, chromosomal instability and cancer in mice. *Nat Genet* 37, 750-755.
- WHO Expert Committee on the Use of Ionizing Radiation and Radionuclides on Human Beings for Medical Research Training, a.N.P.-W.H.O. (1977). *Use of ionizing radiation and radionuclides on human beings for medical research, training, and nonmedical purposes : report of a WHO expert committee [meeting held in Geneva from 1 to 8 March 1977]* (Geneva: World Health Organization).
- Yang, X., Zeng, Z., Hou, Y., Yuan, T., Gao, C., Jia, W., Yi, X., and Liu, M. (2014). MicroRNA-92a as a potential biomarker in diagnosis of colorectal cancer: a systematic review and meta-analysis. *PLoS One* 9, e88745.
- Yap, G. (2002). Affymetrix, Inc. *Pharmacogenomics* 3, 709-711.

Zhang, B., Kirov, S., and Snoddy, J. (2005). WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res* 33, W741-748.